

CSNDSP 2004

School of Electrical, Electronic and Computer Engineering,
University of Newcastle upon Tyne

Steering Committee

Prof R J Clarke	Heriot-Watt University	<i>Chairman</i>
Prof Z Ghassemlooy	Northumbria University	<i>Secretary</i>
Prof A C Boucouvalas	Bournemouth University	
Dr S S Dlay	Newcastle University	<i>Local Organiser</i>

Local Organising Committee

Dr S S Dlay *Chair*

Prof O R Hinton

Dr W L Woo

Mr S Grey

ISBN: **0-7017-0177-3**

National/International Technical Committee

Prof. H Aghvami	Kings College, London, UK
Prof. M Al-Akaidi	De Montfort Univ, UK
Prof. P Ball	Oxford Brooks Univ. UK
Dr B Bashiri	Southampton Univ. UK
Dr M Benaissa	Sheffield Univ. UK
Prof. J L Bihan	Ecole Nation. D'Ingen. De Brest, France
Dr K Brown	Heriot-Watt Univ. UK
Prof. R A Carrasco	Staffordshire Univ. UK
Prof. F Castanie	INPT, France
Dr L Chao	Nanyang Tech. Univ. Singapore
Dr M Connelly	Univ. of Limerick, Ireland
Prof. A Constantinides	Imperial College, UK
Dr S Dlay	Newcastle Univ. UK
Prof. P G Farrell	Lancaster Univ. UK
Dr. Shan Fu	Cranfield Univ. UK
Dr A Gonzalez	Univ. of Valencia, Spain
Prof. L Izzo	Univ. of Napoli, Italy
Prof. K Liu	Reading Univ. UK
Prof. M D Logothetis	Univ. of Patras, Greece
Prof. H Mouftah	Univ. of Ottawa, Canada
Dr J A Robinson	Univ. of York, UK
Dr D Rovira	INPT, France
Prof. J Senior	The Univ. of Hertfordshire, UK
Dr D Smith	Northumbria Univ. UK
Dr J Sodha	University of the West Indies, Barbados
Dr I Soto	Santiago Univ. Chile
Dr U Speidel	Univ. of Auckland, New Zealand
Dr L Stergioulas	Lancaster Univ. UK
Prof. R T Valadas	Univ. of Aveiro, Portugal
Dr. V Vitsas	TEI Thessaloniki, Greece
Prof. M Zervas	Southampton Univ. UK

Contents

Complete Papers by Topic

Communications

Networks

Digital Signal Processing

Circuits

Listing of Papers by Author

CSNDSP 2004

Communications

Title of Paper	Author	
A Comparative Study of 1x1, 2x1 And 2x2 MIMO-OFDM Systems Based on IEE	Mr Micheal Drieberg	1
A Comparison of Complexity and Performance of Turbo Codes over Rayleigh Fadi	Mr Christopher Dennett	5
A Family of Irregular LDPC Codes Constructed Based on Optical Orthogonal Cod	Miss Hong Wen	10
A low complexity joint Huffman and Convolutional decoder employing state count	Dr K M Sunjiv Soyjaudah	14
A Modified Niederreiter Public Key Cryptosystem for Combined Coding and Cry	Mr Sachidananda Govinda	18
A NEW APPROACH TO PREDICTIVE POWER CONTROL OF WCDMA SY	Miss Neslihan Guven	23
A New Non Data-Aided Frequency Offset Recovery Technique for OFDM Syste	Dr A. Al-Dweik	27
A New Robust Frequency-Domain Decision Feedback Equaliser for DMTModem	Mr Daniel Franklin	31
A Novel Detection Technique for Uplink Power Control	Mr Kostas Kastamonitis	35
A Simple Efficient Adaptive Algorithm For Jitter Reduction in TDM Over Packet	Dr Liu Erwu	39
Adaptive MF-TDMA: Burst Length Control as a Rain Fade Countermeasure	Miss Eleni Noussi	43
Adaptive MMSE Receiver Structures, Employing Spread Spectrum and Frequency	Mr Dikaos Konstantakos	47
Adaptive PN Code Acquisition in DS-CDMA Systems	Mr Job Akpofure Obiebi	51
An Adaptive HARQ-II Protocol using RCPC Codes	Prof Jin-Fu Chang	55
Analysis of Boundary Conditions for Real-time Facsimile Transmission over UDP/	Mr Bhumin Pathak	59
Analysis of HARQ Error Control Scheme for TCP Applications based on ATM o	Mr Fabiano Bizinelli	63
Bandwidth Scheduling Techniques for Differentiated Services Support in Ethernet	Mr Dawid Nowak	67
BER Improvement in the Zero-IF Receivers	Mr Vahan Nahapetyan	71
Billing solutions for 3G data services	Dr Peter Ball	75

Title of Paper	Author	
Blind Adaptive Multiuser Detection Based on Subspace Tracking and Kalman Filte	Mr Hongwei Zhou	79
Channel SNR Estimate for Turbo Decoding	Dr Janak Sodha	83
Constellation Shaping for Turbo-Coded Modulation	Dr Stephane Yves Le Goff	87
Constructing Finite Pseudo-Random Strings using a Codeset-Based Entropy Meas	Dr Ulrich Speidel	91
Construction of a Convolutional ElGamal Elliptic Curve Cryptosystem	Mr B Ontiveros	95
Decoding with Arbitrary Pulse-Shape in Optical PPM Communications	Dr Chuanyun Zou	99
Design of a digital BPSK demodulator using Field Programmable Gate Arrays	Mr Faruque Ahamed	103
Dual semiconductor optical amplifier polarisation independent wavelength conversi	Dr David Ian Forsyth	107
Effects of the Prefilter Bandwidth on Symbol Synchronizers	Mr Antonio Reis	111
Estimating The Position and velocity of Mobiles with Neural Network	Dr Mohamad Reza Sabagh	115
Expurgated Tangential Bound of Low Density Parity Check	Mr Vu Duc Ngo	120
Iterative turbo decoder design with convolutional interleavers.	Mr Sina Vafi	124
Low Complexity Residual Phase Tracking Algorithm for OFDM-based WLAN Sy	Mr Suvra Sekhar Das	128
lower order based adaptive receiver for ds-cdma communication system	Dr Shihab Jimaa	132
MODULATION AND PROTOCOL RECOGNITION IN COMMUNICATION	Mr Ali Reza Attar	136
Multifunctional unit for Rayleigh Fading Channels combining Huffman,Convolutio	Dr Sanjeev Soyjaudah	140
Multilevel HARQ on Wireless Channels	Mr Qinglin Luo	144
New Accurate Upperbounds for Maximum Likelihood Decoder of STBC	Dr Vahid Meghdadi	148
New SLA Creation and Optimal Resource Management	Dr Sreko Krile	152
Noise and Crosstalk Analysis of SMZ Switches	Mr Razali Ngah	156
OFDM Peak-to-Average Power Reduction Scheme with Spectral Masking	Mr Himel Asanga Surawe	160
Optical Fibre Refractive Index Reconstruction from Mode Near Field	Mr Xin Qian	164

Title of Paper	Author	
Optimum noise thresholds in decision directed impulse noise mitigation for OFDM	Dr Jean Armstrong	168
Performance Evaluation of Medium Dependence of Turbo Codes	Dr Seema Verma	172
Performance Evaluation of MRC for MC-CDMA Communication Systems over N	Prof Joy Iong-Zong Chen	176
Performance of Adaptive Fuzzy-Based Partial Parallel Interference Canceller for M	Dr Yung-Fa Huang	180
Performance of Authentication and Encryption algorithms in IPSec Tunnelling	Mr Siraj Shaikh	184
Performance of Quantized Turbo Decoding for Decoders	Dr hasan omar	188
Pre-Post-Rake Schemes for DS/SS Mobile Communications	Prof Gaetano Giunta	195
QoS-based Call Admission Control using Multi-code for Voice/Data Traffic in CD	Dr Fariborz Mahmoudi	199
ROBUST BLIND CHANNEL ESTIMATION FOR SPACE-TIME CODED SYS	Mr Stylianos Manioudaki	203
Simulating Heterogeneously Timed Networks in Network Simulator NS	Mr Guang-Yeu Luo	207
Simulation of All-Optical Recirculating Fibre Loop Buffers employing Soliton Puls	Mr Adrian Als	212
Simulation Study of WFQ Packet Scheduling Mechanism for different coding sche	Mr Colin Chai	217
Super-Orthogonal Space-Time Trellis Codes together with Blind Channel Identifica	Prof CANCES Jean-Pierre	221
The combination of hyperelliptic curves and block codes for data encryption	Mr I Jirón	225
The Effect of Timing and Fequency Offsets on OFDM Systems Using One and T	Dr A. Al-Dweik	229
The Evolution of a Wireless RFID and a Cell Phone	Mrs Katariina Maaria Pen	233
The use of clustering methods for combining equalization and decoding in the GSM	Mrs Adina Burian	237
Transmission of M-ary Digital Signals Using Controlled Projective Synchronisatio	Mr Chin Yi Chee	241
Turbo Trellis Coded Modulation with modified transition metric for Rayleigh fadin	Mr Kostas Koutsouvelis	245
Using Invisible Region Wave Vectors for Determining the Properties of Microwave	Dr Dave Smith	248
Wireless Communication System Performance Enhancement Using Adaptive Array	Mr Amin Hamdallah Al-K	252

CSNDSP 2004

Networks

Title of Paper	Author	
A comparative study of combinations of concealment techniques with error correct	Dr K M Sunjiv Soyjaudah	257
A Fast and Accurate Intrusion Detection System	Ms Hashem Davarpanah	261
A Method for Upper Bounding on Growth of Network Access Speed	Mr Thomas Phillip Knuds	267
A Multi-constrained QoS Aware Scheduler for Class-based IP Networks	Prof Paulo Carvalho	271
A New Architecture and MAC Layer for Quality of Service Support in Wireless H	Mr Robert Sizeland	275
A new model for the probability distributions of crosstalk in twisted pair cables	Prof Nils Holte	279
A New Proposed Burstness Factor For Real-time Traffic Pricing across the Interne	Dr Hussein Eissa	283
A New Traffic Engineering Approach for IP Networks	Miss Ljiljana Adamovic	287
A Novel MAC Protocol and Traffic Scheduling for HFC Network with QoS Supp	Mr Char-Long Chen	294
A Prediction Method of Self-Similar Network Traffic Based on the Covariation Ort	Dr Ge Xiaohu	298
A Size-limited Cluster Maintenance Technique for Mobile Ad-hoc Networks	Miss Gayathri Venkatara	302
An Energy-Ecient Clusterhead Assignment Scheme for Hierarchical Wireless Senso	Dr Dajin Wang	306
An Improved Jitter Time-stamp Method for Source Timing Recovery in AAL2	Dr Jacqueline Walker	310
Analysis and Evaluation of Traceback Schemes for Distributed Denial of Service A	Mr Fawad Abbas	314
Bluetooth Interference Suppression in IEEE 802.11b	Mr Erik Björnemo	318
Congestion control of AQM routers by an improved PID controller	Dr Yongji Wang	322
Dynamic Resource Control for QoS-Aware Multimedia Wireless Networks	Dr Shaokai Yu	326
e-learning promotes students' deep learning:	Dr M J Lopez	330
Failure of Adaptive RED in Dynamic Networks	Mr Martin Fricker	334

Title of Paper	Author	
Generalized Max-Min Fairness Policy for Elastic Traffic based on Linear Program	Dr Michael logothetis	338
Gigabit Ethernet over OC-48 Transponder for Metro-Optical Access	Mr Po-Chun Chiang	342
Graphical Simulation Tool for Mobile Networks Routing	Mrs Kamlesh Dutta	346
Graphical User Interface to access RMON probe in Java	Mr Peter Vial	351
IEEE 802.11 WLANs: Performance analysis in presence of bit errors	Mr Periklis Chatzimisios	355
Improving a Simulated Annealing Approach for Network Optimisation using Inexa	Mr Gilberto Flores Lucio	359
Improving performance through optimization of the RTS/CTS mechanism in IEEE	Mr Periklis Chatzimisios	363
Mobility of Wireless Networks for M-Business	Mr Julius, Nkenchior Agu	367
Modelling IrDA Transport Layer TinyTP	Mr Pi Huang	371
Node Discovery and Self Initialization Protocol for Distributed Underwater Acous	Mr Al-Khalid Othman	375
Novel Rotated Channel Assignment Schemes for Non-Line-of-Sight Fixed Wireless	Mr Konstantinos Ntagkou	379
Offering Broadband services via the Telephone system in a Volatile Economic Envi	Mr Alberto Fernandes	383
Optimization of OFDM modulation for Broadband Power Line Communication Im	Prof Xiao-rong Cheng	387
OverTester: A Peer-to-Peer, Modular Toolkit for Network Performance Evaluation	Prof Theodore Apostolop	392
Packet Scheduling Configuration in Proteo Network-on-Chip	Mr David Antonio Sigüen	396
Performance Investigation of a Broadband OFDM System Incorporating Selection	Mr EMMANOUIL KAL	400
Proposal and Evaluation of Dynamic Class Assignment Method: On Simplified On	Mr Kenta Yasukawa	404
Provision of Virtual Line-switched Service over DWDM with PFTS.	Mr Dengyuan Xu	408
Sigmoid Function Based Dynamic Threshold Scheme for Shared-Buffer Switches	Mr Boran Gazi	412
The Antnet Routing Algorithm – A Modified Version	Mr Firat Tekiner	416
The effect of using multiple ARs in MANETs for global connectivity	Mr Vinh Dien Hoang	421
UNIVERSAL QUADRATURE BROADCAST MODULATOR WITH SEVERA	Mr Marcin Nowakowski	425

Title of Paper	Author	
Using Mobile Agents for Data Collection, Data Correlation and Response in an Int	Mr Nattapon Chatprecha	429
Virtual Transmission Based MAC Protocol for Wireless Access	Mr Tamás Radvánszki	433

CSNDSP 2004

Digital Signal Processing

Title of Paper	Author	
A design of softbots for intelligent, reactive intrusion detection systems	Mr Islam Mohamed Hega	437
A Low-Complexity Hybrid FIR/IIR Algorithm for Multi-Channel Line Echo Cance	Dr Amjad Luna	441
A minutiae extraction fingerprint algorithm	Mr Panagiotis Chrisantho	445
A New Image Compression Technique Based on Combining Artificial Neural Netw	Prof Peter E. William	448
A New Method for Linear-Phase FIR Filter Designing with Desired Shape	Dr k Mohamed-pour	452
A New Search Algorithm for Vector Quantization	Mr Seyed Tabatabaian	458
A Novel Fast Search Algorithm For MPEG Standard	Mr Somphob Soongsathit	462
A survey paper on regression tusting techniques	Dr Ajay Rana	466
An efficient codebook design for Speaker Dependent Coding System	Mr Jakkree Srinonchat	471
Analysis of the IrFM Digital Payment System	Mr Pi Huang	474
Arbitrarily Shaped Video Object Compression with Weighted Finite Automata	Mr shumu li	479
ARTIFICIAL NEURAL NETWORK ASSESSMENT OF THE PROGNOSTIC V	Mr Stuart Grey	483
Biometric User Authentication	Mr Raghav Menon	487
Classification of Cancerous Cells Images using Clustered Fuzzy-Neural Machine T	Mr Ephraim Nwoye	491
Coherence-free Photonic Notch Filter	Mr Erwin Chan	495
Combined Spatial and Temporal Domain Wavelet Shrinkage Algorithm for Video D	Mr Eric Balster	499
Demodulation Schemes for Interferometric Optical Fiber Sensors	Mr Barry Griffin	503
FUZZY IMAGE SEGMENTATION CONSIDERING OBJECT SURFACE SIM	Mr Mohammad Ameer Al	507
Information Theoretic Approach to Blind Deconvolution under Impulsive Noise	Mr Chen Wei	511

Title of Paper	Author	
INTERACTIVE SPEECH RECOGNITION SYSTEMS USING COMMITTEE	Mr M SRINIVASAN	515
MORPHOLOGICAL ANALYSES OF HISTOLOGICAL IMAGES USING FRA	Mr Khaled Marghani	520
NESTED WIGNER DISTRIBUTIONS AND THEIR APPLICATION TO SIGN	Mr Saad Ahmed Qazi	524
New Method of Comparison NS-FRLS and NLMS Algorithms for the Acoustic Ec	Mr MD Arezki	528
New method to count vehicles in highway by using some pixels and threshold	Dr Heydar Toossian Shan	533
On nonuniform sampling of one-dimensional signals	Dr Krishna Busawon	536
PARAMETER ESTIMATION& DIGITAL PEAK LOCALIZATION ALGORI	Mr Muhammad Wasim R	541
Pseudo PCA for Cluster Analysis of Time-Series Data	Mr Daewon An	544
Sampling-Reconstruction Procedure of Gaussian Processes with a Finite Number o	Dr Vladimir Kazakov	548
SIGNAL SEPARATION OF NONLINEAR MIXTURES USING FUZZY-GEN	Ms LC Khor	552
Simple expressions for the locations and amplitudes of the extrema of the $(\sin x/x)^r$ f	Prof Jean Le Bihan	556
Symbol Synchronizers Operating at Sub Data Rate	Mr Antonio Reis	560
technical White paper on IDM	Dr Ajay Rana	564
Three New Versions of Numerically Stable Fast Least Squares Algorithms(NS-FR	Mr MD Arezki	576
Three-pass JPEG-LS Image Compression	Dr Mohammad Fathi	580
Wavelet Footprint Representation of ECG Signals	Dr Varun Jeoti	584
Weierstrass Approach to Nonlinear Blind Signal Separation	Ms Pei Gao	588

CSNDSP 2004

Circuits

Title of Paper	Author	
Clock Jitter in Direct Sampling Systems	Mr Theodoros Chalvatzis	592
Demonstration of 1.65 um Optical Time-Domain Reflectometry in On-Line Monit	Mr I-Yu Kuo	596
Efficient Memory Structures for Dynamic Multicast Routing Architecture	Mr Siew Kei Lam	600
Implementation of an RLS Equalizer for OFDM-based WLAN	Mr Gabriel Caffarena	604
Iris Filter: The VLSI Perspective	Dr Chai Soon Lim	608
Local Optimal Algorithm for High Performance Solution of Reconfigurable VLSI A	Dr Jigang Wu	613
Multiple Carries Prediction to Enhance the Performance of a 32-bit Estimated Carr	Mr Esmail Milad Ashmila	617

A Comparative Study of 1x1, 2x1 And 2x2 MIMO-OFDM Systems Based on IEEE802.16a

Micheal Driberg, Yew Kuan Min and Varun Jeoti

Wireless Communication Group, University Teknologi PETRONAS,
Bandar Seri Iskandar, 31750 Tronoh, Perak Darul Ridzuan, Malaysia
E-mail: varun_jeoti@petronas.com.my

Abstract – A comparison of 1x1, 2x1 and 2x2 multi-input multi-output (MIMO) – orthogonal frequency division multiplexing (OFDM) system based on IEEE802.16a, wireless Metropolitan Area Network (MAN) have been carried out by simulation of the systems in Matlab/Simulink. MIMO that refers to the use of multiple antennas at the transmitter and multiple antennas at the receiver is especially powerful in providing diversity to overcome the performance degradation due to channel fades. The model uses the OFDM physical (PHY) layer using QPSK constellations with Reed-Solomon and Convolutional Coding (RS-CC) scheme for forward error correction (FEC) at rate $\frac{3}{4}$. The model has been tested using all 6 Stanford University Interim (SUI) channels. The Space Time (ST) code used is Alamouti's ST Block Code which gives very good performance and is quite simple to implement. Simulation results show that the diversity gain at BER of 10^{-3} when using 2x1 MISO compared to 1x1 SISO is around 10dB for all SUI channels. At the same target BER, the extra diversity gain when using 2x2 MIMO compared to 2x1 MISO range from 3 to 7dB. Also, the extra SNR required for BER at 10^{-6} for all 6 SUI channels are only 5-7dB more than AWGN when using 2x2 MIMO.

Keywords – MIMO-OFDM , IEEE802.16a , ST Codes, fixed Broadband Wireless Access.

I. INTRODUCTION

IEEE802.16 is the standards body for fixed broadband wireless access system. Fixed broadband wireless access systems require high bit-rate, reliable and power efficient transmission systems. In order to reduce the ISI and to avoid using very complex equalizer at high bit rates, orthogonal frequency division multiplexing (OFDM) is used. In OFDM, a single wideband channel is divided into many overlapping but orthogonal narrowband channels [1]. This effectively lengthens the symbol time and reduces the ISI. Equalization is simple to implement in OFDM. Due to the nature of the wireless channel, which fades in time and space, reliable

communication requires a large fade margin so that the deep fades do not disrupt important communication. This limits the range that can be achieved given a specified power and may even cause the cellular deployment to become economically unfeasible. MIMO has been shown to be effective in providing diversity that lowers the fade margins needed to achieve a similar performance [2]. With diversity of MIMO, a more reliable and power efficient system is possible.

In this paper, the simulation model has been developed according to the standards specifications of IEEE 802.16a and uses the OFDM PHY layer. The modulation and coding rate is QPSK RS-CC rate $\frac{3}{4}$. The channel is modeled using the 6 SUI channels that have been adopted by IEEE802.16a to evaluate air interface performance. The rest of the paper is organized as follows: Section II introduces the OFDM system that is followed by SUI channels in section III. Section IV deals with the Alamouti's STBC used in the MIMO. Section V presents simulation results of the 1x1, 2x1 and 2x2 MIMO-OFDM system.

II. OFDM SYSTEM

In this section, we give a summary of the various blocks of the baseband OFDM system. Details of the OFDM system can be found in the IEEE802.16a standard documents [3]. The block diagram of the OFDM system is shown in Fig.1.

Binary data is first grouped together to make up the specified number of bits per frame. The processing is done frame by frame. Each bit frame is then fed into a randomizer whose function is to prevent long runs of bit '0' or bit '1'. The randomizer is implemented by XORing the input with the output of a Pseudo Random Binary Generator (PRBS). The PRBS generator polynomial is $1 + X^{14} + X^{15}$.

Next the bit frame goes through a RS Encoder as a Forward Error Correction (FEC) method to correct bit errors. For QPSK RS-CC rate $\frac{3}{4}$, the RS code used is the shortened and punctured RS(40,36,2) code derived from a systematic RS(255,239,8) code. This is followed by another FEC which is the Punctured Convolutional Encoder. The convolutional code used is the industry (171,133) rate $\frac{1}{2}$ with

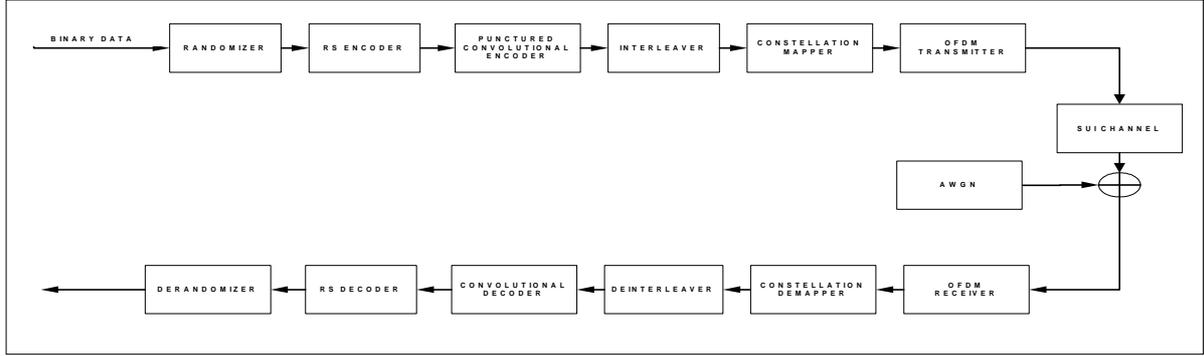


Figure 1: OFDM Baseband System Block Diagram

constraint length $K = 7$. To achieve a rate $5/6$ for the convolutional encoder, it is punctured by a puncturing pattern of (10101, 11010). This gives an overall RS-CC code rate of $3/4$

The bit frame is then fed into the interleaver which spreads out the adjacent bits onto non-adjacent carriers and ensures that adjacent bits get mapped onto less or more significant bits of the constellation to avoid long runs of lowly reliable bits. The constellation mapper maps the bit frame to the constellation specified, which is QPSK in our case. The constellations are gray mapped and normalized to have equal average power.

The symbols frame now enters the OFDM transmitter where each symbol is mapped onto the corresponding carriers which are then transformed into time domain samples by passing them through a 256 point IFFT. Only 200 carriers out of the 256 carriers are used. The unused carriers are the the DC carrier and the side carriers which provide a natural frequency decay. Out of the 200 used carriers, 8 are pilot carriers that are used for tracking purposes. The OFDM samples now undergo the SUI multipath channels with noise added as AWGN.

Equalization is performed at the receiver. The remaining blocks of the receiver side are the blocks that correspond to respective transmitter block described above.

III. SUI CHANNELS

SUI channels are a group of 6 channels that have been adopted by IEEE802.16 to evaluate air interface performance [4]. The 6 channels model the typical channels for 3 types of terrains. Terrain type A is hilly terrain with moderate-to-heavy tree densities while terrain type C is flat terrain with light tree densities. Terrain type B is intermediate between terrain type A and terrain type C. Table 1 shows the terrain type and corresponding SUI channels that represents them. Parameters of each of the SUI channel can be found in [4].

TABLE 1

Terrain Types And Corresponding SUI Channels

Terrain Type	SUI Channels
C	SUI-1, SUI-2
B	SUI-3, SUI-4
A	SUI-5, SUI-6

Each of the SUI channel is modeled as 3 taps. Each tap is simulated according to its delay and power. K-factor is the ratio of the constant component power to the random component power of each tap. We have used the omni antenna parameter and 90% K-factor. The taps are normalized by the corresponding factors to achieve average channel power of 1.

IV. ALAMOUTI'S SPACE TIME BLOCK CODE FOR MIMO SYSTEM

Alamouti presented a simple transmitter diversity scheme for 2 transmitter that improved signal quality at the receiver through frequency flat fading channels by simple processing at the transmitter and linear decoding at the receiver [5]. The scheme proposed has been shown to achieve the same performance as the classical maximal-ratio receive combining (MRRC) scheme which is widely deployed as a form of receiver diversity with a 3dB performance penalty since the total transmitter power is divided by both the transmitters. The scheme for 2 transmitters and 1 receiver is shown in Fig.2.

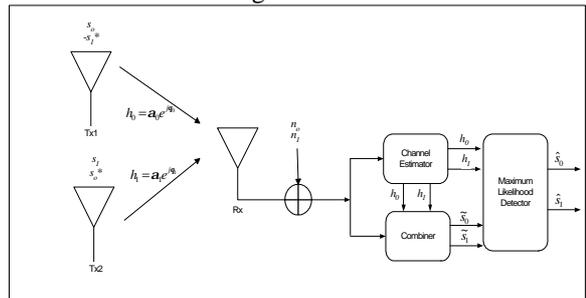


Figure 2: Alamouti's 2Tx and 1Rx STBC Scheme

At any symbol period, 2 symbols are transmitted simultaneously from both transmitters. During the first symbol period, Tx1 transmits symbol s_0 and Tx2 transmits symbol s_1 . At the next symbol period, Tx1 transmits symbol $-s_1^*$ and Tx2 transmits symbol s_0^* . At the receiver, the received signal during the first symbol period and second symbol period is r_0 and r_1 respectively. The combining and decision rule details can be found in [5].

Alamouti also showed that the above scheme achieves the diversity order of 2 which is the number of transmitter antennas times the number of receiver antennas. He also showed that the scheme can be easily extended to M receiver antennas for achieving a diversity order of 2M. For the simulation of 2x2 system, we have used the extension of the scheme for 2x2. The scheme for 2x2 system is shown in Fig.3.

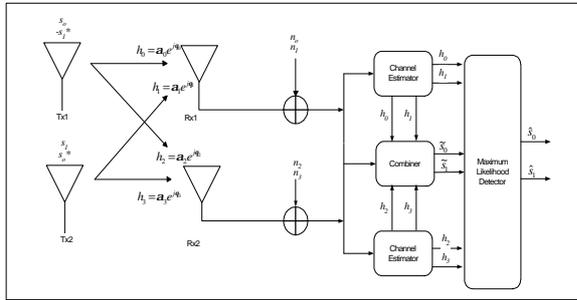


Figure 3: Alamouti's 2Tx and 2Rx STBC Scheme

In OFDM systems, the scheme is applied on each carrier for 2 consecutive symbol periods since each carrier undergoes flat fading

V. SIMULATION RESULTS OF THE 1x1, 2x1 and 2x2 MIMO OFDM SYSTEM

A simulation of MIMO-OFDM system based on IEEE802.16a have been developed in Simulink/Matlab. The model uses the OFDM PHY layer and the particular model is QPSK RS-CC rate 3/4. In this work, we have assumed perfect channel state information and perfect synchronization. We also specify the packet size to be 1kB which allow us to assume quasi-static channel. To get accurate results, the average performance of the system across 10000 independent channel realization is used which means a run of 8.4×10^7 source bits. The performance of the 1x1, 2x1 and 2x2 MIMO OFDM system is shown in Figure 4 – 9 using SUI channel 1-6.

From the perusal of the performance curves, the diversity gain at BER of 10^{-3} when using 2x1 MISO compared to 1x1 SISO is around 10dB for all SUI channels. At the same target BER, the extra diversity gain when using 2x2 MIMO compared to 2x1 MISO range from 3 to 7dB where more gain is attained from SUI-1 to SUI-6 channel due to the

lower K-factor for higher numbered SUI channels. We expect the diversity gains to be even more at higher BER for example at BER of 10^{-6} . At target BER of 10^{-6} , note that the extra SNR required is only 5 – 7dB more than AWGN when using 2x2 MIMO for all the SUI channels.

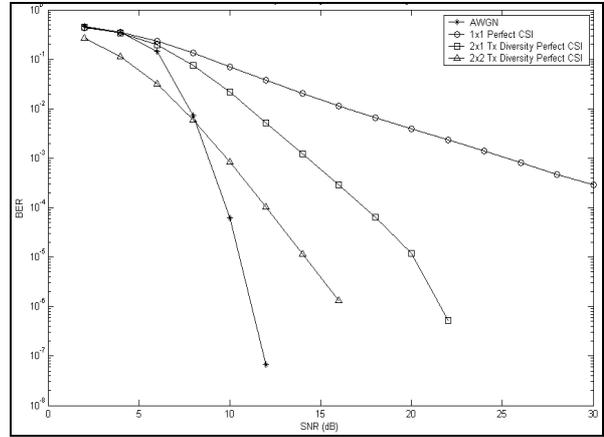


Figure 4: BER vs SNR of MIMO OFDM system under SUI-1 channel

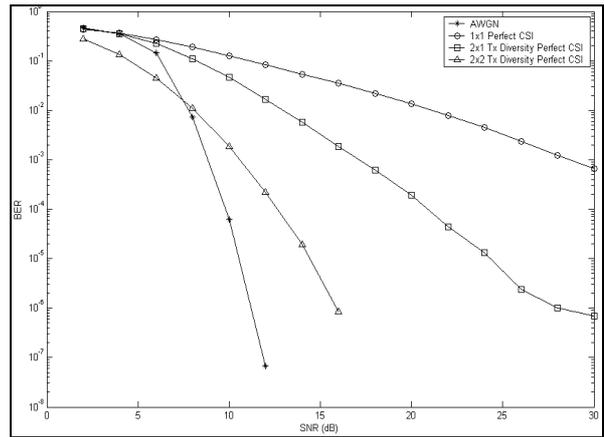


Figure 5: BER vs SNR of MIMO OFDM system under SUI-2 channel

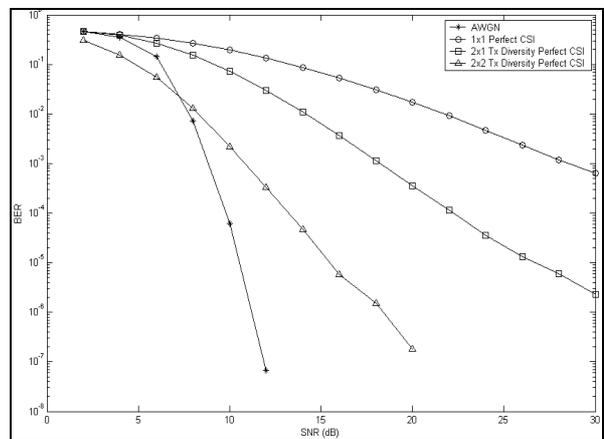


Figure 6: BER vs SNR of MIMO OFDM system under SUI-3 channel

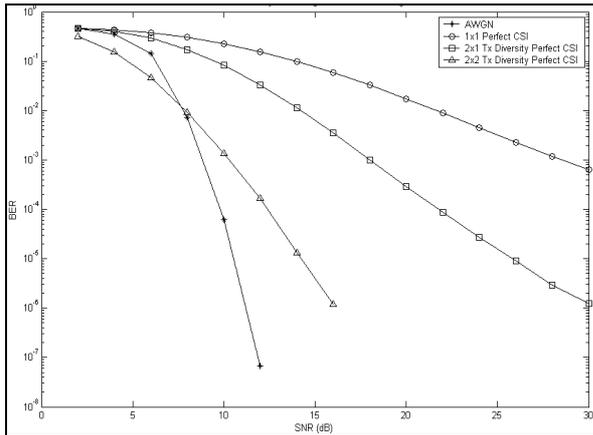


Figure 7: BER vs SNR of MIMO OFDM system under SUI-4 channel

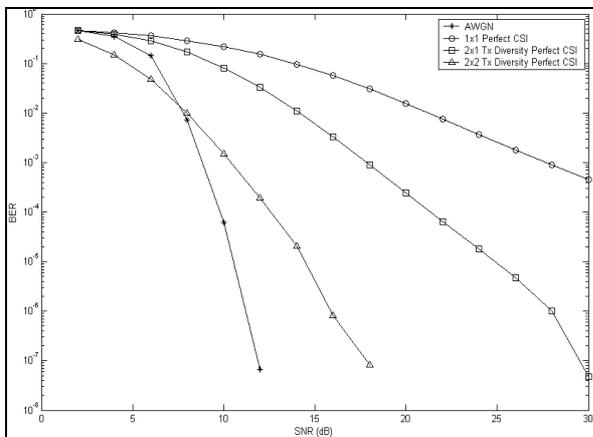


Figure 8: BER vs SNR of MIMO OFDM system under SUI-5 channel

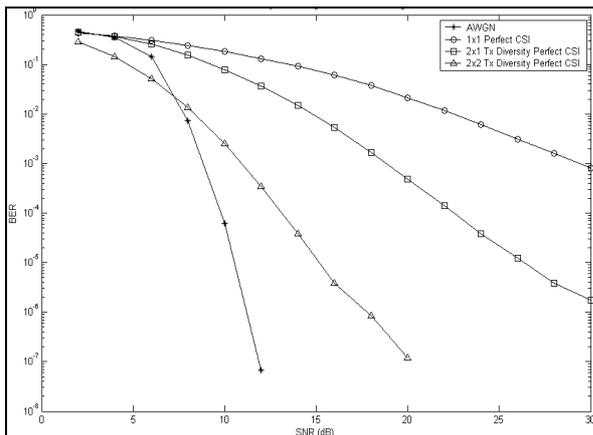


Figure 9: BER vs SNR of MIMO OFDM system under SUI-6 channel

V. CONCLUSION

A simulation of MIMO-OFDM system based on IEEE802.16a has been developed in Simulink/Matlab. We have presented the simulated performance results of the system under all the 6 SUI channels for 1x1 SISO, 2x1 MISO and 2x2 MIMO systems. It is also shown that 2x1 MISO and 2x2 MIMO can improve the performance significantly which is degraded due to channel fades. Simulation results shows that the extra SNR required for BER at 10^{-6} for all 6 SUI channels are only 5-7dB more than AWGN when using 2x2 MIMO. Particularly, using a simple STC like Alamouti's STBC is effective and easy to implement in OFDM systems where the code is applied on the same carrier for 2 consecutive symbol periods. The implementation of 2x2 MIMO is a small price to pay for the significant performance gains attained without any expansion in bandwidth.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Srikanth, AU-KBC, Chennai, Ron Murias, WiLan, Calgary and Prof. M.U. Siddiqi, MMU, Cyberjaya for helpful discussions and IRPA grant (04-99-02-0009-EA001) from the government of Malaysia for their financial sponsorship.

REFERENCES

- [1] Richard V. Nee, Ramjee Prasad, "OFDM For Wireless Multimedia Communications", Artech House, 2000
- [2] G.D. Durgin, *Space-Time Wireless Channels*. Prentice Hall, 2003
- [3] IEEE802.16a, "IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems – Medium Access Control Modifications and Additional Physical Layer Specifications for 2-11 GHz", 2003
- [4] V. Erceg et al, "Channel Models for Fixed Wireless Applications", IEEE802.16 Broadband Wireless Working Group, IEEE802.16a-03/01, Jun 2003
- [5] S.M. Alamouti, "Simple Transmit Diversity Technique for Wireless Communications", *IEEE Journal on Select Areas in Communications*, vol16, pp. 1451-1458, 1998

A Comparison of Complexity and Performance of Turbo Codes over Rayleigh Fading Channels

C. P. Dennett BEng (Hons) Prof. R. A. Carrasco
C.Dennett@wlv.ac.uk

Abstract

Turbo Codes are well documented for their excellent performance over Added White Gaussian Noise (AWGN) channels. Results very close to Shannon's limit are well documented. These performance curves are for very large frame sizes, with high numbers of iterations, using complex decoding strategies. This paper seeks to investigate the performance of less complex decoding strategies on smaller frame sizes, outlining their construction and comparing both the performance differences over Rayleigh fading channels between two of the most common strategies and the differences in operational complexity.

Introduction

Since their introduction in 1993 [1], turbo codes have been the subject of massive amounts of research, reaching to within 0.2dB of Shannon's capacity limit for Added White Gaussian Noise (AWGN) channels [2] [Berroux ref]. Their application of Soft-Input Soft-Output (SISO) algorithms to make use of *a priori* information obtained from previous component decoders results in excellent error control capabilities.

Two primary SISO decoding strategies have been commonly employed throughout the last eleven years, Maximum *A Priori* (the MAP algorithm) and the Soft-Output Viterbi Algorithm (SOVA). While it is without doubt that the former outperforms the latter (by around 0.7dB for a four state component code over an AWGN channel), it is also true to say that the complexity of MAP is also far in excess of its competitor. This led to the application of logarithms to the MAP decoding process, thus replacing expensive multiplicative operations with additions and producing the Log-MAP and Max Log-MAP variants of the algorithm. However, these processes still require at least double the number of operations of SOVA [Vucetic book ref].

The iterative nature of the turbo-decoding process, with its multiple interleaving processes and component decoder recursions brings with it

inherent delays. It should also be noted that an increase in frame size would result not only in better performance, but also in an increase in the number of memory and operations required.

This paper examines the performance of small data frame size turbo codes over Rayleigh fading channels and compares the results of the Log-MAP algorithm with those of SOVA, whilst also examining the differences in operational complexity.

The Turbo Code System

The turbo encoder uses multiple systematic convolutional component encoders, separated by interleavers in parallel concatenation. The system under investigation uses two identical encoders, as described in figure 1. The systematic quality of the component codes implies that only the first encoder's data bit need be transmitted, as all subsequent encoders' data bits are merely interleaved versions of the first.

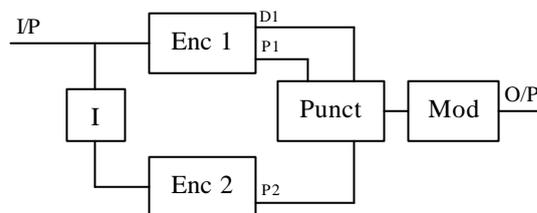


Figure 1: Turbo Encoder with two component codes

In order that the first component decoder's output is immediately of value to the second decoder, and therefore to reduce the number of iterations of the turbo decoder, tail bits are appended to the input data frame, such that the path the codeword takes through the decoder trellis ends at the all-zero state. This process is referred to as *tail biting* and the bits as *tail bits*. Obviously, the interleaver severely reduces the possibility that the second codeword will also terminate at the all-zero state.

Figure 1 also shows a puncturing mechanism, which may be applied if an increase in rate is required. The standard mechanism deletes alternate parity bits, as described in figure 2, in this

system the result is a code rate 1/2 as opposed to 1/3.

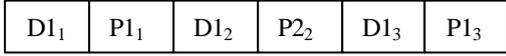


Figure 2: Puncturing Method

The serialised output is then modulated using BPSK prior to transmission.

At the receiver, a serial-to-parallel conversion is applied, to return the transmitted data stream to its three original components. Erasures are then re-inserted where puncturing has taken place, each erasure being replaced with a zero value.

Finally, before decoding can commence, the received information is *weighted*, that is, scaled with an approximation of the transmission channel's reliability. This takes the form $4 \frac{E_b}{N_0} a$,

where $\frac{E_b}{N_0}$ is the signal-to-noise ratio and a is the fading amplitude of the channel (one, for AWGN).

The structure of the turbo decoder is described in figure 3.

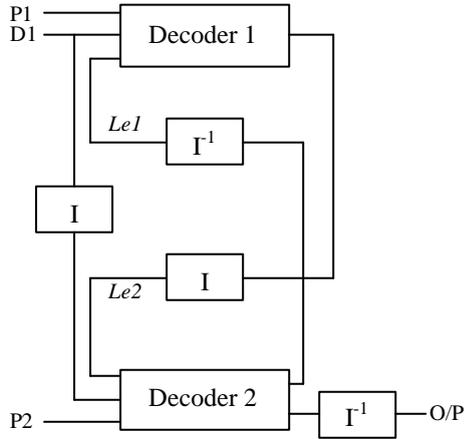


Figure 3: Turbo Decoder

For the system investigated, two component decoders in serial concatenation are necessary. The output of the component encoders is of the form $sign[L(u)]$, where the $sign$ denotes the hard output and $L(u)$ represents the reliability of this decision.

Turbo decoding begins at the first component decoder, no *a priori* information is present therefore the decoder acts solely on the received systematic data and the corresponding parity stream. Once decoding is complete, a *posteriori* information is extracted from the soft outputs, with the intention of only passing new information discovered to the next component

decoder, therefore maintaining statistical independence between bits from one iteration to the next. The information is interleaved (*Le2* in figure 3), to match the data input to the next decoder, which takes as its input the interleaved data stream, the second parity stream and the new *a priori* information. Once this decoder has terminated, a *posteriori* information is de-interleaved (*Le1*) and passed on to the first decoder, where the next iteration can begin.

Decoding Algorithms

In the interests of brevity, only the most pertinent aspects of the two decoding strategies are discussed. For a more detailed analysis, the reader is directed to [1], [3], [4].

The MAP algorithm, also known as the BCJR algorithm [5] minimises bit errors through estimation of the *a posteriori* probabilities of the individual bits in the codeword.

The turbo decoder's decision is whether $u_k = +1$ or -1 . It is taken that $u_k = +1$ if $P(u_k = +1|y) > P(u_k = -1|y)$ and $u_k = -1$ otherwise. Put more simply, it can be said that decision

$$\hat{u}_k = sign[L(u_k)] \quad (1)$$

Where

$$L(u_k) \cong \log \left(\frac{P(u_k = +1|y)}{P(u_k = -1|y)} \right) \quad (2)$$

Where y represents the received information. Re-writing (2), in terms of the code trellis gives:

$$L(u_k) = \log \left(\frac{\sum_{S^+} P(s_{k-1} = s', s_k = s, y_1^N)}{\sum_{S^-} P(s_{k-1} = s', s_k = s, y_1^N)} \right) \quad (3)$$

Where $s_k \in S$ is the state of the code at time k , S^+ is the set of ordered pairs (s', s) corresponding to all state transitions $(s_{k-1} = s') \rightarrow (s_k = s)$ brought about by the input $u_k = +1$, likewise S^- .

For Log-MAP, the final equation becomes:

$$L(u_k) = \log \frac{\sum_{s=0}^{M-1} e^{\bar{a}_{k-1}(s') + \bar{g}_k^1(s',s) + \bar{b}_k(s)}}{\sum_{s=0}^{M-1} e^{\bar{a}_{k-1}(s') + \bar{g}_k^0(s',s) + \bar{b}_k(s)}} \quad (4)$$

Where:

$$\bar{a}_k(s) = \log \sum_{s'=0}^{M_s-1} \sum_{i \in (0,1)} e^{\bar{a}_{k-1}(s') + \bar{g}_k^i(s',s)}$$

$$\bar{a}_0(0) = 0 \text{ and } \bar{a}_0(s) = -\infty, s \neq 0 \quad (5)$$

$$\bar{b}_k(s) = \log \sum_{s'=0}^{M_s-1} \sum_{i \in (0,1)} e^{\bar{b}_{k+1}(s') + \bar{g}_{k+1}^i(s,s')}$$

$$\bar{b}_N(0) = 0 \text{ and } \bar{b}_N(s) = -\infty, s \neq 0 \quad (6)$$

$$\bar{g}_k^i(s',s) = P(y_k | x_k) P(x_k | s_k = s, s_{k-1} = s')$$

$$P(s_k = s | s_{k-1} = s') \quad (7)$$

Log-MAP uses the Jacobian algorithm recursively to exactly evaluate equation (4)

The SOVA [6] is a Maximum Likelihood Sequence Estimator (MLSE). As a MLSE, the SOVA maximises the likelihood function for the whole received codeword, rather than individual bits, hence the increase in probability of error when compared to MAP.

The SOVA searches for a state sequence $S^{(m)}$, or information sequence $u^{(m)}$ that returns the maximum *a posteriori* probability $P(S^{(m)}|y)$. The system documented is for a binary trellis, therefore m is 1 (the surviving path) or 2 (the competing path).

For systematic codes, as used in turbo coding systems, this can be shown to result in the following metric calculation:

$$M_k^{(m)} = M_{k-1}^{(m)} + u_k^{(m)} L_C y_{k,1} + \sum_{j=2}^N x_{k,j}^{(m)} L_{Ck,j} y_{k,j}$$

$$+ u_k^{(m)} L(u_k) \quad (8)$$

This shows that the components of a metric at time k include the previous metric, *a priori* information, and channel reliability. Once decoding is complete, the *a priori* input and the received systematic input must be removed to provide the *a priori* information for the next decoder.

SOVA Decoder Realisation

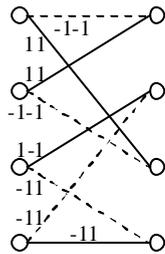


Figure 4: [7;5]₈ trellis structure

Rather than reproduce the mathematical theory of turbo codes, which as already stated are readily available, the authors have chosen to apply the theory in an attempt to further the understanding of those not entirely familiar with the process.

Beginning with a turbo code comprising two [7;5]₈ component encoders with the trellis representation given in figure 4 and a pseudo random interleaver, the data word [0110] is encoded, tailed and transmitted over an AWGN channel ($\alpha = 1$ with $E_b/N_0 = 1$). Upon reception, erasures are inserted and scaling is applied. The received data, with a transmission error at the beginning of P1, can be represented as:

Rx D1	-4	4	4	-4	4	4
Rx P1	0	0	-4	0	-4	0
Rx P2	0	4	0	4	0	4

Table 1: Data After Transmission, Erasure Insertion and Scaling (Highlighted cells = erasure insertion)

Figure 5 illustrates the component code trellis with the received information for D1 and P1 in bold above and the survivor path in bold.

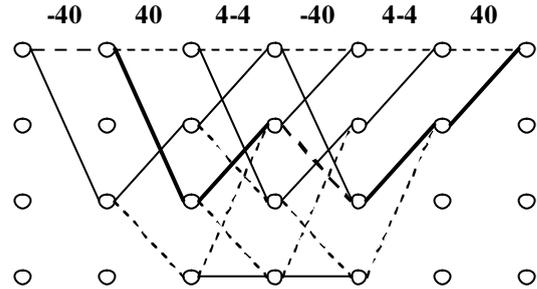


Figure 5: Trellis showing tail and survivor path

The survivor path is determined in a method similar to the standard Viterbi algorithm, according to the equation:

$$M_t = M_{t-1} + (D1_t \times D1_r) + (P1_t \times P1_r) + (D1_t \times L_e) \quad (9)$$

Where the '_t' indicates a bit corresponding to a possible transition of the code's trellis and '_r' indicates a scaled, received bit.

Table 2 gives the calculation of the survivor metrics for the first two time steps. M_{t-1} represents the metric from the previous time step.

T	M_{t-1}	Systematic	Parity	A priori	Sum
X	M_{t-1}	(trellis syst×rx syst)	(trellis par×rx par)	(trellis syst×a priori)	M_t
1	0	(-1×4)	(-1×0)	(-1×0)	4
1	0	(1×4)	(1×0)	(1×0)	-4

2	4	(-1×4)	(-1×0)	(-1×0)	0
2	4	(1×4)	(1×0)	(1×0)	8
2	-4	(1×4)	(-1×0)	(1×0)	0
2	-4	(-1×4)	(1×0)	(-1×0)	-8

Table 2: Metric Calculations for time steps 1 and 2

After completion of the forward recursion of the decoder strategy, the received trellis, with metrics can be shown to be:

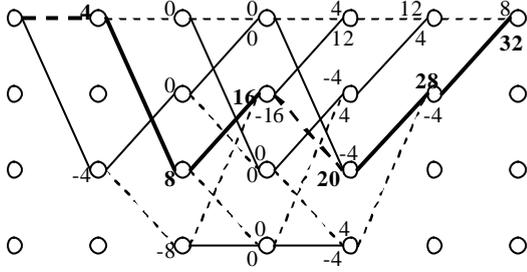


Figure 6: Received Trellis with Survivor Metrics Highlighted

To improve the reliability of the surviving path, Hagenauer *et al* [4] describe a backward recursion for SOVA. Beginning at the first time step where multiple paths merge to a state, store:

$$M_{t,surv} = 0.5|M_{t,surv} - M_{t,comp}| \quad (10)$$

Where $M_{t,surv}$ is the survivor metric and $M_{t,comp}$ is the metric for the competing path at that state. This updated value is then compared with previous survivor metrics, the lowest stored. Wherever there is a transition data bit difference between the survivor and competitor, the survivor metric is updated with this lowest value. Figure 7 shows this metric update procedure for time $t = 3$.

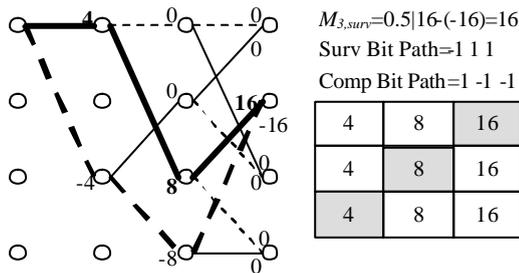


Figure 7: Update Procedure at time $t = 3$

In figure 7, the bold, dashed lines indicate the competitor path, while the bold solid lines indicate the survivor path. The table shows the update procedure. Since all transition bits differ, once the survivor metric at time $t = 3$ has been updated, the survivor metric at time $t = 2$ must be updated. The current survivor metric is lower than that stored, therefore the storage is updated to **8** and the survivor metric is updated to that. The same can be said for the survivor state at time $t = 1$.

Upon completion of this update procedure, the metric values are multiplied by the transition data bit value to give the ‘soft’ output of the decoder. From this, the received, weighted systematic information is subtracted along with any *a priori* information (note, no *a priori* information at the beginning of iteration 1):

Decoder o/p	-4	8	12	-12	12	12
Rx Systematic	-4	4	4	-4	4	4
A priori information	0	0	0	0	0	0
Result	0	4	8	-8	8	8

Table 3: Decoder Output and *A posteriori* Information

This information can now be interleaved and passed to the second component as *a priori* information. Note that the hard decisions of the decoder output are $[-1 \ 1 \ 1 \ -1 \ 1 \ 1]$, which match the input data word, plus tail bits.

Complexity

Table 4 provides a comparison of decoder complexity estimates per unit of time, based on implementation [3]. Log-MAP can be implemented using a look up table, storing eight values of $|\delta_2 - \delta_1|$, proven as optimal, with no improvement for a more defined array [7]. It is assumed that one look-up procedure is as operationally expensive as one addition is made.

Operation	Log-MAP	SOVA
Addition	$6 \cdot 2^k \cdot 2^m + 6$	$2 \cdot 2^k \cdot 2^m + 9$
Multiplication	$2^k \cdot 2^m$	$2^k \cdot 2^m$
Maximisation	$4 \cdot 2^m - 2$	$2 \cdot 2^m - 1$
Look-up	$4 \cdot 2^m - 2$	

Table 4a: Complexity Comparison Equations for (n, k) component codes of memory m

Table 5 gives the results for the turbo codes investigated here.

Component code	Log-MAP	SOVA
(2, 1), $m = 4$	354	136
(2, 1), $m = 2$	90	40

Table 4b: Number of Operations per code

Simulation and Results

Two turbo codes systems have been implemented. One uses a four state rate 1/2, constraint length $K = 3$ code, with generator $[7;5]_8$, while the other uses a 16 state, constraint length $K = 5$ code with generator $[23;33]_8$. These codes were run for a 1024 bit interleaver. The pseudo random interleaver used in all experiments was developed as suggested in [8].

The systems were simulated over a Rayleigh channel of variance 0.5, with ideal channel state information, for both the Log-MAP algorithm and SOVA. Figures 6 and 7 below show the probability of error for each code after eight decoder iterations.

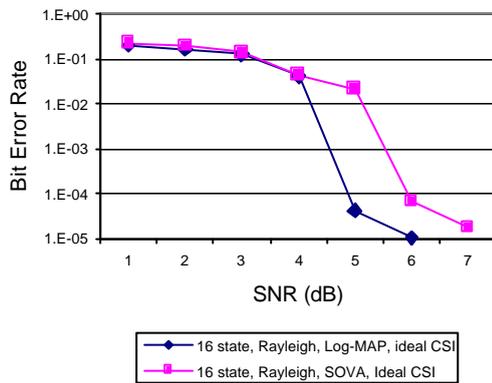


Figure 6: $[23;33]_8$ Turbo Codes over Rayleigh Fading Channel (variance 0.5), Log-MAP versus SOVA

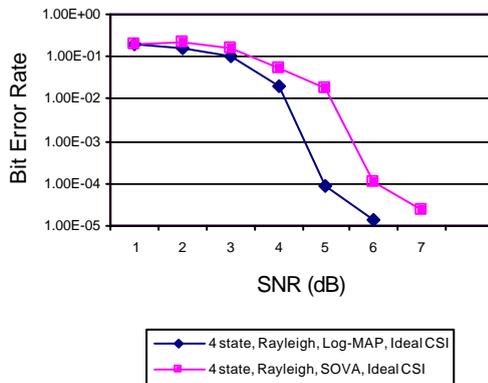


Figure 7: $[7;5]_8$ Turbo Codes over Rayleigh Fading Channel (variance 0.5), Log-MAP versus SOVA

Conclusions

The simulation results both demonstrate the superior decoding of turbo codes with the Log-MAP algorithm. The gain at $1e-4$, for instance, is in the region of 1dB. In fact, over Rayleigh channels, the gain of Log-MAP over SOVA appears to be increased when compared with the gain of Log-

MAP over SOVA when applied to AWGN channels (0.5dB to 0.75dB from simulation for equivalent BER).

On the other hand, the calculations pertaining to the complexity of each algorithm indicate that SOVA has at least half the complexity of Log-MAP.

All though it is true that further iterations will increase the gain of the decoders, study of any of the references will show that gain per iteration reduces severely after approximately eight iterations. However, given the reduced complexity of the SOVA algorithm when compared to the Log-MAP, an increase in the number of SOVA iterations may return a bit error rate equivalent to Log-MAP, while still being lower in complexity.

References

- [1] C.Berrou, A.Glavieux, P.Thitimajshima, "Near Shannon Limit Error Correcting Coding and Decoding: Turbo Codes", Proc. IEEE Int. Conf. Comms, pp. 1064-1070, Geneva, 1993
- [2] C.E.Shannon, "A Mathematical Theory of Communications", Bell System Tech. Journ., Vol.27, pp.379-423, 623-656, July, October, 1948
- [3] B.Vucetic, J.Yuan, "Turbo Codes, Principles and Applications", Kluwer Academic Publishers, ISBN 0-7923-7868-7, 2000
- [4] J.Hagenauer, E.Offer, L.Papke, "Iterative Decoding of Binary Block and Convolutional Codes", Trans. Inf. Theory, Vol. 42, No. 2, pp. 429-445
- [5] L.R.Bahl *et al* "Optimal Decoding of Linear Codes for Minimising Symbol Error Rate", IEEE trans. Inf. Theory, Vol. IT-20, pp. 248-287, March 1974
- [6] G.D.Forney, "The Viterbi Algorithm", Proc. IEEE, Vol. 61, pp. 268-278, 1973
- [7] P.Robertson, E.Villebrun, P.Hoher, "A Comparison of Optimal and Sub-Optimal MAP Decoding Algorithms Operating in the log Domain", Proc. ICC1995, Seattle 1995
- [8] A.S.Barbulescu, S.Spietrobon, "Interleaver Design for Turbo Codes", Electronic Letters, Vol. 30, pp. 2107-2108, 1994
- [9] M.C.Valenti, B.D.Woerner, "Iterative Channel Estimation and Decoding of Pilot Symbol Assisted Turbo Codes Over Flat-Fading Channels", IEEE Journ. Select. Areas Comms., Vol. 19, No. 9, pp. 1697-1705, Sept 2001

Family of Irregular LDPC Codes Constructed Based on Optical Orthogonal Codes

Hong Wen, Fei Hu, Fan Jin

*School of Computer and Communication Engineering, Southwest Jiaotong University,
Chengdu, Sichuan 610031, P. R. of China
E-mail: Sunlike@china.com*

Abstract

In this paper, we present a method for constructing irregular LDPC (low density parity check) codes based on the optical orthogonal codes. The resulting codes are quasi-cyclic codes and can be encoded with low complexity. They perform well with the sum-product iterative decoding.

Key words: LDPC codes, optical orthogonal codes, sum-product iterative decoding

1. Introduction

Low-density parity-check (LDPC) codes were discovered by Gallager [1] in 1962 and had recently been rediscovered [3], [4]. It has been shown that these codes achieve a remarkable performance with iterative decode that is very close to the Shannon limit [4]. Consequently, these codes have become strong competitions to turbo codes for error control in many communication and digital storage systems where high reliability is required.

An LDPC code is specified in terms of a sparse $N \times M$ parity-check matrix H . An (N, j, k) regular LDPC code is a code of length N having H with j 1's in each column and k 1's in each row. Recently, Luby *et al.* extended Gallager's results to consider irregular codes, that is, codes with nonconstant row and column weights in, and showed that these codes are capable of outperforming regular codes [5]. An (N, K) irregular LDPC code is a code of length N -bit and K -bit message. The irregular codes are constructed via a pseudo-random process which usually involves discarding codes which contain 4-cycles. It is known that the existence of short cycles in the Tanner graph prevents an exact error-probability analysis of iterative decoding procedures. Without 4-cycles in the Tanner graph of codes, the iterative sum-product decoding algorithm can be performed well.

The decoding of LDPC codes is linear complexity in the code length. One major criticism concerning LDPC codes has been their apparent high encoding complexity. Encoding is, in general, performed by matrix multiplication and so complexity is quadratic in the code length.

One option for efficient encoding and avoiding 4-cycle is to use algebraic code construction. In the case of regular codes a number of algebraic construction have been presented [2], [7], [8]. Less consideration has been given to structured irregular codes. The aim of this paper is to construct irregular quasi-cyclic codes free of 4-cycles using optical orthogonal codes (OOCs). Our new codes have the advantages of low encoding complexity.

2. Optical Orthogonal Codes

Definition 1[9]: An $(n, \omega, 1)$ Optical Orthogonal Code C (or briefly $(n, \omega, 1)$ -OOC) is a family of binary code vectors of length n and weight ω which satisfy the auto-correlation and cross-correlation property:

1) The Auto-correlation Property:

$$\sum_{i=0}^{n-1} x_i x_{i+\tau} \leq 1, \quad (i+\tau) \bmod n, \quad 0 < \tau < n \quad (1)$$

2) The Cross-correlation Property:

$$\sum_{i=0}^{n-1} x_i y_{i+\tau} \leq 1, \quad (i+\tau) \bmod n, \quad 0 < \tau < n \quad (2)$$

for any $\mathbf{x}=(x_0, x_1, \dots, x_{n-1}) \in C$ and $\mathbf{y}=(y_0, y_1, \dots, y_{n-1}) \in C$ with $\mathbf{x} \neq \mathbf{y}$.

A convenient way of viewing optical orthogonal codes is from a set-theoretic perspective. An $(n, \omega, 1)$ -OOC C can be alternatively considered as a collection of k -sets of integers modulo n , in which each k -set corresponds to a codeword and the numbers in each k -set specify the nonzero bits of the codeword.

Example 1: We illustrate the concept of OOCs by an example. It can be checked that

$$C = \{110010000000, 1010000100000\}$$

is an $(13, 3, 1)$ -OOC with two codewords. In set notation:

$$C = \{C_1, C_2\} \\ = \{\{0, 1, 4\}, \{0, 2, 7\}\} \pmod{13}.$$

An $(n, \omega, 1)$ is said to be optimal if it contains $\left\lfloor \frac{n-1}{k(k-1)} \right\rfloor$ codewords. The OOCs that are used to construct the LDPC in this paper are not necessary to be optimal.

Using (153,5,1)-OOC [10]:

$C_1=\{0, 3, 4, 9, 11\}$, $C_2=\{0, 14, 26, 42, 55\}$
 $C_3=\{0, 25, 48, 75, 99\}$, $C_4=\{0, 36, 70, 108, 143\}$

We can construct following irregular LDPC:

- 2/3, (459, 306) code with the column weigh distribution $\gamma=[5, 3, 2]$:

$A_1=\{0, 3, 4, 9, 11\}$, $A_2=\{0, 26, 55\}$, $A_3=\{25, 75\}$

- a rate 3/4, (612, 459) code with the column weigh distribution $\gamma=[5, 5, 3, 2]$:

$A_1=\{0, 3, 4, 9, 11\}$, $A_2=\{0, 14, 26, 42, 55\}$,
 $A_3=\{0, 48, 99\}$, $A_4=\{36, 108\}$

- a rate 4/5, (765, 612) code with the column weigh distribution $\gamma=[5, 5, 3, 3, 2]$:

$A_1=\{0, 3, 4, 9, 11\}$, $A_2=\{0, 14, 26, 42, 55\}$,
 $A_3=\{0, 48, 99\}$, $A_4=\{0, 70, 143\}$, $A_5=\{36, 108\}$

- a rate 5/6, (918, 765) code with the column weigh distribution $\gamma=[5, 5, 3, 3, 2, 2]$:

$A_1=\{0, 3, 4, 9, 11\}$, $A_2=\{0, 14, 26, 42, 55\}$,
 $A_3=\{0, 48, 99\}$, $A_4=\{0, 70, 143\}$,
 $A_5=\{25, 75\}$, $A_6=\{36, 108\}$

In an AWGN channel the iterative decoding performance of these new codes was shown in fig. 2 and 3. The maximum number of iterations is 50. The simulation results demonstrate that there are 0.15-dB and 0.40-dB performance gain to be made over the regular code and randomly constructed codes with similar parameters. Although it is not expected that the codes presented will outperform randomly constructed optimized irregular codes, they have the advantage of a reduced encoding complexity.

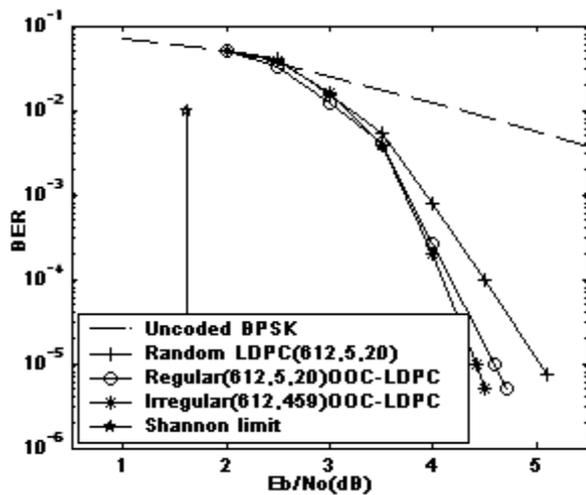


Fig. 2. The rate 3/4, (612, 458) irregular code from OOC with $\gamma=[5, 5, 3, 2]$ is compared to randomly constructed code and regular code from OOC with similar parameters.

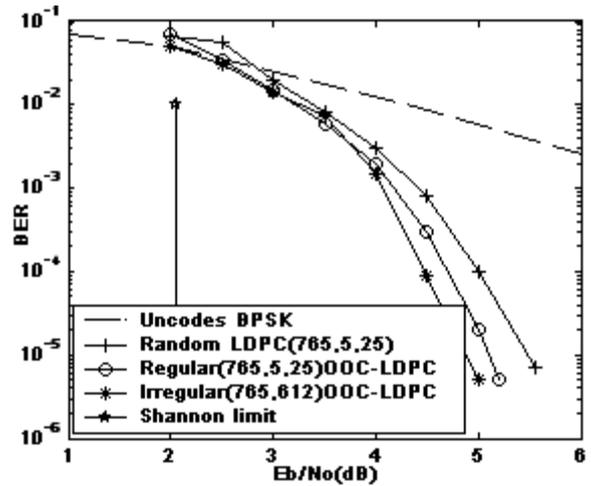


Fig. 3. The rate 4/5, (765, 612) irregular code from OOC with $\gamma=[5, 5, 3, 3, 2]$ is compared to randomly constructed code and regular code from OOC with similar parameters.

References

- [1] R. G. Gallager. Low density parity check codes [J]. IRE Trans on Information Theory, 1962; 8(1): 21 ~ 28.
- [2] D. J. C. Mackay, M. C. Davey. Evaluation of Gallager Codes for Short Block Length and High Rate Application [A]. Proc. of IMA workshop on Codes, Systems and Graphical models, New York, 2000; 113 ~ 130.
- [3] D. J. C. Mackay. Good Error Correcting Codes Based on very Sparse Matrices [J]. IEEE Trans on Information Theory, 1999; 45(2): 399 ~ 431.
- [4] D. J. C. Mackay, R. M. Neal. Near Shannon limit performance of low density parity check codes [J]. Electronic Letters, 1996; 32(18): 1645 ~ 1646.
- [5] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi and D. A. Spielman. Improved Low Density Parity Check Codes Using Irregular Graphs [J]. IEEE Trans on Information Theory, 2001; 47(2): 285 ~ 298.
- [6] S. J. Johnson, Steven R. Weller. A Family of Irregular LDPC Codes with Low Encoding Complexity [J]. IEEE Communications letters, 2003; 7(2): 79 ~ 81.
- [7] B. Ammar, B. Honary, Y. Kou and S. Lin. Construction of Low Density Parity Check Codes: A Combinatoric Design Approach [J]. Proc. of ISIT2002, Lansanne Switzerland, 2002; 311.
- [8] S. J. Johnson, Steven R. Weller. Resolvable 2-Design for Regular Low Density Parity-Check Codes [J]. IEEE Trans. On Communications, 2003; 51(9): 1413 ~ 1419.

- [9] Yan Xun Chang, Ryoh Fuji-Hara and Ying Miao. Combinatorial Constructions of Optical Orthogonal Codes with Weight 4 [A]. IEEE Trans on Information Theory, 2003 ,49 (5) : 1283 ~ 1292.
- [10] Samvel Martirosyan, A. J. Han Vinck, A Construction for Optical Orthogonal Codes with Correlation 1 [J]. IEICE Trans. Fundamentals, 2002; vol. E85-A, no. 1: 269 ~ 272

A low complexity joint Huffman and Convolutional decoder employing state counter algorithm

K.M.S SOYJAUDAH, T.P FOWDUR, M.A HOSANY

ABSTRACT

A state counter algorithm is proposed to perform joint Huffman and Convolutional decoding. An analytical comparison is also made with the algorithm developed by Honary.et.al[1]. The state counter algorithm provides a much greater reduction in complexity than the algorithm proposed in [1]. Simulation results show that there is no loss in performance as compared to the separate scheme for any of the phase shift keying modulation scheme. The only tradeoff is a slight reduction in compression ratio due to the restrictions which have to be imposed in performing Huffman coding so that the state counter algorithm can be applied.

1.INTRODUCTION

In many communication applications data is transmitted in relatively short and finite block lengths. As such the Shannon's Separation theorem[2], which vindicates the fact that optimal performance can be achieved by performing source and channel decoding separately is a questionable statement because it is valid only where sufficiently large block lengths are used. The possibility of merging channel and source decoding has therefore become an interesting scope of research in digital communications.

There are basically two ways of performing Joint Source Channel Decoding (JCS D). In the first one JCS D can be viewed as another approach to combat channel noise by exploiting the residual redundancy of the source encoder at the receiver side. The source redundancy can be utilized as a priori information of the source statistics to help decoding. In this approach, however the source and channel decoders remain separate. On the other hand, JCS D can be used to reduce decoding complexity and decoding time by literally fusing the source and channel decoders into a single joint decoder. These two approaches are illustrated in Figures 1 and 2. In this paper we propose a new algorithm to perform joint Huffman and Convolutional decoding and contrast it with the algorithm proposed in [1], henceforth referred as algorithm1.

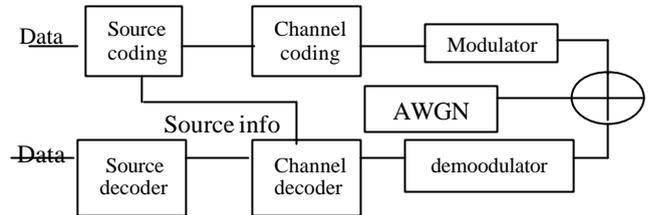


Figure1 : JCS D for performance improvement

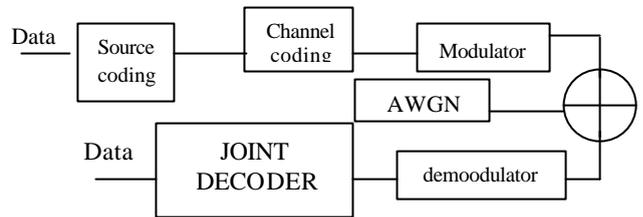


Figure2 : JCS D for complexity reduction

The two algorithms can be illustrated as shown in Figures 3 and 4. In algorithm1, the states of the Huffman decoder are incorporated into the trellis of the convolutional decoder and a joint decoder whose number of states varies according to the number states of the Huffman decoder is obtained. But in the state counter algorithm a Huffman state counter is used, which allows the joint decoder to have a fixed number of states which is independent of the number of Huffman states but which is set by the number of states in the channel decoder.

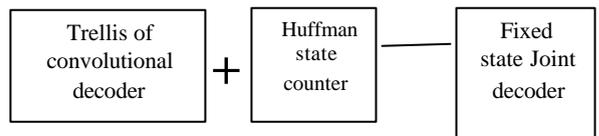


Figure 3: Illustration of algorithm1.

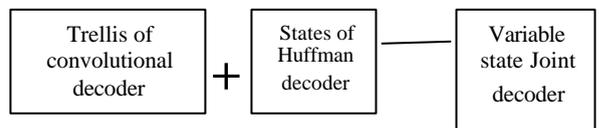


Figure 4: Illustration of state counter algorithm.

In section 2 an analysis of the Huffman algorithm is provided with emphasis on the restrictions required so that the state counter algorithm can be used. In section 3 an overview of joint Huffman and convolutional decoding using algorithm1 is provided. In section 4 the state counter algorithm is given and illustrated with an example.

In section 5 an analytical comparison between the state counter algorithm and algorithm1 is provided. Finally in section 6 the simulation results are given followed by a conclusion in section 7.

2.HUFFMAN CODING

To perform joint decoding with the state counter algorithm two constraints have to be imposed on the Huffman codewords which are as follows:

- 1.The Huffman codeword corresponding to a character with probability of occurrence p_1 must be longer than the codeword corresponding to a character with probability p_2 where $p_2 > p_1$.
- 2.All codewords apart from the codeword for the most probable and least probable character in the transmitted character set must be formed with a series of n 1's followed by one zero where $1 \leq n < K$ and K is the length of the longest codeword . Consider the character set a,b,c,d and e with probabilities 0.4,0.3,0.1,0.1and 0.1 respectively. To satisfy the above two constraints the Huffman codewords must be constructed according to the Huffman tree shown in Figure5.

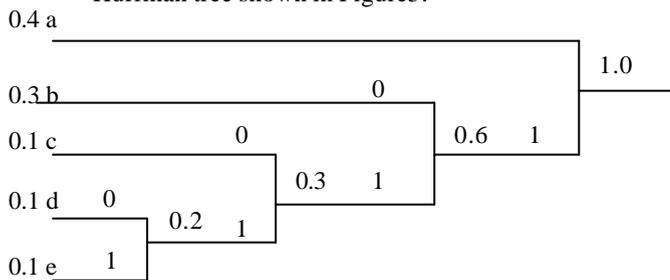


Figure5 : Huffman tree construction.

The codewords obtained are as shown in table1.

Character	Code
a	0
b	10
c	110
d	1110
e	1111

Table1 : Huffman codewords

However, it is true that using this Huffman tree construction for generating codewords, it is more likely that the code word length will be a little more, and this will reduce the compression ratio. But the advantage of generating such codewords is found in joint decoding where they can greatly reduce the decoding complexity using the state

counter algorithm which will be discussed in section4.

3.JOINT DECODING USING ALGORITHM 1

Let us now design a joint decoder for decoding the character set given in section 2 using the codewords of table 1 and following algorithm 1. The Huffman decoder state diagram is as shown in Figure6.

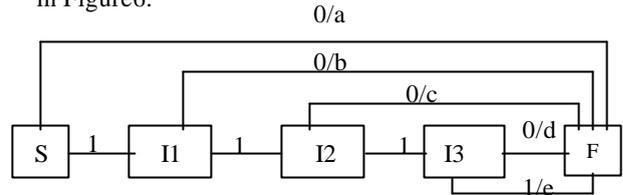


Figure 6: Huffman decoder state diagram.

Using a (2,1,2) rate $\frac{1}{2}$ convolutional encoder with generator sequences, $g_1 = [111]$ and $g_2 = [101]$ and the trellis of the trellis of the Viterbi decoder, the joint trellis obtained using algorithm1 as is shown on Figure 7.

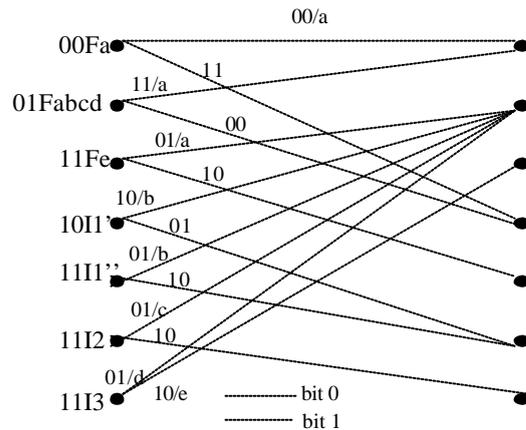


Figure 7: Joint trellis to decode character set.

4.JOINT DECODING USING STATE COUNTER ALGORITHM

The essence of the state counter algorithm is to eliminate all redundant states in a joint trellis formed by algorithm1 so that the number of states of the joint decoder is reduced to only the number of states of the channel decoder. In fact with the state counter algorithm the concept of final and intermediate states is eliminated and instead a state counter which will keep track of the sequence of bits decoded is incorporated into the channel decoder so as to enable the decoder

to determine when the codeword of a character has been detected. The trellis of a joint decoder designed with the convolutional encoder of section 3, with the state counter incorporated into it is as shown in Figure 8.

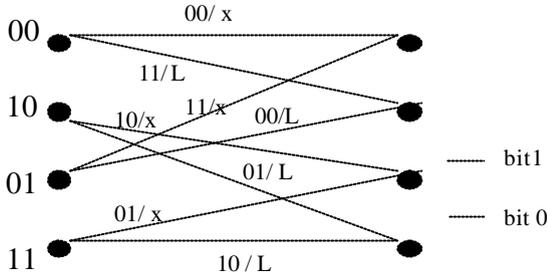


Figure 8: Joint decoder with state counter.

Assuming that at the decoder side an ordered character array is available for the whole character set transmitted in which the most probable character is at position number 1 and the least probable character is at position number L. Then, 00/x means that the character whose position number is given by x is decoded for an input bit 0 where $1 \leq x < L$. 10/L or 01/L means the character at the last position in the array is decoded for an input bit 1 but only when the position number is equal to L, otherwise no character is decoded for an input bit 1.

State counter algorithm :

1. Initialize count = 1.
2. Initialize L = length of longest Huffman codeword + 1.
3. Repeat steps 4, 5 and 6,
4. Determine trellis branch with minimum metric at time t .
5. If decoded bit = 0 ,
 - i) decode character at Position = count .
 - ii) reset count to 1.
6. If decoded bit = 1,
 - i) Increment count by 1.
 - ii) If count = L,
 - decode character at last position
 - reset count to 1.
7. Until all bits have been decoded.

Let us illustrate the principle of this algorithm by considering how the 5 characters of the character set given in section 2 are decoded. First of all if we consider the joint trellis of Figure 7, we can easily notice that the states Fe, I1', I2 and I3, all occur at the same channel decoder state, 11. Hence these states are redundant and can therefore be merged into a single state without affecting the performance of the decoder. We will therefore be left with only 4

states and to do joint decoding we will use a trellis similar to Figure 8. Normally at the decoder side a table in which the characters would be arranged as shown in Table2 will be available, in which the most probable character i.e 'a' is at Position1 and the least probable character i.e 'e' is at Position5.

Position	Character
1	a
2	b
3	c
4	d
5	e

Table2

The joint decoding algorithm using state counter proceeds as follows :

1. Initialize a variable, Position_number to 1 and a variable, Max_length to the length of the longest codeword in the character set + 1 which in this case is $4 + 1 = 5$.
2. Determine the most probable trellis transition at time t using the trellis of the decoder.
3. If the most probable trellis transition at time t corresponds to an input bit 0 then decode the character at Position = Position_number from the table and reset Position_number to 1. Thus if Position_number = 1, decode the character 'a' at time t and if Position_number = 2 , decode character 'b' and so on, but only the character 'e' will not be decoded because its codeword does not end with a 0.
4. If the most probable trellis transition at time t corresponds to an input bit 1:
 - a. increment Position_number by 1
 - b. If Position_number = Max_length then decode the character at Position = Position_number from the table (character 'e' in this case) and reset Position_number to 1.
 - c. If Position_number > Max_length, then go back to step 2 and repeat algorithm for the next time interval.
5. Repeat steps 2, 3 and 4 until all bits have been decoded.

5.COMPLEXITY ANALYSIS

The overall number of states for a separate scheme, Osep is given as

$$Osep = Ns(I+1) . (1)$$

Where N_s = the number of states of channel decoder and I is the number of intermediate states of the Huffman decoder.

The Overall number states for a joint scheme using algorithm1 , O_{jnt1} is given as

$$O_{jnt1} = N_s + I. \quad (2)$$

The Overall number states for a joint scheme using the state counter algorithm, O_{jnt2} is given as :

$$O_{jnt2} = N_s. \quad (3)$$

Equation (3) clearly shows that the state counter algorithm provides a joint decoder with a fixed number of states which is much less than algorithm1.

6.SIMULATION RESULTS AND ANALYSIS

The simulations were carried out on an Intel Pentium IV 2.4 GHz PC using MATLAB 6.1. The performance of the joint scheme using the state counter method and the separate decoding scheme was compared in terms of bit error rate by transmitting a 4 KB text file in an Additive White Gaussian Noise channel and by using BPSK and 8PSK modulation schemes. The decoding time was compared for algorithm 1, the state counter algorithm and the separate decoding scheme after a workspace had been created for all the variables, so that a fair comparison could be made. The Graph of BER against E_b/N_0 for algorithm1, the state counter algorithm and the separate scheme is shown in Figure9. From the performance graph it can be deduced that the same performance in terms of bit error rate is obtained for both the separate and joint decoding schemes irrespective of the modulation scheme employed. But the key fact to note is that the decoding time in seconds for the separate scheme is more than that of the joint schemes and the joint scheme using the state counter algorithm is 4 times less for the case of this text file than that using algorithm1 as shown in Table 3.

Decoding scheme	Decoding time (s)
Separate	17
Algorithm1	12
State counter	3

Table 3 : Decoding times for simulated schemes.

7.CONCLUSION

The complexity of the joint decoding scheme using the state counter method has proved to be much less than the separate scheme or even the decoding scheme using algorithm 1. The only drawback of the state counter algorithm is the restriction which it imposes on the construction of the Huffman codewords which leads to a slight reduction in compression ratio.

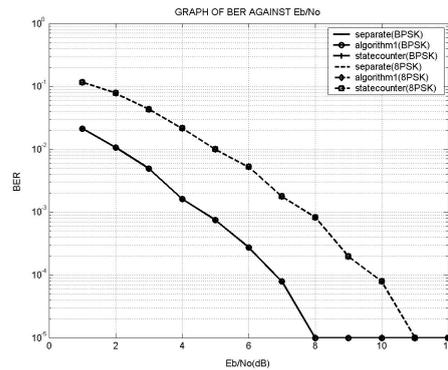


Figure9 : Performance curves for separate, algorithm1 and state counter algorithm.

REFERENCES

- [1] Honary B., Shaw M., Darnell M., "Combined Huffman and convolutional codes: Concatenated form without feedback.", *Electron. lett.*, Vol. 26, No. 6, pp.367-368, 15 Mar.1990.
- [2] Shannon, C. E., Weaver, W "The Mathematical Theory of Communication" *University of Illinois Press*, Urbana Ill, 1949.
- [3] Sklar B., "Digital Communications: Fundamentals and Applications", *Prentice-Hall International Inc.*, London, 1988.

E-mail address :

s.soyjaudah@uom.ac.mu ,
pawanfowdur@yahoo.com ,
m.hosany@uom.ac.mu

Postal address :

Department of Electrical and Electronic Engineering,
 Faculty of Engineering,
 University of Mauritius,
 Reduit,
 MAURITIUS.

A Modified Niederreiter Public Key Cryptosystem for Combined Coding and Cryptography

Govinda S.

Centre for Information Technology and Systems

University of Mauritius

Reduit, Mauritius

Email : mgovinda@uom.ac.mu, Tel : (230) 454-1041

Honary B.², Ackburally M.F.³, Soyjaudah K.M.S³

2. Department of Communication Systems, Lancaster University , U.K
3. Department of Electrical and Electronics Engineering, University of Mauritius, Reduit, Mauritius.

Abstract

Many public-key cryptosystems (PKC) based on error-correcting codes have been proposed as an alternative to the common cryptographic algorithms based on number theory. The Niederreiter PKC based on RS codes was introduced in 1986 but was soon broken by cryptanalysts. In this paper, the Niederreiter is revisited and a modification of the scheme is presented which offers a number of advantages, namely combined encryption and error-control within one single block , better security and better throughput compared with the Rivest Shamir Adleman (RSA) scheme which is the most widely implemented PKC in force nowadays. It will be shown that the Niederreiter PKC is particularly suitable for burst error channels.

Introduction

In 1978, McEliece suggested a public-key cryptosystem [1] which makes use of the existence of a class of error-correcting codes, the Goppa codes. The security of this system is based on the NP-completeness of the general decoding problem for linear codes. There have been no successful attempts to find possible leakages of the structure of Goppa codes and thus to cryptanalyse the McEliece cryptosystem. In 1986, an Austrian mathematician called Harald Niederreiter employed the Reed-Solomon (RS) codes to design the Niederreiter PKC [2]. The reasons for the choice of RS codes were as follows:

1. RS codes form the core of the most powerful known algebraic codes. In this respect, RS codes have excellent error-correcting capabilities.

2. There exists a fast and efficient decoding algorithm (e.g. the Berlekamp-Massey algorithm) for decoding RS codes in $O(n \log n)$.

Since its introduction, the Niederreiter PKC, like many other PKC's based on error-correcting codes, has largely remained of theoretical interest. The main reasons are:

1. They require enormous keys in order to achieve comparable security with currently implemented PKC's based on factoring of prime numbers or the discrete logarithm problem. McEliece cryptosystem has never been implemented commercially because in order to have a comparable security level with the RSA scheme which has today a key size of 1024 bits, the McEliece scheme needed a public key of the order of 500,000 bits.
2. The ciphertext is larger than the plaintext because of redundant bits added in the error-control system.

In this paper, the Niederreiter PKC is modified by inserting an additional private key to cater for channel errors. The resulting cryptosystem can be viewed as a hybrid of the McEliece PKC and the Niederreiter PKC and gives the possibility of combined error-control and cryptography.

It is shown that with the proposed Modified Niederreiter PKC, there is an increase in the work function as compared to the McEliece cryptosystem with comparable public-key size. This proves that the proposed Modified Niederreiter PKC is more secure to brute force attacks.

The Niederreiter PKC

Summary of the Algorithm

- RS codes over $GF(q)$, (n, k, d) , t -error correcting code
- H $[(n-k)$ -by- n] Parity-check matrix
- S $[(n-k)$ -by- $(n-k)$] Non-singular matrix
- P $[n$ -by- n] Permutation matrix

Encryption Procedure:

- Private Key: H, S, P
- Public Key: $H' = SHP$ and t
- Message: n -bit vector u over $GF(q)$ with weight $= t$
- Encryption: ciphertext $c = H'u^T$
- The ciphertext is then sent over the communication channel.

Decryption Procedure:

- Received cyphertext: $c = H'u^T$
- Decryption: Three steps for recovering original message:
 1. Multiplication on inverse non-singular S^{-1} matrix.
Since $c = H' \cdot u^T = (SHP) \cdot u^T$, then
$$\begin{aligned} S^{-1} \cdot c &= S^{-1} \cdot (H' \cdot u^T) \\ &= S^{-1} \cdot ((SHP) \cdot u^T) \\ &= HPu^T \end{aligned}$$
 2. Use a fast decoding algorithm for RS codes to find $u \cdot P^T$
 3. Multiplication on inverse permutation P^{-1} matrix and obtain message matrix u^T

$$u = uP^T \cdot (P^{-1})^T$$

The cryptosystem depends on the simple, but crucial fact that the matrix H yield a mapping from some vector u over $GF(q)$ of length n to another vector c over $GF(q)$ of length $n-k$. This is one-to-one when restricted to vector u of weight $= t$.

The Niederreiter PKC is said to be of knapsack type because the encryption can be viewed as picking t columns from the matrix H' and forming a weighted sum of these t column vectors.

Cryptanalysis of the Niederreiter PKC

A brute force attack on deciphering a specific ciphertext would be based on trying all possible message vectors of weight $= t$. From the original Niederreiter paper [2], the number of such vectors is

$$\sum_j^t \binom{n}{j} (q-1)^j \quad (1)$$

Nevertheless, in [3], Ernest F. Brickell and Andrew M. Odlyzko first described a cryptanalytic attack to the Niederreiter PKC and in 1992, the Niederreiter PKC was completely broken by Sidelnikov and Shestakov [4]. The cryptanalysis of Sidelnikov and Shestakov is based on the fact that the H matrix of the RS codes are constructed using a simple and known procedure and that the S and P matrices being binary scrambles and permutes the H matrix in such a way that it can be recovered in $O(n \log n)$ operations.

It is now recalled that there has not been any successful cryptanalysis on the McEliece scheme [5], although Goppa codes are also constructed using known procedures. The reason for the security of the McEliece scheme is due to the fact that instead of the parity check matrix, the generator G matrix is used. The transmitted codeword is $c = mSGP + e$ where e are intentional random errors introduced at the transmitter with weight less than the error correction capacity of the code. This later feature makes brute force attacks unsuccessful.

The Proposed Modified Niederreiter PKC

The modified version of the Niederreiter cryptosystem includes both cryptography and error-control coding. It resembles the original Niederreiter PKC except that another key has been introduced to cater for errors that might be introduced due to channel noise.

The Proposed Algorithm

- RS codes over $GF(q)$, (n, k, d) , t -error correcting code.

Private keys

1. G^T is the transpose of a generator matrix having dimensions $[(n-k)$ -by- n] of the $(n, (n-k), d^{ch})$ RS code over $GF(q)$, where $d^{ch} = k + 1$ and having an error correcting capability of $t^{ch} = k/2$. G^T is in systematic form

2. S is an $[(n-k)\text{-by-}(n-k)]$ non-singular scrambling matrix which is generated by a random process.
3. H is an $[(n-k)\text{-by-}n]$ Parity-check matrix of some Generalized RS codes. H is in non-systematic form.

It is to be noted that both G^T and H are from the same Galois Field and both the private keys which are in fact a fast decoding algorithm for the RS code.

4. P is an $[n\text{-by-}n]$ non-singular permutation matrix which is also generated using the same random process.

Public key

The public key, H^{new} , is an $[n\text{-by-}n]$ matrix given by

$$H^{\text{new}} = G^T \cdot S \cdot H \cdot P$$

The legitimate user publishes the matrix H^{new} as the public key in some directory, in the hope that it is very difficult to get the secret matrices G^T , S , H and P separately from the product.

Encryption Process

- Message: n -bit vector u over $GF(q)$ with weight = t
- An artificial vector of errors, e , over $GF(2)$ of weight t^{ch} or less, which is randomly chosen and added by the sending party. e will in fact constitute a *deliberate random* error vector in the codeword. This step introduces increased confusion and diffusion into the codeword production stage. This is mainly due to the very random nature of noise vectors.

The ciphertext is calculated as the syndrome c^{new} of a code with the public key H^{new} and the error vector e :

$$\begin{aligned} c^{\text{new}} &= H^{\text{new}} \cdot u^T + e \\ &= G^T S H P \cdot u^T + e \end{aligned}$$

The set of ciphertexts is the set of all the possible syndromes of correctable errors.

Decryption Process

- Received cyphertext:
 $c^{\text{new}} = H^{\text{new}} \cdot u^T + e + n$

Where n is the noise introduced by the channel.

The legitimate user calculates the syndrome by pre-multiplying c^{new} by a parity-check matrix, H^{ch} , of the Generalised RS code. H^{ch} has dimensions $[k\text{-by-}n]$.

$$\begin{aligned} H^{\text{ch}} \cdot c^{\text{new}} &= H^{\text{ch}} \cdot (G^T S H P u^T + e) \\ &= H^{\text{ch}} \cdot G^T S H P u^T + H^{\text{ch}} \cdot e \end{aligned}$$

If the channel has not introduced any error in the ciphertext, then, the syndrome will be equal to $H^{\text{ch}} \cdot e$. The first task of the decoder is to remove the vectors e and G^T using a fast decoding algorithm (e.g. the Berlekamp-Massey algorithm). Otherwise, since e has Hamming weight = t^{ch} , therefore the decoder will be able to decode the ciphertext as long as the sum of the number of errors introduced by the channel and the vector e is less or equal to t^{ch} .

Then, from the result of the decoding, if successful, the first $n-k$ terms are taken which corresponds to the size of the matrix, $c = S H P \cdot u^T$. c is in fact the ciphertext obtained using the original Niederreiter scheme.

After this step the decryption process continues as in the original Niederreiter PKC. It is pre-multiplied by S^{-1} to get the *cipher-syndromes* of the plaintext u .

$$\begin{aligned} S^{-1} \cdot c &= S^{-1} \cdot (H' \cdot u^T) \\ &= S^{-1} \cdot ((S H P) \cdot u^T) \\ &= H P u^T \end{aligned}$$

After applying again the Berlekamp-Massey algorithm, a scrambled version of the plaintext $u \cdot P^T$ is obtained.

Finally, the legitimate user post-multiplies the transpose of the inverse of the permutation matrix to recover the plaintext u .

$$u = u P^T \cdot (P^{-1})^T$$

Simulation and results

Simulations were carried out using the following parameters:

RS codes over $GF(256)$, (255,127,129), 64-error correcting code.

$$\begin{aligned} n &= 255 \\ k &= 127 \\ n-k &= 128 \\ d &= 129 \end{aligned}$$

$$d^{ch} = k + 1 = 128$$

$$t = \lfloor (d - 1)/2 \rfloor = 64$$

$$t^{ch} = \lfloor (d^{ch} - 1)/2 \rfloor = 63 \quad (\text{number}$$

of correctable symbols in the channel)
The dimension k of the ciphertext c should be in a medium range relative to the length n :

- If k is too small, then there are relatively few good codes of dimension k , which make it easier to break the cryptosystem;
- If k is too close to n , this results in short ciphertexts, which again imperils the security of the cryptosystem.

Encryption and Decryption Time and Throughput

For a 500kB file, the encryption time was found to be 13 minutes and 32 seconds whilst the decryption time is 25 minutes and 46 seconds. This means that the total throughput for encryption and decryption is 0.85 kb/s given that a rate $\frac{1}{2}$ algebraic coded cryptosystem is being employed. In [7], it is shown that, on a similar platform, the 1024-bit RSA scheme has a total throughput of 0.15 kb/s. The proposed scheme is therefore six times faster than the RSA scheme.

Cryptanalysis of the proposed Modified Niederreiter PKC

Two kinds of attacks on any public-key cryptosystem can be considered. The first is based on getting a plaintext from an intercepted ciphertext. The second consists of getting the private keys from the public key.

It is obvious, therefore, that there are two elements of security that are basically required by the system:

1. For an opponent, trying to compute the secret key from the public key must be an intractable problem.
2. For an opponent, knowing the public key and a cryptogram C , the search for the corresponding message m , must be an intractable problem for all C .

Getting a plaintext from ciphertext

In the McEliece PKC the average work function can be calculated using the following formula[5]:

$$W(n, k, t) = 3.5k^3 \frac{\binom{n}{k}}{\binom{n-t}{k}} \quad (2)$$

PKC	PARAMETERS	SIZE OF PUBLIC-KEY (BITS)	WORK FUNCTION
McEliece	Binary, $n = 1024, k = 524, t = 50$	Large: 5×10^5	$> 2^{59}$
The Modified Niederreiter	q -ary, $q = 256, n = 255, k = 127, t = 64$	Large: 5×10^5	$> 2^{101}$

Table 1: Workfunction for the Cryptanalysis of the modified Niederreiter PKC and McEliece PKC.

Table 1 shows the results concerning the work function of the modified Niederreiter PKC against that of the McEliece PKC. Using the formula given in equation 2, the work function of the modified Niederreiter PKC is very large compared with the McEliece Scheme.

Getting the secret key from the public key

Using a code of $n = 255$ and $k = 127$ implies a key size of about 500,000 bits. Also, both G^T and H are non-binary. Therefore, it is expected that the product $G^T S H P$ will give a result which is more difficult to reverse back in order to obtain the constituent private keys.

For P , the number of operations is $255!$
For S , the number of operations is $(2^{127} - 2^0)(2^{127} - 2^1) \dots (2^{127} - 2^{126})$

The results for the above would give astronomical values and thus the probability of a successful attack on the secret keys is practically zero.

Combined Error Control Coding and Cryptography

In the proposed Modified Niederreiter cryptosystem, the private key H , is used for encryption while the private key G^T is used for error control. For added security an error vector e , is inserted. The received coded-ciphertext is given by $c^{\text{new}} = H^{\text{new}} \cdot u^T + e + n$. By a carefully studying the channel, the level of intentional noise to be added can be chosen and the added noise in the channel will be correctable as long as the total weight of the e and n vectors is less than t^{th} . We therefore have combined coding and cryptography.

Public-Key Size

The biggest disadvantage of the proposed scheme is the very large public-key size. For the $n = 255, k = 127$, the public-key size is $n \cdot n \cdot 8$ bits, that is about 500,000 bits. Comparing this with a public-key size of 1024 bits for RSA, with practically same security, the key size of the proposed scheme might seem phenomenal. However, one should note that with improvement in microprocessor technology, the handling of a large key can be done very easily and no difficulties were encountered during the simulation phase.

Conclusion

The possibilities offered by the proposed PKC of combined error-control coding and cryptography implies that with the same throughput two important blocks in a communication system will be done at once. The redundancy associated with error control coding is now compensated by the fact that cryptography of an extremely high security level is embedded in the system. Finally since RS codes is being used it can be deduced that the combined encoding/ encryption technique is well suited for burst error correction such as for deep space communication.

References

[1] R.J. McEliece, A Public-Key Cryptosystem Based on Algebraic Coding Theory, Jet Propulsion Laboratory,

California Institute of Technology. DSN Progress Report, 1978.

[2] H. Niederreiter, Knapsack-Type Cryptosystems and Algebraic Coding Theory, Problems of Control and Information Theory, vol. 15 no. 2, 1986.

[3] E.F. Brickell and A.M. Odlyzko, Cryptanalysis: A Survey of Recent Results, Invited Paper, IEEE Proceedings, Vol 76, no 5, 1988.

[4] V.M. Sidelnikov and S.O. Shestakov, On Insecurity of Cryptosystems Based on Reed-Solomon Codes, Discrete Mathematics and Applications, Vol. 2, no. 4, 1992.

[5] E.M. Gabidulin, On Public-Key Cryptosystems Based on Linear Codes: Efficiency and Weakness., Codes and Ciphers, Proceedings of the 4th IMA Conference on Cryptography and Coding, IMA Press, 1995.

[6] K. Gibson, The Security of the Gabidulin Public Key Cryptosystem, Advances in Cryptology—EUROCRYPT 96, vol. 1070, Lecture Notes in Computer Science, 1996, pp 212–223.

[7] B. Schneier, Applied Cryptography, John Wiley & Sons Second Edition, 1996, pg 466

A NEW APPROACH TO PREDICTIVE POWER CONTROL OF WCDMA SYSTEMS

N. Guven, O. Hinton, C. Tsimenidis.
University of Newcastle Upon Tyne, UK

Abstract. In this paper, an application of power control in wideband code division multiple access (WCDMA) systems, is considered. Power control is a unique method of mitigating interference and compensating the fading effects.

To further improve power control methods, a new approach to predictive power control techniques has been developed. The results have shown that the proposed technique would work perfectly with a prediction model which tracks the channel in a precise way and optimum adaptive step sizes.

I. Introduction

WCDMA is the preferred air interface technology for third generation communication systems all over the world. Compared to second generation narrow-band CDMA methods, WCDMA offers significant advantages, such as higher bit rates and better capacity.

In the FDD mode (Frequency Division Duplex) of WCDMA systems, two different RF-bands is assigned for uplink (UPL, from mobile to base station) and downlink (DWL, from base station to mobile). The users in the uplink employ same carrier frequency and the more power than necessary of a user can block a large part of the base station. In this regard, uplink power control will be the focus of this work.

As long as to balance the power of users, to compensate the power variations of a single user coming from path loss, shadowing, multi-path fading, etc, various power control techniques have been developed. These techniques are mainly collected under the title of closed-loop power control.

In the uplink, the base station compares the target and the measures of the criteria depending on the power control scheme and sends commands to adjust the transmission powers of the mobiles. Each command is placed in a slot of the dedicated control channel. If the duration of one time-slot is $666.67 \mu s$, from (a frame duration / the number of slots in a frame = $10 \text{ ms} / 15$), this cycle is executed at a rate of 1500 times per second for each mobile station. However it should be kept in mind that, as the speed of the mobile user increases, this control mechanism becomes unable by conventional techniques [1].

Closed-loop power control techniques can be classified in many different ways.

According to the parameters measured to determine the power control command, power control techniques can be classified into three categories. These are Strength-based, SIR-based (signal to interference ratio) and BER-based (bit error rate) [2].

In the Strength-based scheme, the strength of the received signal is measured to determine whether it is higher or lower than the desired level. The command to reduce or increase the power is made accordingly [2].

The SIR-based has better system performance than the Strength-based. In this scheme, the base station monitors E_b/N_0 (the ratio of energy per bit to the spectral noise density), calculates SIR and compares it to a target SIR. If the measured SIR is higher than the target SIR, the base station will command the mobile station to decrease the power; if it is lower, the base station will command the mobile station to increase the power. The SIR target of the closed-loop power control is set by outer-loop power control according to the needs of individual radio links. If the signal and interference powers are constant, the BER will be a function of the SIR, and in this case the quality of the service will be equivalent [2]. However, in reality SIR is time-variant and thus average SIR will not correspond to average BER. In this case, BER is a better quality measure but is not quick enough. Therefore, SIR-based technique is preferred [2].

According to update strategies, power control algorithms can be classified as fixed step size algorithm and adaptive step size algorithm and reasonably adaptive algorithm would improve the results.

Lastly, techniques using predicted values instead of measured values have been developed and these techniques are known as predictive power control techniques. It has been shown that predictive techniques outperform conventional ones [3].

Recent techniques have started to combine and use the best aspects of the conventional techniques. Sim revealed that a multi-step-size algorithm can perform better than a single-step-size algorithm in his predictive model [3]. It has been proved that a model which is both predictive and adaptive would give better simulation results than previous ones [4].

In this paper, a new adaptive prediction model is proposed which is also SIR and strength based. It is indicated with the simulation results that this model is fast convergent.

II. The Proposed Power Control Technique

The proposed closed-loop model is shown in Figure 1. It has been developed to keep both SIR values and received powers of users constant in WCDMA networks.

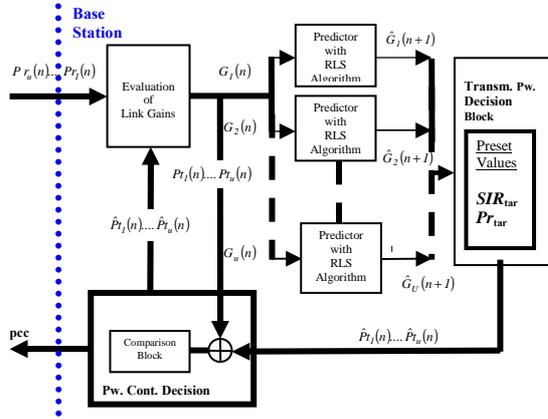


Fig. 1 The Proposed Model

At the base station, the received powers of the users Pr are assumed to stay constant due to unchanged channel conditions during one power control duration which lasts one time slot. Hence, there will be no need for more than one bit information to decide the power adjustment command and moreover the predictors will be used to provide the next duration values to control the powers. Since the transmitted powers of the users, Pt , are also already known, the link gains (G_i for the user i) are easily calculated from,

$$Pr_i = Pt_i G_i \quad (1)$$

By reformulating in dB scale, G_i becomes

$$G_i = Pr_i - Pt_i \quad (2)$$

The link gain described here is the total of the antenna gains of the user equipment (UE), the base station (NODE-B), and the loss from propagation and multi-path effects. Mathematically,

$$G_i = G_{UE} + G_{NODE_B} - Loss \quad (3)$$

The observed link gains at the base station are separately fed to the predictors to obtain the predicted gains called \hat{G}_i as shown in Fig. 1. The predicted gains are only dependent on the current and the past link gains of the same user. To predict the next link gain, the RLS (Recursive Least Square) algorithm has been used as a preferred technique for fast fading [5]. RLS algorithm can be presented as follows:

$$\mathbf{k}_i(n+1) = \frac{\beta^{-1} \Theta_i(n) \mathbf{G}_i(n)}{1 + \beta^{-1} \mathbf{G}_i(n)^H \Theta_i(n) \mathbf{G}_i(n)} \quad (4)$$

$$\hat{G}_i(n+1) = \mathbf{w}(n)^H \mathbf{G}_i(n) \quad (5)$$

$$e_i(n+1) = G_i(n+1) - \hat{G}_i(n+1) \quad (6)$$

$$\Theta_i(n+1) = \beta^{-1} \Theta_i(n) - \beta^{-1} \mathbf{k}_i(n+1) \mathbf{G}_i(n)^H \Theta_i(n) \quad (7)$$

$$\hat{\mathbf{w}}_i(n+1) = \hat{\mathbf{w}}_i(n) + \mathbf{k}_i(n+1) e_i^*(n+1) \quad (8)$$

where β is a pre-selected scalar, M is filter length, $\mathbf{G}_i(n)$ is the input of the predictor for user i at the n th power control period,

$$\mathbf{G}_i(n) = [G_i(n), G_i(n-1), \dots, G_i(n-M+1)]^T \quad (9)$$

$\Theta_i(n)$ is a square matrix of size M , and $\hat{\mathbf{w}}_i(n)$ is the estimated filter weight vector

$$\hat{\mathbf{w}}_i(n) = [\hat{w}_{i1}(n), \hat{w}_{i2}(n), \dots, \hat{w}_{iM}(n)]^T \quad (10)$$

After this stage, the next user powers are to be decided. The proposed algorithm does not include classical eigen-value calculations as other techniques. The signal-to-interference ratio (11) is used instead.

$$SIR_i = \frac{Pr_i}{\sum_{j=1, j \neq i}^u Pr_j} \quad (11)$$

By substituting (1), we get

$$SIR_i = \frac{Pt_i G_i}{\sum_{j=1, j \neq i}^u Pt_j G_j} \quad (12)$$

Accordingly to find the transmitted power for user i , (12) can be changed to

$$Pt_i = \frac{SIR_i}{G_i} \sum_{j=1, j \neq i}^u Pr_j \quad (13)$$

Based on the idea that using the predicted gain, the target SIR value for the user i and the target received powers for the remaining users in the next period would provide the essential transmitted power for user i , the algorithm is obtained by

$$\hat{Pt}_i(n+1) = \frac{SIR_{i_tar}}{\hat{G}_i(n+1)} \sum_{j=1, j \neq i}^u Pr_{j_tar} \quad (14)$$

As clearly shown from this power update equation, the model regulates not only SIR but also the received signal strengths. The simplicity of the proposed power control algorithm is provided by assigning the assumed target values instead of some observed results. No outer loop is required with this approach to decide target SIR based on BER at the base station side. Therefore, unnecessary calculations and transmission of information will be prevented. In [6], Zhu and Dong have demonstrated the benefits of using assumed values on a similar technique while adjusting the coefficient α by preset values.

It is also worth pointing out that the proposed algorithm does not require any observed values of the remaining active users in the network. Since the pre-set common target values are used in the algorithm, the proposed model implements a centralized control mechanism using information globally available in the system. However, it is essentially a distributed model.

When the base station determines the next transmission power level, the difference between the decided and the current transmission powers for the

i th user is mapped to one of the step sizes listed in Table 1.

The power difference is expressed as

$$\Delta_i^n = \hat{P}t_i(n) - Pt_i(n) \quad (15)$$

The step levels have been empirically adjusted in the simulations. The power control command (PCC) is assigned accordingly as three-bit control data. Moreover, if the difference is low enough, PCC may not be sent to transmitter for the sake of saving bandwidth. It is well known that there are six time slots for the uplink dedicated control channel under circumstances of different number of feedback indicator bits (FBI) and/or transport format combination indicator bits (TFCI) included [1]. In the same way, flexible downlink channel structures can be proposed to provide various number of power control commands.

Conditions	pcc	Step
$\Delta_i^n \leq -0.5$	111	-3
$\Delta_i^n \leq -0.5$	110	-0.5
$-0.5 < \Delta_i^n \leq -0.3$	101	-0.3
$-0.3 < \Delta_i^n \leq -0.2$	100	-0.2
$-0.2 < \Delta_i^n \leq -0.1$	011	-0.1
$0.1 < \Delta_i^n \leq 0.2$	010	0.1
$\Delta_i^n \geq 0.2$	001	0.2
$-0.1 < \Delta_i^n \leq 0.1$	000	0

Table 1: Step size list

III. Simulation Model and Results

The simulation model is based on one base station and randomly distributed users. The cell radius is assumed 5km and the velocities of the users are kept at 100km/h. To make a clear explanation, this work will include one communication service and consequently fixed target values. However the proposed technique has had also good performance with multi-service systems by assigning different preset values for each service type, since the target values are service dependent in practical network scenarios.

The Cost-231 Walfish_Ikegami model [7] is used for the propagation loss and a Rayleigh model to reflect the fast fading effect. The RLS (recursive least square) algorithm is set by the initials; $\hat{\mathbf{w}}(0)=0$, $\Theta(0)=\rho^{-1}\mathbf{I}$ where \mathbf{I} is the identity matrix of size M , $\beta=1$, $\rho=0.001$ and $M=3$ are the pre-set values for the predictive model. The parameters of the algorithms used in the proposed model are described in the following tables.

Carrier frequency	f_c	2 GHz
Building separation	B	40 m

Road width	$R_w=b/2$	20 m
Road orientation	Φ	90°
Base station height	h_b	50 m
Mobile height	h_m	2 m
Building height	h_r	20 m

Table 2: Parameters for Cost-231 Walfish-Ikegami Model

Symbol frequency	f_s	15 kbps
Complex-WGN noise	S_{gm}	1
Number of samples	N	1500

Table 3: Parameters for the Rayleigh Model

Power control speed	1500 Hz
The target received power	-30 dBw
The target SIR for 3 users	-3.0103 dB
Min transmitted power of mobile	-30 dBw
Max transmitted power of mobile	3 dBw

Table 4: Parameters for the proposed model

The simulation results were obtained for 3 users in a Rayleigh channel environment. The adjusted transmitted powers of the users are assumed to exhibited a symmetric pattern with link gains presented under imperfect prediction and imprecise step sizes in Fig. 2(a). The predicted and observed link gains are compared in Fig. 2(b). As easily seen, the plots would not be symmetric under deep fading circumstances. Therefore, the results of the received powers and the SIR values may have fluctuations even after the convergence as in Fig. 2(c). The convergence has generally been obtained after approximately 10 iterations; in other words, the 100th bit with the limitation of power amplifiers that are assumed to amend the power 3 dBw at maximum during the analysis. The analysis of the results has also shown that the deviations can be reduced using perfect prediction and efficient step sizes. The obtained plots, applying the accurate amount of the power change determined by the base station and assuming the prediction of the channels are accurate, are presented in Fig. 3. The robustness of the proposed power control algorithm is clearly indicated without any deviations in these plots.

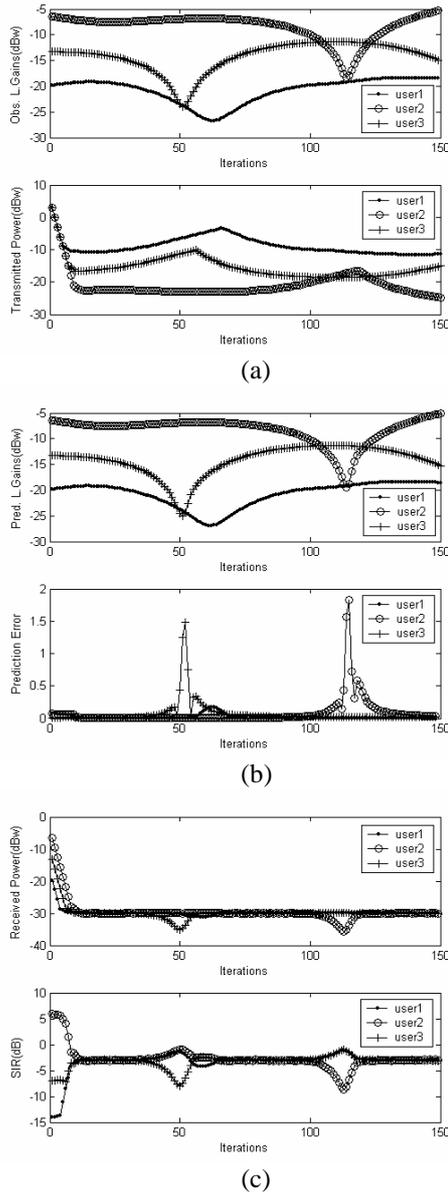


Fig. 2 (a) The transmission powers and the consequently observed link gains; (b) the predicted link gains and prediction errors; (c) the received powers and SIR values of the three users.

IV. Conclusion

A new approach to adaptive prediction based power control has been proposed. Even at high mobility network scenarios, better performance has been attained compared to other techniques. It has been shown that the proposed algorithm is fast convergent however the results could still be improved by updating the prediction model and adaptive step sizes more frequently.

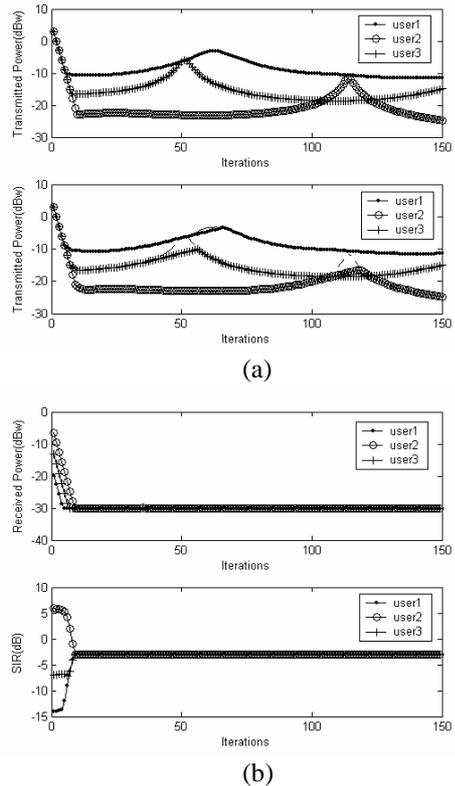


Fig. 3 (a) The transmission powers with perfect prediction vs. imperfect conditions; (b) the received powers and SIR values with perfect prediction.

References

- [1] Blaise F., Elicegui L., Goeusse F., Vivier G., "Power control algorithms for soft handoff users in UMTS", Mot labs, Paris, 2002.
- [2] Novakovic D., Dukic M., "Evolution of the power control techniques for DS-CDMA toward 3G wireless communication systems", *IEEE Com. Surveys & Tutorials*, 2000.
- [3] Sim M.L., Gunawan E., Soong B., Soh C., "Performance Study of Close-Loop Power Control Algorithms for a Cellular CDMA System", *IEEE Trans. Veh. Tech.*, vol. 48, no. 3, May 1999.
- [4] Aldajani M. A., Sayed Ali H., "Adaptive predictive pow. cont. for the uplink channel in DS-CDMA cell. sys.s", *IEEE Trans. Vehicular Tech.*, vol. 52, no.6, Nov. 2003.
- [5] Lau F.C.M., Tam W.M., "Achievable-SIR-Based Predictive Closed-Loop Power Control in CDMA Mobile System", *IEEE Trans. Vehicular Tech.*, vol. 51, no.4, July 2002.
- [6] Lv L., Zhu S., Dong S., "Fast convergence distributed power control algorithm for WCDMA systems", *IEE Proc.-Commun.*, vol. 150, no.2, April 2003.
- [7] Stuber G., *Principles of mobile communication*, Kluwer Academic Publishers, 2nd edition, 1997.

A New Non Data-Aided Frequency Offset Recovery Technique for OFDM Systems

A. J. Al-Dweik

Communications Engineering Department, Etisalat University
Sharjah, UAE, P.O.Box 980, (e-mail: dweik@fulbrightweb.org)

ABSTRACT

A new technique to recover the carrier frequency offset of OFDM systems is proposed. The proposed technique does not require cyclic prefix, pilot signals, training symbols, or any other supplementary data, this will significantly improve the system power and bandwidth efficiency. The proposed algorithm exploits the interference that results due the loss of orthogonality between subcarriers where the variance of the resulted interference is proportional to the carrier frequency offset. The proposed technique does not require prior fine symbol synchronization and it is capable of estimating the carrier offset at E_b/N_0 as low as -3 dB with symbol timing offset that is less than or equal to 10% of the symbol duration

I. Introduction

Orthogonal frequency division multiplexing (OFDM) a modulation technique employs k orthogonal subcarriers modulated by k parallel data streams [1]. Compared to single carrier systems, OFDM shows significant immunity in multipath and impulsive noise channels, and possesses high bandwidth efficiency due to the overlapping of the subcarriers spectra. Due to these advantages, OFDM was adapted as the modulation scheme for audio digital broadcasting (ADB) system [2] and is also proposed as the terrestrial HDTV transport in Europe [3], and is being investigated for other wireless applications such as digital broadcast television and mobile communication systems, as well as for broadband digital communication on existing copper networks.

The main limitation for using OFDM is its high sensitivity to synchronization errors and the interference introduced by multipath channels. An accepted means of avoiding the interference and preserving orthogonality between subcarriers is to copy the last K samples of the body of the OFDM symbol and append them as a preamble, the cyclic prefix (CP), to form the complete OFDM symbol [1]. The effective length of the OFDM symbol as transmitted is this CP plus the body of the OFDM symbol. Performance degradation resulted from synchronization errors can be minimized using robust synchronization techniques. Most of the existing synchronization techniques are designed to exploit the redundant data in the CP [4]. The main disadvantage of these methods is the need for high E_b/N_0 [4]. An other method is to transmit training symbols (known symbol pattern used to

assist the demodulator) [5] [6], such techniques are usually more robust except that it comes on the expense of the system resources.

It is also important to emphasize that not all wireless channels are necessarily fading channels. In this case, the channel does not introduce ISI for which using the CP will result in a significant loss in the transmission power and bandwidth efficiency. Traditionally, carrier recovery in such systems is achieved by transmitting a pilot signal as one of the OFDM channels [7], CP can also be used in AWGN channels and it can be used in the carrier recovery process [4].

In this work, we propose a non data-aided (NDA) frequency offset recovery algorithm for OFDM based systems where pilot signals, training symbols, or any other supplementary data are not required.

II. OFDM Systems

Consider an OFDM system with N ($2L$) complex subcarriers $C_{-L} \dots C_{L-1}$, modulated by N complex parallel data streams with symbol rate (R_s). The equivalent complex baseband representation of the transmitted signal can be expressed as [8]

$$S(t) = \sum_{i=-L}^{L-1} a_i e^{j2\pi \frac{i}{T_s} t}, \quad 0 \leq t \leq T_s \quad (1)$$

where $T_s = 1/R_s$, a_i is the complex data symbol modulating the i th subcarrier during one OFDM symbol duration, $a_i = a_{i,I} + ja_{i,Q}$. The real part of $S(t)$ represents the inphase component and the imaginary part represents the quadrature component. The k th symbol of subchannel- i ($\hat{a}_{i,k}$) can be extracted from (1) by correlating the received signal ($r(t)$) with C_i ,

$$\hat{a}_{i,k} = \frac{1}{T_s} \int_0^{T_s} r(t) C_i(t) dt, \quad kT_s \leq t \leq (k+1)T_s \quad (2)$$

notice that \hat{a}_i has an analog value while a_i has a discrete value, $\hat{a}_i = a_i$ if and only if $r(t) = S(t)$.

Assume initially that the channel is perfect and all signal parameters are known except the carrier frequency which has an offset (Δf), the received signal with frequency offset can be expressed as

$$r(t) = \sum_{i=-L}^{L-1} a_i e^{j\frac{2\pi}{T_s}(i+\mu)t}, \quad 0 \leq t \leq T_s \quad (3)$$

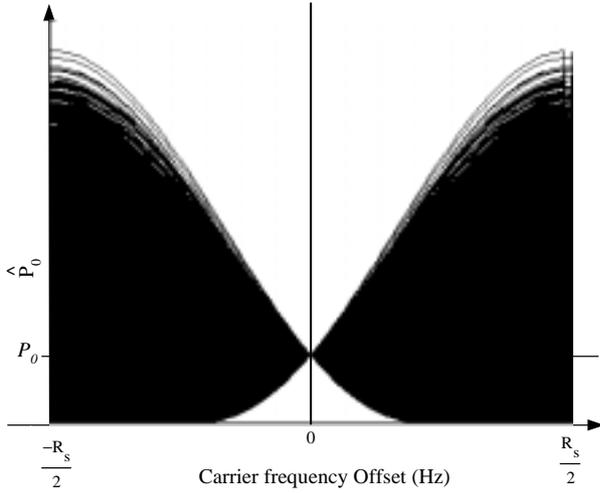


Fig. 1. OFDM symbol power as a function of frequency offset

where μ , the relative carrier frequency offset is equal to $\Delta f/R_s$. Without loss of generality we will assume that subchannel-0 is the subchannel of interest, thus

$$\hat{a}_{0,k} = \sum_{i=-L}^{L-1} a_{i,k}\alpha_i + a_{i,k}^*\beta_i \quad (4)$$

where $a_{i,k}^*$ is the complex conjugate of $a_{i,k}$, $\alpha_i = \frac{\sin(2\pi\mu)}{2\pi(i+\mu)}$, and $\beta_i = \frac{1-\cos(2\pi\mu)}{2\pi(i+\mu)}$. The term that corresponds to $i=0$ can be separated from the summation in (4),

$$\hat{a}_{0,k} = a_{0,k}\alpha_0 + a_{0,k}^*\beta_0 + ICI \quad (5)$$

The interchannel interference (ICI) is similar to (4) except that the terms that correspond to $i=0$ are omitted.

III. The Proposed Carrier Recovery Algorithm

If the magnitude of the received symbol is calculated $\{\text{Re}^2(\hat{a}_{0,k}) + \text{Im}^2(\hat{a}_{0,k})\}$, the power of the received symbol is obtained

$$\hat{P}_0 = P_0 \text{sinc}(\mu) + a_{0,I,k}a_{0,Q,k}\alpha_0\beta_0 + \dots \quad (6)$$

where $\text{sinc}(x) = \sin(\pi x)/\pi x$. Due to the randomness of the transmitted data, all the terms in (6) will be random except the first term which is deterministic for a fixed value of μ . As an example, the value of \hat{P}_0 as a function of μ is shown in Fig. 1 in the range $(-0.5, 0.5)$. This figure is generated in a fashion similar to a regular eye-diagram except that the symbol amplitude is replaced by the symbol power and the time is replaced by the relative frequency offset. An important point to be concluded from this figure is the proportionality between the randomness of \hat{P}_0 and μ , this randomness will decrease by decreasing μ until $\hat{P}_0 = P_0$ when $\mu = 0$.

As a conclusion, only one deterministic point corresponds to $\mu = 0$ do exist in fig. 1, all other points are random. The frequency offset can be estimated by searching

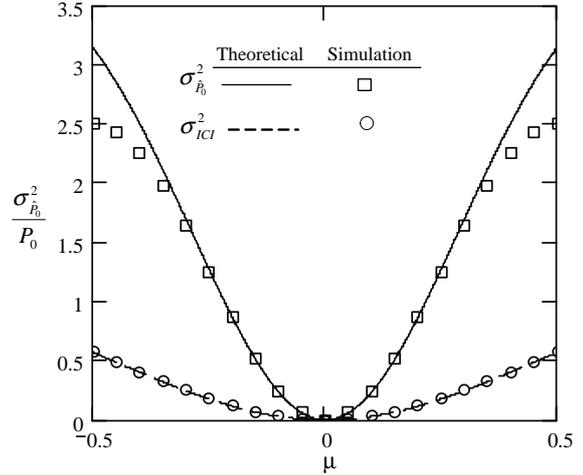


Fig. 2. Normalized $\sigma_{\hat{P}_0}^2$ and σ_{ICI}^2 as a function of μ

for that deterministic point. In the presence of noise and fading the optimum point is the one that posses the minimum amount of randomness, or equivalently the point with minimum variance since it is the optimum metric to measure the dispersion of \hat{P}_0 [9].

For an accurate estimation for the variance of the received symbols power, a large number of symbols is required. If only subchannel-0 is used in the estimation process, a large time delay will result. Fortunately, all subchannels are approximately affected similarly by the ICI except for the subchannels at the band edges which can be excluded [1]. Thus, for an OFDM system with large number of subcarriers the variance estimation can be accomplished by calculating the variance of the symbols power at the output of all subchannels except the ones at the band edges. The overall time required for the offset estimation depends on the technique used to search for the minimum variance point.

Due to the constellations symmetry of most modulation techniques used in OFDM systems such as MPSK and QAM, we can assume that $E\{ICI\} = 0$, the ICI variance (σ_{ICI}^2) can be represented by

$$\sigma_{ICI}^2 = P \sum_{\substack{i=-L \\ i \neq 0}}^{L-1} \left(\frac{\sin(\pi\mu)}{\pi(i+\mu)} \right)^2 \quad (7)$$

Since all subcarriers usually have equal average power, $P_i (E\{a_{i,k}^2\})$ was taken out of the summation [5].

Using (7) and Fig. 2 it is clear that σ_{ICI}^2 is proportional to μ , hence, the relation between the variance and μ can be exploited to indicate the position of the correlator window with respect to the spectrum of any subchannel. For subchannel-0, $\hat{P}_0 = \hat{a}_{0,I}^2 + \hat{a}_{0,Q}^2$, the variance of the \hat{P}_0 ($\sigma_{\hat{P}_0}^2$) at the receiver can be expressed as [9]

$$\sigma_{\hat{P}_0}^2 = \sigma_{\hat{a}_{0,I}^2}^2 + \sigma_{\hat{a}_{0,Q}^2}^2 + 2\text{cov}\{\hat{a}_{0,I}^2, \hat{a}_{0,Q}^2\} \quad (8)$$

where the variances of $\hat{a}_{0,I}^2$ and $\hat{a}_{0,Q}^2$ are identical and can be expressed as

$$\sigma_{\hat{a}_{0,I}}^2 = \sigma_{\hat{a}_{0,Q}}^2 = (F_0 - P^2)(\alpha_0^4 + \beta_0^4) + 2\sigma_{ICI}^4 + 4P [P\alpha_0^2\beta_0^2 + \sigma_{ICI}^2 \text{sinc}^2(\mu)] \quad (9)$$

The covariance of \hat{a}_I and \hat{a}_Q can be expressed as

$$\text{cov}\{\hat{a}_{0,I}^2, \hat{a}_{0,Q}^2\} = (2F_0 - 6P^2) (\alpha_0^2\beta_0^2) \quad (10)$$

where $F_0 = E \{a_{0,I,k}^4\} = E \{a_{0,Q,k}^4\}$. The behavior of (8) is shown in Fig. 2 as a function of μ for an OFDM system with 16 subchannels all are using QPSK modulation.

IV. Performance of the Proposed Technique

In practical operating conditions, the channel is not perfect and several impairment parameters will affect the received signal. In this work we will investigate the performance of the proposed technique in channels with additive white Gaussian noise (AWGN) and non perfect symbol timing. Consider an N subchannels OFDM system, we assume that the symbol duration T_s has been selected so that $T_s \gg T_g$ where T_g is the width of the CP, as a consequence we can make the following assumptions [10]. First, it is assumed that the impulse response of the channel does not change during transmission, this corresponds to static-fading [5], second, the effect of ISI is negligible. Therefore, the useful part of the received signal can be expressed as

$$r(t) = \sum_{i=-L}^{L-1} a_i e^{j\frac{2\pi}{T_s}(i+\mu)t} + n(t), \quad 0 \leq t \leq T_s \quad (11)$$

where $n(t)$ is a zero-mean complex AWGN with $N_0/2$ power spectral density. This model for the received signal can be used even if the CP is completely not used in the OFDM system [11]. The received signal will also have a random time delay (ε). In the absence of information about symbol timing, the correlator window of the receiver will be working partially on two consecutive symbols and the received k th symbol can be expressed as

$$\hat{a}_{i,k} = \frac{1}{T_s} \int_{\varepsilon}^{T_s} r_1(t) C_i(t) dt + \frac{1}{T_s} \int_{T_s}^{T_s+\varepsilon} r_2(t) C_i(t) dt \quad (12)$$

The signals $r_1(t)$ and $r_2(t)$ represent the OFDM signal during the time periods of the symbols $a_{i,k}$ and $a_{i,k+1}$ respectively. Without loss of generality we will assume that subchannel-0 is the subchannel of interest, then $a_{0,k}$ can be recovered by substituting $C_0 = 1$ in (12). Thus

$$\hat{a}_{0,k} = \sum_{i=-L}^{L-1} a_{i,k} \chi_{1,i} + a_{i,k}^* \chi_{2,i} + a_{i,k+1} \chi_{3,i} + a_{i,k+1}^* \chi_{4,i} + n \quad (13)$$

where $\chi_{1,i} = \frac{\sin(2\pi\mu) - \sin(2\pi\Delta(i+\mu))}{2\pi(i+\mu)}$, $\chi_{2,i} = \frac{\cos(2\pi\Delta(i+\mu)) - \cos(2\pi\mu)}{2\pi(i+\mu)}$, $\chi_{3,i} = \frac{\sin(2\pi\Delta(i+\mu))}{2\pi(i+\mu)}$,

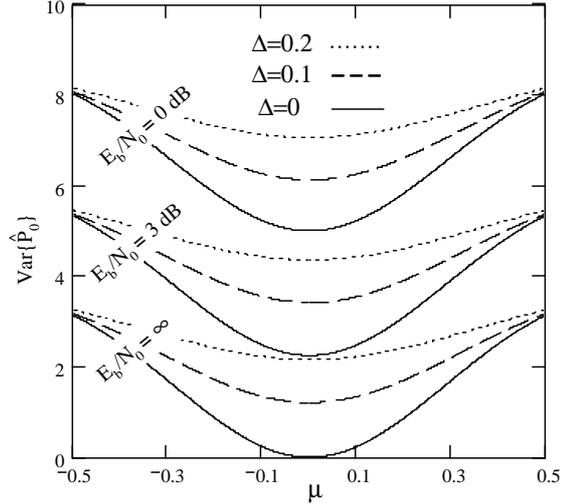


Fig. 3. $\sigma_{\hat{P}_0}^2$ as a function of μ for different values of N_0 and Δ with and without fading.

$\chi_{4,i} = \frac{1 - \cos(2\pi\Delta(i+\mu))}{2\pi(i+\mu)}$, $\Delta = \frac{\varepsilon}{T_s}$ and n is a complex additive white Gaussian noise with variance that is equal to $\frac{N_0}{2}$. Equation (13) can be rewritten as

$$\hat{a}_{0,k} = a_{0,k} \chi_{1,0} + a_{0,k}^* \chi_{2,0} + a_{0,k+1} \chi_{3,0} + a_{0,k+1}^* \chi_{4,0} + d \quad (14)$$

where $d = ICI + n$, ICI is similar to (13) except that n and the terms that correspond to $i = 0$ are omitted. The term d has zero-mean and variance (σ_d^2) that can be expressed as

$$\sigma_d^2 = P \sum_{\substack{i=-L \\ i \neq 0}}^{L-1} \sum_{j=1}^4 \chi_{j,i}^2 + \frac{N_0}{2} \quad (15)$$

Notice that $\hat{a}_{0,I}$ and $\hat{a}_{0,Q}$ are correlated, thus $\sigma_{\hat{P}_0}^2$ can be calculated according to (8), where

$$\sigma_{\hat{a}_{0,I}}^2 = \sigma_{\hat{a}_{0,Q}}^2 = (F_0 - P^2) \sum_{i=1}^4 \chi_{i,0}^4 + 2\sigma_d^4 + 4P \times \left\{ P \left(\begin{array}{l} \chi_{1,0}^2 \sum_{i=2}^4 \chi_{i,0}^2 + \chi_{2,0}^2 \times \\ \sum_{i=3}^4 \chi_{i,0}^2 + \chi_{3,0}^2 \chi_{4,0}^2 \\ \sigma_d^2 \sum_{i=1}^4 \chi_{i,0}^2 \end{array} \right) + \right\} \quad (16)$$

and the $\text{cov}(\cdot)$ can be expressed as

$$\text{cov}\{\hat{a}_{0,I}^2, \hat{a}_{0,Q}^2\} = (2F_0 - 6P^2) (\chi_{1,0}^2 \chi_{2,0}^2 + \chi_{3,0}^2 \chi_{4,0}^2) \quad (17)$$

The effects of AWGN and timing offset (Δ) are shown in Fig. 3, AWGN will increase the variance of \hat{P}_0 . Fortunately, this increment is not a function of μ , this will result in a fixed shift to the variance profile without any change in its shape. Theoretically speaking, AWGN will have no effect on the system performance as long as the variance is calculated over a sufficient number of symbols enabling the system to operate at very low values of E_b/N_0 , Fig. 3. The effect of Δ is quite different since

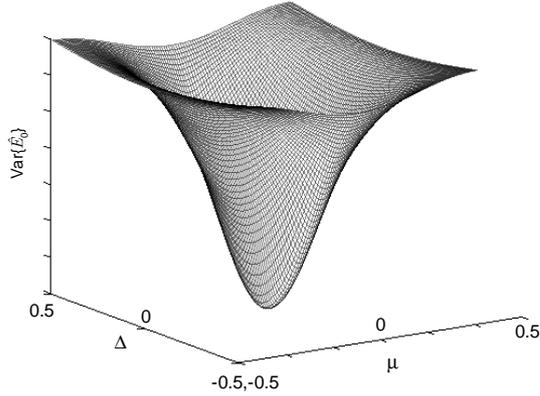


Fig. 4. $\sigma_{\hat{E}_0}^2$ as a function of μ and Δ , no fading and $N_0 = 0$.

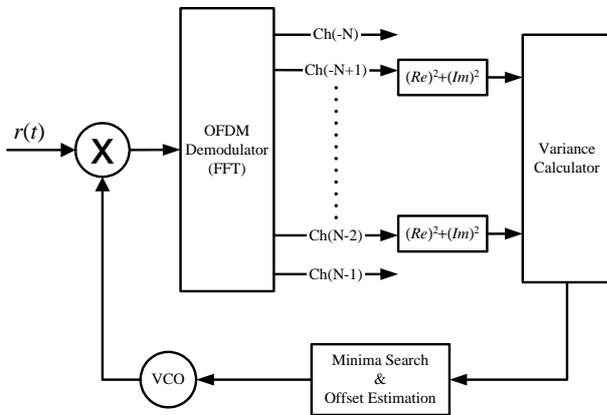


Fig. 5. Block diagram of the proposed system

it affects the shape of the variance profile by introducing flatness in the variance profile, this will reduce the dependency of the variance on μ and will reduce the estimation accuracy. This effect is shown in Fig. 3, and in Fig. 4 which shows that the optimum variance profile (maximum curvature) can be obtained at $\Delta = 0$.

As we can see from Fig. 5, the Voltage Controlled Oscillator (VCO) output represents the estimated carrier offset. The accuracy of our estimation can be measured by calculating the standard deviation of the VCO output from the actual carrier offset. Using simulation, the standard deviation ($\sqrt{E\{(VCO_{output} - \Delta f)^2\}}$) for the VCO output was measured for $\mu = 0.2$, $\Delta = 0$ and $1/32$, Fig. 6.

V. Conclusion

In this work, the carrier frequency offset estimation of OFDM systems was considered. A new technique was proposed to estimate the frequency offset based on the variance of the power for the received symbols. The large number of parallel subchannels has enabled a rapid and accurate estimation.

The effect of non perfect symbol timing on the frequency offset estimation was also addressed in this work.

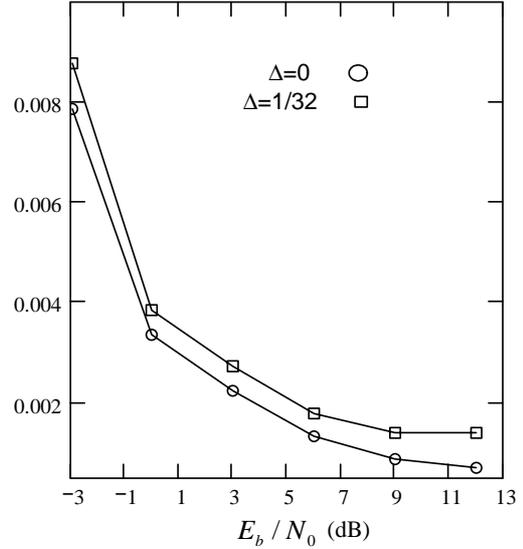


Fig. 6. Relative frequency offset estimation versus E_b/N_0

Analytical and simulation results have shown that timing offsets will create flatness in the variance profile making the search for the minimum point much harder. Hence, symbol timing has to be achieved first for accurate frequency estimation.

References

- [1] R. Nee and R. Prasad, *OFDM Wireless Multimedia Communications*. Boston: Artech House, 2000.
- [2] Radio broadcasting systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers, European Telecommunication Standard 300 401, 2 nd ed., European Telecommunication Standard Institute, Valbonne, France, 1997.
- [3] "Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television" European Standard EN 300 744 V1.1.2 (1997-08).
- [4] J. van de Beek, M. Sandell, and P. Börjesson, "ML estimation of timing and frequency offset in OFDM systems," *IEEE Trans. Signal Proc.*, vol 45, pp. 1800-1805, July 1997.
- [5] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Tans. Commun.*, vol. 42, no. 10, Oct. 1994.
- [6] E. Larsson, G. Liu, J. Li and G. Giannakis, "Joint symbol timing and channel estimation for OFDM based WLANs," *IEEE Commun. Lett.*, vol.5, pp. 325-327, Aug. 2001.
- [7] F. Tufvesson, M. Faulkner, P. Hoeher and O. Edfors, "OFDM time and frequency synchronization by spread spectrum pilot technique," in *proc. IEEE Commun. Theory Mini Conference*, 1999.
- [8] J. Proakis, *Digital Communications*, New York: McGraw-Hill, 1989.
- [9] S. Ross, *Introduction to Probability Models*, CA: Academic Press, 1997.
- [10] S. Bulumulla, S. Kassam, and S. Venkatesh, "A Systematic Approach to Detecting OFDM Signals in a Fading Channel," *IEEE Trans. Commun.*, Vol. 48, NO. 5, MAY 2000, pp. 725-728.
- [11] S. Ho Kim and S. Wu Kim, "Frequency-Hopped Multiple-Access Communications with Multicarrier On-Off Keying in Rayleigh Fading Channels," *IEEE Trans. Commun.*, Vol. 48, No. 10, pp. 1692-1701, October 2000.

A New Robust Frequency-Domain Decision Feedback Equaliser for DMT Modems

Daniel Franklin, Ian Burnett and Joe Chicharo

Faculty of Informatics
University of Wollongong
Wollongong 2522 NSW AUSTRALIA
E-mail: daniel_franklin@uow.edu.au

Abstract—Communications systems using Discrete Multi-Tone (DMT) modulation, such as those used in current wireless data networks (802.11a/g, HIPERLAN/2), often need to deal with time-varying channels. When sudden changes in channel impulse response occur in such systems, the classical adaptive equaliser scheme - Frequency-Domain Decision Feedback Equalisation (FD-DFE) - has difficulty in correctly tracking the changed channel conditions. In this paper we present a simple scheme which can detect and correct this problem, without seriously degrading throughput.

1. Introduction

DMT/OFDM systems are now widely deployed over a variety of static channels, such as ADSL and powerline modems, and dynamic channels such as digital audio/video broadcasting (ETSI DVB-T, DVB-C, DVB-S) and wireless networking systems (IEEE 802.11a/g, HIPERLAN/2). While the problem of equaliser design is relatively simple for the case of static channels, time-varying channels require some form of continuously adaptive equalisation.

In most DMT systems it is convenient to use Frequency-domain Equalisation (FEQ), since the demodulation process already involves a time-to-frequency domain transformation [1]. In this case, it is assumed that the frequency response of each subchannel is approximately flat, and thus individual subchannels may be equalised by a complex single-tap filter. DMT systems operating over a dynamic channel must constantly update their initial estimate of equaliser coefficients (obtained via transmission of a training sequence known to both the transmitter and receiver) to track changes in the channel impulse response (CIR). For OFDM/DMT systems, the most common scheme for tracking the channel is a form of Frequency-Domain Decision Feedback Equaliser (FD-DFE) [2], [3], [4], [5], [6]. This equaliser scheme works well when the channel changes relatively slowly. Unfortunately, there are a number of realistic scenarios where the channel can change rapidly and cause FD-DFE systems to fail.

2. FD-DFE Channel Tracking Problems

Since the FD-DFE relies on an estimate of the actual symbol transmitted, it can be thought of as

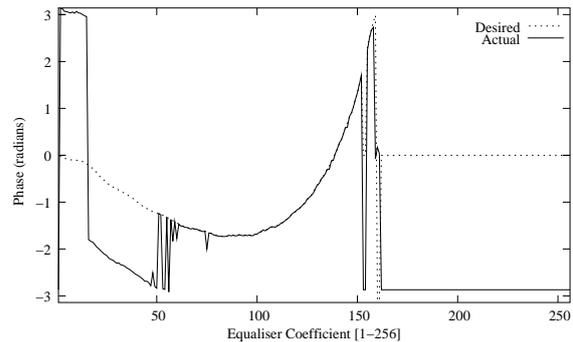


Fig. 1. Desired and Actual Equaliser Coefficients

a conventional FEQ system which estimates its own training sequence each time a data symbol is received. To increase immunity to noise, it is necessary to use a moving-average filter to retain memory of the past equaliser characteristics. Unfortunately, sustained bursts of noise or discontinuous changes in Channel Impulse Response (CIR) still result in a number of undesirable effects in an FD-DFE system.

If the phase response in a given subchannel has changed by more than $\pm 45^\circ$, the equaliser is very likely to converge on a value with the correct magnitude but with the phase in error by $\pm 90^\circ$ or $\pm 180^\circ$. Similar effects are caused by powerful noise bursts (in particular, impulsive noise events) which persist over a large number of symbols, since the equaliser coefficients may be greatly disturbed, and may settle back to values with an incorrect phase. This 'phase-flipping' is a result of the 4-way rotational symmetry which is present in the QAM constellations used in DMT/OFDM modems.

The desired and actual phase response of the equaliser in a DMT modem which is suffering from the phase-flipping condition is shown in Figure 1. In this case, the effect has been caused by sudden changes in the loading and termination of a simulated powerline OFDM modem. In practise, such changes often result from the connection, disconnection, or activation of inductive loads on the powerline.

Note that in the results shown in Figure 1, subchannels 153, 154 and 162-255 are unused due to unacceptable SNR. Subchannels 1 and 256 are also

unused (as they are the DC and Nyquist frequencies - while it is possible to use these subcarriers, it is simpler to avoid using them altogether). A phase error of approximately 180° is evident in subchannels 2-15, -90° in 16-50, 53, 54 and 56, while the rest of the (used) subchannels are correctly equalised.

3. Robust FD-DFE

We propose an architecture for frequency-domain DFEs which is able to detect and rapidly correct the problem of phase-flipping. In addition, our design has a number of desirable characteristics:

- Detection of the phase-flip problem requires no additional overhead (e.g. per-subchannel check-sums);
- Correction of phase-flipping does not require an interruption to modem operation, only a temporary, modest degradation of throughput;
- This correction can be made as fast or as slowly as desired, with the level of throughput degradation scaling gracefully.

In normal operation, the FD-DFE operates as a standard proportional decision feedback equaliser (exactly as described in [2]).

The algorithm of the new equaliser scheme is as follows:

- Transmit training sequence, train FEQ, tracking channel changes using FD-DFE;
- Compute the mean-squared error (MSE) between the output of the decision device and the output of the equaliser:

$$E = \frac{1}{N} \sum_{i=0}^{N-1} \left| Y_k Z_k - \hat{X}_k \right|^2 \quad (1)$$

Where N is the number of equaliser coefficients, Y_k is the k -th received symbol, Z_k is the k -th complex equaliser coefficient, and \hat{X}_k is the k -th detected symbol. This is a measure of how well the equaliser is tracking the current channel characteristics;

- Monitor the Bit Error Rate (BER) using the Forward Error Correction (FEC) built into the transmission;
- If the BER exceeds 2.5 times its long-term average (moving-average filter with a memory factor of 0.95) while the MSE is less than 2.5 times its long-term average, the equaliser is suffering from phase-flip;
- In this case, signal the transmitter to switch to rotating pilot mode - depending on the speed of adaptation desired, a small number of subcarriers transmit equaliser training symbols in each macrosymbol period.

- When the raw (pre-FEC) BER drops back to an absolute value of 5×10^{-2} , this implies that the phase-flipping condition has ended, therefore the transmitter is signalled to return to full-rate.

A complete block diagram for the system is shown in Figure 2.

The FD-DFE coefficient update function for under normal operation is given by

$$Z_{i+1}(k) = M_f Z_i(k) + (1 - M_f) \frac{\hat{X}_i(k)}{Y_i(k)} \quad (2)$$

where i is the symbol number and k is the subchannel index. The moving-average filters are simple IIR filters with a transfer function given by

$$H(z) = \frac{1 - M_f}{1 - M_f z^{-1}} \quad (3)$$

where M_f is the memory factor. When phase-flipping has been detected, Equation 2 changes to

$$Z_i(k) = \frac{T_i(k)}{Y_i(k)} \quad (4)$$

where $T_i(k)$ is the training symbol in the i -th iteration on the k -th subchannel, for those values of k corresponding to subchannels bearing training symbols.

The various factors and constants in this algorithm have been determined by experimentation; the threshold values constitute a hysteresis and are not critical. It should also be noted that phase-flip tends to result in a much higher BER than other effects.

The rotating-pilot phase of the algorithm allows the training overhead to be smoothly scaled up from zero training micro-symbols to 100% training micro-symbols, depending on the desired speed of training.

4. Results

The new FD-DFE has been tested under a various channel-transition and noise conditions. The equaliser is shown to converge rapidly after a disturbance, when a simple DFE would be unable to maintain equalisation. We examine three distinct CIR transition cases. The modem is a DMT system with 256 subchannels and 512 bits/macrosymbol for a data rate of 2 Mb/s. The sample rate is 2 MSamples/s.

A. Simple Abrupt CIR Transitions

In the 221st symbol period, the CIR abruptly changes in time-scale, the final impulse response being 20% shorter than the initial impulse response.

Figure 3 shows the BER resulting from the conventional and robust FD-DFE designs. In this case, when phase-flipping has been detected, 16 equaliser coefficients at a time are ripple-trained, resulting in only a 10% reduction in channel capacity. In contrast, the conventional FD-DFE fails to converge on the

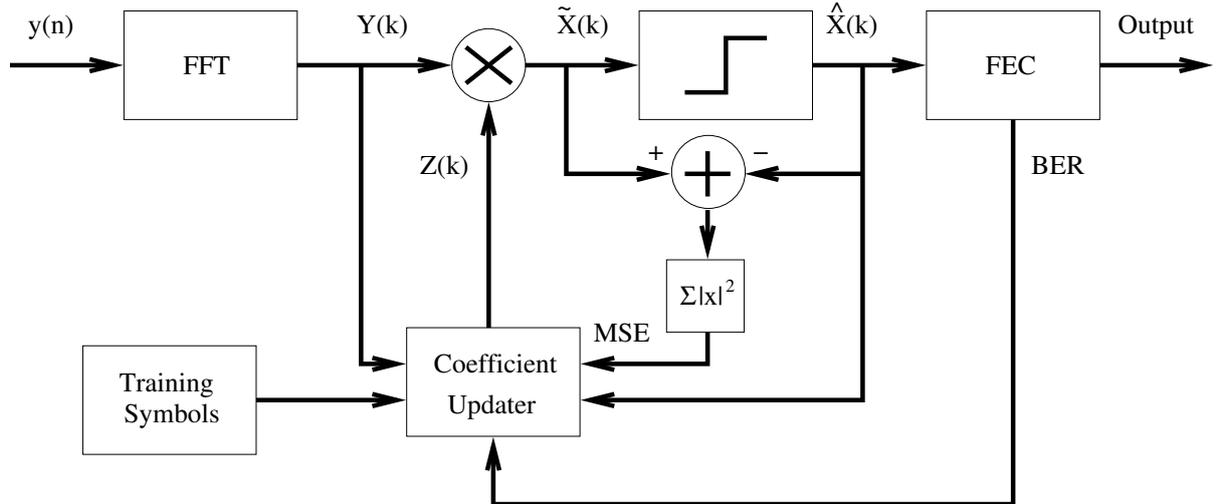


Fig. 2. Block diagram of robust FD-DFE

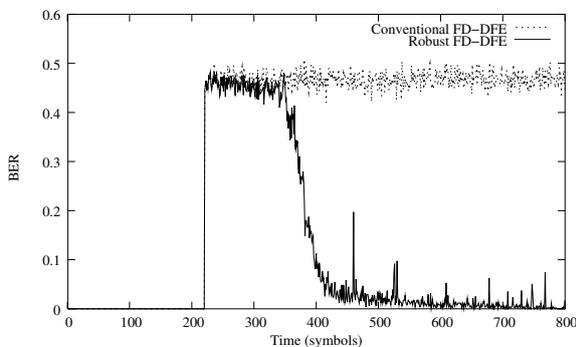


Fig. 3. Bit Error Rate, for conventional and robust FD-DFE, Long-to-Short CIR transition

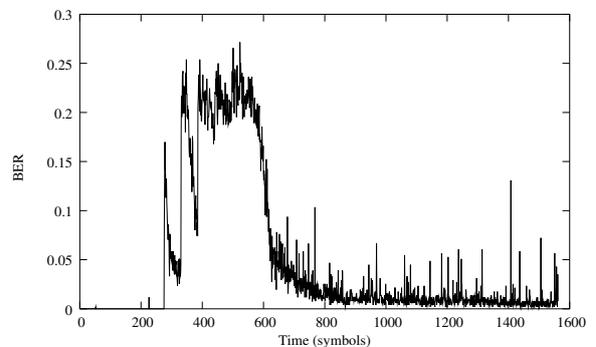


Fig. 4. Bit Error Rate for robust FD-DFE, smooth CIR transition

correct equaliser coefficients, resulting in the worst-case BER of 0.5 (i.e., no information could be extracted from the signal). For the new system, the BER rapidly drops back from the worst-case value of 0.5 to 3.2×10^{-4} . In comparison, under identical conditions, the conventional FD-DFE would need to interrupt data transmission for re-training, which, using the same modem parameters, would take more than twice as long to yield the same BER.

The small residual BER results from the bit allocation no longer being optimal. This is effectively a degradation in SNR, as can be seen by the small increase in MSE after the disturbance. The residual BER may be corrected independently either using some form of water-pouring algorithm or a BER-based bit allocation scheme such as the Inverse Robin Hood algorithm [7].

Reversing the order of the change in impulse response results in a similar outcome. Once again, convergence is achieved, although the residual BER is much higher at 1.3×10^{-1} . However, given the now sub-optimal bit allocation table, this is the best achievable performance under the new channel conditions. By contrast, the an unaided FD-DFE algorithm is

unable to cope with the change, resulting in a sustained BER of 0.5 after the transition. Full re-training of the equaliser is required in this case.

B. Smoothly Time-Varying Channels

Figure 4 shows the change in BER as the CIR changes gradually between those used in Section 4-A; this time in five intermediate steps instead of a single jump. Since the transitions are close together, the spikes in MSE and BER introduced by each small change tend to accumulate. Thus, although a single small transition of this nature will not, in isolation, result in phase-flipping, a second transition will result in a BER high enough to start development of the condition. Approximately 600 symbols after the start of the transition, the re-training algorithm has succeeded in restoring an acceptable BER.

For cases where the CIR changes more slowly, this accumulation does not occur and can be dealt with adequately by conventional FD-DFE implementations.

C. Reversible Channel Changes

Figure 5 shows the BER when the robust modem operates over a channel (with minimal noise) which undergoes two abrupt changes in CIR. The BER spikes

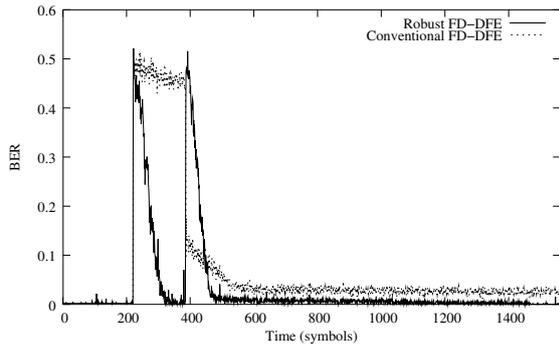


Fig. 5. Bit Error Rate for robust and conventional FD-DFE, with two CIR changes

momentarily at each change, then rapidly drops back down. By comparison, the unaided FD-DFE under identical conditions is unable to recover from the initial change. The new algorithm is thus able to obtain a greater period of useful transmission under these conditions.

5. Noise Tolerance

To establish its robustness, the model needs to be tested under realistic noise conditions. All noise simulations have been conducted with a mixture of AWGN and impulsive noise (with an identical noise PDF in each case), using the channel model presented in [8].

A. Effects of Channel SNR

The new algorithm is insensitive to noise up to a certain threshold, after which performance degrades significantly. For SNR better than 12 dB, convergence occurs within a predictable timeframe, taking approximately 200 symbols from the time of the disturbance for the DFE to completely stabilise. However, as the SNR deteriorates to below 10 dB, the algorithm is unable to converge at all within the simulation time. At this point, the algorithm offers no advantage over the conventional FD-DFE since full re-training would be required in both cases.

B. Transient Noise Bursts

Although SNR is a limiting factor in the effectiveness of this algorithm, short-term noise bursts only cause a temporary degradation in BER. Figure 6 shows BER curve of a modem with the robust FD-DFE system subjected first to a sudden change in channel conditions (as before) followed by a sudden noise burst (lasting for approximately 40 symbols). The long-term average SNR is 20 dB, which drops to 0 dB during the noise burst.

At the onset of the burst, the BER immediately increases. Despite the presence of a severe noise burst, the MSE does not increase as much as at the point where the CIR has changed. This characteristic allows the phase-flipping phenomenon to be distinguished from noise burst events. Once the noise burst ends,

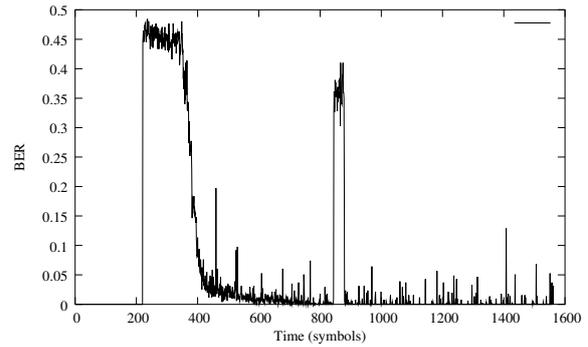


Fig. 6. Bit Error Rate for robust modem, with very powerful noise burst at $t = 850$ symbols

both the MSE and BER both rapidly drop back to the previous levels, demonstrating that the new algorithm is able to handle extreme short-term noise conditions very well. Under the same conditions, a conventional FD-DFE will eventually converge, although it will take approximately twice as long to do so.

6. Conclusion

In this paper we have presented a frequency domain DFE scheme which exhibits excellent robustness in rapidly time-varying channels subject to noise, provided that the SNR is better than 12 dB. It has also been shown to operate correctly on more slowly time-varying channels, and behaves well when the channel is affected by serious short-term noise events. The robust FD-DFE scheme also allows for partial re-training of equaliser coefficients, and is therefore able to avoid a complete shut-down of the connection during this interval.

References

- [1] John A. Bingham, "Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come," *IEEE Communications Magazine*, May 1990.
- [2] Jukka Rinne and Markku Renfors, "Equalisation of Orthogonal Frequency Division Multiplexing Signals," in *Proceedings of IEEE International Conference on Communications*, Nov. 1994, vol. 1, pp. 415–419.
- [3] Martin V. Clark, "Adaptive Frequency-Domain Equalization and Diversity Combining for Broadband Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1385–1395, Oct. 1998.
- [4] Steffen Trautmann, Tanja Karp, and Norbert J. Fliege, "Frequency Domain Equalization of DMT/OFDM Systems with Insufficient Guard Interval," in *ICC 2002*, 2002, vol. 3, pp. 1646–1650.
- [5] Martin V. Clark, "Adaptive Frequency-Domain Equalization and Diversity Combining for Broadband Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, Oct. 1998.
- [6] Kostas Berberidis and Jacques Palicot, "A Frequency-Domain Decision Feedback Equalizer for Multipath Echo Cancellation," in *Proceedings of the Global Telecommunications Conference*, Nov. 1995, pp. 98–102.
- [7] Daniel Franklin and Ian Burnett, "A New Adaptive Algorithm for Bit Allocation in DMT Modems," in *Proceedings of the 2nd WSEAS Int. Conf. on Signal, Speech and Image Processing (ICOSSIP 2002)*, Aug. 2002, vol. 1.
- [8] Daniel Franklin and Jiangtao Xi, "An Improved Channel Model for ADSL and VDSL Systems," in *2000 International Conference on Communications Technology Proceedings*, Aug. 2000, vol. 1, pp. 30–33.

A Novel Detection Technique for Uplink Power Control

Kostas Kastamonitis, Boris Grémont, Misha Filip

Microwave Telecommunication Systems Research Group

University of Portsmouth, Department of Electronic & Computer Engineering,

Anglesea Road, Portsmouth, PO1 3DJ, U.K. Tel: +44(0)23-92846028,

Email: kkastas@ee.port.ac.uk, {[boris.gremont](mailto:boris.gremont@port.ac.uk), [misha.filip](mailto:misha.filip@port.ac.uk)}@port.ac.uk

Abstract

We present a novel detection technique for uplink power control that avoids frequency scaling. The proposed technique is applicable to real-time detection/compensation of rain attenuation. The obtained results are proven to be extremely accurate compared to those using classical frequency scaling models.

1 Introduction

The atmospheric fading becomes very significant for satellite systems operating at higher frequency bands especially at Ka/V-band. This results in a degradation of the Quality-of-Service (QoS) and/or the availability of the satellite communications.

In this paper, we focus on a wideband outbound satellite link using adaptive up-link power control (ULPC) in order to counteract for the up-link rain attenuation, [1].

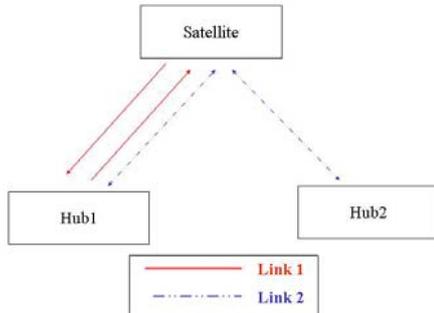


Figure 1. Return broadcasting satellite topology

The scenario in Figure 1 consists of the Hub1 station transmitting at 50 GHz. The signal is then down-converted and transmitted back to the earth stations at 40 GHz. We assume that Hub1 and Hub2 lie in the same downlink footprint so that the two hubs can receive all the bursts they transmit after a round trip time. We call the Hub1-Sat-Hub1 transmission the *loop-back* link for link1. We are interested in applying ULPC on the up-link Hub1-to-Satellite. We will focus on the loop-back link Hub1-Sat-Hub1 (see solid line Figure 1). Hub2 is primarily used to feed Hub1 periodically with a particular set of measured fading data. These data will then be used in order to detect the uplink attenuation for link1, [2]. We assume that the two hub stations are strictly identical. In particular both

can measure the *CNR* ratios, CNR_1 and CNR_2 on the loop-back links for Hub1 and Hub2 respectively. They can also measure respectively the downlink attenuation a_{d1} and a_{d2} from the 40 GHz beacon signal. If effective, ULPC should compensate for the rain attenuation affecting the uplink of link1. We note that Hub1 can only:

1. estimate the total CNR_1 variations on the received modulated down-link from Hub1
2. measure the downlink attenuation, a_d using a beacon receiver
3. estimate the sky-noise temperature using the co-located radiometer
4. Hub1 periodically receives updates of the total CNR variations, (cnr_2) and downlink attenuation, a_{d2} (Figure 1). cnr_2 is measured by Hub2 for the link Hub1-Sat-Hub2. Thus it shares the same uplink as that of the loop-back link1.

The co-located radiometer is introduced primarily as a means of improving the robustness of the feeder link from Hub1 to satellite. The independent passive monitoring of the up-link with the radiometer will prove useful:

- 1) in case there is a problem with the beacon sub-system
- 2) for some necessary pre-processing (namely the separation of rain attenuation from other effects) and the online determination of the clear-sky level, [4].

For these reasons, the utility of the radiometer outweighs its cost since its main function is to make sure that ULPC can be applied even in case when beacon measurements are unavailable, [4].

2 Signal Quality Detector

We describe a simple subsystem for measuring the received *CNR* on a modulated QPSK symbol stream. This subsystem is critical for the implementation of the final detection algorithm to be described in Section 4.

The overall carrier-to-noise ratio estimation is performed using the *second and fourth moments* (M_2M_4) estimator, [5]. The primary advantage of the in-service M_2M_4 estimator is that carrier phase synchronisation is not required. This simplifies the hardware complexity of the estimation unit.

Since the demodulated output, y_n of the satellite channel is complex, the derivation of the complex detector in [6] is chosen, given by:

$$\left(\hat{CNR}\right) = \frac{\sqrt{2M_2^2 - M_4}}{M_2 - \sqrt{2M_2^2 - M_4}}, \quad (1)$$

where M_2 and M_4 are the second and fourth moments of the estimator, respectively given by:

$$M_2 \approx \frac{1}{N_{sym}} \sum_{n=0}^{N_{sym}} |y_n|^2, \quad (2)$$

and

$$M_4 \approx \frac{1}{N_{sym}} \sum_{n=0}^{N_{sym}} |y_n|^4, \quad (3)$$

The estimation accuracy of the M_2M_4 estimator is proportional to the depth N_{sym} of the processed symbols, y_n , in equations (2) and (3).

We have simulated this estimator for QPSK demodulation. We have considered that 1000 symbols are processed for evaluating the moments in (2) and (3). A comparison of true CNR (blue trace) with achieved CNR (red trace) is illustrated in Figure 2.

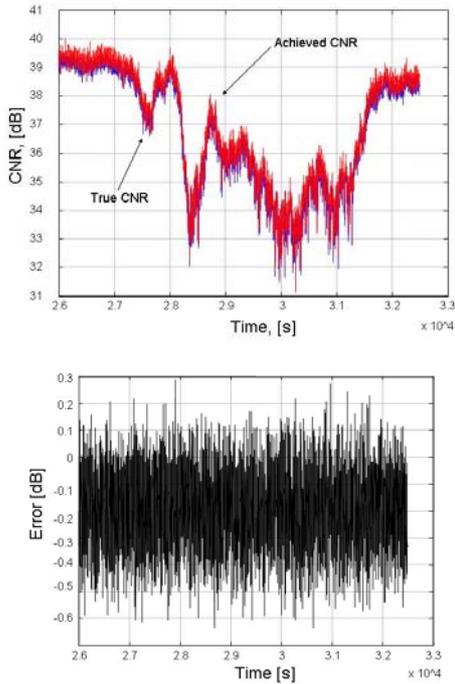


Figure 2. Comparison of results using the M_2M_4 estimator

The mean estimated error in Figure 2 is approximately equal to 0.2011 dB and its variance is 0.0146 dB. This indicates that the chosen estimation scheme gives reasonable results.

3 Downlink Beacon Detector

A novel online algorithm of a beacon/radiometric detection has been presented in [2]. The algorithm is capable of estimating every second the gaseous attenuation. It also corrects for the rain scattering blindness of the collocated radiometer. As the radiometer cannot see scintillations, it is possible after suitable processing to subtract the beacon from radiometric measurements, [3]. Figure 3 illustrates the online clear-sky correction of the downlink beacon attenuation. The difference between the original beacon time series (blue trace) and the corrected beacon time series (red trace) is quite minimal, especially during deep fade. This is however quite crucial for an accurate estimate of scintillation with zero mean, [3],[4]. Furthermore, Figure 4 illustrates the correction of the estimated radiometric measurements (red trace) with respect to the original radiometric measurements (blue trace). It can be seen that the dynamic range of the radiometer has significantly increased. This adds to the actual robustness of the beacon detection unit.

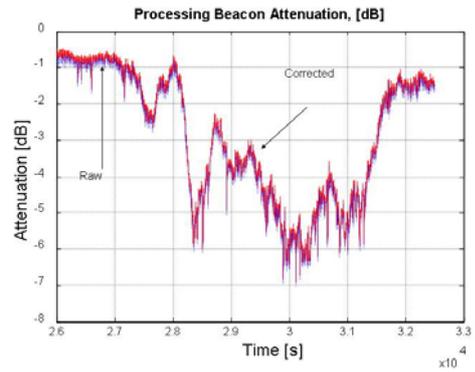


Figure 3. On-line clear sky beacon signal correction

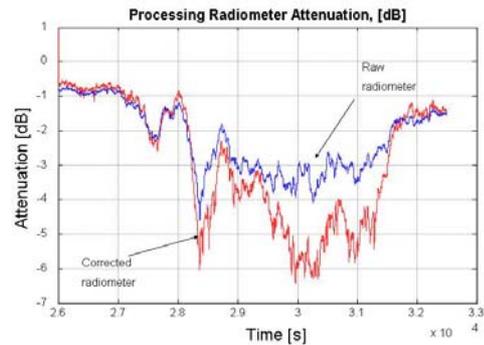


Figure 4. Comparison of the estimated radiometric attenuation

4 Estimating Uplink Attenuation

In order to estimate the up-link rain attenuation on Hub1, the relevant parameters, collected from the two links (red and blue trace) in Figure 1, are needed. In particular, Hub1 measures its own loop-

back CNR , $\hat{z}_1 = 1/(\widehat{cnr})_{t1}$ and the downlink attenuation, $\hat{y}_1 = \hat{a}_{d1}$ in ratio. In addition, Hub2 periodically feeds Hub1 with information about its own measured values of CNR , $\hat{z}_2 = 1/(\widehat{cnr})_{t2}$ and downlink attenuation $\hat{y}_2 = \hat{a}_{d2}$. A communication protocol has to be established between the two Hub stations for the periodic and reliable exchange of such information. It is assumed that the overall carrier-to-noise ratio over link1 can be approximated from [1] by:

$$\hat{z}_1 = \frac{1}{A \cdot x_1} + \frac{1}{B \cdot \hat{y}_1}, \quad (4)$$

The hats denote quantities that are measured by the Hub. In a similar way, the overall carrier-to-noise ratio over link2 is given by:

$$\hat{z}_2 = \frac{1}{A \cdot x_1} + \frac{1}{B' \cdot \hat{y}_2}, \quad (5)$$

The parameters $A = (cnr)_{u1}$, $B = (cnr)_{d1}$, denote the up and down clear-sky CNR s on the link Hub1-Sat respectively and $B' = (cnr)_{d2}$ denotes the down-link clear-sky CNR on Sat-Hub2 link. x_1 is the up-link attenuation and needs to be determined. A method to measure the values of A and B is first described. Once A and B are determined (i.e. once we have a reasonable model of the clear sky link budget of loop-back link (Hub1-Sat-Hub1), we then estimate the up-link attenuation x_1 as this is required for the deployment of the ULPC fade mitigation technique.

4.1 Determination of A and B

Clearly, the downlink clear-sky CNR , $B' = (cnr)_{d2}$ in Equation (5) from Sat-Hub2 is different from the clear-sky, $B = (cnr)_{d1}$ from Sat-Hub1. However, it is reasonable to assume that $B' \approx B$, provided the two hubs are not too far away from each other. If this is indeed the case, the parameter B is easily determined by combining Equations (4) and (5) giving:

$$B = \frac{1}{\hat{y}_1 \cdot \hat{y}_2} \cdot \frac{\hat{y}_2 - \hat{y}_1}{\hat{z}_1 - \hat{z}_2}, \quad (6)$$

Once B is known, we can calculate the clear-sky up-link CNR A by considering the measured clear-sky overall CNR , \hat{z}_{01} on the Hub1-Sat-Hub1 link:

$$\hat{z}_{01} = \frac{B + A}{A \cdot B}, \quad (7)$$

Rearranging Equation (7), we get:

$$A = \frac{B}{B \cdot \hat{z}_{01} - 1}, \quad (8)$$

4.2 Estimating x_1

The calculated values of A and B can finally be inserted in Equation (4) to yield the estimated up-link rain attenuation, x_1 as:

$$x_1 = \frac{B \cdot \hat{y}_1}{A \cdot (B \cdot \hat{y}_1 \cdot \hat{z}_1 - 1)}, \quad (9)$$

This completes the process for estimating x_1 without using the scaling factor of rain attenuation. An example of the estimated up-link attenuation can be seen in Figure 5. It can be seen that the algorithm gives a close to perfect estimation for the up-link rain attenuation. The accuracy of the up-link estimation will depend strongly on (i) the accurate estimation of the overall CNR for both links and (ii) the FMT communication protocol between the two Hubs and (iii) the precise estimation of both the beacon downlink attenuated signals for Hub1 and Hub2.

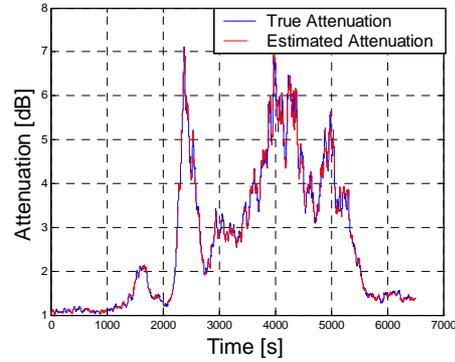


Figure 5. Estimated and true uplink rain attenuation

5 Operational Constraints

We have assumed that the two Hub stations are located in a close proximity to each other so that their down-link CNR s are equal, (i.e. $B \approx B'$). We will identify the exact distance $d_{1,2}$ between the two Hubs (see Figure 6) that satisfies this assumption.

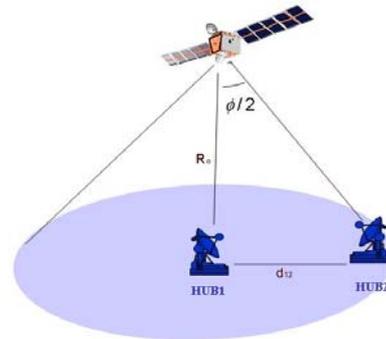


Figure 6. Loop-back geometrical assumptions

The impact of the differential Free Space Loss is irrelevant to this study so, we will only consider the impact of the antenna pattern. For this, we consider a Gaussian antenna, the gain of which is given by,[7]:

$$G(\phi) = 10 \log_{10} \left[\eta \left(\frac{\pi D}{\lambda} \right)^2 \exp \left(-2.76 \left(\frac{\phi}{\phi_{3dB}} \right)^2 \right) \right] \quad (10)$$

where, $D = 2m$ is the satellite antenna diameter and $\eta = 0.7$ the antenna efficiency. ϕ is the angle of boresight of the antenna, λ is the wavelength and $\phi_{3dB} = 70 \cdot c / (f_{down} \cdot D)$ [deg] is the 3dB beamwidth of the antenna, [8].

We assume that Hub1 lies at the centre of the satellite footprint (see Figure 6). Then, the maximum satellite antenna gain, G_{max} is at the direction of Hub1 (i.e. direction of the antenna boresight, $\phi = 0$ deg). Then, the distance, $d_{1,2}$ in Figure 6, can be calculated as a function of the differential gain, ΔG as:

$$\Delta G = G_{max} - G(\phi) \text{ [dB]} \quad (11)$$

where G is the satellite antenna gain for different angles, ϕ . We can represent ϕ as a function of the antenna gain by inverting Equation (10):

$$\phi(G) = \phi_{3dB} \sqrt{\frac{\ln(G\lambda^2 / (\eta\pi^2 D^2))}{(-2.76)}} \text{ [deg]} \quad (12)$$

Then, using simple trigonometry (see Figure 6), we can calculate $d_{1,2}$ as a function of the angle ϕ [deg] using:

$$d_{1,2} = R_o \cdot \tan(\phi / 2) \text{ [km]} \quad (13)$$

where $R_o = 35786$ km is the geostationary satellite altitude.

Figure 7 shows $d_{1,2}$ as a function of the differential antenna gain, ΔG . It can be seen that the maximum distance between Hub1 and Hub2 is 178 km. Therefore, Hub2 can lie in a circle centre at Hub1 with a maximum radius of 178 km.

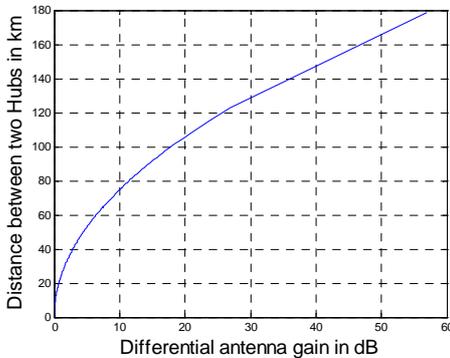


Figure 7. Distance between Hub1 and Hub2

6 Conclusions

We have described a novel ULPC method to counteract rain attenuation on the uplink of a satellite communication link. The scenario considers high frequency feeder links at V-Band. The main advantage of the proposed ULPC scheme is that it by-passes the use of scaling factor models, [2]. This adds stability and precision since frequency scaling models of rain attenuation are known to be inaccurate. The proposed technique can be used when two or more identical Hub feeder stations try and counteract rain using ULPC. Finally, the distance between the two Hub stations has been calculated.

7 References

- [1]. Ippolito L.J., “*Radiowave Propagation in Satellite Communications*”, Van Nostrand Reinhold Comp., New York, 1986.
- [2]. Kastamonitis K., Grémont B., Filip M., “Loop-Back Detection Technique for Uplink Power Control”, *22nd AIAA Intern. Comm. Sat. Systems Conf. and Exhibit*, 2004-3198, 9-12 May 2004, Monterey California.
- [3]. Catalan C., “*An Investigation of Simultaneous Fade Dynamics and Radiometry in a 40 GHz Satellite Down-Link*”, PhD Thesis, University of Portsmouth, January 2000.
- [4]. Kastamonitis K., Grémont B., Filip M., “On-line Extraction of Scintillations for Satellite Links with Up-Link Power Control”, *1st Int. Workshop COST280*, PM3-015, Qinetiq, Malvern, July 2002.
- [5]. Benedict T. R., Soong T. T., “The Joint Estimation of Signal and Noise from the Sum Envelope”, *IEEE Trans. Inform. Theory*, Vol. IT-13, pp: 447-454, July 1967.
- [6]. Pauluzzi D. R., Beaulieu N. C., “A Comparison of SNR Estimation Techniques for the AWGN Channel”, *IEEE Trans. Comm.*, Vol. 48, No. 10, October 2000, pp: 1681-1691.
- [7]. Gagliardi R. M., “*Satellite Communications*”, 2nd Edition, New York, Van Nostrand Reinhold, 1991.
- [8]. Maral G., Bousquet M., “*Satellite Communications Systems, Systems Techniques and Technology*”, 4th Edition, John Wiley & Sons, 2002.

A Simple Efficient Adaptive Algorithm For Jitter Reduction in TDM Over Packet Switching Networks

Liu Erwu, Shen Gang, Lin Changsheng, Jin Shan, Zhang Kaibin, Gui Luoning
Research & Innovation Center,
Alcatel Shanghai Bell Corp, Shanghai, 201206, P.R.China
Email: Erwu.liu@alcatel-sbell.com.cn

Abstract:

Packet Switching Networks (PSN) are being increasingly used to transport TDM services. Jitter reduction or absorption is a key issue in TDM over packet switching networks (TDMoPSN). To achieve good jitter reduction performance in real network environments where network delay and jitter keep changing all the time, one has to implement adaptive jitter control. This research proposes a simple efficient adaptive algorithm to be used in TDMoPSN networks for efficient jitter absorption, ordered packet delivery and delay spike detection. Finally, we evaluate the performance of the proposed algorithm. Simulations shown that the proposed algorithm works very well in different network scenarios. Compared with existing histogram-based methods, the proposed one is simple and efficient in that it only needs packet-late indication and maximum/minimum filling level for dynamic adjustment purpose.

Keyword: TDMoPSN, Adaptive jitter buffer, Delay spike.

1. Introduction

Recently, important progress has been done on TDMoPSN [1][2][3]. Though there are many other reasons to migrate TDM services to packet switching networks, certain key problems must be solved first. The paper is organized as follows, in Section II, Delay jitter issue in TDMoPSN is outlined and some techniques are described. Section III describes our packet-late indication based (PLIB) adaptive jitter buffer techniques. Related simulations for different distribution delay models are presented in Section IV. Section V concludes the paper.

2. Delay Jitter

Delay jitter, is mainly interpreted as difference in the total end-to-end delay of two consequent packets in the flow. To support the transmission of TDM traffic over PSN, receiver should remove packet delay jitter introduced by PSN. Common method for jitter removal is to use receiver jitter buffer. Jitter buffer operates by introducing additional end-to-end delay, holding packets that arrive before their scheduled playback time, and discarding those that arrive late. To achieve ideal playout, each packet must be buffered for as much time as the longest delay experienced by any packet in the packetized TDM stream.

There are three main techniques using receiver jitter buffer for jitter absorption known today: fixed length jitter buffer, constantly increasing jitter buffer[4] and adaptive jitter buffer.

3. Proposed Algorithm

Though widely implemented in TDMoPSN product, fixed length jitter buffer often introduce larger delay than preferred. In 2002, Zarlink[4] implemented constantly increasing jitter buffer in its product to achieve better jitter absorption performance, but constantly increasing jitter buffer can only increase its size to the maximum jitter experienced but never decrease its size even though the network jitter decreases as it is still not intelligent enough to track the keep-changing delay jitter in real network environments. For real network environments where network delay and jitter keep changing, adaptive jitter buffer has to be implemented.

[5] proposed an *queue monitoring* (QM) based adaptive jitter buffer, whose fundamental idea is that, instead of measuring delay differences, one can

directly measure the impact of the delay jitter on the receiver by observing the occupancy of the jitter buffer over time. Common QM-based methods usually need lots of counters and thresholds. Moon et al [6] proposed a *histogram*-based adaptive algorithm, in [6], the delays of the last K packets are recorded and the distribution of delays is updated with each incoming packet. The frequency of each delay is maintained in a histogram. The delay distribution is computed using a cumulative sum of the frequencies. The algorithm calculates a given percentile point of the delay in the distribution function and uses it as a playout delay. Histogram-based method needs lots of historical packets delay information for good performance.

In order to reduce computational complexity, we present a very simple efficient algorithm. In the algorithm, we use packet-late indication to adjust jitter buffer. By observing the occupancy of the jitter buffer and packet-late indication over time, our proposed algorithm implements the adaptive jitter buffer in a very simple efficient way for jitter buffer adjustment purpose.

The main idea of PLIB algorithm is depicted in Fig.1. When network jitter increases, there must be some packets arriving too late to be delivered, the PLIB algorithm uses this indication to increase jitter buffer filling level from A to B. When network jitter decreases, the maximum/minimum filling level in next adjustment interval must fall into the (minimum filling level, maximum filling level) in last adjustment interval, the PLIB algorithm uses this indication to decrease filling level from B to a configurable value C, and will quickly increase to D.

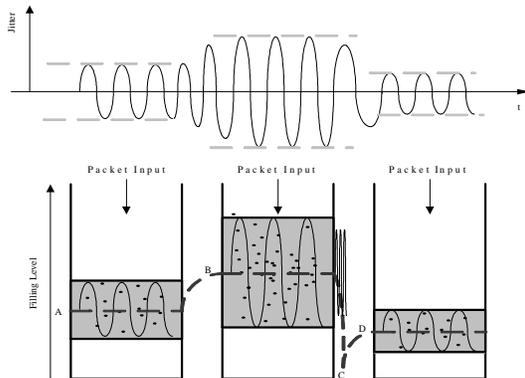


Fig.1. PLIB adjustment algorithm

3.1 PLIB Algorithm

The PLIB algorithm is composed of 4 major processes: *Packet Arrival Process* (PAP), *Packet Leave Process* (PLP), *Buffer Monitoring Process* (BMP) and *Buffer Adjustment Process* (BAP).

3.1.1 PAP Procedure

We describe PAP procedure in Pseudo-C code as follows,

```

Get_pk_from_network(pk);
Get_JB_filling_level();
Update_curr_min_max_filling_level(); /*
update current minimum/maximum PLIB filling level
for current adjustment interval; */
Insert_pk(pk); /* inserts the jittered,
out-of-order packet into PLIB jitter buffer */

```

3.1.2 PLP Procedure

We describe PLP procedure in Pseudo-C code as follows,

```

If(Get_pk_by_sn(next_sn,pk)==FALSE) /* packet
with next serial number not in the buffer */
Send_pk_late_indication(); /* send
packet_late indication to BAP process to
increase buffer */
Else Remove_and_send(pk); /* remove
and send the packet */

```

3.1.3 BMP Procedure

BMP monitors minimum/maximum filling level of PLIB jitter buffer in a constant interval time. We describe BMP procedure in Pseudo-C code as follows,

```

if (last_min < curr_min && last_max >
curr_max) /* current (minimum, maximum) is in the
range of last (minimum, maximum) */
send_decrease_indication(); /* send
indication to BAP process to decrease buffer filling
level */
Else /* update last and current
minimum/maximum jitter buffer filling level */
Update_last_min_max_filling_level();
Update_curr_min_max_filling_level();

```

3.1.4 BAP Procedure

We describe BAP procedure in Pseudo-C code as follows,

```

If(pk_late_indication)
    Insert_one_pk(); /* insert one packet into
    the buffer to increase filling level by one */
Else if(decrease_indication)
    JB_decrease_to(filling_level_threshold);
Else no action;

```

4. Simulations

We modeled the PLIB algorithm in OPNET[6]. PLIB adaptive jitter buffer is modeled as follows in OPNET. The model builds upon Ethernet MAC model. The packet transmission delays are assumed to follow Geometric distribution to model discrete network delay.

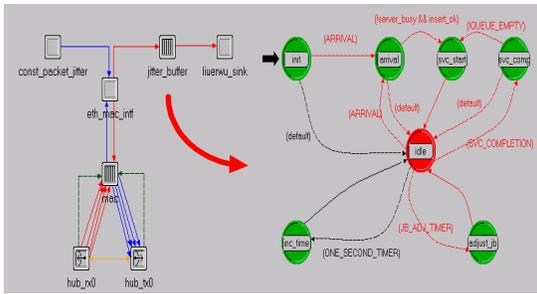


Fig.2. PLIB Adaptive Jitter Buffer Model implemented in OPNET

The discrete model is based on a model developed by Bolot [8], stating that the inter-arrival process of packets emitted at regular intervals \mathbf{d} fits with a Geometric distribution:

$$\text{Pr ob}\{t_r^{i+1} - t_r^i = k \cdot \mathbf{d}\} = p \cdot (1 - p)^k$$

Where t_r^i is the transmission delay of the i -th packet, k is non-negative integer, parameter p represents network QoS.

Due to the paper length limitation, we show here only the simulation results for Geometric distribution delay model in the following sections. The simulation results for continuous delay model (Normal and Pareto distribution delay) will appear in the extension version of the paper.

4.1 Simulations Configurations

For all simulations, the mean network delay is 0.5s, packet size is 500bytes, the sender sends packet at rate of 400pks/s. To verify PLIB algorithm, we evaluate its performance under real network environment with varying jitter and under network environment with delay spikes.

4.1.1 Real network environment

with varying jitter

Network jitters in Fig.3 are modeled by changing QoS over time to represent real network environments. Fig.2 also shows the simulation results for such network environment.

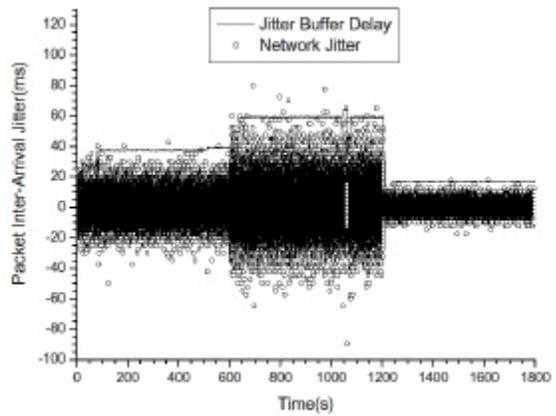


Fig.3. PLIB performance for network environment with varying jitter

We conclude that PLIB algorithm works very well under such network environments.

We then evaluated our PLIB algorithm in real network scenario. The simulation result is shown in Fig.4. We can see from Fig.4 that PLIB algorithm is robust enough to adapt delay estimates in the presence of spikes. As a comparison, Fig.5 is the simulation result for histogram-based method copied from [9]. We can conclude from Figures 4 and 5 that the PLIB algorithm proposed by us performs as well as, or better than existing histogram based adaptive algorithm as it usually needs to store lots of historical packets delay information for good performance. In fact, we can see that our method produce smaller end-to-end delay.

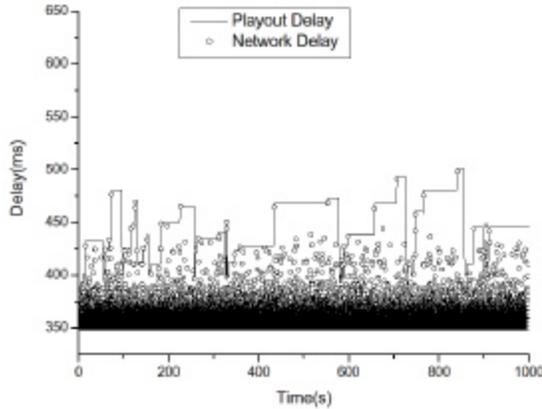


Fig.4. PLIB-based method

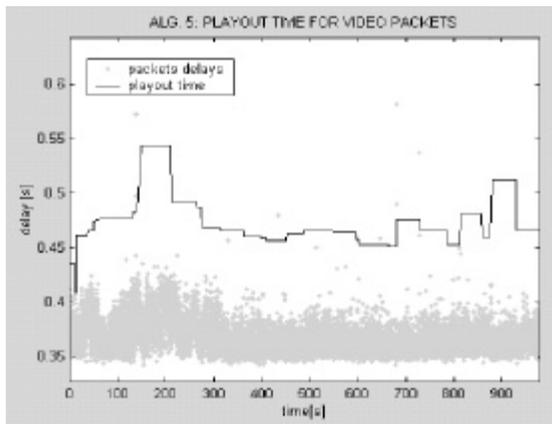


Fig.5. Histogram-based method

5. Conclusions

A very simple adaptive algorithm has been presented to effectively track and absorb the keep-changing delay jitter in real networks environment. Simulation results show that the proposed PLIB algorithm helps improve quality of service of TDM services in packet switching networks in that the algorithm can achieve very low end-to-end jitter, and can also guarantee ordered packet delivery even under networks with delay spikes. Due to the paper length limitation, we show here the PLIB algorithm performance for Geometric distribution delay model (discrete model). In fact, we have verified that the PLIB algorithm can still work very well for Normal and Pareto distribution model (continuous model).

References:

[1]. T. Johnson, TDM Circuit Emulation over Packet

(CEP-TDM), work in progress, February 2003, draft-ietf-pwe3-tdm-00.txt.

- [2]. M. Riegel, Requirements for Edge-to-Edge Emulation of TDM Circuits over Packet Switching Networks (PSN), work in progress, December 2003, draft-ietf-pwe3-tdm-requirements-01.txt
- [3]. MEF document, Circuit Emulation Service Definitions, Framework and Requirements in Metro Ethernet Networks v3-6, work in progress, July 2003
- [4]. Zarlink Semiconductor, MT90880/1/2/3 TDM to Packet Processors Data Sheet, December 2002.
- [5]. D-L. Stone, and K. Jeffay, An empirical study of a jitter management scheme for video teleconferencing, *Multimedia Systems*, 1995, vol.2, no.2.
- [6]. S-B. Moon, J. Kurose, and D. Towsley, Packet Audio Playout Delay Adjustment: Performance Bounds and Algorithms', *Multimedia Systems*, Jan 1998, vol.6, pp. 17-28.
- [7]. OPNET Modeler Documentation 9.0.A, OPNET Technologies, Inc., Bethesda, Maryland, Sep 2002.
- [8]. J-C. Bolot, Characterizing end-to-end packet delay and loss in the Internet. In *Proceedings of ACM SIGCOMM '93*, pp.289-299, September 1993.
- [9]. M. Narbutt, and L. Murphy, Adaptive Playout Buffering for Audio/Video Transmission over the Internet, *Proc. of the IEE 17th UK Teletraffic Symposium*, Dublin, Ireland, May 16-18, 2001, p. 27/1-27/6.

Adaptive MF-TDMA: Burst Length Control as a Rain Fade Countermeasure

Eleni Noussi, Boris Grémont, Misha Filip

Microwave Telecommunication Systems Research Group

University of Portsmouth, Department of Electronic & Computer Engineering,

Anglesea Road, Portsmouth, PO1 3DJ, U.K. Tel: +44(0)23-92846028,

Email: noussie@ee.port.ac.uk, {boris.gremont, misha.filip}@port.ac.uk

Abstract

Satellite services are migrating towards the higher part of the spectrum, where implementation of fade countermeasure becomes necessary. This paper proposes a methodology on how to design a Fade Mitigation Technique (FMT) suitable for application to Ka/V band Multi-Frequency TDMA networks. The technique provides compensation for rain attenuation at the expense of capacity. The attenuation protection as a function of spare resource (extra timeslots) is investigated. Both the advantages and costs of its deployment are demonstrated, with the intention to meet the user-specified Quality of Service requirements when considering a large number of links simultaneously affected by rain.

1 Introduction

A result of the need to accommodate high-rate transmission is to push into increasingly higher frequency bands, namely Ka band (30/20 GHz) and V band (50/40 GHz). This trend is explained by the large segments of frequency spectrum required for supporting the high data rates planned in newer systems, [1]. Most VSAT and Digital Broadcasting Satellite TV systems in operation today use portions of the Ku band.

A major drawback to the use of higher frequencies is significant rain attenuation, which increases rapidly with increasing carrier frequency, [2], [3], [4]. It can cause serious signal quality degradation of earth-space communication links, having a major impact not only on individual links, but also on the global network.

MF-TDMA allows a group of Return Channel Satellite Terminals (RCSTs) to communicate with a Gateway using a set of carrier frequencies. In order to guarantee a Quality of Service, it is important that the Network Control Centre (NCC) gets to know the actual needs of each of the active RCSTs of the network. Therefore each station needs to monitor and measure its specific traffic requirements that are then communicated to the NCC. If resources are available, the NCC generates a new Terminal Burst Time Plan (TBTP) accommodating the needs of all its active stations at a superframe/frame level, so that each RCST knows

what timeslots have been assigned to it. This mechanism can also be used to drive a Fade Mitigation Technique, i.e. for the purposes of FMT slot reservation when a terminal needs more resources under rain conditions, [8].

2 Coding & Burst Length Control

Within an adaptive MF-TDMA fade countermeasure scheme, a portion of MF-TDMA timeslots is reserved as a shared resource, which can be distributed to any stations within the network subject to rain fading. When a burst within the MF-TDMA frame is subject to fading, it is allocated some extra timeslots into which it can expand. Since the user data rate is not changed, this expansion results in an increase in average power (or energy/bit) of the signal and so counteracts the effect of the fade, [5].

If G_c is the coding gain obtainable by coding the message, and the burst duration is increased H times (i.e. the station is allocated H extra FMT timeslots), then the original power margin, M_o , of the station is increased to, [6]:

$$M = M_o + G_H + G_C = M_o + 10 \log_{10} H + G_C \quad (1)$$

Every time the burst duration is doubled in time (or the symbol rate is halved) there is an increase in average transmitted power of 3 dB.

In the event of a fade, the energy/bit could be increased in either of the following ways: (i) the duration of the transmitted bits could be increased (symbol rate reduction), or equivalently, each bit could be repeated, or (ii) the entire block of bits could be repeated after appropriate expansion:

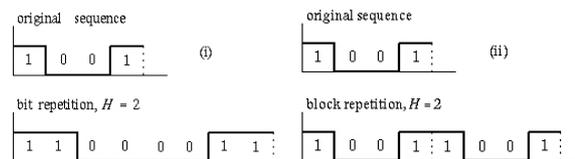


Figure 1. Example of (i) contiguous bit (ii) data block repetition

Comparison between the two techniques has shown that (i) is more efficient in terms of the enhancement that can be obtained for a particular utilisation of the shared resource time within the frame.

A baseline coding rate of 7/8 provides a 4 dB protection for a BER threshold of 10^{-8} (see Figure 2).

In general, if S is the original number of slots/frame being transmitted, reducing the code rate from the baseline code rate r_1 (decided by the clear-sky link budget and the QoS (BER) objective) a new code rate r_2 , results in expanding the transmission over $\lceil (r_1/r_2) \times S \rceil$ timeslots. For example, if one slot was originally used and the code rate switched from a baseline clear-sky code rate $r_1=7/8$ to a lower rate $r_2=1/2$ to combat fading, then the total number of slots required would be $\lceil (7/8)/(1/2) \times 1 \rceil = \lceil 1.75 \rceil = 2$ slots, i.e. one extra timeslot needed to fit the coded bits. For a BER objective of 10^{-8} this would translate into an extra 2 dB protection against fading (Figure 2).

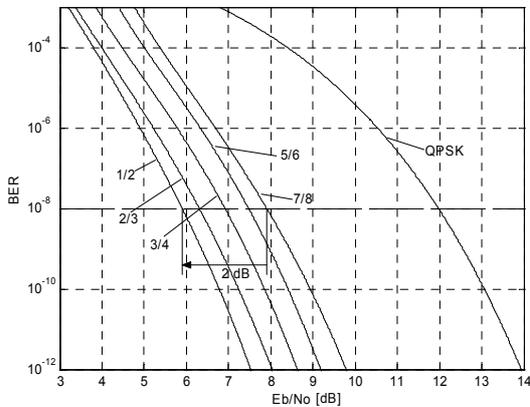


Figure 2. BER vs. E_b/N_0 performance of a convolutional coding system with QPSK, for various code rates

If, additionally, burst length control was employed, an extra 3 dB protection (total of 5 dB) would be provided by transmitting the information bits at half the speed, over twice as many slots: $\lceil (r_1/r_2) \times S \times 2 \rceil = \lceil 3.5 \rceil = 4$ slots, i.e. 3 extra timeslots in total.

2.1 Return Link Analysis

Radio frequency link analysis provides the means to calculate the quality of the information contents delivered to the data link control layer. The quality of digital information is measured by the BER, which depends on the type of modulation and coding performed, and on the carrier to noise power spectral density ratio, C/N_0 , at the input of the receiver, considered as a quality measure of the link quality. Comparison of what the satellite system can provide in terms of power, noise etc. with the requirements, i.e. comparing the E_b/N_0 value achieved with the required value for a specific QoS target, leaves a link margin which should be positive but not excessively so, otherwise the link would be over-engineered. This positive margin

represents all those components which it has not been possible to calculate in the budgets but which may affect the link quality (examples may be phase noise, AM/PM conversion in amplifiers, non-Gaussian interference etc.).

Considering the above, QPSK with convolutional code rate of 7/8 is chosen as the system baseline modulation/coding scheme, leaving a positive link margin of 2.73 dB (Table 1). The attenuated values were obtained considering 15-dB rain attenuation (threshold value for 99.91% availability, [7]) on the uplink.

	UPLINK	DOWNLINK	
Central Tx frequency, f	29.66	19.7	GHz
Transmitted power, P_{Tx}	2	1.07 (TWTA)	W
Antenna Tx gain, G_{Tx}	49.7	42.6	dB
EIRP	49.8 (ES)	27.02 (SAT)	dBW
Figure of merit, G/T	15 (SAT)	32.7 (HUB)	dB/K
C/N_0	79.74	77.16	dBHz
Mod. scheme, coding	QPSK, convol. $\rho = 7/8$		
OVERALL LINK			
Total clear-sky $(C/N_0)_{T,cs}$	75.25		dBHz
Achieved $(E_b/N_0)_{T,cs}$	12.14		dB
Modem implement. losses	1.5		dB
Link Margin, $LM_{(cs)}$	2.73		dB
Total $CNR_{T(rain)}$	64.5		dBHz
Total $(E_b/N_0)_{T(rain)}$	1.39		dB

Table 1. Return link analysis (RCST to Hub)

In the event of rain in some portion of the footprint, the drop in QoS on individual links will require the allocation of spare timeslots to counteract the rain attenuation. Figure 3 illustrates an example of employing coding and bit rate reduction to counteract fading.

Note that the original fixed power margin of the station of 2.73 dB is due to the convolutional code rate of 7/8, chosen as a fixed baseline for minimum redundancy and maximum throughput.

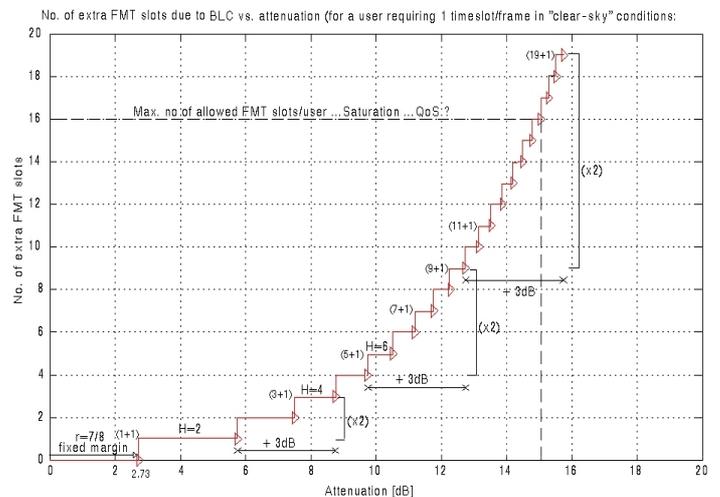


Figure 3. No. of extra FMT slots due to Burst Length Control vs. attenuation protection

Attenuation protection is thereafter provided in steps at the expense of capacity; every time the burst duration is doubled (hence the number of timeslots is doubled), an additional 3-dB attenuation protection is provided.

Figure 4 shows an example of carrier-to-noise ratio attenuation for a 30/20 GHz link. The signal degradation in terms of carrier to noise ratio due to rain can be seen; the clear-sky values (available $C/N_0=75.25$ dBHz, achieved $E_b/N_0=12.14$ dB) were calculated through the return link (RCST to Hub) budget analysis (Table 1).

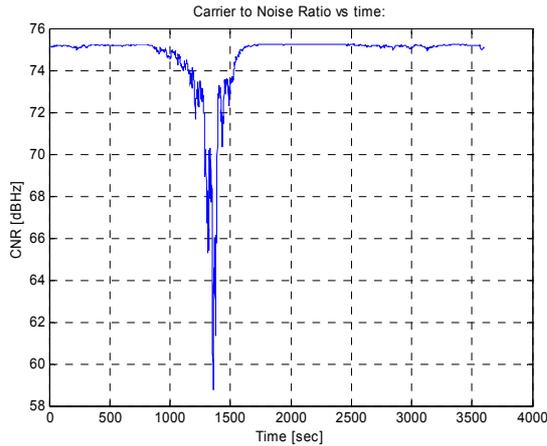


Figure 4. Simulating rain – an example of CNR degradation for a 30/20 GHz link

Figure 5 illustrates the corresponding BER in time during that event, showing the drop in QoS on the particular link during rainy conditions.

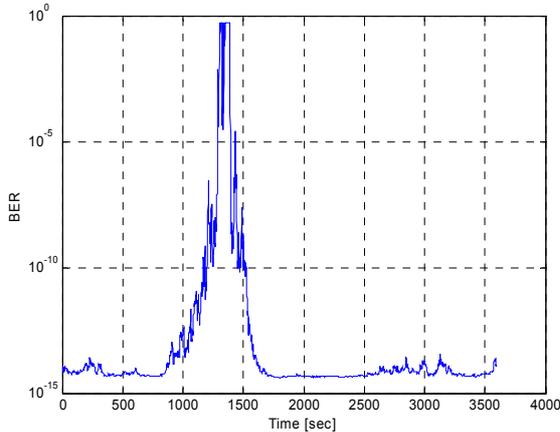


Figure 5. Bit error rate during rain event

Depending on the attenuation level, allocation of extra FMT slots will provide the desired compensation (Figure 6 and Figure 7).

We note the advantages of such an adaptive system over a fixed one that would constantly allocate a fixed number of slots (16 extra slots for a 15 dB attenuation threshold).

In the case of an adaptive system, the number of slots will vary in time depending on the actual requirement, as shown in Figure 6. This results in a much more efficient use of spare FMT slots, which can therefore be allocated to other users.

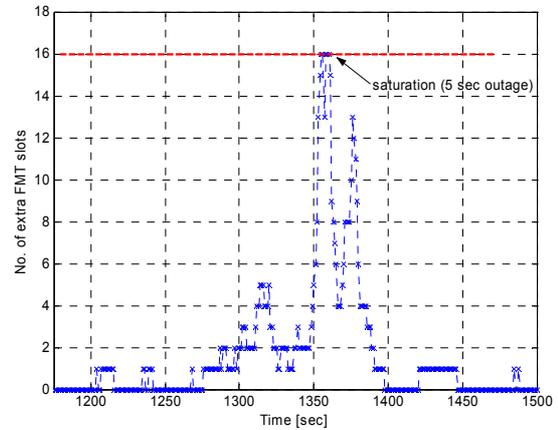


Figure 6. No. of extra FMT slots for one user during rain

Maintaining the BER target to achieve long-term availability would then only be limited by the dynamic range of the FMT, i.e. the maximum number of extra slots (16 slots in our example, see also Figure 3); if more slots were required, the system would not perform satisfactorily and there would be network outage as shown in Figure 6. The FMT provides compensation during the fade event, but not for this 5 seconds outage period, the only period during which the QoS threshold of 10^{-8} cannot be maintained (Figure 7):

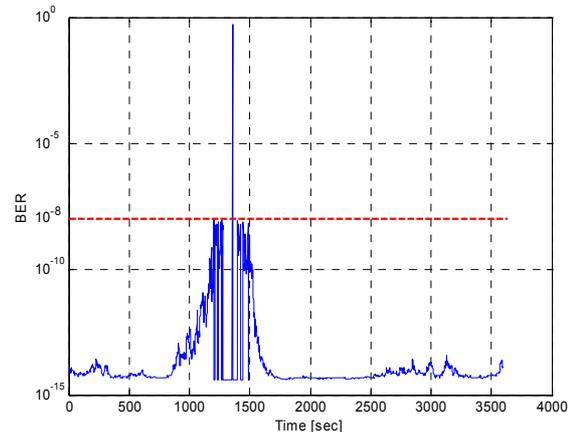


Figure 7. BER during a fade event with FMT

It becomes obvious, however, that this adaptiveness of the system should also be controlled, in order to avoid unnecessary “switchings” in short periods of time (see Figure 6) due to rain dynamics (fast changes in attenuation levels), i.e. the number of requests should probably be reduced in order to minimise signalling.

Assuming a population of N users, the NCC must be able to learn quickly about the links affected, for effective allocation of communication and FMT slots; service communication can be achieved through out-of-band (OBR) and in-band requests (IBR), [8].

A major issue is that MF-TDMA must be able to support all the active users. This implies that if the number of FMT spare slots is limited, saturation can occur when many users request slots at the same time. There will therefore be a finite blocking probability of not meeting the QoS parameters originally agreed. Alternatively, the requests may be queued, in which case, the limited capacity in the presence of rain will result in longer delays or a re-negotiation of QoS between RCST and NCC could be performed depending on the service level agreement.

Figure 8 shows an example of rain cells over areas with active and non-active users inside the coverage of area of a satellite.

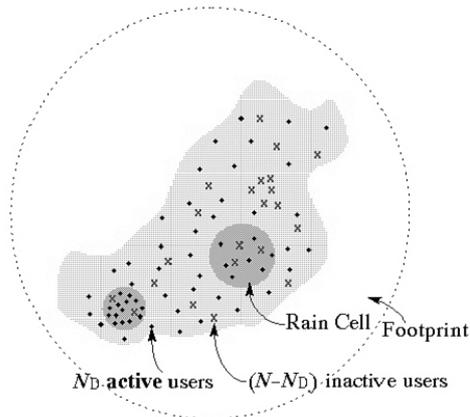


Figure 8. Example of rain cells over areas with active and non-active users inside the footprint of a satellite

It becomes apparent that the total number of required FMT slots depends on, [8]: (i) users' density and location (the greater the concentration of users the greater the number of required FMT slots); (ii) space and time characteristics of rain (for rain over large areas, more FMT slots are required); (iii) magnitude of the rain attenuation; (iv) actual traffic characteristics of the user stations; (v) QoS parameters.

Furthermore, depending on the hour of the day or the month, the rain conditions may be quite severe resulting in a more extensive use of FMT slots. This would be detrimental if severe fades occurred at times when user traffic is high.

3 Conclusions & Further Work

This paper has proposed a methodology on how to design a Fade Mitigation Technique suitable for application to Multi-Frequency TDMA networks. It

can provide compensation for rain attenuation at the expense of capacity; the attenuation protection as a function of spare resource in the form of extra timeslots has been investigated. The advantages, costs and issues, when it is used, have been demonstrated, with the intention to meet the user-specified QoS requirements.

The design of MF-TDMA networks with Burst Length Control requires a detailed study of the impact of rain attenuation on the satellite footprint. Long-term analysis will allow investigations into the share between traffic and FMT slots. The probability of blocking under severe rain conditions needs to be investigated as a measure of the grade of service. Short-term analysis will permit the study of the resource allocation algorithms that the NCC needs to deploy for a rapidly converging and fair management of the MF-TDMA channel.

4 References

- [1] M. H. Hadjitheodosiou, A. Ephremides, D. Friedman, "Broadband Access via Satellite", Technical Research Report, CSHCN T.R. 99-2, Centre for Satellite and Hybrid Communication Networks, University of Maryland, 1999, available at http://techreports.isr.umd.edu/TechReports/CSHCN/1999/CSHCN_TR_99-2/CSHCN_TR_99-2.pdf
- [2] L. J. Ippolito, "Radio propagation for space communications systems", IEEE Proceedings, vol. 69, no. 6, pp. 697-727, U.S.A., June 1981.
- [3] S. I. E. Touw, "Analyses of Amplitude Scintillations for the Evaluation of the Performance of Open-Loop ULPC Systems", MSc Thesis, Eindhoven Univ. of Technology, Eindhoven, The Netherlands, 1994.
- [4] R.A. Nelson, "V-band: Expansion of the Spectrum Frontier", Via Satellite, February 1998, pp. 66+.
- [5] D. J. Emerson, R. M. Nelhams, B. L. Clark, "Adaptive TDMA for Fade Countermeasures", Olympus Utilisation Conference, ESA SP-292, 12-14 April 1989, Vienna.
- [6] F. Carassa, "Adaptive Methods to Counteract Rain Attenuation Effects in the 20/30 GHz Band", Space Communication & Broadcasting 2, pp. 253-269, 1984, North Holland.
- [7] E. Noussi, "Satellite Communication Systems with Up-Link Power Control", MSc Thesis, University of Portsmouth, September 2002.
- [8] E. Noussi, B. Grémont, M. Filip, "Centrally Managed MF-TDMA Broadband Network with Fade Mitigation built upon DVB-RCS", 9th Ka & Broadband Communications Conference, p. 215-222, 5-7 November 2003, Island of Ischia, Italy.

Adaptive MMSE Receiver Structures, Employing Spread Spectrum and Frequency Diversity Techniques for Highly Dynamic Multipath Channels.

Dikaios P. Konstantakos, Bayan S. Sharif and Alan E. Adams

Underwater Acoustics Group, School of Electrical, Electronic and Computer Engineering,
University of Newcastle upon Tyne, NE1 7RU, United Kingdom.

Tel. +44 (0)191 2227273, Fax. +44 (0)191 2228180, email: dikaios.konstantakos@ncl.ac.uk

Abstract—Two adaptive receiver architectures are proposed in order to achieve reliable communications in a multipath-fading, highly time-variant channel, such as the shallow-water acoustic environment. Each structure uses spread spectrum and frequency diversity techniques, employing a combined adaptive, single-user equalizer and carrier phase estimator, based on minimum mean-square error (MMSE) filtering for every branch of the system. Additionally, a second stage of adaptive filters which update their parameters on a chip-rate basis may be included, in order to achieve more accurate parameter estimation, at the cost of increased complexity.

I. Introduction

Recently, the increased military and commercial demand for the development of a reliable, high data rate communication links between the nodes of a shallow water acoustic network has encouraged the employment of spread spectrum techniques and more sophisticated receiver structures. The underwater acoustic (UWA) medium is characterised by a limited transmission bandwidth and its multipath nature, due surface and bottom scattering of the acoustic signal. The dynamic nature of the UWA channel, in conjunction with the motion of the transceivers results in spectral broadening of the acoustic signal.

In this paper, a multicarrier, direct-sequence (DS) code-division multiple-access (CDMA) system is adopted in order to compensate for the previously mentioned phenomena. The receiver structures combine the operations of de-spreading, maximum ratio combining, equalization and phase carrier estimation. Furthermore, they suppress intersymbol interference (ISI), multiple access interference (MAI) and inter-carrier interference (ICI).

II. System Description

The modulation scheme utilized is Quadrature Phase Shift Keying (QPSK) multicarrier (MC) DS-CDMA [1]. The information sequence is coded into QPSK symbols and then each symbol is copied to M different, orthogonal carriers. Different users transmit using the same set of subcarriers, but with different spreading codes. The k th user's transmitted signal is represented as

$$s_k(t) = \text{Re} \left\{ \sum_{i=-\infty}^{+\infty} \sum_{m=1}^M \sqrt{\frac{2P_k}{M}} d_k[i] \cdot c_k[m] \cdot \cos \left(2\pi \left[f_c + m \frac{1}{T_s} \right] t + \varphi_{m,k} \right) \cdot \text{rect}(t - iT_s) \right\} \quad (1)$$

where $d_k[i]$ is the symbol stream on the parallel M branches and c_k is the spreading code of the k th user. P_k and f_c are the k th user average transmitted signal power, which is divided equally amongst the M carriers, and the initial carrier frequency, respectively. $\varphi_{m,k}$ is the random phase of the m th carrier for the k th user, exhibiting a uniform distribution between $[0, 2\pi]$. Additionally, $\text{rect}(t - iT_s)$ is the pulse waveform, which is unity in the interval $[0, T_s]$ and zero otherwise.

The shallow water channel is modeled by employing a direct path and a number of paths reflected at the channel boundaries. The transmitted signal passes through a frequency-selective, multipath-fading channel, with independent fading at the subcarriers. This response is described by the equation

$$h_{m,k}(t, \tau_{m,k}, \vartheta_{m,k}) = \sum_{i=1}^L a_{m,k}^i(t) \delta(t - \tau_{m,k}^i) \cdot e^{j\vartheta_{m,k}^i} \quad (2)$$

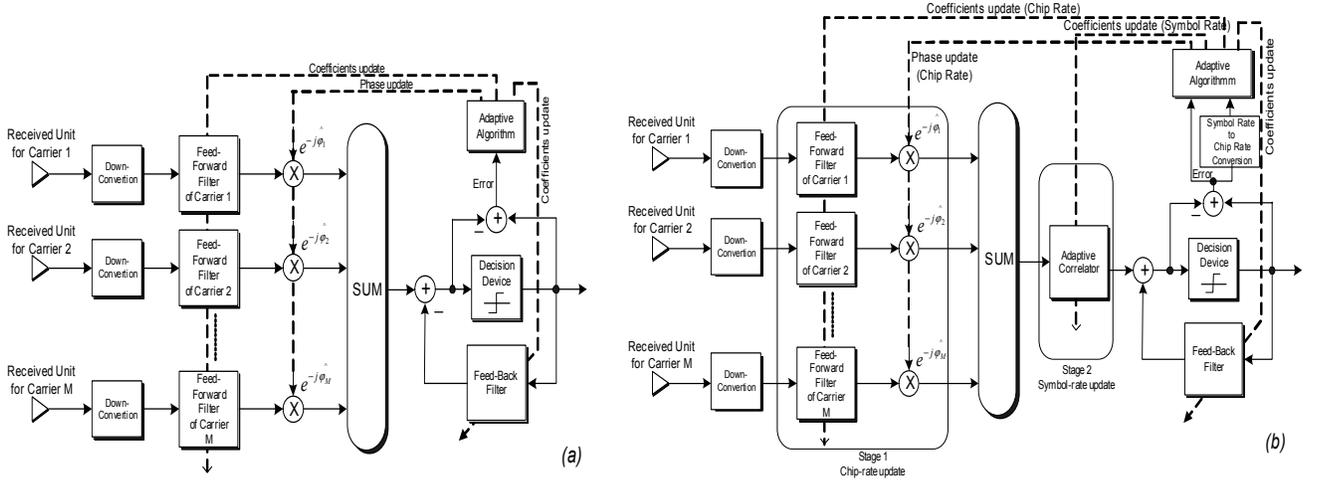


Fig. 1. Adaptive Multicarrier Receiver Structures: (a) simplified MC-DS-CDMA
(b) 2-stage/mixed update rate MC-DS-CDMA

where L is the number of propagation paths, $\{a_{m,k}^i(t)\}_{i=1}^L$, $\{\tau_{m,k}^i(t)\}_{i=1}^L$ and $\vartheta_{m,k}$ are the time-varying channel magnitudes, the corresponding propagation-path delays and the phases of the m th carrier and the k th user, respectively.

Equation (2) is commonly encountered in RF communications and is referred to as a tapped-delay line model of a frequency selective channel. Nevertheless, it can be used to model the UWA medium, on condition that wide-sense, stationary, uncorrelated scattering is assumed. Each individual path is represented as a complex-valued, random, time-varying process, whereby the rate of change of the process and, accordingly, the Doppler broadening of the signal, is controlled using a third-order infinite impulse response (IIR) filter whose frequency response has a Doppler-shaped spectrum.

Extensive studies [2, 3] of the statistics of the acoustic signal that propagates through an UWA medium, have stated that the amplitude of the received signal is Rayleigh distributed and its phase can be approximated as uniformly distributed.

The received MC-CDMA received signal is thus described as

$$r(t) = \text{Re} \left\{ \sum_{k=1}^K \sum_{i=-\infty}^{+\infty} \sum_{m=1}^M \sum_{l=1}^L \sqrt{\frac{2P_k}{M}} \rho_{k,m} d_k[i] c_k[m] \cdot \cos \left[2\pi \left(f_c + m \frac{1}{T_s} \right) t + \vartheta_{k,m} \right] \cdot \text{rec}(t - iT_s - \tau_{l,k,m}) \right\} + n(t) \quad (3)$$

where L denotes the number of multipath signals and $\tau_{l,k,m}$ is the propagation delay for the l th ray of the m th subcarrier for the k th user. The Additive White Gaussian Noise (AWGN) is denoted by $n(t)$.

The acquired signal is fed through M parallel band-pass filters to acquire the M different carriers and then the signal of each branch is down-converted, chip-matched filtered and sampled in order to be processed through an interpolator that is controlled by an early-late gate loop, used for timing estimation and as a Doppler tracking mechanism. The timing estimator has significant importance in an UWA communication link, since the effect of Doppler spread is not limited only to a frequency shift of the carrier, but also to a time expansion/compression of the waveform itself.

III. Receiver Architectures

The first receiver structure considered, shown in Fig. 1(a), uses a combined adaptive, single user equalizer and carrier phase estimator, based on minimum mean-square error (MMSE) filtering for every frequency band that is used. Coefficient estimation is performed jointly for all the branches of the system. The filter outputs are summed using a Maximum Ratio Combiner (MRC) and then passed through a hard-decision quantizer. A feed-back filter (FBF) can be added, employing previous symbol decisions to eliminate the ISI on the current symbol being detected, transforming the linear receiver into a multicarrier decision-feedback equaliser (DFE). The number of the taps in the feed-forward filter (FFF) session of each carrier depends on the spreading gain.

The 2-stage/mixed update rate MC-DS-CDMA receiver, shown in Fig. 1(b), consists of two stages of feed-forward filters (FFF). The first stage comprises a bank of adaptive filters, followed by a bank of decision-directed digital phase locked loop (DPLL) structures. The number of the FFF and the DPLL

depends on the number of the utilized carriers. The role of this stage is to find the optimum weights of the MRC and to track the carrier phase accurately. The main feature is that both the filters and the phase estimators update their parameters on a chip-rate basis, in contrast to the parameter update of the second stage, which is performed at a symbol-rate. The second stage consists of one adaptive filter which is used as a despreader and MAI canceller. The chip-rate updating procedure is performed by spreading the error signal with the signature code of the desired user. Both of the receivers use the MMSE criterion in order to optimize parameters. Mathematically expressed as

$$J_{\min}(n) = \min_{\mathbf{w}(n)} E \left\{ |e(n)|^2 \right\} \quad (4)$$

where $\mathbf{w}(n)$ denotes the filters taps vector, $E\{\}$ indicates the expectation process and $e(n)$ the error signal at the n th symbol interval. In this paper, the parameters of the adaptive MMSE filters are updated according to the normalized least-mean square (NLMS) algorithm as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \Delta_w \mathbf{u}^*(n) e(n) \quad (5)$$

$$\Delta_w = \frac{a}{\beta + \|\mathbf{u}(n)\|^2} \quad (6)$$

where $a \in (0,2)$ can be adjusted to control the speed of convergence, and β is a small positive constant that guarantees stability of the algorithm when the power of the received signal is low. Additionally, $\|\mathbf{u}(n)\|$ denotes the norm of the input signal vector $\mathbf{u}(n)$.

IV. Simulation Results

Due to the limited availability of bandwidth in the UWA channel, and the high dynamic nature of the environment which results in excessive Doppler

Simulation Parameters	
Frequency Band	8-12 kHz
Number of Carriers	2
Central Carrier Frequency	10 kHz
Sampling Frequency	48 kHz
Delay Spread	3 ms
Range	3 km
Doppler Spread	2 Hz

Table 1. Scenario parameters

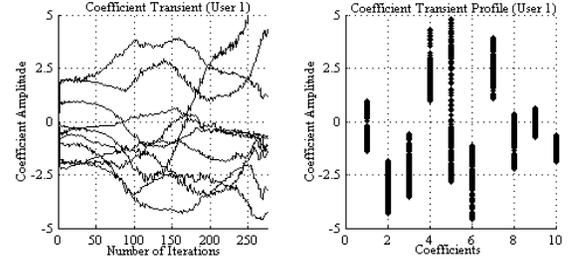


Fig.2. Weight Transient Profile of the second stage of the mixed-update rate MC-DS-CDMA receiver

broadening of the signal, the spreading sequences ought to have short length, in order to maintain a high data rate and a rapid adaptation rate. Furthermore, because of the long multipath effect, each signature must have good autocorrelation properties. The employed signature sequences are complex-valued, pseudo-random codes of 10 chips [4], and the examined scenario parameters are indicated in Table 1.

Fig.2 illustrates the adaptive weights transient profile of the second stage of the mixed update rate MC-DS-CDMA receiver. Note that this filter operates as a despreader, estimating the spreading code, even when the environment conditions vary rapidly. Fig.3

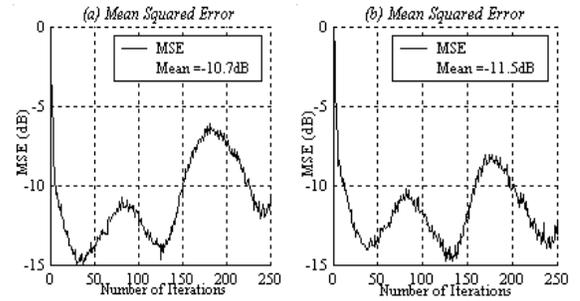


Fig.3. MSE of (a) simplified MC-DS-CDMA and (b) mixed-update rate MC-DS-CDMA receiver

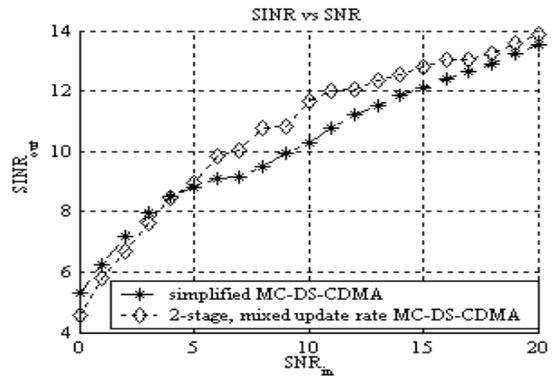


Fig.4. SINR versus SNR of the two receiver structures

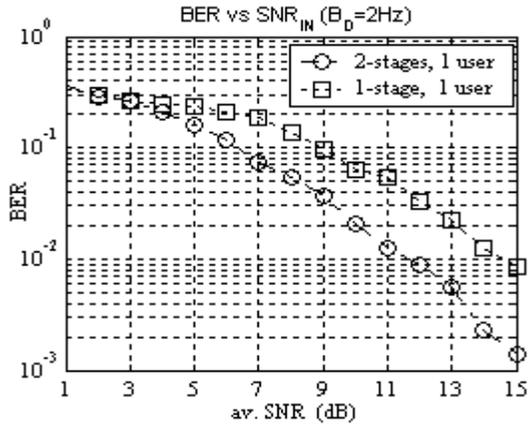


Fig. 5. BER versus SNR of the two receiver structures

and Fig. 4 demonstrate the MSE and the output Signal-to-Interference-Noise Ratio (SINR) versus the input Signal-to-Noise Ratio (SNR) of the examined receivers, respectively. The simplified MC-DS-CDMA receiver has a mean value of convergence of -10.7 dB, while the more complex structure exhibits a gain of 1 dB. The level of convergence represents an approximation of the SINR at the output of the detector, mathematically expressed [5]

$$SINR_{dB} \cong 10 \log_{10} \frac{1 - J_{\min}}{J_{\min}} \quad (7)$$

Fig. 5 shows the performance comparison between the receivers considered. The performance gain for single user transmission is approximately 3 dB, at a BER level of 10^{-2} when the 2-stage structure is employed.

V. Conclusion

Two adaptive receiver structures that employ both spread spectrum and frequency diversity techniques has been proposed for reliable communication in a dynamic underwater environment. Computer simulations showed the effectiveness of the system in a multipath-fading and highly time-variant channel. The 2-stage/mixed-update rate receiver outperformed the simplified MC-DS-CDMA structure at the cost of increased computational complexity.

Acknowledgments

The authors greatly acknowledge the support of the EU for this research, under contract No. EVK3-CT-2000-000039 (ACME).

References

- [1] S. Hara, R. Prasad, "Overview of multicarrier CDMA", *IEEE Communication Magazine*, pp. 126-133, Dec. 1997
- [2] R. A. Iltis and A. W. Fuxjaeger, "A digital DS spread-spectrum receiver with joint channel and Doppler shift estimation," *IEEE Trans. On Communications*, vol. 39, No. 8, pp. 1255-1267, Aug. 1991
- [3] T. C. Gallaudet and C. P. de Moustier, "Acoustic Backscatter fluctuations," *J. Acoust. Soc. Am*, vol. 114, No. 2, Aug. 2003
- [4] I. Oppermann and B. S. Vucetic, "Complex Spreading Sequences with a Wide Range of Correlation Properties", *IEEE Trans. On Communications*, vol. 45, no. 3, pp. 365-375, Mar. 1997
- [5] J. G. Proakis, *Digital Communications*, McGraw Hill, New York, 3rd edition, 1995

Adaptive PN Code Acquisition in DS-CDMA Systems

Job A. Obiebi, Mohammad Y. Sharif and John Sharp.

School of Engineering, Napier University, 10 Colinton Rd, Edinburgh, EH10 5DT

j.obiebi@napier.ac.uk

The conventional method of using fixed thresholds in the analysis of PN code acquisition circuits will not be able to adapt to changing noise power in mobile environment. In this paper an adaptive double dwell matched filter PN code acquisition circuit is proposed. It is analysed in a Rayleigh fading channel using a Cell Averaging Constant False Alarm Rate (CA-CFAR) technique. Expressions are derived for detection probability and false alarm rate and it is shown that the scheme is more robust to changing noise conditions than the conventional method. In terms of performance and hardware requirements, the technique is compared with the method of using a reference noise power estimator. Numerical results are presented to demonstrate that the performance of the technique compares favourably. It is shown that the proposed technique can be implemented with reduced hardware.

I. Introduction

An important implementation area for the use of Spread Spectrum (SS) for commercial mobile communication is the Pseudo-random Noise (PN) code synchronisation. Synchronisation is achieved by code acquisition (coarse alignment) and code tracking (fine alignment). Code acquisition is usually the more difficult of the two and thus most research is centred on it [1]-[8]. In SS mobile communication, the code synchronisation process has to be done as fast as possible and with high probability. The code acquisition process mainly entails the correlation of the received PN code with a local PN code. This correlation process could be active or passive [1]. The passive correlation using matched filter (MF) is considered here.

Since an estimate of the phase of the incoming PN code is usually not known in advance, the full spreading code length is taken as the uncertainty region which is divided into sub regions (cells) determined by the step-size in which the search will be carried out. These cells can be searched in a parallel or serial manner and hardware complexity favours the serial search [1]. The time spent in searching a code phase is usually known as the dwell time. Dwell time could be single or multiple dwell. There are many incorrect phases during the code search process. Using multiple dwell, these incorrect phases could be discarded quickly and thus provide significant reduction in time spent searching the cells [7]. Significantly more improvement is obtained by going from single to double dwell than to higher dwells [7]. Thus, in this work, a double dwell is used. The second dwell can be implemented by using a Coincidence

Detector (CD) [1], a longer integration time in the form of Post Detection Integration (PDI) [3] or a second MF of longer length [6].

An important issue in the implementation of the code acquisition circuit is the selection of the threshold that is used in the decision process. Various methods of selecting threshold are analysed in [2]. These methods range from fixed threshold, optimal to CFAR methods. The CFAR method is at the core of adaptive threshold methods [2], [4]-[6]. In this work, a CA-CFAR method is used because of the simplicity of implementation. This CA-CFAR technique is analysed in [5] and [6] for a Rayleigh fading and non-fading Additive White Gaussian Noise (AWGN) channel respectively. In [5], the emphasis is on the probability of detection as a function of PDI. In [6], a MF of longer length is used in the second dwell and it is also analysed for a non-fading AWGN channel.

In this work the performance of a MF-based double dwell PN code acquisition using CA-CFAR technique is analysed in a Rayleigh fading channel. The purpose here is two-fold: firstly, instead of increasing the length of MF in the second dwell, PDI is used to increase the dwell time and to improve the Signal-to-Noise Ratio (SNR) for a given MF length restricted by the amount of expected frequency offset and secondly, to analyse the performance of the circuit in a frequency non-selective Rayleigh fading channel with Multiple Access Interference (MAI). The rest of the paper is organised as follows. Section II describes the proposed acquisition circuit. In section III the channel model is described and relationships are derived for the proposed method. Numerical results are presented in section IV and conclusions drawn in section V.

II The Proposed CA-CFAR Method

The proposed acquisition circuit is shown in Figure 1. The MF is of short code length restricted by the expected frequency offset. It is a double dwell immediate rejection circuit. In the second dwell, a much longer integration time is used implemented as PDI. The output of MF is stored in the CFAR processor as shown in Figure 2 with $n+1$ registers. Prior to acquisition, out of phase correlation (noise) samples from the MF are stored in the n windows. The output of the MF for the cell under test is stored in the first window as Y . A scaling factor, T_r ($r=1, 2$ dwells) is used to set the threshold in the CFAR processor. This fixes the false alarm rate of the system. Thus the operation of the circuit is as follows. If the first dwell indicates that the output of MF exceeds the threshold set by T_1 , second dwell is enabled and if after L PDIs, the output is found to exceed another threshold set by T_2 , then acquisition is declared and tracking follows. It

is seen that as the noise power in the CFAR processor varies, the threshold also changes to reflect the environmental condition, because of the fixed scaling factor, T_r . Thus the circuit can adapt to changing noise power.

III. Performance Analysis

Assumptions: *There is only one cell corresponding to the correct code phase; the MF is sampled at the chip rate; oscillator instability specification is ± 1 to ± 5 ppm corresponding to 1.8 to 9 kHz frequency offset in IS-95; the Rayleigh fading is slow enough that the amplitude and phase remain constant during each correlation interval but fast enough that successive correlation intervals are independent [3]; the channel is wide sense stationary.*

The channel model

The channel model is a frequency non-selective Rayleigh fading channel with K asynchronous users. The received Binary Phase Shift Keyed (BPSK) signal at the receiver is thus given as:

$$r(t) = \sum_{k=1}^K \sqrt{2S_k} c_k(t + \zeta_k T_c) [x_k(t) \cos(\omega_0 t + \theta_k) - y_k(t) \sin(\omega_0 t + \theta_k)] + n(t) \quad (1)$$

where S_k is the signal power of the k th user, ζ_k is the k th user delay in the received code normalised to T_c , $x_k(t)$ and $y_k(t)$ are the in-phase and quadrature Rayleigh fading components of the signal of the k th user respectively, $c_k(t)$ is the PN sequence of length W chips and $n(t)$ is AWGN of double-sided power spectral density given as $N_0/2$. The operation of the circuit is that of deciding between two hypothesis denoted as H_i , ($i=0, 1$). H_1 indicates the situation where the correct code phase is present and the output of the MF exceeded a threshold with a probability of detection, P_d and tracking mode enabled. If the delay is more than T_c , denoted as H_0 , the output will be less than the threshold and the code search process continues by advancing the code phase by ΔT_c where $\Delta=1$. It is also possible that during the H_0 state, the output could be greater than the threshold and the circuit will falsely declare acquisition with a false alarm probability, denoted here as P_{fa} . The false alarm penalty time is given as T_{fa} .

Statistics of the MF output

The output of MF due to the fading signal plus AWGN (for $i=0, 1$) is given as:

$$u_{i,I} = \Phi_{i,I} + \lambda_{i,I} + \phi_{i,I} + \sigma_{n,I}^2 \quad (2a)$$

$$u_{i,Q} = \Phi_{i,Q} + \lambda_{i,Q} + \phi_{i,Q} + \sigma_{n,Q}^2 \quad (2b)$$

where $\Phi_{i,I} = \Phi_{i,Q}$ is the variance of the signal due to the desired user, $\lambda_{i,I} = \lambda_{i,Q}$ is the effect of partial correlation due to the fact that the length of MF is much less than the spreading code, $\phi_{i,I} = \phi_{i,Q}$ is the

contribution of MAI and $\sigma_{n,I}^2 = \sigma_{n,Q}^2$ is the zero-mean independent identically distributed Gaussian random process due to the AWGN with common variance $\sigma_n^2 = NN_0 T_c / 2$. I and Q represent the in-phase and quadrature terms. N is length of MF.

$$\sigma_{s,i}^2 = D(0, \sigma_i^2) \quad (3)$$

Where σ_i^2 is the common variance of $u_{i,I}$ and $u_{i,Q}$.

That is,

$$\sigma_i^2 = u_{i,I} = u_{i,Q} \quad (4)$$

$D(m, \sigma^2)$ represents the Gaussian distribution with mean m and variance σ^2 .

Let x_i denote the i th chip component of the fading process with variance $E[x_i^2] = \sigma_g^2$ and autocorrelation

function given from [8] as $E[x_i x_j] = \rho_{|i-j|} \sigma_g^2$. Where

$\rho_{|i|} = J_0(2\pi f_D i T_c)$ and J_0 is the Bessel function of order 0. $E[x]$ denotes the expectation of x . Under H_1 the variance at the output of MF due to the desired user is given as:

$$\Phi_{1,I} = \sigma_g^2 ST_c^2 G(1-p)^2 \quad (5a)$$

$$G = N + 2 \sum_{l=1}^{N-1} (N-l) \rho_l$$

$$\lambda_{1,I} = \sigma_g^2 ST_c^2 N p^2 \quad (5b)$$

where $p \in 0,1$ is the residual code offset. Similarly under H_0 :

$$\Phi_{0,I} = 0 \quad (6a)$$

$$\lambda_{0,I} = \sigma_g^2 ST_c^2 [1-2p+2p^2] \quad (6b)$$

and it can be shown that the MAI term is given as:

$$\phi_{1,I} = \phi_{0,I} = \frac{(K-1)\sigma_g^2 NST_c^2}{3} \quad (7)$$

Using (3), for H_1 and H_0 and setting $p=0$ we obtain:

$$\sigma_{s,1}^2 = \frac{NN_0 T_c}{2} \left(\frac{G \bar{E}_c}{NN_0} + \frac{(K-1)\bar{E}_c}{3N_0} + 1 \right) \quad (8)$$

$$\sigma_{s,0}^2 = \frac{NN_0 T_c}{2} \left(\frac{\bar{E}_c}{N_0} + \frac{(K-1)\bar{E}_c}{3N_0} + 1 \right) \quad (9)$$

$$\sigma_s^2 = \frac{\sigma_{s,0}^2}{\sigma_{s,1}^2} = \frac{\bar{E}_c / N_0 + (K-1)\bar{E}_c / 3N_0 + 1}{G \bar{E}_c / NN_0 + (K-1)\bar{E}_c / 3N_0 + 1} \quad (10)$$

where $\bar{E}_c / N_0 = 2\sigma_g^2 ST_c / N_0$ is the Rayleigh faded SNR per chip. Frequency offset between transmitter and receiver leads to degradation in SNR. The amount of degradation is given as [3]:

$$SNR_{loss} = \mu = \sin^2(N \Delta f T_c) \quad (11)$$

Writing $\Xi_c / N_0 = \mu \bar{E}_c / N_0$, for the degraded case, (10) becomes:

$$\sigma_s^2 = \frac{\Xi_c / N_0 + (K-1)\Xi_c / 3N_0 + 1}{G \Xi_c / NN_0 + (K-1)\Xi_c / 3N_0 + 1} \quad (12)$$

Probability of false alarm and detection

From (4) and Figure 1, it is seen that summing the in-phase and quadrature components of the output of the MF after squaring will result in a probability distribution function that is chi-square distributed with two degrees of freedom. Thus after some algebra, it can be shown that the probability of false alarm of the proposed circuit for r dwells is:

$$P_{fa_r} = \sum_{q=0}^{L-1} 1/q! \frac{(Ln-1+q)!(T_r)^q}{(Ln-1)!(1+T_r)^{Ln+q}} \dots\dots\dots (13)$$

and the detection probability is:

$$P_{d_r} = \sum_{q=0}^{L-1} 1/q! \frac{(Ln-1+q)!(T_r \sigma_s^2)^q}{(Ln-1)!(1+T_r \sigma_s^2)^{Ln+q}} \dots\dots\dots (14)$$

From (13) it is seen that P_{fa} is independent of the SNR, thus achieving a CFAR. L is 1 for $r=1$ and 4 for $r=2$.

IV. Numerical Results

Mean Acquisition Time

The mean acquisition time, $E\{T_{acq}\}$ is the average time that elapsed while searching the code phases before the declaration of acquisition. Using the uniform serial search [1], $E\{T_{acq}\}$ of the proposed circuit is:

$$E\{T_{acq}\} = \frac{T_c}{P_{d1}P_{d2}} \left((1 + P_{d1}L + (W-1)(1 + P_{fa1}L + T_{fa}P_{fa1}P_{fa2})(1 - P_{d1}P_{d2}/2)) \dots\dots\dots (15) \right)$$

MATLAB language programming is used to produce the results and the following parameters are used: the PN code length is 32767; search step-size is ΔT_c , $\Delta=1$; then the uncertainty region is $W=32767$; $T_c=10^{-6}/1.2288$ sec; the MF length is 64; in the second dwell $L=4$ is used for the PDI; the penalty time $T_{fa}=10000$ chips. While the proposed circuit maintains a CFAR, the nonadaptive (conventional) circuit's false alarm rate varies according to the SNR. Thus it cannot detect the desired code position as fast as the proposed scheme. This is shown in Figure 3 where a sub-optimum threshold that gives $P_{fa}=10^{-3}$ at -5 dB is used for the nonadaptive circuit. Figure 4 shows the detection probability as a function of SNR for different values of normalised Doppler shift, $f_D T_c$. It can be seen that as the Doppler shift increases, the fading in the channel increases. Unless otherwise stated, $f_D T_c=10^{-4}$ is used in the results that follow.

Figure 5 compares the performances of the proposed method, the nonadaptive method, the method of using a MF of longer length in the second dwell denoted here as longer integration (LI) and the Reference MF (R-MF) method [4]. An equivalent MF length of 256 is used in the second dwell of the LI method and the other three methods use PDI in the second dwell. The proposed technique gives more than 9dB performance improvement than the nonadaptive method. Also, it can be seen that as the number of window cells in the CFAR processor is increased, the performance of the proposed technique approximates that of using an R-

MF. When $n=20$ the difference between the two methods is less than 1dB. Thus the reduced number of delay lines used in the proposed scheme more than compensates for this small loss in performance. However, at lower SNR, the LI performs better than the other methods.

Figures 6 and 7 shows the effect of frequency offset on the proposed, nonadaptive, R-MF and LI methods. When the frequency offset is less than 2 kHz and at lower SNR ($\Xi_c/N_o < -14$ dB), using the LI method yields better results than the proposed technique. From our assumptions, the frequency offset can be up to 9 kHz. When the offset is greater than 2 kHz, the proposed scheme performs better than the LI method. This confirms the validity of using PDI. However, for the nonadaptive method, apart from the degradation in P_{d_r} , the P_{fa} is also increased for obvious reason. This translates into increased $E\{T_{acq}\}$.

Effect of MAI

All the plots generated so far assume that there is only one user, i.e. $K=1$. In Direct Sequence (DS) CDMA systems, the number of users has an impact on the performance of the acquisition system. Further motivation to determine the effect of MAI stems from the soft capacity of CDMA systems. It is seen that as the number of users increases, the $E\{T_{acq}\}$ also increases but more in the nonadaptive method. This is shown in Figure 8.

V. Conclusions

An adaptive double dwell PN code acquisition circuit suitable for DS-CDMA systems has been proposed and analysed in Rayleigh fading channel. The scheme is compared to an acquisition circuit that uses an R-MF for noise power estimation. It is seen that the performance of the proposed scheme compares favourably with the R-MF circuit as the window size n of the CFAR processor is increased. It is also more robust to frequency offset than the method of using LI in the second dwell. When compared to a non-adaptive scheme, the considerable reduction in mean acquisition time shows that the scheme can provide more robustness to the varying noise in mobile environment.

VI. References

- [1] A. Polydorus and C. L. Webster, "A unified approach to Serial Search Spread-Spectrum Code Acquisition" – Parts 1& 2: *IEEE Trans. On Comms*, vol. 32, no. 5, May. 1984, pp542-560
- [2] J. Iinatti, "On the Threshold Setting Principles in Code Acquisition of DS-SS Signals", *IEEE Journal on Selected Areas in Comm.*, vol. 18, no. 1, Jan. 2000, p. 62-72.
- [3] A.J Viterbi, "CDMA: Principles of Spread Spectrum Communication", Addison Wesley Publishing Company, 1995
- [4] B. Ibrahim & H. Aghvami, "Direct Sequence Spread Spectrum Matched Filter Acquisition in Frequency Selective Rayleigh Fading Channels", *IEEE Journal on Selected Areas in Commun*, vol. 12, no. 5, June 1994, p. 885-890.
- [5] C. Kim, H. Lee and H. Lee, "Adaptive Acquisition of PN Sequences for DSSS Communications", *IEEE Trans. On Comms*, vol. 46, no. 8, Aug. 1998, pp993-996

- [6] H. Oh, D. Han, C. Lim and K. Park, "Adaptive double-dwell PN Code Acquisition in Direct-Sequence Spread-Spectrum Systems", MILCOM 2000 – IEEE Military Communications Conference, no. 1, Oct. 2000, pp139-143.
- [7] D. M. DiCarlo and C. L. Weber, "Multiple Dwell Serial Search: Performance and Application to Direct Sequence Code Acquisition", *IEEE Transactions on Communications*, vol. Com-31, no. 5, May 1983, p. 650-659.
- [8] E. Sourour and S. C. Gupta, "Direct-Sequence Spread-Spectrum Parallel Acquisition in Nonselective and Frequency-Selective Rician Fading Channels", *IEEE Transaction on Communications*, vol. 10, no. 3, April 1992, p. 535-544.

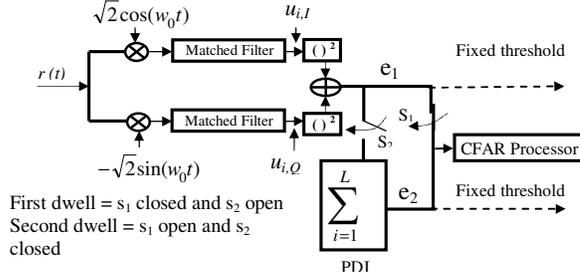


Figure 1. Proposed circuit. (Dotted lines represent path for non-adaptive method.)

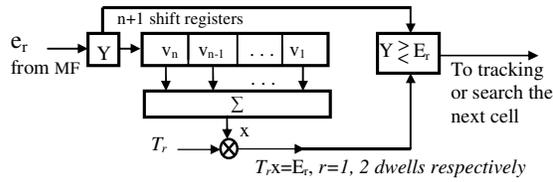


Figure 2. CA-CFAR processor.

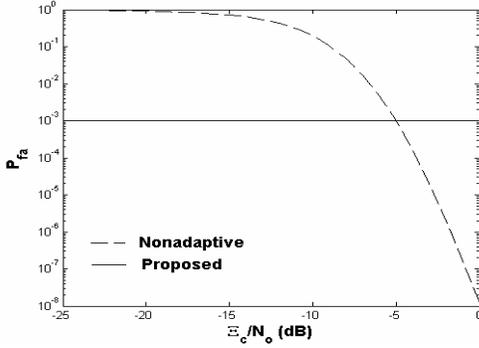


Figure 3. P_{fa} versus E_c/N_0 .

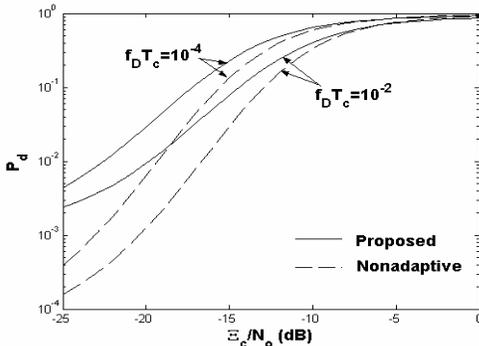


Figure 4. P_d versus E_c/N_0 for different Doppler shift ($P_{fa}=10^{-3}$ a parameter no frequency offset).

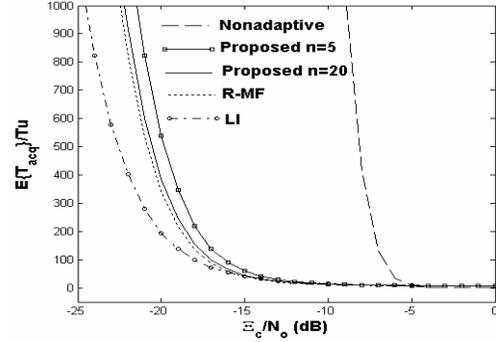


Figure 5. Normalised $E\{T_{acq}\}$ versus E_c/N_0 ($P_{fa}=10^{-3}$ at $E_c/N_0=-5$ dB. $T_u=W*T_c$).

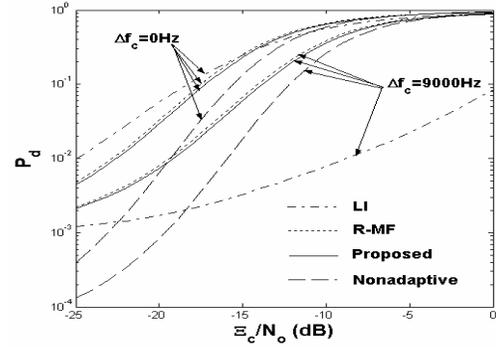


Figure 6. Effect of frequency offset on P_d ($P_{fa}=10^{-3}$).

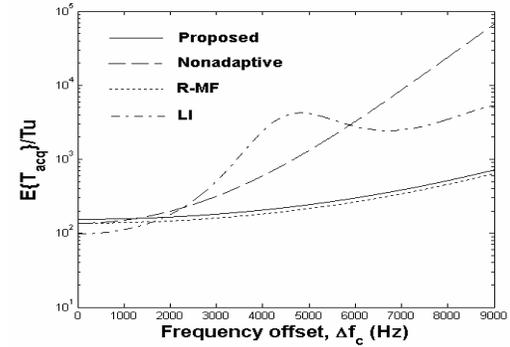


Figure 7. Effect of Frequency offset on normalized $E\{T_{acq}\}$ ($P_{fa}=10^{-3}$ at $E_c/N_0=-18$ dB).

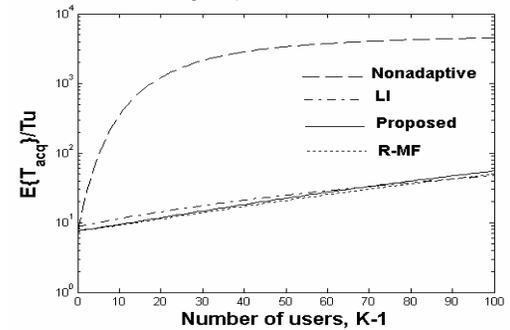


Figure 8. Effect of MAI on normalised $E\{T_{acq}\}$ ($P_{fa}=10^{-3}$ at $E_c/N_0=-5$ dB and no frequency offset).

An Adaptive HARQ-II Protocol using RCPC Codes¹

De-Yi Chen² and Jin-Fu Chang³

Abstract

Effective error control techniques are highly desired to combat the omnipresent errors in mobile radio channels. These errors are the result of background noise plus fading. The errors have an additional nature of being time varying. In other words, the quality of a channel could be poor at one time but good at another. When channel quality is good, more information should be transported; or less redundancy should be invested in error control. Therefore error control techniques are anticipated to have a new dimension of feature dubbed “adaptive”. The type II hybrid ARQ (HARQ-II) protocols are chosen for consideration due to their many facets of attraction that researchers have demonstrated. The family of RCPC (rate compatible puncture code) codes have the feature of adjustable code rate and is picked to make a HARQ-II adaptive. The purpose of this study is to demonstrate the feasibility of realizing an adaptive HARQ-II protocol through RCPC codes.

1. Introduction

Error control technique is a means to combat channel errors in the transportation of information-carrying

¹ The work reported in this paper is supported by the National Science Council, Taiwan (NSC-2001-2213-E-260-011)

² Institute of Communications Engineering, National Taiwan University, Taipei, Taiwan 107; currently employed by Quanta Computer Inc., Wen-Hwa 2nd Rd., Keuishan, Taoyuan, Taiwan 333; Terence_Chen@quantatw.com.

³ Department of Electrical Engineering and Institute of Communication Engineering, National Chi-Nan University, Puli, Nantou, Taiwan 545; jfchang@ncnu.edu.tw. Also with Department of Electrical Engineering and Institute of Communications Engineering, National Taiwan University, Taipei, Taiwan 107.

messages to ensure the integrity of a message. Error control in general is realized at the cost of injecting redundant bits into a message block through specific rules. This is the spirit of traditional forward error correcting (FEC) codes. Automatic repeat request (ARQ) has opened up a new branch of error control techniques. In the development of ARQ protocols, a hybrid between FEC and ARQ has come into existence. A properly designed Hybrid ARQ (HARQ) protocol is able to meet a specified performance goal involving both throughput and delay. There are two basic types of HARQ: HARQ-I and HARQ-II.

In this paper we focus our work on HARQ-II.

The quality of a radio channel employed in mobile communications is hurt not only by the background noises but also by a well recognized phenomenon called fading. User mobility further makes condition of the channel become time varying and location dependent. In other words, error control coding faces the challenge of not only a much noisier but also a time varying channel. This justifies the proposal of an adaptive HARQ protocol and the rate compatible puncture code (RCPC) is our choice to materialize the ideal of adaptation.

2. An RCPC-based HARQ-II

The design proposed in this section benefits from the discussions in [1]-[3]. In our design, C_0 is any ordinary (n,k) linear block code and C_1 is a RCPC generated from a 1/3-rate $(3,1,m)$ convolutional code. In other words, a k -bit information block is first converted to an n -bit word of C_0 for its initial attempt. In the meantime these n bits are padded with m 0's to

become an $(n+m)$ -bit word and fed into the $(3,1,m)$ convolutional encoder to generate a codeword of $3(n+m)$ bits to be used as a base in any subsequent retransmissions.

To deal with retransmissions, three puncturing tables are employed. Shown in Figure 1 is an example to demonstrate mechanism in dealing with retransmissions. The generator matrix G of C_1 and the three associated puncturing tables P_1 - P_3 are given in Figure 1. These tables are taken from [3]. When a retransmission is requested due to the failure of a codeword from C_0 in its initial attempt, the RCPC word generated from P_1 is then sent. The $(n+m)$ -bit word received is first de-convoluted by the $g(1)$ in G to get an n -bit word and checked against C_0 to see whether it is error-free. If yes, then the entire process terminates here by returning an ACK to the sender. Otherwise, a NAK is returned to request the transmission of the $n+m$ bits generated from P_2 . If the n -bit word converted from $g(2)$ is found to have contained no errors then the process stops here; otherwise the received word and the earlier received $n+m$ bits (corresponding to P_1) are combined through P_1 and P_2 to become a word of a $1/2$ -rate RCPC. If the resultant word is decodable, that is if the number of errors is containable within the capability of the $(2,1,m)$ code, then the decoded result is checked against C_0 to see if the m -bit message-carrying block is valid. If this round still fails in the end, then the $n+m$ bits created from P_3 is sent and the process is repeated except that if decoding of the $(3,1,m)$ is necessary the input needs to be generated together with the traces of the earlier two rounds through the applications of P_1 - P_3 . Notice that the rate of the underline RCPC now becomes $1/3$. If the attempt fails again at this stage then the entire process has to be restarted.

In general an $1/N$ -rate convolutional code can be

employed in HARQ-II as C_1 and a larger N allows to create more RCPCs with differing rates.

3. An Adaptive HARQ-II Derived from the RCPC

Data packets to be transmitted are prioritized according to their time urgency. A more urgent packet is protected by lower rate RCPC, and vice versa. In materializing the above type of channel adaptation, we also attempt to piggyback new information through unused space in the course of conducting retransmission for a data block.

C_0 again is assumed to be an (n,k) systematic linear block code and C_1 is an RCPC derived from a half-rate $(2n,n,m)$ convolutional code. Four puncturing tables are employed to generate four different RCPCs. We further assume that each codeword of C_1 consists of $2(n+m)$ bits and needs two time slots to transmit. See Figure 2 for an example of G and its four puncturing tables P_0 to P_3 .

Referring to Figure 2, we notice that G together with P_0 generate a full rate code of $(n+m)$ -bit word and this code is the C_0 to be used in the HARQ-II protocol. We further notice that G together with P_1 , P_2 , and P_3 jointly produce the remaining $n+m$ bits of the code. In other words, G and each of P_1 to P_3 generate a code of word length $(n+m)/3$.

From Figure 2 we observe that $P_0 \oplus P_1$, $P_0 \oplus P_1 \oplus P_2$, and $P_0 \oplus P_1 \oplus P_2 \oplus P_3$ together with G produce a RCPC of rate $3/4$, $3/5$, and $1/2$, respectively.

For the example we lay out in Figure 2, time urgency of data packets are divided into three priorities 1,2,3 with 3 representing the most urgent. When a codeword of C_0 is transmitted, retransmission is requested only when the transmission of this codeword fails. Suppose, for instance, time urgency of this data packet equals 1 and the $(n+m)/3$ bits generated from G and P_0 are sent during

retransmission. In this case the $2(n+m)/3$ unused bits in the same time slots can be used to piggyback the transmission of the two thirds of a new data block. If time urgency equals 2, then the $2(n+m)/3$ bits generated from G together with P_0 and P_1 are sent and the unused $(n+m)/3$ bits are used to piggyback one third of a new data block. Finally, if time urgency equals 3, then the full $n+m$ bits produced from G and P_0, P_1, P_3 are sent. Whatever received at the receiver for retransmission are combined together with the earlier received $n+m$ bits from the codeword of C_0 to go through a decoding procedure and to hope to recover the valid codeword of C_0 .

Let us at this place give an example in prioritizing a packet's time urgency. Two thresholds T_1 and T_2 , with $T_1 < T_2$, are chosen to differentiate priorities. If a data packet must be transmitted immediately, it is of course given the highest priority in time urgency. If the packet can tolerate a deadline of no larger than T_1 , it is given the second highest priority in time urgency. Finally, if the deadline of a packet may be upto T_2 , then it is given the lowest priority 1. To be more specific, for instance, we may set $T_1=2$ and $T_2=6$ time slots.

4. Numerical Experiments

For the (conventional) HARQ-II introduced in Section II, we assume $n=252$ and $k=242$. Thus we choose the $(252,242)$ systematic linear code for C_0 and the codes in Figure 1 for the RCPC. For the adaptive HARQ-II, we employed the code specified in Figure 2.

Regarding the radio channel, we employ the well cited Gilbert-Elliot channel under a roundtrip propagation delay of 2 time slots. The experiments are conducted for an aggregation of 300 data blocks. As for traffic model, we assume packets arrive in batch according to a Poisson process.

Figure 3 shows the result for an environment in which time urgency is not imposed. In this figure the dotted curve represents the result of the conventional HARQ-II. For the adaptive HARQ-II, we show three curves labeled by 1,2, and 3. For level 1, it acts as if each packet have the time urgency 1. In other words, during retransmission only 1/3 of the time slot is used to send the parity check and the rest 2/3 are used to piggyback new information. It is reasonable to see that level 1 exhibits the best performance in throughput. It is also reasonable to see that level 3 and conventional HARQ-II show results so close to each other.

We shall now in Figures 4-5 demonstrate the performance of our adaptive HARQ-II. The throughput efficiency and reliability performance for batch size equal to 10. Here "reliability" is defined to be the ratio of the number of successfully received and in-time packets over the number of successfully received packets. Three values of deadline are considered. In these figures we witness that adaptive HARQ-II indeed performs better than the conventional. Furthermore, when deadline is more relaxed, better throughput and reliability can be expected.

5. References

- [1] S. Lin and Y. M. Wang, "A modified selective-repeat type-II hybrid ARQ system and its performance analysis," IEEE Trans. on Communications, Vol. 31, pp.593-608, May 1983.
- [2] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," IEEE Trans. on Communications, Vol. 36, pp.389-400, April 1988.
- [3] H. Ma and M. El Zarki, "Transmission of MPEG-2 encoded VoD services over wireless access networks using type-II hybrid ARQ schemes with

RCPC codes,” IEEE Int. Conf. on System Sciences, p.8040, January 2000.

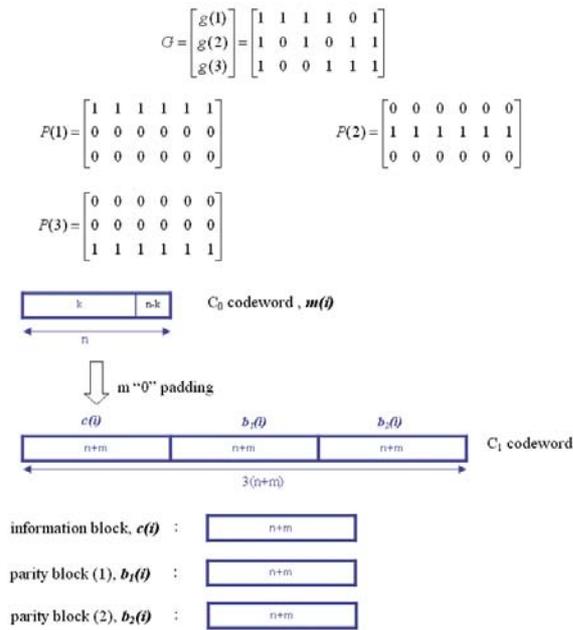


Figure 1. An example to demonstrate the usage of RCPC in HARQ-II.

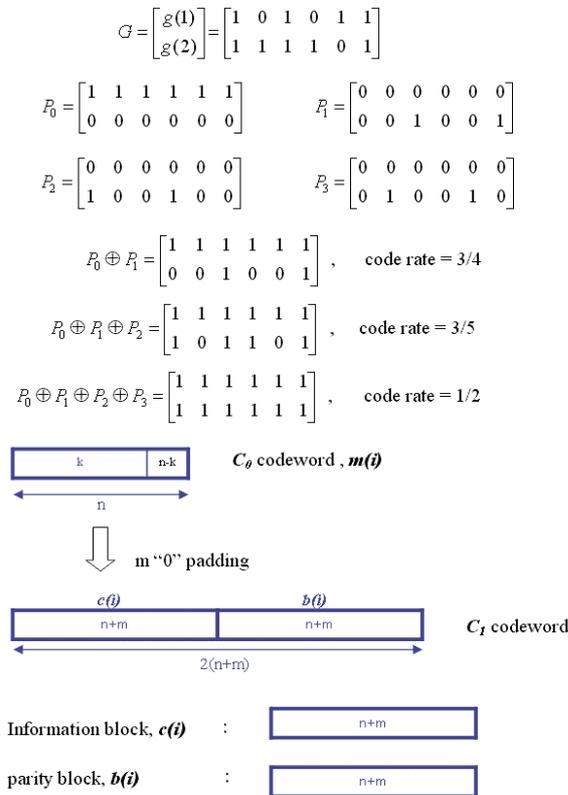


Figure 2. An example to demonstrate the

construction of C_0 and RCPC to be used in an adaptive HARQ-II.

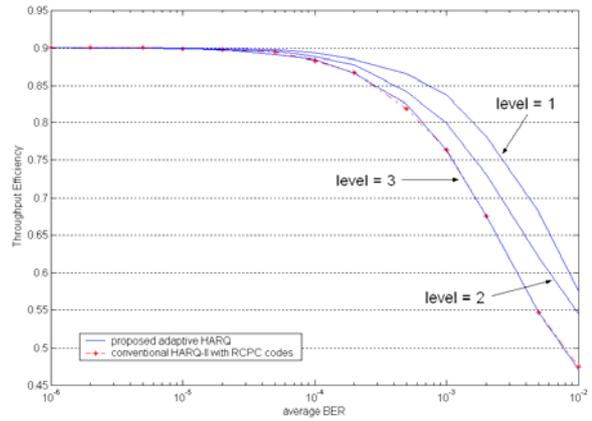


Figure 3. Comparison between conventional HARQ-II and adaptive HARQ-II when time urgency is not imposed.

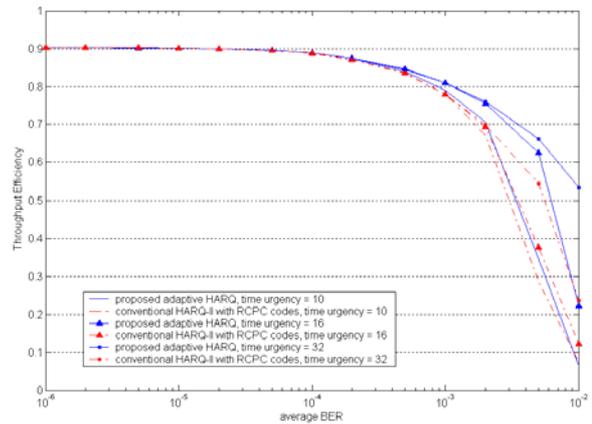


Figure 4. Comparison in throughput between conventional HARQ-II and adaptive HARQ-II when batch size equals 10

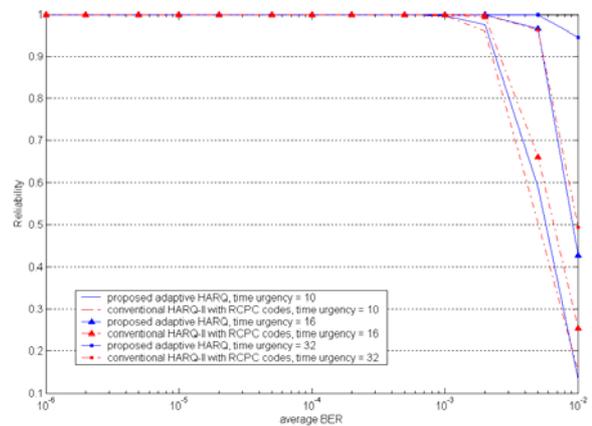


Figure 5. Comparison in reliability between conventional HARQ-II and adaptive HARQ-II when batch size equals 10

Analysis of Boundary Conditions for Real-time Facsimile Transmission over UDP/IP

Bhumin H. Pathak (bpathak@brookes.ac.uk), Geoff Childs (gnchilds@brookes.ac.uk)
and Maaruf Ali (mali@brookes.ac.uk)
Communications Research Group,
School of Technology
Oxford Brookes University, Oxford, OX3 0BP, U.K.

Abstract

A protocol structure, which provides interworking of traditional fax machines with packet networks, provides tremendous opportunities for cost savings by fax calls over packet networks. The scope of this paper is to analyse the boundary conditions under which the traditional fax transmission can be successful over a packet based network where UDP is used as a transport layer protocol. A real-time implementation of traditional fax transmission over a packet-network is carried out, and at the end of the experiments boundary conditions of successful transmission in terms of delay, packet-loss and packet arrival errors are computed. When fax is transmitted using legacy fax terminals and IP, on the transport network, it is important to understand the limitations imposed by timers of different circuit-switched and packet-switched protocols. A conclusion is drawn on the basis of the results obtained by the practical implementations of the actual fax over IP transmission.

1. Description of Fax Services over Packet-Switched Networks

There are basically two types of services available for Fax over packet-switched networks –

- Store-and-Forward method (ITU-T T.37)
- Real-time method (ITU-T T.38)

These two approaches differ in terms of delivery and the method of receipt confirmation.

- **Store-and-Forward Method:** This simple fax service provides fax gateways, which are specially designed for this purpose; with the capabilities to take calls from G3 fax machines. These messages are then converted into simple e-mail messages by the fax gateway. The converted e-mail messages are transmitted over the Internet as a traditional e-mail service. At the receiving end, a fax gateway performs the reverse task. The original fax message is regenerated from the e-mail message and is sent to the destination G3 fax device. The original fax document can be sent as an attachment with the e-mail service in the form of a TIFF image. Although the lack of an immediate confirmation message is a disadvantage, store-and-forward fax has several advantages, including

delivery at off-peak hours, sophisticated retry-on-busy algorithms, and the ability to broadcast a single fax to multiple receiving fax machines.

- **Real-time Method:** In this mode, the real-time fax gateways can deliver faxes to remote fax machines while the sending fax machines are still processing pages. At the source, the gateway receives analogue fax signal and demodulates it into its digital form using a fax modem. Packets are formed out of this digital fax signal and sent over the IP network using the T.38 protocol. At the receiving end, the fax gateway demodulates the digital fax packets into T.30 analogue fax signals to be sent to the destination fax machine through a gateway modem.

Where most of the fax transmission infrastructure is already in place, integrating existing analogue T.30 transmission with T.38 fax gateways, can provide easy migration to the packet-based service.

T.38 fax gateways provide the following functions:

- Demodulate incoming T.30 fax signals at the sending gateway;
- Translate T.30 fax signals into T.38 Internet Fax Protocol (IFP) packets;
- Exchange IFP packets between the sending and receiving T.38 gateways;
- Translate T.38 IFP packets back into T.30 signals at the receiving gateway.
- Two implementations of ITU-T T.38 are possible:

In addition to this, T.38 can also be configured to provide following implementation options.

- **Fax Relay** – Network delay becomes a factor of consideration when real-time fax is deployed. Use of fax-relay can reduce this effectively.
- **Real-time Faxes with Spoofing** – Spoofing techniques are used to extend the delay tolerance of fax machines. These techniques add to the T.30 protocol used by fax machines to communicate, keeping them on line beyond their normal T.30 timeout intervals.

2. Experimental Set-up

A real-time implementation of traditional fax transmission over packet-based network was set-up in the lab. The basic architecture for this fax transmission over packet-based networks is shown in Figure 1. As can be seen from the figure, a legacy fax terminal is used to transmit the fax in a traditional way over a PSTN cloud.

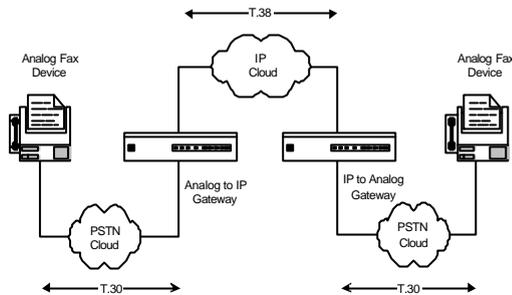


Figure 1. Real-time Implementation for Fax Transmission over IP-Cloud.

The actual fax transmission in this scenario, where T.30 data and signalling is transmitted on T.38 protocol, is described below:

1. The analogue T.30 protocol tries to establish a fax call between the transmitting source fax machine and the gateway. This session is very well understood by all entities as a traditional T.30 fax call at both ends.
2. The gateway converts the received T.30 fax signalling and data to its corresponding T.38 control signalling and data. The fax image data transmitted by T.30 is converted into Internet Fax Protocol (T.38) packets. These packets are then transmitted over packet-based network using either TCP or UDP as transport means.
3. At this point several other protocols also play an important rôle in establishing a session between two gateway entities. The Session Initiation Protocol (SIP) or H.323 with e-num service or static routing can be used for fax-call session establishment and call routing between gateways.
4. Received T.38 is converted back again to T.30 signalling and data, by the receiving gateway, which then is transmitted on a circuit-switched analogue network to the destination fax device.

The following protocol stack diagram shown in Figure 2, illustrates the different protocols working at the different layers of this architecture.

T.38 fax protocol works above the transport layer service provided by TCP/IP or UDP/IP. Use of TCP is more effective when the bandwidth for fax communication is limited, or for communication between Internet Aware Fax

devices, since TCP provides flow control. UDP is more effective when the bandwidth for Fax communication is sufficient. The following part of the paper discusses the analysis of the limitations when UDP/IP is used as a transport media.

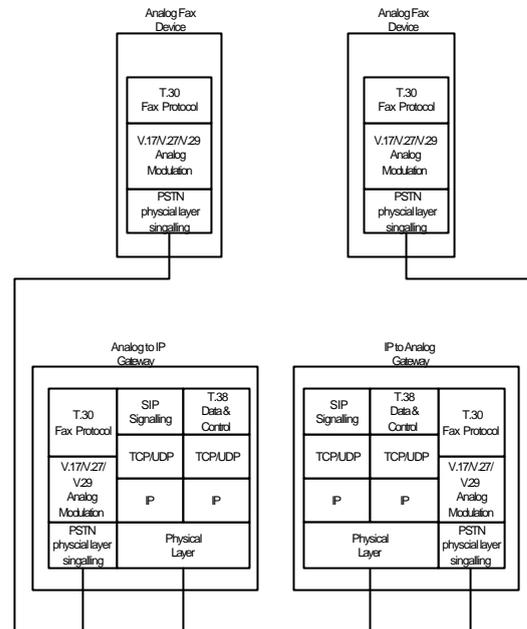


Figure 2. Protocol Stack for T.30 and T.38 Combination.

3. Behaviour and Boundary Conditions with UDP

Figure 3 shows the encapsulation of IFP (T.38) packets into UDP and IP transport.

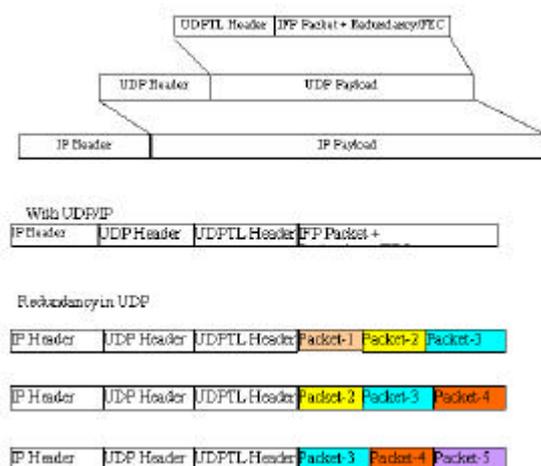


Figure 3. Layered Model IFP/UDP/IP Packet.

These frames with included redundancy for possible error correction are transmitted with each new packet of UDP. So, each UDP packet contains one new and one or more redundant

frames. This is particularly of concern in the case of low-bit rate channels.

4. Sensitivity to Delay

It is observed during several practical implementations that with minimum error protection and correction scheme, T.38 fax session with UDP at transport layer transmits at least one packet of around 214 octets every 20 ms (at the rate of 85.5 kbps approximately) in the ringing phase of a call where minimum packet exchange is expected. This data rate increases with transmission of data packets and with message confirmation phase. As T.38 protocol with UDP provides error correction with redundant transmission of the same packet, most of these packets transmitted have a major portion of the same information contained as packets transmitted before. The amount of redundancy to be included is decided by T.38 during the session establishment. This redundancy is also observed in all consecutive frames captured and analysed during this transmission. For low-bit rate channel this causes considerable amount of bandwidth loss.

According to ITU-T T.30 recommendation, protocol T.30 can sustain delays of up to 3 seconds with $\pm 15\%$ of variation. The sensitivity of T.30 protocol timers plays an important rôle in defining the boundary conditions related to delay and bandwidth parameters on the T.38 part of the above-mentioned architecture. As the delay on packet-based network increases, packet queuing builds up on different nodes. This may cause certain packets critical to T.30 timers on both circuit-switch ends being delayed beyond the time limits defined in the standard or being dropped from a queue as a result of a specified TTL in the packet. This may result in termination of the fax call without successful completion.

Let us assume that the probability of a packet being delayed is uniformly distributed in a certain range over the transmission period, as shown in Figure 4. Which means probability of a packet being delayed by X seconds is the same as the probability of packet being delayed by Y seconds where X and Y both are assumed to be in a certain bracket of minimum to maximum possible delay. In this case, with the help of the experimental set-up designed in the lab as described before, fax over UDP/IP network is transmitted with uniformly distributed delay over a packet-based network. From the various practical observations it is concluded that maximum sustainable delay on this combination of protocols (T.30 and T.38) is around 1800-1850 ms, each-way. But this applies only when the delay is uniformly distributed throughout the transmission period.

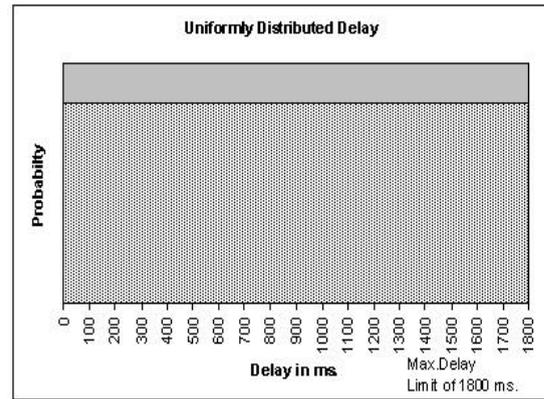


Figure 4. Transmission with Uniformly distributed Delay.

It is also interesting to see the behaviour of this combination of protocols with variable delay parameters. By variable delay it is meant that probability of packet being delayed by X seconds is not the same as packet being delayed by Y seconds, where X and Y both are assumed to be in a certain bracket of possible delay values. If the delay is variable it implies more restrictions on capabilities. When delay is variably distributed between certain values over the transmission period, the behaviour of this protocol architecture actually depends on the range of the delay variation for the delay. If the bracket is stretched beyond some limit, the transmission can be unsuccessful. Figure 5 shows these situations.

This combined architecture transmits successfully when the delay upper limit is 1800 ms and the lower limit is 700 ms. Transmission fails if the delay bracket is stretched beyond the 1100 ms window of operation.

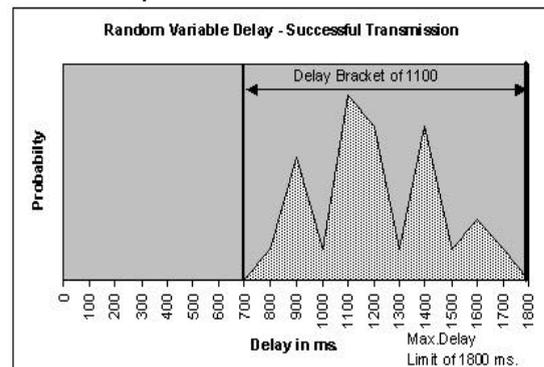


Figure-5. Transmission with Variable Delay-A.

Hence if the maximum delay is set to 1700 ms and the minimum to 600 ms, the transmission can still be successful, as the delay bracket remains unchanged at 1100 ms, though the ultimate upper limit is set to 1800 ms, in any case.

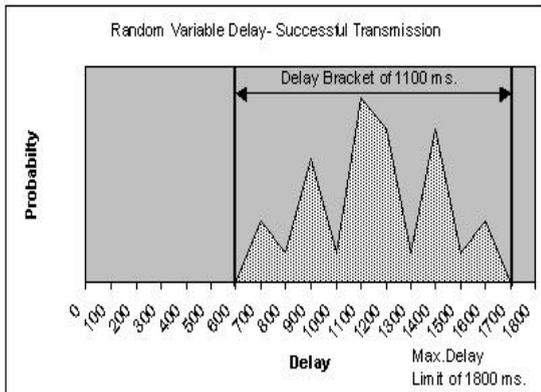


Figure 6. Transmission with variable Delay-B.

Figure 7 indicates the case where the delay bracket is stretched beyond the limit of 1100 ms. In this case transmission is unsuccessful. The call is dropped during the training phase.

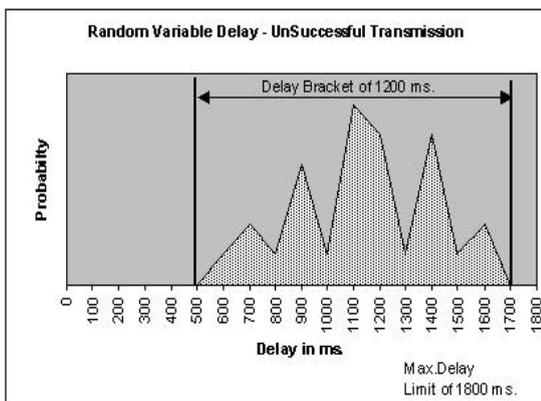


Figure 7. Transmission with Variable Delay-C.

Here in case of variable delay an important observation is that, the protocols break at the data transmission period. This is unlike breaking at initial call establishment and training session in the case of uniformly distributed delay of beyond 1800 ms.

Again when the delay is variable with Gaussian like distribution over the transmission period, this variation plays an important rôle in the capabilities of the protocols. It can be observed that keeping 1800 ms as a mean, and varying delay with standard deviation of 50 ms can make the protocols unsuccessful. The result is the same even when the mean value is decreased to around 1200-1300 ms. This clearly indicates the sensitivity of T.30 & T.38 combined architecture towards variable delay.

5. Sensitivity to Packet Loss

Similar to most of the protocols, combination of T.30 and T.38 is sensitive to packet loss errors. With UDP on the transport layer, it is always a trade-off between the amount of redundancy included into packets and bandwidth limitations. The packet loss is of significant importance when it increases beyond 3-4% of the total packets transmitted (one way with no delay). When coupled with delay parameters at maximum delay of 1800 ms (one way) this decreases to 1-2% of the total packets transmitted.

6. Conclusion

From the above boundary conditions analysis, it is apparent that when T.38 protocol is used in conjunction with T.30, T.30 timers limit the delay boundary condition. As UDP has minimum error correction and protection scheme, it uses redundant transmission to overcome the transmission errors to the detriment of significant bandwidth loss. Also this combination is sensitive to variable delay conditions as described in the paper. Alternatively the TCP protocol can be selected as the transport media for low-bit rate transmission requirements. The architecture can be modified by introducing fax-relay or spoofing techniques to sustain more delay variations.

References

- [1] ITU-T recommendations T.38 – Procedures for real-time Group 3 facsimile communication over IP networks.
- [2] ITU-T recommendations T.37 – Procedures for the transfer of facsimile data via store-and-forward on the Internet.
- [3] A Simple Mode of Facsimile Using Internet Mail, RFC2305.
- [4] Internet FAX Gateway Functions, Internet-Draft: draft-ietf-fax-gateway-protocol-08.txt
- [5] Fax over Packet, White Paper by Texas Instruments, January 1998, Version 2.1, SPEY002.
- [6] A Simple Mode of Facsimile Using Internet Mail, RFC2305, draft-ietf-fax-service-v2-05.txt

Acknowledgement

The author would like to thank Mr. Paul Febvre and Mr. Tom McHugh at Inmarsat Ltd., London, for their continued guidance during the period of research for this paper.

Analysis of HARQ Error Control Scheme for TCP Applications based on ATM over ADSL

Fabiano Bizinelli, Walter Godoy Junior
Federal Center of Technological Education of Parana
Av. Sete de Setembro, 3165, bloco D, 2º andar
Curitiba-PR – CEP 80230-901 – BRAZIL

fabiano@citec.cefetpr.br, godoy@cefetpr.br <http://www.natec.citec.cefetpr.br>

Abstract

This paper analyzes an improvement on the error control scheme for Asymmetric Digital Subscriber Line (ADSL) access network. This scheme avoids TCP congestion avoidance and consequently throughput reduction when acknowledgments are not received due to ADSL frame errors. Simulations were run in NS-2 to verify the performance of the presented scheme in contrast to the standard system for TCP applications based on ATM transport.

1. Introduction

Broadband communication now is an essential service that telecommunications providers must offer to their customers. The need for broadband communication technologies has risen from the idea of integration of different applications, such as voice, video, and file transfer, in a unique method for transmission.

Voice telephony companies can provide broadband service through Digital Subscriber Line (DSL) technologies, a group of wired digital transmission technologies. They hinge on digital transmission over the physical medium used for voice telephony, the copper twisted pair. Asymmetric DSL (ADSL) is just one of these solutions; it relies on different data rates, usually downstream rate much higher than upstream rate. For example, video application, which was the primal target of ADSL, has traffic patterns described by this asymmetry.

However, the boom of broadband market has taken place by the time of the fast Internet access. Since its deployment, Internet applications have been urging for more bandwidth as its contents evolved from simple texts with up to tens of kilobytes in size to multimedia web pages with hundreds or even thousands of kilobytes. Internet is based on a set of protocols defined by the Internet Engineering Task Force, IETF. This group of protocols is known as TCP/IP architecture, which represents the two most important protocols of it: Transmission Control Protocol (TCP) and Internet Protocol (IP).

While IP is a best-effort connectionless network protocol, TCP is a connection-oriented transport protocol that provides a reliable communication channel for user applications. This guarantee is realized by a positive acknowledgement with retransmission

mechanism, together with sliding window for flow control. Moreover, TCP includes congestion avoidance mechanisms that improve performance during congestion periods. TCP congestion avoidance mechanisms reduces throughput, allowing the cleaning of the buffers of all routers through the network.

Any kind of data packet, including TCP/IP, is transported over ADSL in one of two ways: Synchronous Transfer Mode (STM) or Asynchronous Transfer Mode (ATM). ATM is preferred since it permits scalability, service integration and quality of service. This transport mode is based on the routing of small fixed-size frames, the ATM cells, by virtual connection indicators through point-to-point links.

Notwithstanding, ATM and ADSL present some features that interfere on TCP/IP data communication. The interference of twisted pair impairments on TCP was deeply studied in a number of papers [1], [2], [3], [4], [5], [6], [7]. Asymmetry may also decrease throughput of TCP in that it reduces the round trip time [4], [8], [9], [10]. Finally, segmentation of TCP/IP packets in ATM cells also turns into bad performance [4], [10], [11].

This paper approaches the problem of TCP packet losses and consequently throughput decrease due to errors in ADSL frames. It presents a new error control scheme based on type-I hybrid ARQ, using the same Reed-Solomon code defined in ADSL standards. This paper is organized as follow. Chapter 2 discusses some ADSL, ATM, and TCP characteristics. In chapter 3, the new error control scheme is presented. Simulation considerations and results are shown in chapter 4. Conclusions and future work are the final topics covered by chapter 5.

2. ADSL, ATM, and TCP

2.1. ADSL

Asymmetric Digital Subscriber Line (ADSL) allows high-speed data communication at copper local loop simultaneously with Plain Old Telephony Service (POTS). Discrete Multitone (DMT), the standard ADSL modulation scheme, is a multicarrier modulation technique where a 1.104 MHz bandwidth is divided into sub-channels or tones of 4.3125 kHz, which is benefic to simplify reception filters. QAM modulation is applied in each sub-channel where constellation

size depends on the signal-to-noise ratio (SNR) of each tone. Thus, ADSL data transmission rate is adaptive to loop conditions – line length, line gauge, bridged taps, induced noises and other impairments.

Before transmission, data is coded by a forward error correction (FEC) code to improve transmission performance over noisy lines. This code is based on Reed-Solomon coding for GF(256). Part of data may also be interleaved, which implies in better protection against impulsive noise, but higher latency. In this way, during processing, two types of data can coexist in one ADSL frame: interleaved data (high latency, high protection) or fast data (low latency, low protection). Also an optional Trellis Coded Modulation (TCM) should be used.

2.2. ATM

ATM is a transfer technique of special format packets, called cells, that employs asynchronous time division multiplexing (ATDM). ATM cells are 53 bytes long with a 5-byte header. ATM transmits point-to-point information in virtual channels, determined by labels in each ATM cell header.

ATM adaptation layer (AAL) is responsible for the adaptation of higher layer packets into ATM cells. Because these packets usually have variable and large size, in contrast with the fixed and small size of ATM cells, AAL layer has a sub-layer entitled segmentation and reassembly (SAR). This sub-layer is responsible for putting one packet over ATM cells for transmission and rebuilding it from receiving cells.

There are five types of AAL. Each one represents a service class. Connectionless services, such as IP, can be provided by AAL3/4 or AAL5. The last one is simpler, so it is the most common option. Higher layer packets are identified in AAL5 by a logical link control (LLC) header. Afterwards, the encapsulated packet receives an 8-byte trailer plus additional padding bytes to form an AAL5 frame whose length is zero divided by 48, the length of ATM cell payload.

2.3. TCP

The function of TCP is to provide connection-oriented, peer-to-peer, and highly reliable communication between machines in packet-commuted networks and in interconnection systems between these networks. This protocol provides reliability by a positive acknowledgment with retransmission mechanism, working with timers, sequential numbers, and piggy-backing acknowledgment. Sliding window is provided for flow control.

In addition to guarantee correct reception of data, TCP also has mechanisms that contribute with the network during congestion collapse. These are termed Slow Start Additive Recovery and Multiplicative Decrease Congestion Avoidance. The former says that, after congestion period, or as the connection is created, congestion window starts with one segment, and it is

incremented for each received acknowledgment. Congestion window is jointly evaluated with receiver window advertisement to determine throughput. The Multiplicative Decrease algorithm states that it reduces congestion window as source identifies congestion.

Besides, there are two other mechanisms that try to better the efficiency of the two algorithms above: Fast Retransmit and Fast Recovery. Fast Retransmit works to avoid retransmission of a segment that may be delayed, while Fast Recovery cancels execution of Slow Start algorithm after Fast Retransmit.

3. HARQ Error Control Scheme

3.1. Problem Identification

Concerning TCP performance over ATM networks based on ADSL access networks, principal problems detected in other works are related to the interaction between TCP acknowledgement mechanism and ADSL asymmetry, ATM segmentation and reassembly and errors generated by noisy lines. Simultaneously papers covering error control in wireless systems were studied. Some publications discussed the implementation of automatic repeat request (ARQ) as a mechanism to overcome packet losses. Inasmuch as these studies, we proposed a hybrid ARQ error control scheme for ADSL lines.

3.2. Description of the Scheme

The presented error control scheme relies on a type-I hybrid ARQ where error detection and correction capabilities are provided by the standard ADSL Reed-Solomon code [12].

Even that RS code is the same, and that RS encoder does not need to be changed as to implement HARQ in ADSL, a decoder with detection and correction capabilities is necessary. It was chosen a bounded-distance errors-and-erasures correction decoder, which can correct and detect errors for any maximum distance code (MDC) [13].

In the standard operation, if an ADSL frame processed by the RS decoder is received with a correctable error or with no error, decoder recovers the correct word and hands it down. This outcome is termed correct decoding. Otherwise, if a frame is received with an uncorrectable error, decoder chooses a word different from the one transmitted, so higher layers must manage those wrong data. This decoder result is called a decoding error.

On the other hand, HARQ can improve performance by detecting some of the uncorrectable errors and asking for retransmission. When a received frame has uncorrectable but detectable errors, it is said that a decoding failure has occurred. In such a case a negative acknowledgment (NAK) is included in each reverse frame, i.e. frame traveling from the destiny to the source of the concerning traffic. Besides, subsequently received frames are buffered, so that data ordering is

unchanged – it is said that receiver is in blocking state. When a NAK is processed at the other end, the source of concerning data, the correspondent frame is re-transmitted. When this retransmission is received at the destiny, it is processed and passed on in case of correct decoding (or error decoding), together with the buffered frames, unblocking the receiver.

During blocking state, if a frame is received with a detectable error, it is not immediately processed. Indeed, it is still buffered as the other frames. Retransmission request of this frame will only be sent when it is processed, which takes place after the reception of the awaited retransmission. In this case, receiver keeps blocked, but now it waits for the retransmission of a different frame.

Retransmission requests require frame identification. This is achieved intrinsically by frame position beneath ADSL superframe structure. This structure is formed by 68 ADSL data frame, numbered from 0 to 67, and one synchronization frame that carries no data, numbered as frame 68. Sequence numbers limits the retransmission window. In other words, if a retransmission of the requested frame has not been correctly received along 67 frames, it is discarded and subsequent buffered frames are passed on.

One innovative feature of the scheme presented here is the Frame Retransmission Interval (FRI), which avoids overload in local loop [14], [15]. The source does not answer to all NAKs received; instead, one NAK in each FRI sequentially received NAKs is processed. The FRI may be defined according to the propagation delay of the line [14]. Inasmuch as this delay is fewer than the duration of a single frame ($\approx 250 \mu\text{s}$), the range of FRI does not need to be high. A typical value for FRI is 4 [15].

ADSL HARQ scheme may be inserted before CRC block in the transmitter reference model, and consequently after the correspondent block in the receiver.

4. Simulation and Results

4.1. Simulation

Simulations were run in NS-2. New classes were created as to include ATM and ADSL characteristics – segmentation and reassembly, cell multiplexing in ADSL frames, idle cells, RS coding/decoding, byte error models, cell delineation.

Since the smallest unit of transmission in NS-2 is a packet, ATM cell delineation operation principle was based on the identification of first and last bytes in the receiver buffer and in each frame so as to calculate header position and to verify header error correction (HEC) field.

Simulation of bounded-distance decoder was taken as described at Ref. [16]. RS code was specified to operate with two redundancy bytes, which implies in one byte correction and two byte detection capabilities.

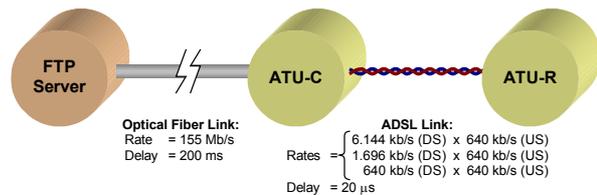


Fig. 1. Topology used during simulations.

Network topology is illustrated at Fig. 1. Three downstream data rates were considered: 6,144 kb/s, 1,696 kb/s, and 640 kb/s. ADSL propagation delay was $20 \mu\text{s}$, simulating a line with approximately feet length. TCP data was generated by a FTP transfer where the server was connected to the ATU-C by a lossless link of 155 Mb/s and 200 ms of delay. The duration of the simulation was 100 seconds.

ADSL features that were not considered during these simulations include: trellis coded modulation (TCM), CRC, interleaving and interleaved buffer, DMT modulation and line impairments. Line impairments were not directly generated. Instead, error generation was only based on bit error rate (BER) from 10^{-7} to 10^{-3} .

4.2. Results

Fig. 2 illustrates relative throughput performance of standard ADSL versus ADSL with HARQ and FRI from 1 to 6, both connected at a downstream rate of 6,144 kb/s and an upstream rate of 640 kb/s. The first observed detail indicates that throughput starts to decrease at bit error rate of $2.5\text{E}-6$. Throughput of ADSL with HARQ is higher than throughput of standard ADSL in the range from $6.3\text{E}-6$ to $2.5\text{E}-4$, reaching up to 10% of difference. From $2.5\text{E}-4$ is similar the throughput of all systems.

In Fig. 3, relative throughput of standard ADSL versus ADSL with HARQ and FRI from 1 to 6, both connected at a downstream rate of 1,696 kb/s and an upstream rate of 640 kb/s, are compared. ADSL with HARQ performs better than standard version from $1.0\text{E}-5$, and difference in relative throughput is up to 30% at $1.0\text{E}-4$.

Fig. 4 shows the outcome of the relative TCP throughput of standard ADSL versus ADSL with HARQ and FRI from 1 to 6, both connected at a downstream and an upstream rate of 640 kb/s. Again, ADSL with HARQ throughput is superior from $1.6\text{E}-5$ in up to 40% for a BER approximately equals to $2.0\text{E}-4$. One can notice that, at a BER of $1.0\text{E}-3$, while TCP throughput of ADSL with the presented error control scheme is between 15% and 20%, traditional ADSL almost cannot carry TCP segments.

For low BER, system performance for both the ADSL with HARQ and the standard ADSL is very similar. In fact, ADSL with HARQ presents a relative TCP throughput a little bit lower than that of standard ADSL. This is caused by the overhead due to an extra byte included in each frame to carry NAK information.

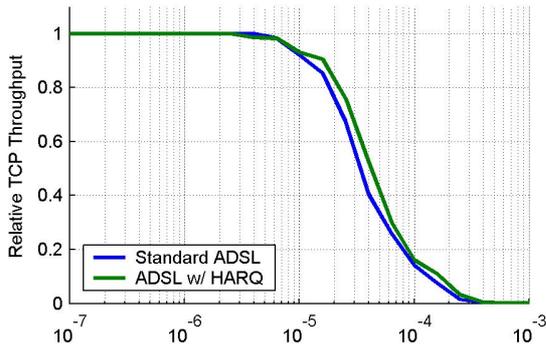


Fig. 2. Relative TCP throughput for transmission rates of 6,144 kb/s \times 640 kb/s.

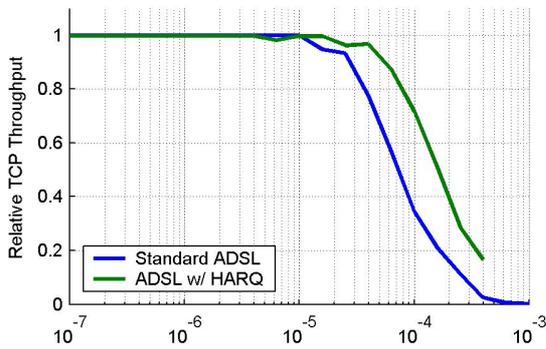


Fig. 3. Relative TCP throughput for transmission rates of 1,696 kb/s \times 640 kb/s.

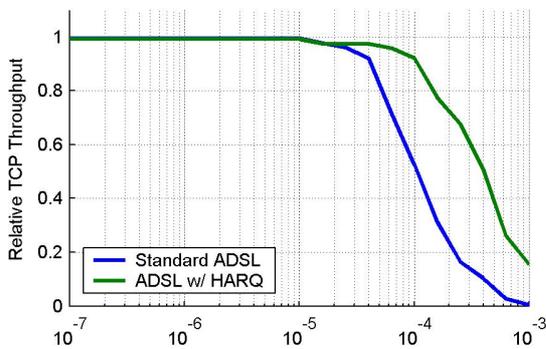


Fig. 4. Relative TCP throughput for transmission rates of 640 kb/s \times 640 kb/s.

5. Conclusions and Future Work

5.1. Conclusions

For high BER, the lower the downstream rate, the higher the advantage of ADSL with HARQ against standard ADSL in terms of relative TCP throughput. This is because of the adaptive characteristic of ADSL: the higher the rate, the higher the number of bytes in each ADSL frame. For a frame that is received with error, the number of damaged cells is proportional to the data rate. If at least one cell that contains bytes from an AAL5 packet presents error, the whole packet

will be discarded. Thus, the probability that a cell with user data is received with error is proportional to the number of cells in each frame.

Nonetheless, variations in relative TCP throughput are higher for systems with higher data rate. This means that the difference in relative TCP throughput for downstream rate of 6,144 kb/s may imply in an absolute difference higher than the difference for lower downstream rates.

5.2. Future Work

In this work we show that HARQ error control scheme for ADSL improves data throughput in TCP/IP networks. We aim to find results for quality of service characteristics in order to verify the operation of the scheme while carrying real-time data.

References

- [1] N. Nedeve, "ATM Cell Error Performance of xDSL Under Impulse Noise," in *Proc. 4th IEEE Intl. Conf. Communications*, Helsinki, June 2001, vol. 4, pp. 1254-1258.
- [2] N. Nedeve, "Data Errors in ADSL and SHDSL Systems due to Impulse Noise," in *Proc. IEEE Intl. Conf. Acoustics, Speech, and Signal Processing*, Orlando, May 2002, vol. 4, pp. 4048-4051.
- [3] D. Laurenson, "Performance of TCP/IP over ADSL subjected to impulse noise," in *Proc. 13th Intl. Symposium Services and Local Access*, Stockholm, June 2000, Section 8:1.
- [4] G. Lu, "The Performance of TCP over ATM on Lossy ADSL Networks," in *Proc. 25th IEEE Conf. Local Computer Networks*, Tampa, FL, October 2000, pp. 418-427.
- [5] X. He, "TCP Performance Analysis of a DMT-based ADSL System," in *Proc. 25th IEEE Conf. Local Computer Networks*, Tampa, FL, October 2000, pp. 439-440.
- [6] X. He, and H. Che, "TCP Performance Analysis and Optimization Over DMT Based ADSL Systems," in *Proc. 4th IEEE Intl. Conf. Communications*, Helsinki, June 2001, vol. 9, pp. 2882-2886.
- [7] X. He, and H. Che, "TCP Performance Analysis and Optimization over DMT Based ADSL Systems," *Computer Communications*, vol. 25, no. 3, pp. 322-328, February 2002.
- [8] H. Balakrishnan, "A Comparison of Mechanisms for Improving TCP Performance over Wireless Links," *IEEE/ACM Trans. Networking*, vol. 5, no. 6, pp. 756-769, December 1997.
- [9] H. Obata, "TCP Performance Analysis on Asymmetric Networks Composed of Satellite and Terrestrial Links," in *IEEE Intl. Conf. Network Protocols*, Osaka, November 2000, pp. 199-206.
- [10] J. Grotz, and G. A. Cope, "Modelling and Design of TCP/IP Based Services over ADSL and ATM," in *IEE Telecommunications Conf.*, no. 451, 1998, pp. 138-148.
- [11] R. J. Gurski, and C. L. Williamson, "TCP over ATM: Simulation Model and Performance Results,"
- [12] F. Bizinelli, *Contribution for the Study of New Error Control Techniques in ADSL Access Networks*, M.Sc. Thesis, Curitiba, CEFET-PR, 2003 [in Portuguese].
- [13] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Upper Saddle River: Prentice Hall, 1995.
- [14] F. Bizinelli, and W. Godoy Jr., "Proposal of a New Algorithm to Improve TCP Performance over ADSL," in *7th IASTED Intl. Conf. Internet and Multimedia Systems and Applications*, Honolulu, August 2003, pp. 324-327.
- [15] F. Bizinelli, and W. Godoy Jr., "Analysis of a Proposed Error Correction Algorithm over ADSL for TCP Applications," in *2nd IASTED Intl. Conf. Communication Systems and Networks*, Benalmádena, September 2003, pp. 223-226.
- [16] A. A. Daraiseh, and C. W. Baum, "Decoder Error and Failure Probabilities for Reed-Solomon Codes: Decodable Vectors Method," *IEEE Trans. Communications*, vol. 46, no. 7, pp. 857-859, July 1998.

Bandwidth Scheduling Techniques for Differentiated Services Support in Ethernet Passive Optical Networks

Dawid Nowak, Philip Perry
Performance Engineering Laboratory
School of Electronic Engineering
Dublin City University
Ireland
Email: {nowakd, perry} @eeng.dcu.ie

John Murphy
Performance Engineering Laboratory
Department of Computer Science
University College Dublin
Ireland
Email: murphyj @eeng.dcu.ie

Abstract—Passive Optical Networks are thought to be the next step in the development of Access Networks and providing broadband access in the "last mile" area. Ethernet PONs (EPON) gain the most attention from the industry as they offer highly flexible, cost effective solution. In this paper we propose algorithms that provide Dynamic Bandwidth Allocation and shift the burden of queue management from the customer to the network, this results in less complicated and more generic equipment used on the customer's premises. We show the results of simulations to validate the effectiveness of algorithms presented.

I. Introduction

With the increasing popularity of the Internet the traffic generated by domestic and small business users has been growing constantly over the last couple of years. Various technologies have been deployed to provide broadband access to the network in the area known as the "last mile". As network operators strive for cost efficiencies, Passive Optical Network (PON) seem to be the next step in the development of Access Networks (AN).

A PON is a point-to-multipoint all optical network with no active elements in the path between the signal source and the destination. On the network side there is an Optical Line Terminator (OLT) unit. The OLT is usually placed in the local exchange and it acts as a point of access to the Wide or Metropolitan Area Network. On the customer's side there is an Optical Network Unit (ONU). An ONU can be placed either in the curb, building or home and its primary task is to convert data between optical and electrical domains.

Two protocols, Asynchronous Transfer Mode (ATM) and Ethernet, have been proposed as the transmission protocol in PONs. In recent years EPONs have gained more attention from the industry. The architecture of an Ethernet network is simple yet extremely efficient. Interoperability between old and new networks can easily be maintained and legacy solutions can be used as EPON data is carried in standard Ethernet frames.

Typically EPON networks are connected in a tree topology with multiple ONUs attached to a single OLT by means of optical splitters as shown in Figs. 1 and 2.

In a downstream (network→user) transmission the OLT uses all the available bandwidth to broadcast

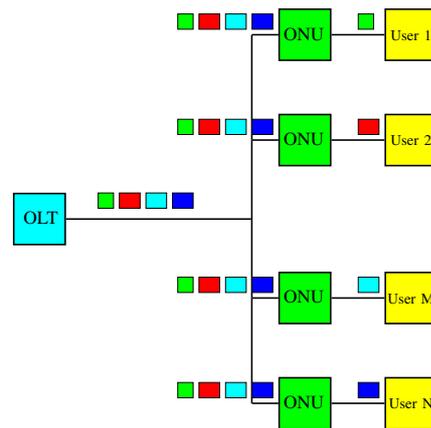


Fig. 1. Downstream transmission in EPON.

packets through the splitter/coupler to every ONU. Each ONU extracts packets from the stream based on the Medium Access Control (MAC) address.

In the upstream direction packets sent by an ONU can only reach the OLT as optical splitter prevents an ONU from receiving packets from other ONUs. In order to avoid collisions between frames from different ONUs at the optical splitter available bandwidth must be shared among all ONUs. The OLT is responsible for assigning a non-overlapping time-slot to each ONU, and ONUs can only transmit during that time-slot. During an off period packets are buffered and when the time arrives send in a burst using all the available bandwidth.

One of the key features of EPON networks is their ability to support Differentiated Services (DiffServ) [1] architecture and offer various levels of quality of service (QoS). Generally three classes of traffic can be distinguished: Expedited Forwarding (EF), Assured Forwarding (AF), and Best Effort (BE). EF services (primarily voice and video) have very strict requirements and demand a constant, low end-to-end delay and jitter. AF services tend to be less sensitive to packet delay but require a guaranteed amount of bandwidth. BE traffic is generated by applications that have no strong requirements regarding traffic properties.

In this paper we analyze various Dynamic Bandwidth Allocation (DBA) algorithms that support

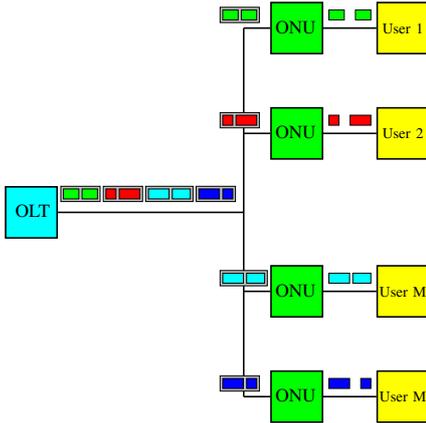


Fig. 2. Upstream transmission in EPON.

DiffServ architecture. In comparison to algorithms already presented in literature, the OLT should be responsible for granting time on a per class rather than per ONU basis, so no intra-ONU scheduling exists. This ensures that the equipment on the customer's side is kept as simple and inexpensive as possible.

II. Background

In EPON networks ONUs cannot transmit at the same time. It is the OLT's responsibility to divide the available bandwidth between ONUs. To achieve that the OLT assigns a non overlapping time-slot to every ONU. In a static bandwidth allocation (SBA) a fixed amount of time is assigned to every ONU. Algorithms with and without support for differentiated services based on a static bandwidth allocation were studied in [2], [3]. The obvious disadvantage of SBA is that bandwidth can not be utilized efficiently. This is especially true in the case where the difference between bandwidth requested by and bandwidth assigned to the source is large.

In [4], a DBA algorithm called "Interleaved Polling with Adaptive Cycle Time (IPACT)" was presented. This algorithm allocates time slots based on information received from ONUs during a polling cycle. IPACT provides statistical multiplexing and gives good bandwidth utilization but due to a variable polling cycle, delay sensitive services are hampered.

Dynamic bandwidth allocation combined with priority scheduling algorithms were studied in [4], [5]. In both papers OLT is responsible for granting time slots to ONUs. Every ONU assigns packets to different queues based on their QoS demands. The main disadvantage of this approach is that to fully support DiffServ, an ONU has to have knowledge about the SLA between a customer and the network provider.

Here, we propose a different approach where all packet scheduling is done at the OLT and the ONU's functionality is limited to the minimum. The bandwidth is allocated per class of traffic rather than per ONU. For every class the OLT informs an ONU about

the allocated transmission window with a Multi-Point Control Protocol (MPCP) [6] GATE message.

Removing the scheduling mechanism from an ONU has two benefits. Firstly the ONU becomes a very simple unit that is easy to manufacture and is inexpensive to maintain. Secondly greater flexibility is achieved as the ONU becomes independent from the OLT. Various scheduling algorithms can be deployed on the OLT without any modifications at the client side. Hence Service Level Agreements (SLA) can be introduced, changed and modified at any time.

III. Dynamic Bandwidth Allocation with DiffServ Support

An efficient bandwidth allocation algorithm is the key to enabling PONs that support DiffServ. Before transmission windows are assigned, various parameters must be taken into account. As for the EF class, delay and jitter are the priority. AF class usually demands various levels of certainty that packets reach the destination and BE class has no strong requirements regarding QoS. The algorithm must balance these factors to achieve the optimal utilization of available bandwidth.

Two MPCP messages are involved in the exchange of control information. The REPORT message is used by an ONU to periodically inform the OLT about the length of its queues. On the other side the OLT issues GATE messages to notify each ONU about transmission times assigned to every class.

A granting cycle is a time in which all active ONUs should have a chance to transmit their data. An increase in the duration of a granting cycle leads to larger delays experienced by packets, as ONU must wait for a longer period of time for its opportunity to send its data. Conversely making a granting cycle too short leads to more bandwidth being wasted to guard intervals that are necessary to separate transmission from two ONUs.

There are two categories of bandwidth allocation algorithms. In static allocation every ONU/class is assigned its fixed share of bandwidth. In dynamic allocation bandwidth is assigned proportionally to the reported queue length. Data that can't be sent during a granting cycle has to wait for the next opportunity.

Here, we want to present new algorithms that could be used with EPONs supporting different classes of service. Let the system have N ONUs with q queues at each ONU. Also, let $B_{n,q}$ be a percent of the total bandwidth requested/assigned to a queue q at ONU n .

A. DBA with Priority Transmission Order - DBA-P

In this algorithm at the first stage every class is assigned bandwidth that is proportional to the bandwidth requested.

$$B_{n,q}^{assigned} = \frac{B_{n,q}^{request}}{\sum_{n,q} B_{n,q}^{request}} \quad (1)$$

To improve the performance of this algorithm for high priority traffic every ONU has a chance to transmit its EF packets at the beginning of the granting cycle.

B. DBA with a Guaranteed Minimum bandwidth - DBA-GM

This approach is a tradeoff between static and dynamic allocation schemes. Every class is assigned some minimal amount of bandwidth that was agreed in the SLA between the network provider and a customer. The amount of bandwidth assigned is dependent on the priority of the traffic.

$$B^{min} = \sum_{n,q} B^{min}_{n,q} \text{ and } B^{min} < 1 \quad (2)$$

The remaining bandwidth is assigned to all classes proportionally to the bandwidth required.

$$B^{avail} = 1 - B^{min} \quad (3)$$

$$B^{excess} = \sum_{n,q} B^{request}_{n,q} - B^{min}_{n,q} \quad (4)$$

$$B^{assigned}_{n,q} = B^{min}_{n,q} + \frac{B^{avail}}{B^{excess}} (B^{request}_{n,q} - B^{min}_{n,q}) \quad (5)$$

As in DBA-P in order to minimize the delay and jitter experienced by EF classes, these classes are assigned bandwidth at the beginning of the granting cycle. The modified version of DBA-GM scheme is shown as DBA-GM-P.

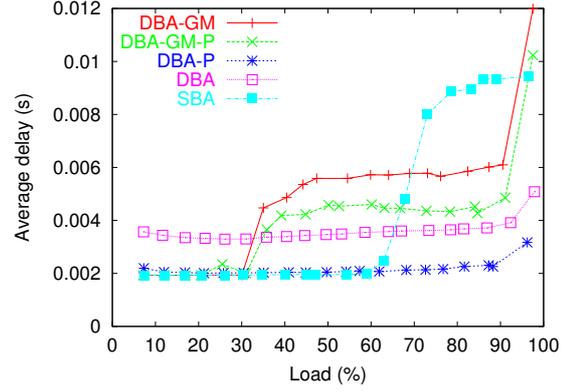
IV. Performance Evaluation

To measure the performance of each bandwidth allocation algorithm we designed an event-driven C++ based EPON simulator. In our research we used 16 ONUs connected in a tree topology to a single OLT operating at a speed of 1Gb/s. Each ONU has three queues with an independent buffering space. The guard time between transmissions from different ONUs is set to $1\mu s$ and the value of Inter-Frame Gap (IFG) between Ethernet packets is 96 bits. In our simulator the length of a granting cycle stays the same throughout the simulation.

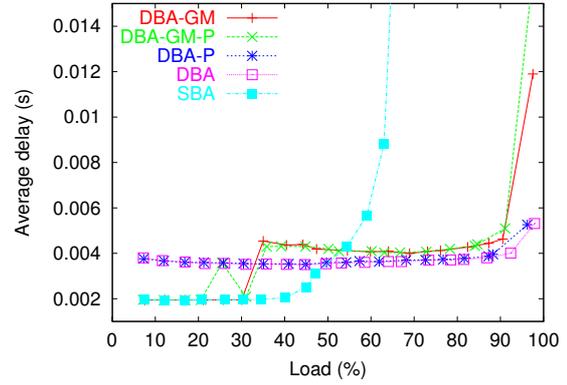
It has been shown that most network traffic (i.e., http, ftp and VBR services) is best characterized by self-similarity and long-range dependence [7]. To model a high priority EF class of traffic (e.g., voice applications) a Poisson distribution is generally used.

In our simulator we used a class of a high priority traffic with a fixed packet length of 70 bytes. A length of packets for AF and BE classes was uniformly distributed between 64 and 1518 bytes. We ran our simulations for various proportions in the volume of EF, AF and BE traffic. Average and maximum packet delay were measured during experiments.

We compared the performance of algorithms proposed in this paper with the performance of SBA and DBA algorithms. The results for 20% of EF, 40% of



(a) EF - average delay



(b) AF - average delay

Fig. 3. Algorithms performance comparison. EF - 20%, AF - 40% and BE- 40%.

AF and 40% of BE traffic are presented in Figs. 3 and 4, and for 30% of EF, 60% of AF and 10% of BE in Fig. 5.

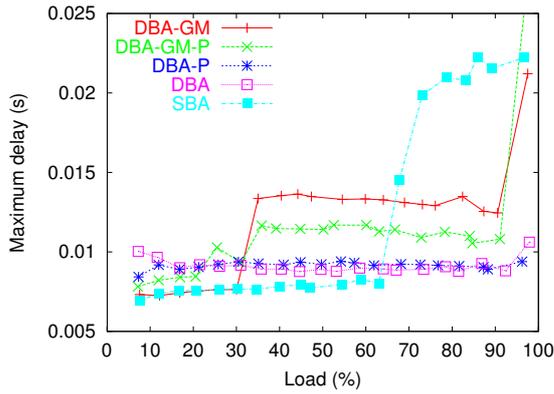
For light loads the SBA scheme showed better performance compared to other algorithms. On the other hand for heavy loads average and maximum delays are much longer than for other algorithms.

DBA algorithm showed a steady performance under various conditions and the difference in the average delay for small and large loads was not larger 2ms.

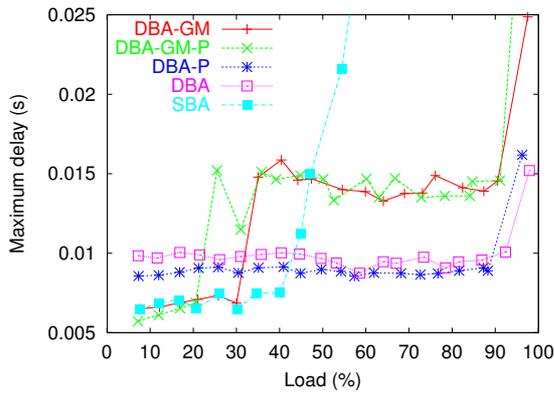
The DBA-P scheme gave good results for both EF and AF classes of traffic. The values of an average delay for light loads are similar to SBA. For heavy loads DBA-P outperforms all other algorithms.

The DBA-GM algorithm showed properties of both static and dynamic allocation. Under low traffic conditions the average delay was as low as for the SBA. If traffic offered was average or heavy it behaved as DBA, although its performance was worse as some bandwidth was statically allocated.

The improved versions of those algorithms (DBA-GM-P) produced better results, due to the fact that high priority packets were sent at the beginning of every granting cycle.



(a) EF - maximum delay



(b) AF - maximum delay

Fig. 4. Algorithms performance comparison. EF - 20%, AF - 40% and BE- 40%.

V. Conclusion

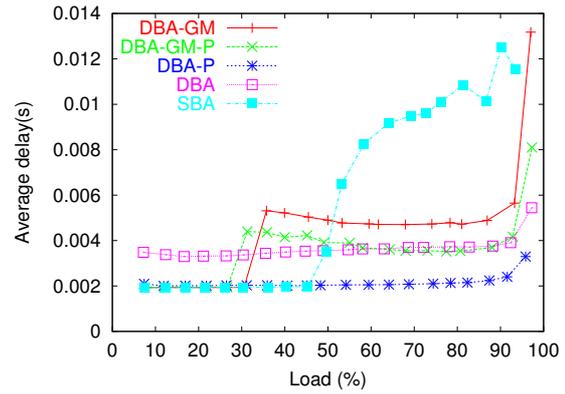
In this paper we addressed the problem of the support for DiffServ in EPON. We proposed algorithms that shifted the responsibility of Access Control and queue management from the ONU to the OLT, as this creates a possibility of developing more generic, less complicated hence cheaper equipment.

We introduced new algorithms that supported that scheme. We ran detailed simulation experiments to analyze their performance.

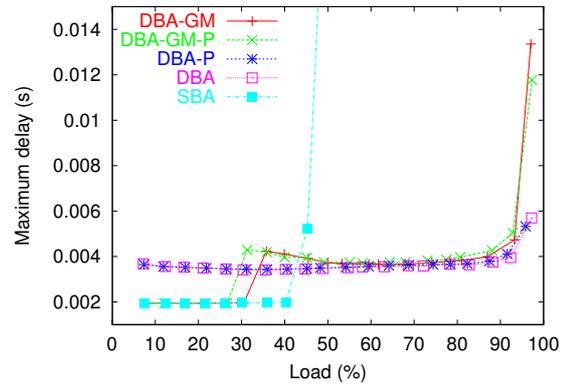
A novel DBA-P scheme showed the best performance for EF and AF class of traffic. The disadvantage of this approach was that as the bandwidth was assigned proportionally to the reported length of queues there was no mechanism to guarantee that the user was allocated as much bandwidth as it was promised in the SLA.

We addressed that problem and proposed an DBA-GM algorithm, where a protection of parameters was achieved by static assignment of a certain amount of bandwidth agreed in the SLA.

The results showed that DBA-GM performance was comparable but not as good as the DBA scheme. Considerable improvement in the values of average delay for EF classes was achieved when a mechanism of priority transmission was applied.



(a) EF



(b) AF

Fig. 5. Average delays comparison for EF - 20%, AF - 60% and BE- 10%.

References

- [1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W.Weiss, "RFC-2475 - An Architecture for Differentiated Services," 1998, <http://www.ietf.org/rfc/rfc2475.txt>.
- [2] G.Kramer, B.Mukherjee, and A.Maislos, "Ethernet Passive Optical Networks (EPON): a missing link in an end-to-end optical internet," S. Dixit, Ed. John Wiley & Sons, Inc, Mar. 2003, ch. 8.
- [3] G.Kramer and B.Mukherjee, "Ethernet PON: design and analysis of an optical access network," *Photonic Network Communication*, vol. 3, no. 3, pp. 307-319, July 2001.
- [4] G.Kramer, B.Mukherjee, S.Dixit, Y.Ye, and R.Hirth, "On supporting differentiated classes of service in EPON-based access network," *OSA Journal of Optical Networking*, vol. 1, no. 8/9, pp. 280-298, 2002.
- [5] Ch.M.Assi, Y.Ye, S.Dixit, and M.A.Ali, "Dynamic Bandwidth Allocation for Quality-of-Service Over Ethernet PONs," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1467-1477, Nov. 2003.
- [6] "IEEE 802.1Q," 1998.
- [7] W. Leland, M. Taquq, W. Willinger, and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1-15, Feb. 1994.

BER Improvement in the Zero-IF Receivers

Vahan Nahapetyan
Yerevan State University, Armenia
E-mail: vahan_n@freenet.am

Arsen Hakhoumian
Yerevan State University, Armenia
E-mail: arsen@irphe.am

Abstract

The double balance mixer scheme in-phase signal input and double-quadrature Local Oscillator is suggested for application on IQ direct conversion receivers. LO leakage influence on DC component and BER of the digital signal is shown. A model of the system is designed which is illustrating the improvement of the receiver sensitivity in case of usage the suggested scheme.

1. Introduction

The increasing interest in the direct conversion or zero-IF receiver architecture (DCR) is based on several qualities of this type of reception which makes it very suitable for integration as well as multi-band multi-standard operation [1]. The main advantage of DCR, contrary to the most widely used heterodyne reception, is an achievement of high image rejection avoiding the use of expensive bulky off-chip filters. Among the problems existing in DCR, limiting their wide application, DC offset is one of the most serious.

To solve this problem the low-IF single conversion receiver architecture was proposed. Unfortunately, the low-IF receivers have comparatively insufficient suppression of an adjacent channel, especially in case of I/Q imbalance [2].

An alternative way of the DC problem solution is design of the DCR structures, where high LO-RF isolation is achieved [3].

DC offset caused by various phenomena contribute to the creation of DC signals. These phenomena can be separated on the three main groups, such as a) LO leakage to the LNA and mixers input due to substrate coupling, ground bounds, bond wire radiation etc., b) LO leakage to the antenna through the mixers and LNA due to their non-sufficient isolation, c) even-order nonlinearity of LNA and mixer. The strength of DC caused by the phenomena of group a) is influenced by chip technology and can be reduced by careful layout or by suitable post processing digital signal processing (DSP) at baseband. DSP removes the DC offset in a way that using cannot be duplicated in the analog domain. For the cases of DC caused by the

phenomena of groups b) and c) reductions in DC signal can be achieved in the analog domain by special circuit design [4]. This paper presents one such circuit.

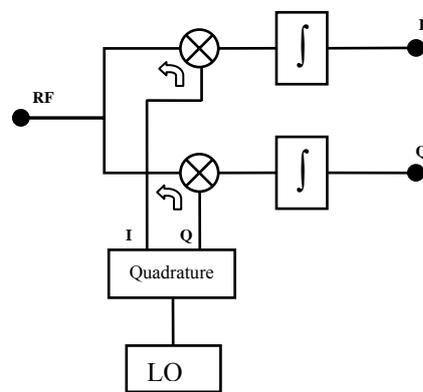


Figure 1. Direct conversion receiver

2. Four-quadrant multiplier DCR

The basic diagram of the conventional IQ DCR is given in Fig. 1. The possible DC offset due to LO self-mixing can be estimated using LO level in the mixers, LO-to-RF isolation and reflection from the RF interface mismatch. Take into account the typical value for LO-to-RF isolation of the mixer -20dB, reverse gain of the LNA -20 dB, corresponding to mismatch SWR=1.1 reflection -20 dB and required LO level approximately 0 dBm, the LO leakage -60 dBm at the LNA input can be obtained. This value of LO leakage power is 30 dB higher than required sensitivity threshold of the receiver. Leakage power after amplification of LNA and self-mixing with LO produced at the output of the mixer DC offset on the order of 10 mV which is enough high to saturate the following circuits.

To solve this kind of DC offset problem various compensative and balance architectures have been suggested [1], [5]. To reduce the LO leakage and DC offset we suggest to use the scheme of analogue double balanced mixers presented in Fig. 2. As distinct from well known scheme with double-quadrature division of input RF and LO [6], here we use in-phase division of RF and quadrature-antiphase

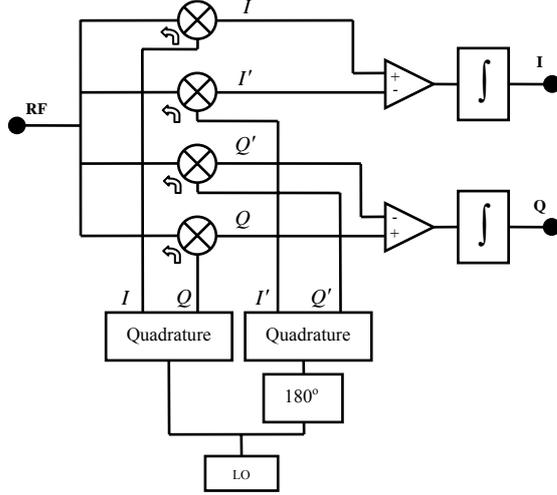


Figure 2. Double balanced DCR

division of LO. Such construction allows suppressing LO leakage to the front end, because at the entry of in-phase splitter the two equal antiphase components have been summed. Presence of differential pairs (I, I') and (Q, Q') enables suppression of residual DC offsets at the I and Q outputs.

Double Quadrature LO generation is however a critical part of this system due to gain and phase mismatches. To create an influence of the gain (Δ) and phase (Θ) mismatches on behavior of this scheme let us write the input RF signal as follows

$$LO_I = (1 + \Delta_1) \cos(\omega_c t + \Theta_1) \quad (2a)$$

$$LO_{I'} = -(1 - \Delta_1) \cos(\omega_c t - \Theta_1) \quad (2b)$$

$$LO_Q = (1 + \Delta_2) \sin(\omega_c t + \Theta_2) \quad (2c)$$

$$LO_{Q'} = -(1 - \Delta_2) \sin(\omega_c t - \Theta_2) \quad (2d)$$

Hence the LO leakage at RF port will be presented as

$$LO_L = L_R \cdot L_M (LO_I + LO_{I'} + LO_Q + LO_{Q'}) \quad (3)$$

The improvement of the LO leakage suppression (L_{4IQ}) in the suggested scheme compared to those in the non-balanced scheme (L_{IQ}) which was presented in Fig. 1 is shown in Fig. 3.

The output of the mixers after filtering of the double-frequency components are the following

$$I = \frac{1}{8} (1 + \Delta_1) [a(t) \sin \Theta_1 + b(t) \cos \Theta_1] + DC_I \quad (4a)$$

$$I' = -\frac{1}{8} (1 - \Delta_1) [-a(t) \sin \Theta_1 + b(t) \cos \Theta_1] + DC_{I'} \quad (4b)$$

$$Q = \frac{1}{8} (1 + \Delta_2) [a(t) \cos \Theta_2 + b(t) \sin \Theta_2] + DC_Q \quad (4c)$$

$$Q' = -\frac{1}{8} (1 - \Delta_2) [a(t) \cos \Theta_2 - b(t) \sin \Theta_2] + DC_{Q'} \quad (4d)$$

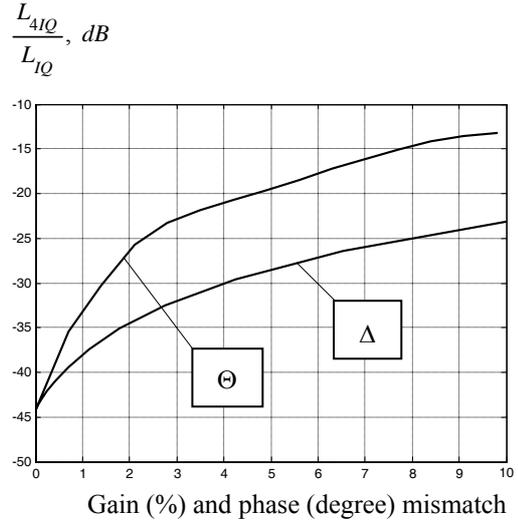


Figure 3. Leakage suppression at RF port in double balanced DCR

where $DC_I, DC_{I'}, DC_Q, DC_{Q'}$ are DC offsets at corresponding branches and can be calculated using (2) and (3). The improvement of the DC offset (DC_{4IQ}) in the suggested scheme compared to those in the non-balanced arrangement (DC_{IQ}) is shown in Fig. 4.

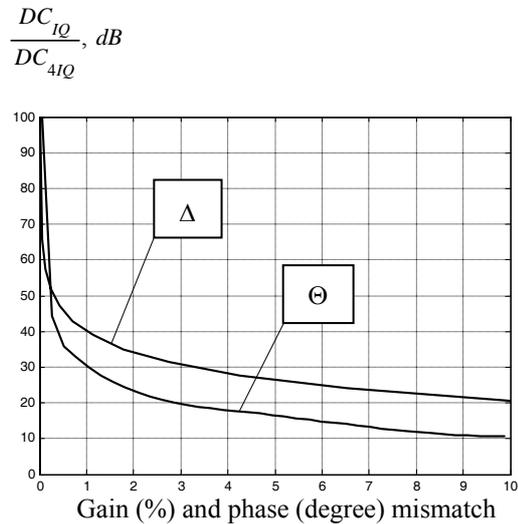


Figure 4. DC offset improvement in double balanced DCR against the gain and phase mismatch

As it can be seen from (2) I, I' and Q, Q' are represented as differential pairs. Hence the wanted signal can be obtained in the following way

$$I(t) = I - I'$$

$$Q(t) = Q - Q'$$

The imbalance errors can be extracted as a sum of the I, I' and Q, Q' pairs

$$\varepsilon(I) = I + I'$$

$$\varepsilon(Q) = Q + Q'$$

These values can be used for imbalance compensation at the source of the LO phase splitter, or digitally after A/D conversion.

As it is seen from the above discussion, LO leakage, DC offset and I/Q imbalance strongly depend on the gain and phase mismatches of the LO splitter. The use of a cascaded four-branch RC polyphase network [7] or polyphase oscillator [8] makes it possible to reach 0.5° phase error and 0.5 dB amplitude error. For such mismatches the suggested scheme provides LO leakage suppression at the RF input of more than 90 dB which is sufficient for many applications of DCR.

3. Digital link simulation model

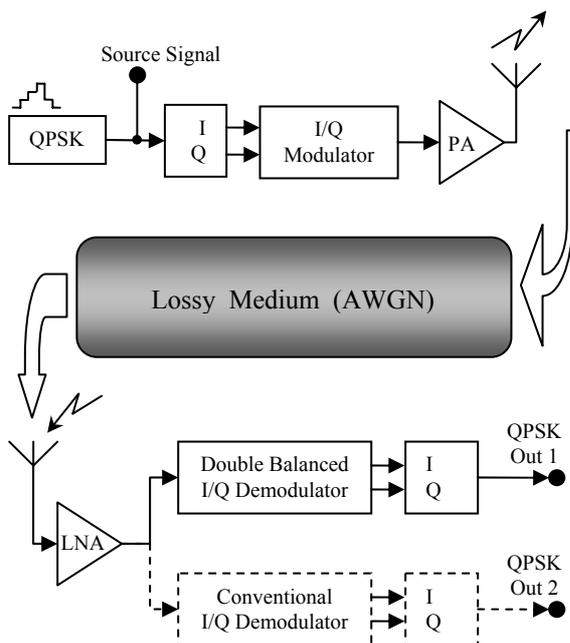


Figure 5. Block diagram of the QPSK digital link model

In the latest decades the major interest of the radio wave propagation is related to the wireless communication especially the digital signal transmission and reception (cell phones, wireless computer network, satellite communication, military issues etc.). The main requirements for the digital

transmitters/receivers are the possibility of easy integration with digital signal processing units, small size and simplicity.

A model in the Matlab 6.5 software environment is designed for the direct conversion receiver behavior consideration using well known QPSK signal for traditional and suggested schemes.

The block diagram of the model is shown in Fig. 5. Four-level baseband digital signal is being divided into two binary streams. These streams are connected to the I and Q entries of the digital I/Q modulator.

Modulated and propagated signal is being passed through the noise channel simulator. After the reception it was amplified by the low noise amplifier and demodulated separately by double balanced receiver and conventional I/Q direct conversion receiver. There are integrators in the demodulator blocks (see Fig. 1 and Fig. 2), one per each of I and Q channels, which are accumulating the energy during one bit duration, comparing it with the neutral value and giving the value of the demodulated bit. Then two binary streams are combined into one four-level stream.

The outputs of double balanced and conventional demodulators are being compared with initial source signal to determine the error rate of detected symbols stream. SER calculator block compares each received symbol with original one and in case of mismatch increments its internal error counter. The output of SER calculator block is the ratio of errors count and the count of the total symbols received.

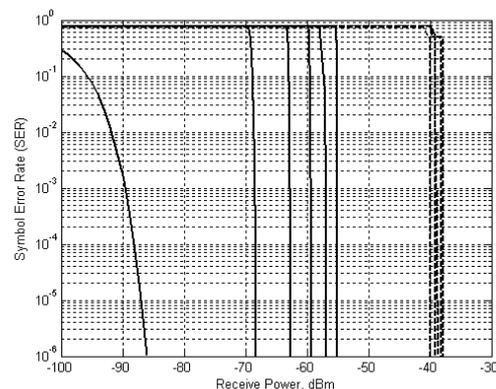


Figure 6. Symbol error rate dependencies from input signal power for the double balanced receiver (solid lines) and conventional receiver (dashed lines). Appropriate phase mismatches are 0, 1, 2, 3, 4, 5 degrees from left to right

4. Simulation results

The noise level of LNA was kept unchanged ($E_s/N_0=10$ dB, when input signal power equals -60 dBm), while the input signal power was varying within wide range for both receivers.

For the different phase mismatch values the dependency between received signal power and symbol error rate was computed and shown in Fig. 6.

In case of zero degree phase mismatch for the suggested scheme we have the first solid line from left and the first dashed line for the conventional DCR. As we can see the first curve is describing well known relationship between BER and signal power. Thus, even big LO leakage does not distort the characteristics of the receiver in case of ideal LO and phase shifters. Increasing the phase mismatch deforms this characteristics and minimal required power of the signal grows up to -55 dBm i.e. the LO leakage becomes more essential.

On the other hand the behavior of the conventional DCR is much worse. The dependency of symbol error rate from signal power is steeply inclined and the sensitivity has a weak dependency from the phase mismatch.

Anyway, as we can note from Fig. 7, even in case of sufficient (5 degree) phase mismatch the dynamic range of the double balanced DCR is about 20 dBm higher compared with conventional one.

Assuming the acceptable symbol error rate 10^{-6} the minimum required signal power was calculated for different phase mismatches. The dependency between the phase mismatch and required power level is shown in Fig. 7.

The slope of the curve for double balanced DCR is sufficient when the phase mismatch is little and it

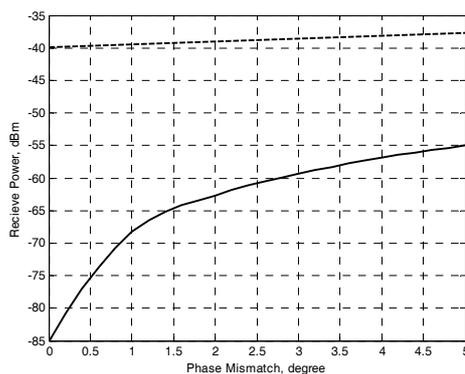


Figure 7. Receiver sensitivity dependencies from phase mismatch in case of fixed acceptable SER= 10^{-6} for double balanced DCR (solid line) and conventional receiver (dashed line)

becomes less growing when the mismatch is high.

As we can note again from Fig. 7 the sensitivity of the double balanced receiver is much higher

compared with conventional one even in case of significant phase imbalance. In case of 0.5 degree phase mismatch we have about 35 dB sensitivity improvement.

5. Conclusions

The usage of suggested double balanced mixer in the direct conversion receivers results DC offset suppression and therefore increased sensitivity and dynamic range of whole receiver. This technique is quite simple and has low cost. It can be realized within one integral circuit. Based on the cost and performance analysis, it is believed that there will be a number of successful applications of this technique, especially to design transceivers for the systems like GSM, Bluetooth and 3G/4G wireless networks.

References

1. A. Parssinen, "Direct Conversion Receivers in Wide-Band Systems", Kluwer Academic Publishers, Boston, USA, 2001.
2. J. Glas, "Digital I/Q Imbalance Compensation in a Low-If Receiver", Bell Labs, Lucent Technologies, NJ, USA, 1998.
3. F. Xiaohui Li and G. Corsetto, "Multi-Order Cancellation Technology", *Microwave Journal Vol. 40, No. 10*, Horizon House, USA, 1997, pp. 84-91.
4. V. Nahapetyan, A. Hakhoumian and A. Hakhumyan, "DC Offset Suppression in Double Balance Direct Conversion Receivers", *Ninth International Conference on HF Radio Systems and Techniques*, IEE, University of Bath, UK, June 2003, pp. 78-80.
5. B. Razavi, "Analog and Digital Signal Processing", *IEEE Transactions on Circuits and Systems—II Vol. 44 No. 6*, 1997, pp. 428-435.
6. J. Crols and M.S.J. Steyaert, *IEEE J. Solid-State Circuits Vol. 31*, 1995, pp. 1483-1492.
7. M.J. Gingell, *Electrical Comman Vol. 48 No. 1-2*, 1973, pp. 21-25.
8. A.W. Buchwald and K.W. Martin, *Electron Lett. Vol. 27 No. 4*, 1991, pp. 309-310.

Billing solutions for 3G data services

Venkata Kapavaparu and Peter Ball

Communications Group
School of Technology
Oxford Brookes University
Gypsy Lane, Oxford, OX3 0BP

pball@brookes.ac.uk

Abstract

Third generation mobile networks will support a range of new data services and the method used to charge for these services will influence the take up and the resultant revenue streams. This paper reviews the possible charging schemes and compares the revenue that can be obtained using flat rate and packet based charging. It shows that the customer spend profile influences which of these schemes generates the most revenue. Packet based charging encourages customers to try out new services and this may be an important factor in kick starting the 3G business. Longer term a hybrid charging solution may emerge incorporating combinations of flat rate and packet based charging, differential charges for QoS classes, content and advertising charges.

1. Introduction

Third generation mobile networks have the capability to support a wide range of new services including video-conferencing, on-line banking, stock trading, sports updates, web browsing and e-mail, mobile commerce and streamed video. 3G operators have invested heavily in establishing the new networks required to support these services and their success depends on the ability to generate strong revenue streams. The method selected for charging will have a strong influence on the take up of the services and the revenue that can be generated from the services.

Voice telephony has traditionally been charged on a time/distance basis. This is appropriate for a circuit switched service as customers pay for what they use. Pre-pay schemes in conjunction with time/distance based charging have encouraged the take up of service in second generation mobile systems because they provide a low cost entry and protect the customer from unexpectedly large bills.

The new services in third generation mobile systems are data services delivered as packets similar to the Internet. With these services there is no longer a direct relationship between the length

of a call and the number of packets delivered, as the time taken to deliver a service may vary due to congestion on the network. Packet based services therefore need a different charging scheme.

In this paper charging schemes for data based services are reviewed. This is followed by an analysis of the revenue that can be obtained by flat rate and packet based charging and consideration of possible hybrid schemes.

2. Charging models

A number of different approaches to charging have been proposed for third generation mobile systems [1,2] these include:

- Flat rate charging
- Usage-based charging (based on packet count)
- QoS-based charging
- Content based charging
- Advertisement charging

Flat rate charging:

Flat rate charging has been adopted for broadband Internet services. This system has the advantage that it is easy for the customer to understand and customers know what they will be billed so there are no surprises. For power users it has the advantage that there is unlimited access at a fixed price.

A disadvantage of this scheme is that unlimited access could lead to network congestion. Operators are then under pressure to install more infrastructure with no corresponding increase in revenue. In addition increasing prices does not control the use of bandwidth.

Packet based charging:

With packet based charging, users only pay for what they use. In addition, because there is no minimum fee, it encourages new users to try new services. It has been successfully adopted as the main charging strategy for DoCoMo's i-mode service [2]. It has the benefit that customers are

charged proportional to their use of the network which avoids unnecessary network usage. Pricing also gives the operator control of network usage.

The disadvantage is that customers may not understand charging on the basis of packets. Moreover packets do not necessarily reflect the value of the content or service.

QoS based charging:

3rd generation mobile networks have four QoS classes: conversational, streaming, interactive, and background classes [1]. Different charges can be made for each of these classes.

An alternative method of QoS charging is the so-called Paris Metro charging scheme [3, 4]. This scheme takes its name from the two-class system formerly used on the Paris Metro, where customers can pay a premium to have a priority service. Under this scheme customers that choose to pay a higher rate are given access to a separate channel. Because this channel is more expensive it attracts fewer customers and therefore the quality of service is higher. The disadvantage of this approach is that it does not use the available bandwidth to the maximum efficiency.

Content based charging:

Third generation mobile services could include a charge for the content delivered. This could be charged directly by the content supplier or the network operator could pay the content supplier and then issue a single bill to the customer for both the content and the network access. The advantage of including a charge for content is that the price would provide a mechanism for controlling access to very popular material. Charging for content also provides a financial reward to content developers and an incentive to develop new material.

Advertising:

Mobile operators have unique access to their customers that is potentially of considerable value to retailers. Operators can charge retailers to advertise their products directly to mobile phone users. The key feature of mobile phones is that the network is aware of the location of the phone user and this allows advertising to be location specific and this may be of greater value to both the phone user and the advertiser.

3. Packet rate v flat rate charging

In this section the flat rate charging and packet based charging are considered in more detail. In particular a comparison is made of the revenue that can be generated from each method of charging.

If customers are charged based on the number of packets received and sent, then the average revenue per user will have a certain distribution. Some users will spend more, some less. We have considered four possible distributions, Gaussian distributions with a mean (M) of 12 units and standard deviations (SD) of 2 and 4 units, a linear distribution and a negative exponential distribution as shown in Figure 1. A range of values has been taken from 0 – 24. These are arbitrary units and represent a range, however it can be considered as the average revenue per user per month in Euros in line with the values given for the i-mode revenues given in [5]. A mean value of 12 has been taken for the Gaussian distribution again corresponding to the value in [5].

The total revenue for a packet based charging system is given by the product of the revenue per user and the number of users.

We make the assumption that this distribution represents the desired spend pattern of the customers. So if a flat rate charge is used, then those customers that would choose to spend more than the flat rate under a packet based scheme would benefit and are likely to adopt the service. Those that would have spent less must decide if they want to spend extra to adopt the service with the incentive of unlimited access. The resultant revenue from a flat rate charge therefore depends on how many customers take up the service.

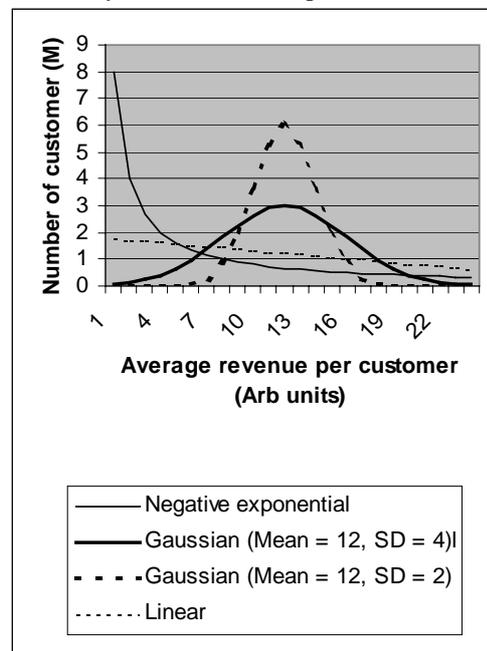


Figure 1 Customer spend distributions

We have calculated the revenue obtained using flat rate charging as a function of the parameter P. Customers that have a spend value of P less than the flat rate or greater are included in the flat rate

revenue. This is shown in Figure 2 for the case of a flat rate of 13 and $P = 4$. All customers with an average spend of 9 or more are included in the flat rate revenue.

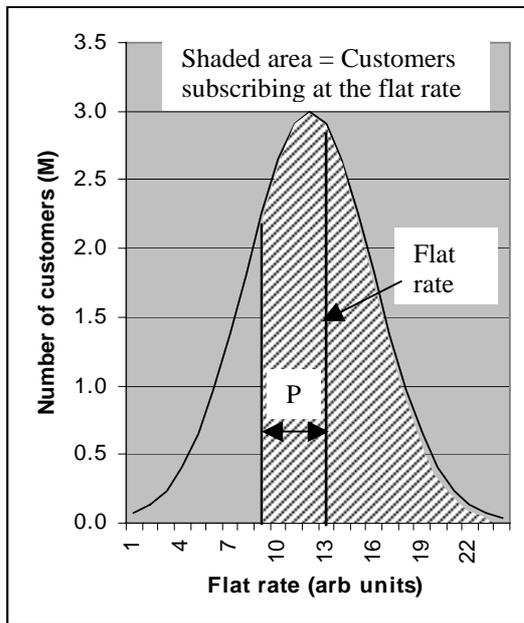


Figure 2 Customers subscribing at flat rate

Figure 3 shows the ratio of the flat rate revenue to the packet based revenue as a function of the flat rate, for values of P from 0 to 12.

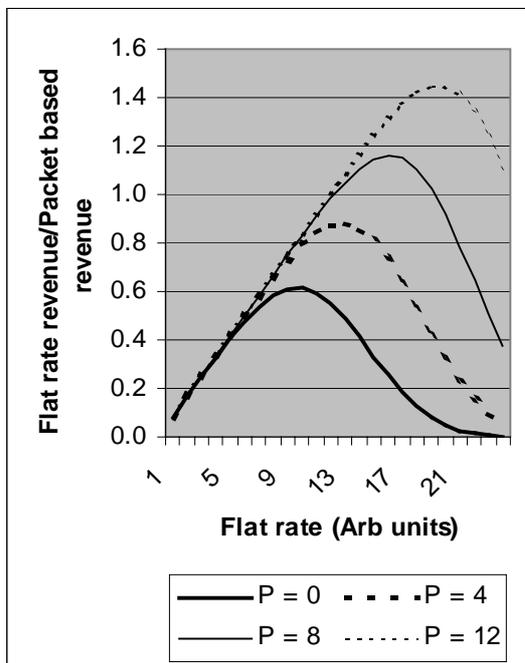


Figure 3 Flat rate/packet based revenue v flat rate for Gaussian distribution profile

For example, if a flat rate of 13 is selected then the curve for $P = 4$ shows the revenue that would be obtained if customers on the Gaussian profile in

Figure 1 with an average revenue of $(13-4) = 9$ and above are included. It can be seen that the flat rate revenue in this case is about 90% of the packet based revenue

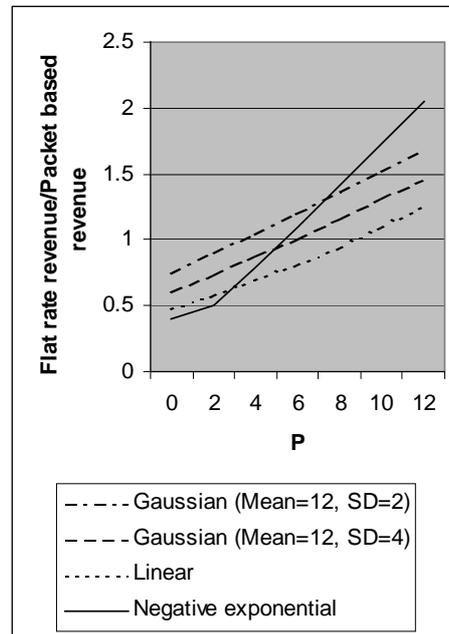


Figure 4 Flat rate/packet based revenue v P

Figure 4 shows the ratio of flat rate revenue to packet based revenue as a function of P for each of the profiles in Figure. For the flat rate revenue to exceed the packet based revenue then P must be greater than 3 for the Gaussian profile with a (SD of 2) (which corresponds to about 25% of the flat rate) and greater than 8 for the linear profile (corresponding to about 50% of the flat rate).

This analysis shows that the customer spend distribution therefore has a strong influence on the revenue that can be obtained from flat rate charging and hence determines whether packet or flat rate charging generates more revenue. Understanding the customer spend profile will help operators to identify the best charging strategy.

Packet based charging has the advantage that customers only get billed for what they have used. This encourages customers to experiment with new services because there is no minimum charge, which should encourage the take up of new services. In addition packet based charging imposes control on the use of network resources that are scarce in a mobile network. Packet based charging may therefore be the preferred solution when 3G networks are initially rolled out.

4. Hybrid charging solutions

The previous section has shown that packet based charging should encourage new customers to

experiment with new 3G services. This scheme however does not reward or encourage heavy users of the services, in particular business customers, to spend more. A flat rate could be offered to these power users but this could lead to excessive use of the network bandwidth and would not generate increased revenue from the large bandwidth users. An alternative is to offer a flat rate for a specified number of packets and then to use packet based charging for any additional use of the network. The rate for these additional packets can be reduced to optimise the package benefits for both the customer and the operator. This bundled approach is used in the i-mode FOMA (Freedom of mobile Multimedia Access) service [5]. Figure 5 shows the cost to the customer as a function of the number of packets used for the following cases:

1. Packet based charging only – 0.003/packet
2. Flat rate of 18 including use of 6000 packets plus 0.001/packet for additional packets
3. Flat rate of 36 including use of 12000 packets plus 0.0005/packet for additional packets

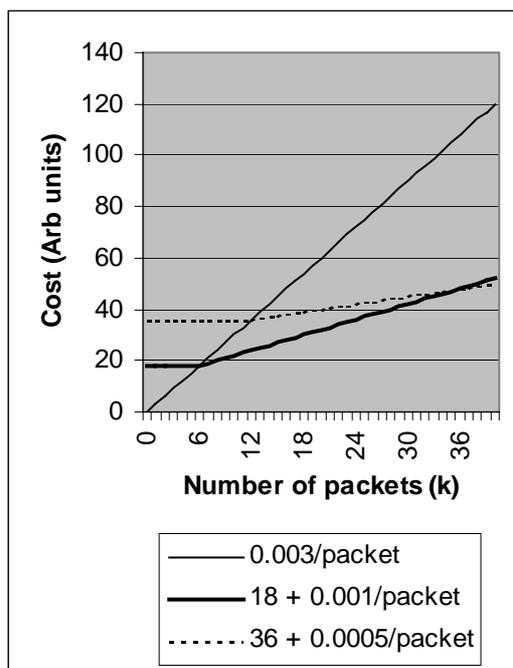


Figure 5 Flat rate and packet based charging - cost to customer v number of packets used

It can be seen that for customers using less than 6000 packets Scheme 1 is most cost effective. For customers using between 6k and 40k packets Scheme 2 is cheapest and Scheme 3 benefits customers using greater than 40k packets.

For the operator this approach guarantees a minimum revenue from the flat rate but continues to generate additional revenue when the number of packets used exceeds the allocation for the flat rate

fee. Customers that are heavy users of the network benefit by receiving a lower overall price per packet.

In order to maximise revenue it is likely that a hybrid charging system will eventually emerge that incorporates some or all of the following: flat rate and packet based charging, differential charges for QoS classes, content charges and charges for advertising.

Conclusions

Third generation mobile networks will introduce a wide range of new data services and the charging mechanism adopted will have a strong influence on the take up and revenue obtained from these services. In this paper we have reviewed the different charging mechanisms for data services and have compared the revenue generated by flat rate and packet based charging. The customer spend profile will determine which of these approaches will generate the most revenue.

Packet based charging has the advantage that it encourages customers to become familiar with the new services offered by the 3G networks and provides control of the network resources and therefore may be the preferred solution in the near term.

A hybrid system may eventually emerge that incorporates the best features from several different schemes with the aim of maximizing revenue streams for the operator and maintaining customer satisfaction.

References

- 1) C.A.Fischer, K.Han and Z.Wang, 'Accounting solutions in the UMTS core network', Bell Labs Technical Journal, Vol 6, No 2, 2002, pp 65-73
- 2) T.T.Ahonen, 'm-Profits: making money from 3G services', Wiley, 2002
- 3) Odlyzko, A, Paris metro pricing for the internet, Proceedings of the first ACM conference on Electronic commerce, pp140-147, November 1999
- 4) Henri Koskinen, 'Paris Metro Pricing' pricing' Seminar on Internet pricing, Helsinki University of Technology, 2003, www.netlab.hut.fi/opetus/s38042/k03
- 5) Sandro Grech, 'i-mode pricing' Seminar on Internet pricing, Helsinki University of Technology, 2003, www.netlab.hut.fi/opetus/s38042/k03

Blind Adaptive Multiuser Detection Based on Subspace Tracking and Kalman Filtering Estimation

H. Zhou, W.L. Woo and B. Sharif

School of Electrical, Electronic and Computer Engineering,

University of Newcastle upon Tyne,

NE1 7RU Newcastle, United Kingdom

Email: hongwei.zhou@ncl.ac.uk, w.l.woo@ncl.ac.uk and bayan.sharif@ncl.ac.uk

Abstract: *Blind multiuser detection for the CDMA communication is considered. A new blind adaptive multiuser detection scheme based on subspace tracking and Kalman filtering estimation is proposed. The new detector has an anchored structure in the signal subspace with the subspace tracked by PASTd algorithm and the coefficients estimated by Kalman filtering algorithm. In the whole detection scheme, the only a prior knowledge is the signature waveform and timing of the desired user. Numerical simulation results verify that the new algorithm has outstanding SINR performance and tracking ability in time-varying environments.*

1. Introduction

Blind adaptive multiuser detection technologies are of special interest in the mitigation of multiple-access interference (MAI) in that the only prior knowledge required is the signature code and timing of the desired user [1]. This makes them particularly attractive in downlink implementation in the code-division multiple-access (CDMA) systems because in a dynamic system, it is difficult for mobiles to obtain accurate information of other users, and a training sequence-based method requires extra channel bandwidth, moreover, reliance on training sequence can be very cumbersome because whenever there is drastic change in the system, the data transition must be temporarily and a fresh training procedure must begin. Intense researches in blind multiuser detection have been motivated since Honig *et al* [1], proposed a blind version of minimum-mean-square-error (MMSE) detector which did not require training sequence or signature/timing information about the interfering users. Motivated by this work and by exploiting the property that the optimum mean-output-energy (MOE) detector exists in signal subspace, recently Roy [2] proposed a subspace-based MOE detector which demonstrated better convergence performance than [1].

Kalman filter is well known in statistical estimation and optimal control theory [3]. Mathematically, it is an optimum estimator for solving linear quadratic Gaussian problem, inherently owning the advantage of extremely fast convergence compared with standard least-mean-square (LMS) methods [4]. In [5] Kalman filter was used for signal detection and channel estimation in an asynchronous wireless communication system and showed performance gain over MMSE detector and recursive least square (RLS) detector. However this approach is not a blind detection strategy in the sense that it needs to know the signature waveforms and the time delays of all users. Recently, Kalman filter has been applied to obtain a blind multiuser detector with lower steady-state excess output energy [6], however, its practicality is impaired by its intensive computational complexity. In this paper, by exploiting the property of the signal subspace that it has the capacity to effectively decrease the computational complexity and improve the convergence rate, we propose a new subspace-based blind multiuser detection scheme which tracks the change of the signal subspace by using PASTd algorithm and estimate detector coefficients by using the Kalman filtering algorithm. The rest of the paper is organized as follows. We describe the signal model in section 2, and then present the proposed blind multiuser detector for time-varying system in section 3. In section 4, computer simulation results are shown, and the conclusion is given in section 5.

2. Signal Model

Consider an antipodally-modulated synchronous DS-SS-CDMA system with K users. After passing through the chip-rate correlator and the chip-rate sampler, the received signal within a symbol interval T can be expressed in discrete-time vector format as:

$$\mathbf{r} = \sum_{k=1}^K A_k b_k \mathbf{s}_k + \mathbf{on} \quad (1)$$

where A_k , $b_k \in \{+1, -1\}$, $\mathbf{s}_k = [s_k(1) \ s_k(2) \ \dots \ s_k(N)]^T$ denote the received signal amplitude, information bit,

and signature waveform of the k^{th} user, respectively. N is the length of the signature waveform (processing gain). \mathbf{n} is the additive white Gaussian noise (AWGN) vector with mean $\mathbf{0}$ and covariance matrix \mathbf{I}_N which denotes the $N \times N$ identity matrix. σ is the standard deviation of the noise. Without loss of generality, it is assumed that $\{b_k\}_{k=1}^K$ are independent equiprobable random variables, and the signature waveform vectors $\{\mathbf{s}_k\}_{k=1}^K$ of the K users are mutually linearly independent, and each has unit energy, i.e. $\|\mathbf{s}_k\|^2 = 1, k=1, 2, \dots, K$.

3. New Subspace-Based Blind Multiuser Detection Scheme

Applying an eigendecomposition to the autocorrelation matrix \mathbf{R} yields

$$\mathbf{R} = E\{\mathbf{r}\mathbf{r}^T\} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T = \mathbf{U}_s\mathbf{\Lambda}_s\mathbf{U}_s^T + \mathbf{U}_n\mathbf{\Lambda}_n\mathbf{U}_n^T \quad (2)$$

where $\mathbf{U} = [\mathbf{U}_s \ \mathbf{U}_n]$, $\mathbf{\Lambda} = \text{diag}(\mathbf{\Lambda}_s \ \mathbf{\Lambda}_n)$. $\mathbf{\Lambda}_s = \text{diag}(\lambda_1 \ \lambda_2 \ \dots \ \lambda_K)$ contains the K largest eigenvalues of \mathbf{R} in descending order, and $\mathbf{U}_s \in R^{N \times K}$ contains the corresponding orthonormal eigenvectors. $\mathbf{\Lambda}_n = \sigma^2 \mathbf{I}_{N-K}$ contains another $N-K$ eigenvalues of \mathbf{R} and $\mathbf{U}_n \in R^{N \times (N-K)}$ contains the corresponding $N-K$ orthonormal eigenvectors. The column vectors of \mathbf{U}_s and \mathbf{U}_n span two orthogonal subspaces, namely, signal subspace and noise subspace, respectively, and $\mathbf{U}_s^T \mathbf{U}_n = \mathbf{0}$. The range spaces spanned by $\mathbf{S} = [\mathbf{s}_1 \ \mathbf{s}_2 \ \dots \ \mathbf{s}_K]$ and \mathbf{U}_s are identical, i.e. $\text{rang}(\mathbf{S}) = \text{rang}(\mathbf{U}_s)$.

Without loss of generality, assume that user 1 is the desired user. By constraining the detector to be in the signal subspace and using an anchored structure, the proposed detector \mathbf{c}_1 is given by

$$\mathbf{c}_1 = \mathbf{s}_1 + \mathbf{s}_{1null} \mathbf{w}_1 \quad (3)$$

where \mathbf{s}_1 is the signature vector of user 1 and is assumed to be known. All the column vectors of matrix $\mathbf{s}_{1null} \in R^{N \times (K-1)}$ are in the signal subspace and span the null space of \mathbf{s}_1 , i.e. $\mathbf{s}_1^T \mathbf{s}_{1null} = \mathbf{0}$. $\mathbf{w}_1 \in R^K$ is a weight vector. The decision of a signal bit is given by

$$\hat{b}_1 = \text{sgn}(\mathbf{c}_1^T \mathbf{r}) \quad (4)$$

In a dynamic system such as the mobile networks, users can enter and exit at any time, thus the signal

subspace is time-varying. Then one problem in this detection scheme is to track the signal subspace \mathbf{S}_{null} in time. Furthermore, another problem is to adaptively determine the coefficient vector \mathbf{w}_1 in a sample-by-sample manner. We consider tracking \mathbf{S}_{1null} first, and then treat adaptive estimation of \mathbf{w}_1 .

In this paper, we adopt the extended projection approximation subspace tracking with deflation (PASTd) algorithm [7, 8] for tracking both rank K and subspace basis \mathbf{S}_{null} . The advantages of this algorithm include almost sure global convergence to the signal eigenvectors, low computational complexity ($O(NK)$) and rank tracking capability. It tracks the signal subspace based on deflation technique as follows. Before we start, it is assumed that there are \hat{K} users in the system, and the value of \hat{K} will be updated in the adaptation. During the operation, the eigenvalues are to be nonincreasingly ordered. At the beginning, the first eigencomponent is dominant. It is updated first and then the projection of the current data vector $\mathbf{r}(n)$ onto the first eigenvector is removed from $\mathbf{r}(n)$. Now the second signal eigencomponent becomes the dominant one in the updated data vector and can be extracted in the same way. This procedure is applied repeatedly until all the \hat{K} signal eigencomponents are estimated given by $\{\hat{\mathbf{u}}_i, \hat{\lambda}_i\}, i=1, \dots, \hat{K}$. Furthermore, assuming the noise to be white, all the eigenvalues in the noise subspace can be estimated and are identical, which are denoted as $\sigma^2(n)$. Based on all the eigenvalue estimates (both in the signal subspace and the noise subspace) given by

$$\hat{\lambda}_1(n), \dots, \hat{\lambda}_K(n), \hat{\lambda}_{K+1}(n) = \hat{\sigma}^2(n), \dots, \hat{\lambda}_N(n) = \hat{\sigma}^2(n) \quad (5)$$

and using information theoretic criterion such as Akaike information criterion (AIC), the rank of the signal subspace (\hat{K}) can be estimated adaptively. AIC is defined by

$$AIC(k) \triangleq (N-k)L \ln(\alpha(k)) + k(2N-k) \quad (6)$$

where L equals the effective length of the exponential window $L = 1/(1-\beta)$, where β is the forgetting factor. $\alpha(k)$ is defined by

$$\alpha(k) \triangleq \frac{\left(\sum_{i=k+1}^N \hat{\lambda}_i \right) / (N-k)}{\left(\prod_{i=k+1}^N \hat{\lambda}_i \right)^{1/(N-k)}} \quad (7)$$

The estimate of the rank \hat{K} is updated by the value of k for which the AIC criterion is minimized, i.e.

$$\hat{K}(n) = \arg \min_k (AIC(k)) \quad (8)$$

\mathbf{S}_{null} can now be obtained as follows:

Form matrix

$$\mathbf{Z} = [\mathbf{s}_1 \quad \mathbf{u}_1 \quad \dots \quad \mathbf{u}_{\hat{K}-1}] \quad (9)$$

Apply Gram-Schmidt method to \mathbf{Z} to obtain an orthonormal matrix

$$\mathbf{Y} = [\mathbf{y}_1 \quad \mathbf{y}_1 \quad \dots \quad \mathbf{y}_{\hat{K}-1}] \quad (10)$$

then \mathbf{s}_{null} is given by

$$\mathbf{s}_{null} = [\mathbf{y}_1 \quad \dots \quad \mathbf{y}_{\hat{K}-1}] \quad (11)$$

Now to determine \mathbf{c}_1 , the only problem left is how to find the optimal weight vector \mathbf{w}_1 . It can be estimated by Kalman filtering algorithm. In a slowly time-varying scenario, the state-space model required for Kalman filtering method can be written as

$$\mathbf{w}_{1opt}(n) = \mathbf{w}_{1opt}(n-1) \quad (12)$$

$$y(n) = \mathbf{h}^T(n) \mathbf{w}_{1opt}(n) + e_{1opt}(n) \quad (13)$$

where $\mathbf{w}_{1opt}(n)$ is the optimum of \mathbf{w}_1 when n is sufficiently large, $y(n) \triangleq \mathbf{s}_1^T \mathbf{r}(n)$ and $\mathbf{h}^T(n) \triangleq -\mathbf{r}^T(n) \mathbf{s}_{null} \cdot e_{1opt}(n) \triangleq \mathbf{c}_{1opt}^T(n) \mathbf{r}(n)$ is the measurement noise which has zero mean and covariance $\varphi_{min} = A_1^2 + \mathcal{E}_{min}$, where \mathcal{E}_{min} is the minimum mean-square-error. Finally the detection scheme is illustrated in figure 1 and the algorithm is summarized in Table 1.

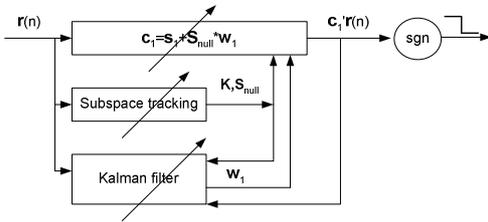


Fig. 1: Diagram of the proposed blind multiuser detection algorithm based on subspace tracking and Kalman filtering estimation.

TABLE 1: Subspace blind adaptive multiuser detection based on PASTd subspace tracking and Kalman Filter estimation

FOR	$t=1,2,\dots$	DO
	Updating eigenvalues and eigenvectors	
	$\{\lambda_k, \mathbf{u}_k\}_{k=1}^{\hat{K}(t-1)}$	
	Updating the rank of signal subspace using AIC criteria:	
	$\hat{K}(t) = \arg \min_{0 \leq k \leq \hat{K}(t-1)} AIC(k) + 1$	

```

IF  $\hat{K}(t) < \hat{K}(t-1)$  THEN
    remove  $\{\lambda_k(t), \mathbf{u}_k(t)\}_{k=\hat{K}(t)+1}^{\hat{K}(t-1)}$ 
ELSE IF  $\hat{K}(t) > \hat{K}(t-1)$  THEN
    Add  $\{\lambda_{\hat{K}(t)}, \mathbf{u}_{\hat{K}(t)}\}$ 
ELSE
     $\mathbf{Z} = [\mathbf{s}_1 \quad \mathbf{u}_1 \quad \dots \quad \mathbf{u}_{\hat{K}-1}]$ 
    Applying Gram-Schmidt to  $\mathbf{Z}$ 
    yields:
     $\mathbf{Y} = [\mathbf{y}_1 \quad \mathbf{y}_1 \quad \dots \quad \mathbf{y}_{\hat{K}-1}]$ 
     $\mathbf{s}_{null} = [\mathbf{y}_1 \quad \dots \quad \mathbf{y}_{\hat{K}-1}]$ 
    Updating  $\mathbf{w}_1(t)$  using Kalman
    filter
END
END

```

4. Simulation Results

In this section, we provide two simulation examples to study the convergence performance and tracking ability of the proposed blind multiuser detection algorithm, respectively. It assumes a synchronous DS-CDMA system with Gold codes as the signature codes (processing gain $N=31$). In implementing the PASTd algorithm, the forgetting factor is $\lambda = 0.999$, the eigenvector matrix are initialized to be identity matrix and the eigenvalues are all 1. In each example, initially there are 8 simultaneous users ($K=8$) of which the user 1 is the desired user. All signal energies are measured in decibels relative to the ambient noise variance σ^2 . Taking $\sigma^2=0.01$, $A_1^2=1$ and $A_2^2=\dots=A_8^2=100$, user 1 has $SNR=10\log(A_1^2/\sigma^2)=20dB$, user2-user 8 have $SNR=10\log(A_k^2/\sigma^2)=40dB$, $k=2,\dots,8$. In measuring the MAI level in the soft output of the detector, the following time-averaged $SINR$ is applied

$$SINR(n) = 10 \log \frac{\sum_{l=1}^M (\mathbf{c}_{1l}^T \mathbf{s}_1)^2}{\sum_{l=1}^M [\mathbf{c}_{1l}^T (\mathbf{r}_l(n) - b_{1l}(n) \mathbf{s}_1)]^2} \quad (14)$$

where M is the number of independent runs. In each run, the random signals b and \mathbf{r} , and the detector \mathbf{c}_1 are independent from those in other runs. The subscript l indicates that the associated variable depends on the particular run. In this paper, M is set to be 600.

Example 1: In this example, the convergence of the new proposed subspace-based Kalman filtering

detector is studied with the assumption that the system is stationary. The time-averaged SINR performance versus iteration number is plotted in Fig.2. Significant performance gain can be observed by comparing the desired user's signal at the entrance of the receiver, where its power is 20 dB lower than the interfering signal, and the one in the output of the receiver, where it is 15 dB higher than the combined power of the interfering signal and the background noise. This verifies the efficiency of the proposed algorithm in suppressing MAI. The difference between the SINR value at convergence and the theoretic value comes from the fact that the eigenvectors obtained by the PASTd algorithm are not identical to the real ones. Applying eigenvalue decomposition (EVD) to the autocorrelation matrix of the received signals can produce accurate estimation of the eigenvectors and thus lead to the almost ideal converged SINR, however, its computational complexity is fairly costly.

Example 2: This example studies the tracking ability of the proposed detector in a time-varying environment. The varying rate should allow the detector to converge to a reasonably stable level. At time $n=600$, one interfering user enters the system with the interfering power 38 dB. At time $n=1200$, one 40-dB user exits the system. Every time the change of the number of users is detected, the eigenvectors and weight vector are reinitialized. The time-averaged SINR performance versus iteration number is plotted in Fig.3. It illustrates that detector can readjust its structure when the number of users changes, so that the converged SINR value almost maintains the same. The adaptation speed is decided by the subspace adaptation process and detector coefficient vector adaptation process.

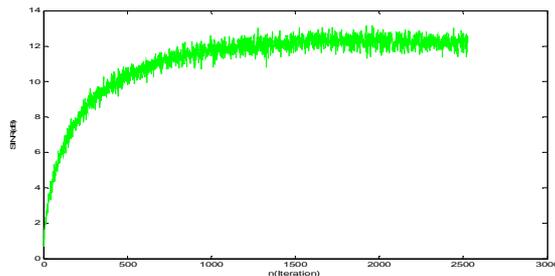


Fig.2. Time-averaged SINR versus iteration in a stationary environment.

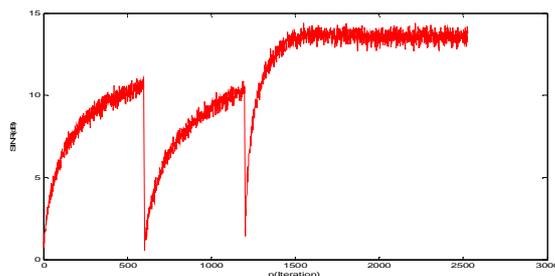


Fig.3. Time-averaged SINR versus iteration in time-varying environment where users enter and exit.

5. Conclusion

In this paper we have developed a new subspace blind multiuser detection scheme that combines the subspace tracking technology and Kalman filter estimation algorithm to solve the adaptive detection problem in time-varying application scenarios. Basically it is a two-stage process, first detecting the rank and the basis of the subspace and then estimating the coefficient vector of the detector. Numerical simulation results show that the new detector has outstanding convergence performance and tracking ability in time-varying environments. Because Kalman filter is a fast-converging algorithm, the tracking speed is mainly determined by the subspace tracking algorithm. Furthermore, the PASTd algorithm has limitation that it can detect the change of only one user in one iteration. More powerful subspace tracking algorithm can improve this in the future. Mathematical analysis of the performance is also contained in the future work.

6. REFERENCES

- [1] M. L. Honig, U. Madhow, S. Verdu, *Blind adaptive multiuser detection*, *IEEE Trans. Information Theory*, vol. 41, no. 4, pp. 944-960, July 1995.
- [2] S. Roy, "Subspace blind adaptive detection for multiuser CDMA," *IEEE Transactions on Communications*, Vol.48, No.1, pp.169-175, Jan, 2000.
- [3] J. B. Burl, *Linear optimal control : H2 and H[infinity] methods*, Addison Wesley Longman, Inc., 1999.
- [4] S. Haykin, *Adaptive filter theory*, 4th ed., NJ: Prentice Hall, 2002.
- [5] T. J. Lim, L. K. Rasmussen, and H. Sugimoto, "An asynchronous multiuser CDMA detector based on the Kalman filter," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1711--1722, Dec. 1998.
- [6] X. Zhang and W. Wei, "Blind adaptive multiuser detection based on Kalman filtering," *IEEE Transactions on Signal Processing*, Vol.50, No.1, pp. 87-94, Jan., 2002.
- [7] B. Yang, "An extension of the PASTd algorithm to both rank and subspace tracking," *IEEE Signal Processing Letters*, Vol.2, No.9, pp.179-182, Sept., 1995.
- [8] X. WANG and H.V. POOR: "Blind Multiuser Detection: A Subspace Approach," *IEEE Trans. on Information Theory*, 44(2), pp.677-690, March 1998.

Channel SNR Estimate for Turbo Decoding

Janak Sodha and Charles Lee*

Department of Computer Science, Mathematics, and Physics
University of the West Indies
Cave Hill, Barbados
e-mail: jsodha@ieee.org

*Department of Electronics
Macquarie University
Sydney NSW2109
Australia
e-mail: cl@ics.mq.edu.au

Abstract— The termination probability of a modified Maximum A Posteriori (MAP) decoder within the overall Turbo decoder is used to estimate the channel SNR to within 1 dB. Frame synchronization [1] and the channel SNR can jointly be established as part of any preamble communication. Simulation results are presented for the standard rate 1/2 (7, 5, N) Turbo code.

Keywords— Channel SNR, Turbo Codes, Error Control Coding

I. INTRODUCTION

To maximize the error-control capability of a turbo decoder, it is essential that the channel signal-to-noise ratio (SNR) $\frac{E_b}{N_0}$ is estimated as accurately as possible. A method to estimate the channel SNR by utilizing the *termination probability* [1] of the turbo decoder operating under the incorrect synchronization state is the subject of this paper.

We begin in section II with a brief review of a rate 1/2 turbo decoder and its termination probability. The termination probability is utilized in section III to develop a method to estimate the channel SNR. The corresponding simulation results are presented in section IV. Finally, we end the paper with conclusions in section V.

II. TURBO DECODER TERMINATION PROBABILITY

Consider a turbo code encoder which consists of the parallel concatenation of two rate 1/2 recursive systematic convolutional (RSC) code encoders (E1 and E2), and an interleaver of size N [2]. Let D1 and D2 represent the two modified MAP decoders within the corresponding turbo decoder. For

a given RSC encoder (E1 or E2) at time k , let $x_k = (x_k^{(s)}, x_k^{(p)})$ represent the transmitted codeword, where $x_k^{(s)}$ is the systematic bit and $x_k^{(p)}$ is the parity bit. Furthermore, let $y_k = (y_k^{(s)}, y_k^{(p)})$ represent the corresponding noisy received codeword. Let $x_1^N = \{x_1, x_2, \dots, x_N\}$ represent the codeword sequence generated by a RSC encoder from time $k = 1$ to N , and let $y_1^N = \{y_1, y_2, \dots, y_N\}$ represent the corresponding received sequence. If $S_1^N = \{s_1, s_2, \dots, s_k, s_{k+1}, \dots, s_N\}$ represents the sequence of encoder states from time $k = 1$ to N , where s_k is the state s of a RSC code encoder at time k , then as in [3], we define the modified probabilities $\tilde{\alpha}_k(s) = \frac{P(s_k=s, y_1^k)}{P(y_1^k)}$ and $\tilde{\beta}_k(s) = \frac{P(y_{k+1}^N | s_k=s)}{P(y_{k+1}^N | y_1^k)}$.

The termination probability P_T defined by $P_T = \sum_s \tilde{\alpha}_k(s) \tilde{\beta}_k(s)$ can be shown [1] to be the probability that D1 of Turbo decoder will terminate in the zero state. A key feature of P_T is that it is independent of k and a good indicator of the relative number of errors in the turbo decoded frame. Since a turbo decoder is able to dramatically reduce the probability of an error in a binary digit beyond 3 dB, P_T would be essentially 1 at a such a relatively high SNR. Thus P_T is a useful indicator at low SNRs only. Refer to [1] for further information on the properties of P_T .

III. CHANNEL SNR ESTIMATE

The channel SNR can be determined by a slight modification of the frame synchronization algorithm in [4] for a turbo decoder. A brief review of this algorithm is as follows. A counter increments by c_a steps if $P_T < 0.5$ and decrements by c_b steps if

$P_T \geq 0.5$. Let x represent the counter value. To ensure the counter only accumulates evidence which rejects the existing synchronization state, a reflecting barrier [5] is placed at $-c_b$ and an upper threshold T . If $x \leq -c_b$, then x is reset to zero. If $x \geq T$, then sufficient evidence has been accumulated to reject the existing synchronization state. At this point, the synchronization state of the decoder is changed and x is reset to zero.

Since P_T for a given frame does not depend on the previous value, the operation of the counter is random walk in which the counter increments with a probability p and decrements with a probability $q = 1 - p$. If the counter is forced to begin from the origin for which $x = 0$, it can be shown [6] that the average number of steps, or equivalently, the *average run-length* (ARL) to reach the threshold is given by $ARL = \frac{1}{(qc_b - pc_a)} \left\{ \frac{c_b(\lambda^T - 1)}{(1 - \lambda^{c_b})} - T \right\}$, where λ satisfies the characteristic equation $1 = p\lambda^{c_a} + \frac{q}{\lambda^{c_b}}$.

The performance of this synchronization algorithm is characterized by the short-ARL and the long-ARL [6], which are the average number of counter steps to reach the threshold T if the decoder is in the off-sync and in-sync state, respectively, with the counter value initially set to zero. Each counter step corresponds to a single decoded frame of size N information bits. Thus, for a given ARL, the average number of information bits processed is equal to $(N * ARL)$. As in [4], we shall restrict ourselves to the simple rate 1/2 (7, 5, 16) turbo code to illustrate the channel SNR estimate technique.

In [4], 10000 frames were decoded to estimate the probability p that $P_T < 0.5$ over a range of SNRs under the in-sync and off-sync synchronization states for the rate 1/2 (7, 5, 16) turbo code. In each case, the turbo decoder was iterated six times before checking whether $P_T < 0.5$. These values for p under the off-sync state are listed in Table I. Using a counter with $c_a = c_b = 1$, Fig. 1 shows the excellent correspondence between theory and simulation for the two SNRs of 0 dB and 4 dB.

To determine the channel SNR, consider the following simulation. The turbo decoder operating in the off-sync mode is used to drive the counter, which starting from zero is allowed to operate until it crosses at least once, the thresholds $T_{\min}, (T_{\min} + T_d), (T_{\min} + 2T_d), \dots, T_{\max}$. As the counter crosses a given threshold, the corresponding

SNR (dB)	p
0	0.865
1	0.817
2	0.767
3	0.754
4	0.727

TABLE I
PROBABILITY OF COUNTER INCREMENTING

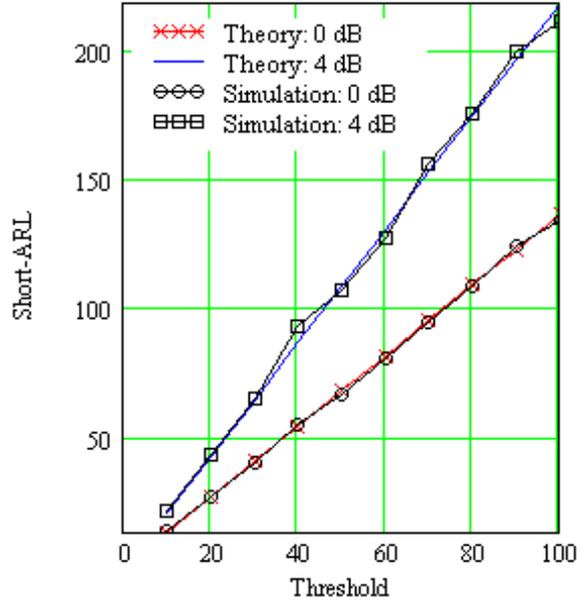


Fig. 1. Short-ARL performance.

run-length is recorded. After crossing all the thresholds in the range T_{\min} to T_{\max} , the counter is reset to zero and the process repeated a further N_{est} times. Thus, for a given threshold, we have N_{est} run-length estimates. Averaging over these estimates, a plot of the ARL versus the threshold is shown in Fig. 2 for $T_{\min} = 50, T_d = 5$ and $T_{\max} = 100$. In this specific case, the channel SNR was set at 1 dB with $N_{est} = 100$. The theoretical short-ARL performance for a channel SNR of 0 dB to 4 dB are also presented in Fig. 2. If we assume the channel SNR can only be either 0, 1, 2, 3 or 4dB, then visually from Fig. 2, given that the points lie along the 1 dB line, we obviously estimate the channel SNR to be 1 dB. More precisely, by calculating the root mean square (rms) error over the threshold range T_{\min} to T_{\max} between the simulation points and each theoretical short-ARL line in turn, the channel SNR estimate is taken to correspond to the minimum rms error.

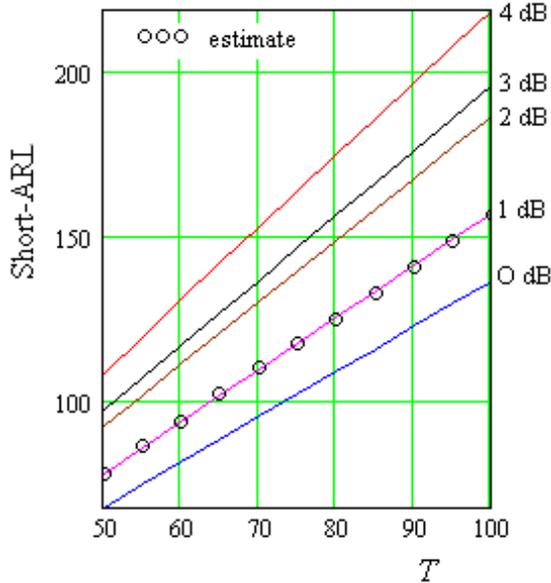


Fig. 2. Theoretical short-ARL performance for 0 to 4 dB overlaid with the channel SNR estimate average run-length.

IV. PERFORMANCE

Simulation results to determine the probability P_{snr} of incorrectly estimating the channel SNR are shown in Fig. 3 with $N_{est} = 20$. For a given channel SNR, the number of channel estimate errors were insured to be at least 10.

Referring to the relative angle separation of the

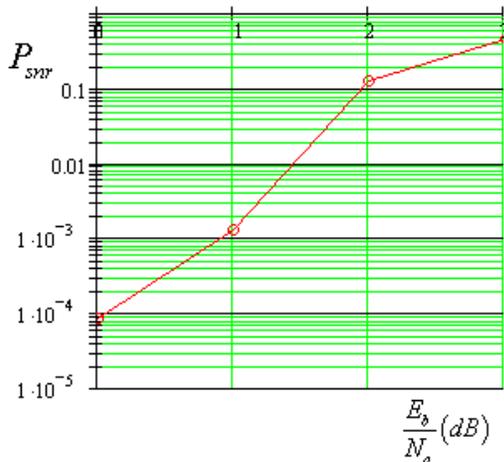


Fig. 3. Probability of channel SNR estimate error for the rate 1/2 (7, 5, 16) code.

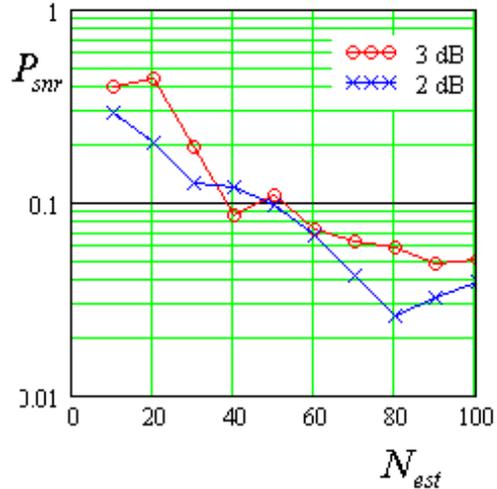


Fig. 4. Reduction in P_{snr} by increasing N_{est} .

theoretical short-ARL lines in Fig. 2, it is not surprising that P_{snr} is worse for a channel SNR of 2 dB and greater. Also, it would be difficult to estimate the channel SNR to within 0.5 dB. Given that the channel SNR is assumed to be either 0, 1, 2, 3 or 4 dB, P_{snr} for 0 dB is the smallest because the SNR choice is only between 0 and 1 dB. Conversely for a channel SNR of 3 dB, P_{snr} becomes worse because the choice is now between 2, 3 and 4 dB with only a small angle of separation between the lines corresponding to 2 and 3 dB. The probability P_{snr} can be reduced by increasing N_{est} as evident from Fig. 4. Another option is to increase the threshold range T_{min} to T_{max} .

V. CONCLUSIONS

For a given threshold T , the average number of information bits processed is given by ($N * \text{short-ARL}$). Plotting the ratio of the short-ARL to its corresponding threshold, as shown in Fig. 5, we note that the short-ARL is less than $2T$ for a channel SNR in the range 0 to 3 dB. Thus, the number of information bits processed to estimate the channel SNR is approximately given by $N(2T_{max})N_{est}$. For example, for $N = 16$, $T_{max} = 100$, and $N_{est} = 100$, the number of bits processed to estimate a low channel SNR is approximately 40 k bytes. In comparison, the frame synchronization algorithm in [4] would typically establish the correct synchronization state in under $N(2T_{max})$ bits. Since both the frame synchronization and channel SNR estimate method utilize the probability P_T to drive a counter, the frame

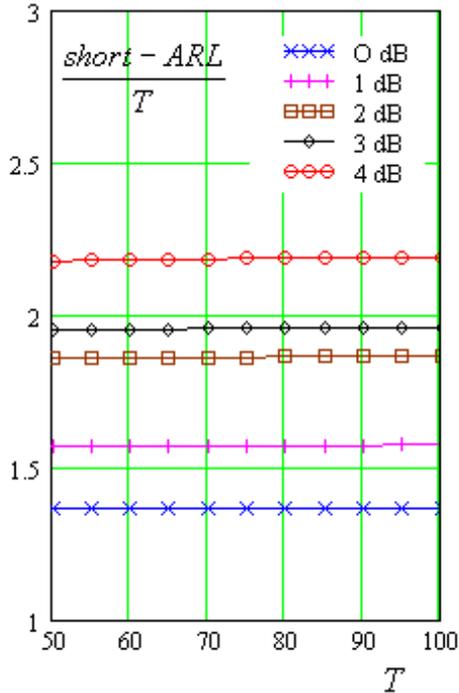


Fig. 5. Short-ARL relative to the set threshold T .

synchronization algorithm can be used initially to establish the synchronization state, before resetting the counter to process the received noisy codeword sequence under the off-sync state to determine the channel SNR.

Ongoing research is to consider the effect of increasing the frame size N , and the influence of a channel SNR estimate error on the performance of the rate $1/2$ $(7, 5, N)$ turbo decoder.

REFERENCES

- [1] E. Hueffmeier, J. Sodha and S. Wicker, "Termination Bits in the BCJR Algorithm", Third International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), Staffordshire University, UK, 15-17 July 2002, pp. 101-104.
- [2] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes", *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261-1271, Oct. 1996.
- [3] W. E. Ryan, "A turbo code tutorial", New Mexico State University, Box 30001 Dept. 3-O, Las Cruces, MN 88003, <http://www.ece.arizona.edu/~ryan/>.
- [4] J. Sodha, "Turbo code frame synchronization", *Signal Processing Journal*, Elsevier, Jan. 2002, Vol. 82, pp. 803-809.
- [5] W. Feller, "An introduction to probability theory and its applications", vols. I and II, New York, Wiley, 1968, 1971.
- [6] G. Lorden, R. J. McEliece and L. Swanson, "Node synchronization for the viterbi decoder", *IEEE Trans. Commun.*, vol. 32, no. 5, pp. 524-531, May. 1984.

Constellation Shaping for Turbo-Coded Modulation

Stéphane Y. Le Goff¹, Bayan S. Sharif², and Shihab A. Jimaa³

¹ Department of Physics and Electronic Engineering, University of Waikato, Hamilton, New Zealand. Email: s_legoff@hotmail.com

² School of Electrical, Electronic and Computer Engineering, University of Newcastle upon Tyne, Newcastle upon Tyne, United Kingdom.

³ Department of Communication Engineering, Etisalat University, Sharjah, United Arab Emirates.

Abstract – We present a new, simple method for combining constellation shaping and turbo-coded modulation over Gaussian channels. By considering the example of a 3-bit/dim scheme using 16-PAM constellation, it is shown that this technique can provide shaping gains of 0.77 dB, and error performance within 1.58 dB of the continuous-input channel capacity limit is achieved.

Index Terms – Bit-interleaved coded modulation, constellation shaping, shaping code, turbo code.

I. INTRODUCTION

For bandwidth-limited applications over Gaussian channels, it has recently been proposed to combine constellation shaping and turbo-coded modulation by employing either a multi-level coding approach [1], [2] or a bit-interleaved coded modulation (BICM) approach [3]. In this paper, we present a novel, simple method for combining shaping and turbo-coded modulation.

The proposed system is designed using the BICM approach in conjunction with a shaping technique introduced by Calderbank and Ozarow in [4]. Throughout this paper, this scheme will be termed ‘bit-interleaved turbo-coded modulation with shaping’ (BITCM-S). The shaping method consists of partitioning the basic constellation into several equal-sized sub-constellations of increasing average energy. A shaping code is then used to specify the sequence of sub-constellations so that low-energy signals are transmitted more frequently than high-energy signals.

The partitioning method preserves the Gray mapping, and is therefore suitable for BICM design, provided that only one level of partition is considered, i.e. the basic constellation is divided into two sub-constellations. This compatibility between shaping and Gray mapping constitutes a crucial point since it is well known that BICM schemes perform well only if Gray mapping is used to label signal points of the basic constellation [5], [6]. Throughout this work, we only consider 2^m -ary one-dimensional (1-D) constellations, hereafter referred to as 2^m -PAM constellations.

II. BITCM-S TRANSMITTER

The generic structure of a BITCM-S transmitter is shown in Fig. 1. The sequence of information bits is encoded by a rate- R_c

binary turbo encoder (turbo-ENC). The corresponding sequence of coded bits is, after interleaving (π), broken into blocks of N bits that are further divided into m binary vectors by a serial-to-parallel (S/P) converter. The first vector M_1 , composed of k bits, is fed into a shaping encoder (S-ENC) which generates a corresponding codeword C_1 of n bits, with $n > k$. The rate of this encoder is thus $R_s = k/n$. The other $(m-1)$ vectors $C_j, j \in \{2, \dots, m\}$, present at the S/P converter output, are composed of n bits each. Finally, a vector $(c_{i,1}, c_{i,2}, \dots, c_{i,m})$, where $c_{i,j}$ denotes the i th bit of vector C_j , is mapped onto a signal point of a 2^m -PAM constellation, called S, according to Gray labeling.

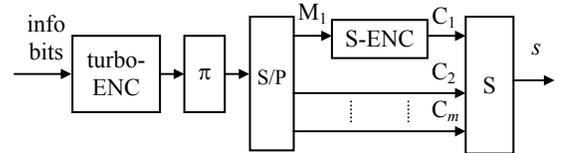


Fig. 1 - Structure of a generic BITCM-S transmitter.

Therefore, the transmission of N encoded bits is performed by emitting n successive 1-D signal points $s_i, i \in \{1, \dots, n\}$, i.e. an n -D signal point denoted s . We can show that the data rate R obtained with such system is

$$R = R_c(R_s + m - 1) \text{ bits/dim}, \quad (1)$$

which is less than the rate $R' = mR_c$ bits/dim obtained with an equivalent BICM scheme without shaping code. Hence, the use of a shaping encoder results in a loss in data rate that can be compensated for by, for example, increasing the turbo code rate R_c .

The constellation S is partitioned into two sub-constellations S_0 and S_1 so that S_0 contains the 2^{m-1} signal points of lowest energies, whereas S_1 is composed of the 2^{m-1} signal points of highest energies [4]. The Gray mapping is performed in such a way that bits $c_{i,1}$, which are the bits generated by the shaping encoder, are used to select one of these sub-constellations.

Assuming that $c_{i,1} = 0$ leads to the selection of S_0 , the shaping encoder is designed so that $\Pr\{c_{i,1} = 0\} > \Pr\{c_{i,1} = 1\}$, i.e. S_0 is emitted more frequently than S_1 . Since the average energy of signal points in S_0 is lower than that in S_1 , this results in an

energy saving at the transmitter output with respect to the case where sub-constellations S_0 and S_1 are used equiprobably (no shaping). Note that the actual energy saving, called shaping gain, is obtained after the loss in data rate has been compensated for.

III. BITCM-S RECEIVER

The block diagram of a generic BITCM-S receiver is depicted in Fig. 2. The received n -D signal r is a vector of n channel samples $r_i, i \in \{1, \dots, n\}$, expressed as $r_i = s_i + v_i$, where v_i is a Gaussian noise sample with zero mean and variance σ^2 .

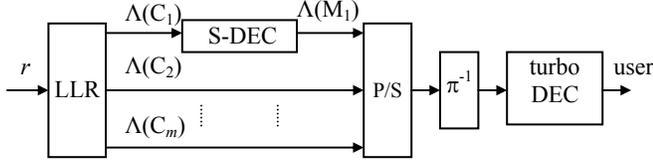


Fig. 2 - Structure of a generic BITCM-S receiver.

From sample r_i , the logarithm of likelihood ratio (LLR) $\Lambda(c_{i,j})$ associated with each bit $c_{i,j}, j \in \{1, \dots, m\}$, is computed using standard, well-known expressions, such as the simplified equations proposed for 2^m -PAM constellations in [7].

Note that $\Pr\{c_{i,1} = 0\}$ is not equal to $\Pr\{c_{i,1} = 1\}$. In the corresponding LLR computations, we take this into account by adding, to the standard expressions obtained for $\Pr\{c_{i,1} = 0\} = \Pr\{c_{i,1} = 1\}$, a term in the form $\log(\Pr\{c_{i,1} = 1\}/\Pr\{c_{i,1} = 0\})$. For example, if we choose to use the simplified equations derived in [7], the LLRs $\Lambda(c_{i,1}), i \in \{1, \dots, n\}$, are simply given by

$$\Lambda(c_{i,1}) = |r_i| - 2^{m-1} + K_i, \quad (2)$$

where the symbol $|\cdot|$ denotes ‘absolute value of’, and K_i is a constant computed as follows:

$$K_i = \frac{\sigma^2}{2} \ln \left[\frac{\Pr\{c_{i,1} = 1\}}{\Pr\{c_{i,1} = 0\}} \right]. \quad (3)$$

For each received n -D signal r , the LLR computation block produces m vectors $\Lambda(C_j) = (\Lambda(c_{1,j}), \Lambda(c_{2,j}), \dots, \Lambda(c_{n,j})), j \in \{1, \dots, m\}$. The soft-input, soft-output (SISO) shaping decoder (S-DEC) uses the MAP algorithm to decode vector $\Lambda(C_1)$ and generate an estimate $\Lambda(M_1) = (\Lambda(m_{1,1}), \Lambda(m_{2,1}), \dots, \Lambda(m_{k,1}))$ of the vector $M_1 = (m_{1,1}, m_{2,1}, \dots, m_{k,1})$ encoded at the transmitter side. The MAP decoding algorithm consists of evaluating, for $q \in \{1, \dots, k\}$, the expression

$$\Lambda(m_{q,1}) = \log \left[\frac{\Pr\{\Lambda(C_1), m_{q,1} = 1\}}{\Pr\{\Lambda(C_1), m_{q,1} = 0\}} \right], \quad (4)$$

where $\Pr\{\Lambda(C_1), m_{q,1} = 1\}$ designates the probability to obtain an LLR value equal to $\Lambda(C_1)$ and have, at the same time, $m_{q,1} = t, t \in \{0, 1\}$. We notice that

$$\Pr\{\Lambda(C_1), m_{q,1} = t\} = \sum_{C_1 \in \Omega_{q,t}} \Pr\{\Lambda(C_1), C_1\}, \quad t \in \{0, 1\}, \quad (5)$$

where $\Omega_{q,t}$ denotes the set of all codewords C_1 obtained after encoding of the vectors M_1 for which $m_{q,1} = t$ ($|\Omega_{q,1}| = |\Omega_{q,0}| = 2^{k-1}$). Applying Bayes’ rule and combining (4) and (5), we obtain

$$\Lambda(m_{q,1}) = \log \left[\frac{\sum_{C_1 \in \Omega_{q,1}} \Pr\{\Lambda(C_1), C_1\}}{\sum_{C_1 \in \Omega_{q,0}} \Pr\{\Lambda(C_1), C_1\}} \right], \quad q \in \{1, \dots, k\}, \quad (6)$$

where $\Pr\{\Lambda(C_1), C_1\}$ is the probability to obtain an LLR value equal to $\Lambda(C_1)$, given the transmission of a particular codeword C_1 . Since codewords C_1 are composed of n bits $c_{i,1}, i \in \{1, \dots, n\}$, we can write

$$\sum_{C_1 \in \Omega_{q,t}} \Pr\{\Lambda(C_1), C_1\} = \sum_{C_1 \in \Omega_{q,t}} \prod_{i=1}^n \Pr\{\Lambda(c_{i,1}), c_{i,1}\} \quad t \in \{0, 1\}. \quad (7)$$

At this stage, we need to find an expression for the term $\Pr\{\Lambda(c_{i,1}), c_{i,1}\}$ defined as the probability to obtain an LLR value equal to $\Lambda(c_{i,1})$, given the transmission of bit $c_{i,1}$. This expression depends on the algorithm used to compute $\Lambda(c_{i,1})$. Assuming that LLRs $\Lambda(c_{i,1})$ are obtained by applying equation (2), one can demonstrate that the probability density function (PDF), $P(y)$, of $\Lambda(c_{i,1})$ for a given value of bit $c_{i,1} \in \{0, 1\}$ is approximately the sum of 2^{m-2} Gaussian PDFs, with variance σ^2 and means λ_x expressed as

$$\lambda_x = (2c_{i,1} - 1)(2x - 1) + K_i, \quad x \in \{1, \dots, 2^{m-2}\}. \quad (8)$$

Finally, by combining (6) and (7), we show that, when a vector $\Lambda(C_1)$ is received, the calculation of LLRs $\Lambda(m_{q,1}), q \in \{1, \dots, k\}$, can be performed using expression

$$\Lambda(m_{q,1}) = \log \left[\frac{\sum_{C_1 \in \Omega_{q,1}} \prod_{i=1}^n \sum_{x=1}^{2^{m-2}} \exp \left\{ -\frac{(\Lambda(c_{i,1}) - \lambda_x)^2}{2\sigma^2} \right\}}{\sum_{C_1 \in \Omega_{q,0}} \prod_{i=1}^n \sum_{x=1}^{2^{m-2}} \exp \left\{ -\frac{(\Lambda(c_{i,1}) - \lambda_x)^2}{2\sigma^2} \right\}} \right] \quad (9)$$

in which parameters λ_x are given by (8). This expression shows that the MAP decoding algorithm can become complex to implement for high values of k and n . In particular, the number of additions required for the evaluation of (9) increases exponentially with k since $|\Omega_{q,1}| = |\Omega_{q,0}| = 2^{k-1}$. One should bear this in mind when selecting appropriate shaping codes for practical applications.

However, for small values of parameters k and n , the implementation of (9) remains of reasonable complexity. This is a crucial point since it is, fortunately, possible to obtain

significant shaping gains when using codes with small values of k and n . To illustrate this statement, we have plotted in Fig. 3 the variation of the asymptotic energy saving (not the actual shaping gain) as a function of the dimension n , when a rate-1/2 shaping code is combined to 16-PAM according to the configuration considered in this paper.

The results in Fig. 3 show that the value of the energy saving “saturates” very quickly as n increases. This indicates that, in practice, there is not much to be gained by increasing the dimension n of the shaping code beyond a certain value. For instance, the asymptotic energy saving achieved by using the moderate-complexity ($R_s = 1/2$, $n = 14$, $k = 7$) code is 3.30 dB, which is only 0.18 dB less than that obtained with the much more complex ($R_s = 1/2$, $n = 32$, $k = 16$) code.

Finally, a parallel-to-serial (P/S) converter combines successive vectors $\Lambda(M_i)$ and $\Lambda(C_j)$, $j \in \{2, \dots, m\}$, to produce a sequence of LLRs which is, after de-interleaving (π^{-1}), used by the binary turbo decoder (turbo-DEC). At this stage, it is worthwhile mentioning an interesting feature of the method described in this paper: as it only consists of inserting a single shaping encoder/decoder into the classical BITCM structure, one is able to preserve most of the simplicity and flexibility that have made BITCM schemes very attractive for many practical applications.

We recall that, in the context of this paper, the term ‘flexibility’ refers to the ability of BITCM systems to accommodate various data rates while keeping unchanged both the modulation scheme and the basic turbo encoder/decoder.

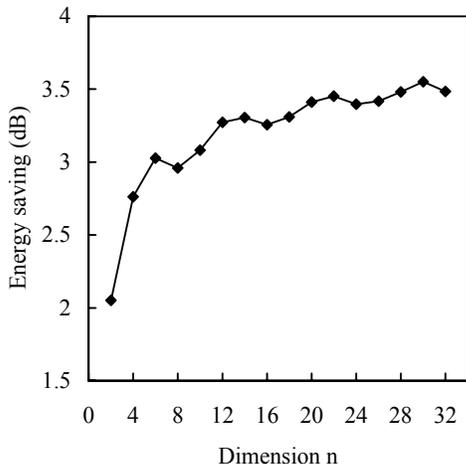


Fig. 3 – Asymptotic energy saving as a function of the dimension n of the code, for a 16-PAM constellation combined to a rate-1/2 shaping code according to the configuration considered in this paper (one level of partition).

IV. EXAMPLE AND SIMULATION RESULTS

In this Section, we consider the example of a 3-bit/dim BITCM-S scheme combining 16-PAM constellation ($m = 4$) and the ($R_s = 1/2$, $n = 14$, $k = 7$) shaping code. The Gray mapping for 16-PAM constellation is indicated in Fig. 4. The

complexity of implementation of the shaping decoder can be considered as reasonable, mainly because there are only $2^7 = 128$ possible codewords ($|\Omega_{q,1}| = |\Omega_{q,0}| = 64$ in equation (9)).

Signal	-15	-13	-11	-9	-7	-5	-3	-1	1	3	5	7	9	11	13	15
$c_{i,1}$	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1
$c_{i,2}$	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
$c_{i,3}$	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1
$c_{i,4}$	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1

Fig. 4 - Gray mapping used for 16-PAM constellation. The bit $c_{i,1}$ is used to partition the constellation into two sub-constellations S_0 and S_1 of average energy $E_0 = 21$ and $E_1 = 149$, respectively.

A computer program was written so as to generate various shaping codes with different rates R_s and dimensions n , and then determine the asymptotic energy saving that can be achieved in each case. This search was aimed at finding the codes that present the best tradeoff between asymptotic energy saving (which determines the shaping gain), rate R_s (which determines the loss in data rate due to shaping), and implementation complexity. We found that the ($R_s = 1/2$, $n = 14$, $k = 7$) shaping code considered earlier in this paper was one of the several suitable choices. The use of such code in association with 16-PAM results both in an asymptotic energy saving of 3.30 dB and a decrease in the data rate from 4 bits/dim to 3.5 bits/dim (in the absence of channel coding).

The turbo code is a rate-6/7 code built from two parallel-concatenated 16-state recursive and systematic convolutional (RSC) codes with polynomials (23, 31) [8]. The MAP algorithm is used for the decoding of each RSC code. In all cases, Gray mapping is such that information bits and parity bits at the turbo-encoder output are equally protected, and LLRs $\Lambda(C_j)$, $j \in \{1, \dots, m\}$, are computed using the simplified equations proposed for 2^m -PAM constellations in [7].

Fig. 5 shows graphs of BER versus E_b/N_0 obtained with such BITCM-S scheme, when the size of the pseudo-random interleaving embedded in the turbo code is equal to either 32768 bits or 2048 bits, and turbo decoding is performed in 8 iterations. For comparison sake, the BER curves obtained with an equivalent 3-bit/dim 16-PAM BITCM system without shaping are also depicted. From Fig. 5, it is seen that the use of a shaping code results in a significant error performance improvement for both interleaving sizes. At a BER of 10^{-5} , we obtain shaping gains equal to 0.52 dB and 0.77 dB, for interleaving sizes of 32768 bits and 2048 bits, respectively.

We have also depicted in Fig. 5 the BER curve obtained with the same system when the turbo code uses 32768-bit interleaving and 18 decoding iterations. It is interesting to compare the error performance of this system to that displayed by the 3-bit/dim 16-PAM scheme recently proposed by Raphaeli and Gurevitz [3]. This scheme combines turbo coding and shaping by using a mapping technique that converts equiprobable binary words generated by the turbo-encoder into non-equiprobable constellation signal points. The turbo code employed in [3] is based on 32768-bit interleaving and 18 decoding iterations.

We observe that, at a BER of 10^{-5} , the scheme presented in [3] outperforms our system by 0.38 dB. This performance difference can be easily explained by the fact that, in our method, the basic 16-PAM constellation can only be divided into 2 sub-constellations, whereas the technique in [3] partitions it into 4 sub-constellations.

The capacity limit of the continuous-input Gaussian channel for a 3-bit/dim application is 10.21 dB. Fig. 5 indicates that, when a BER of 10^{-5} is taken as a reference, our BITCM-S scheme is able to perform within 1.58 dB of this capacity limit. Such error performance is approximately 0.5 dB away from that obtained with the more complex 32-D multi-level turbo-coded modulation scheme described in [2] which is, at the time of writing, the most powerful 3-bit/dim 16-PAM coding system ever designed.

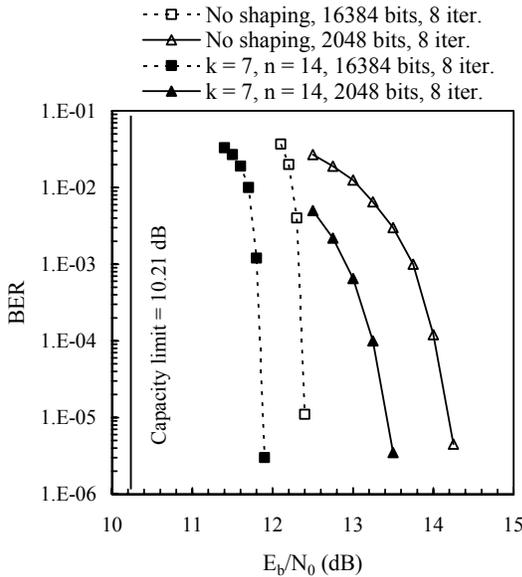


Fig. 5 - Performance comparison over Gaussian channel between several 3-bit/dim 16-PAM BITCM schemes, for 32768-bit and 2048-bit interleaving. (BITCM-S: rate-6/7 turbo code, rate-1/2 shaping code with $k = 7$ and $n = 14$, 8 or 18 decoding iterations; BITCM schemes: rate-3/4 turbo code and 8 decoding iterations).

V. CONCLUSIONS

We have introduced a simple technique to combine constellation shaping and turbo-coded modulation. Simulation results show that a 3-bit/dim 16-PAM scheme designed using this technique can achieve a shaping gain of 0.77 dB, and perform within 1.58 dB of the channel capacity. It is important to mention that our technique preserves most of the simplicity and flexibility that have made BITCM systems so attractive for many practical applications. Finally, we believe that the error performance of BITCM-S schemes could be further improved by incorporating the shaping decoder inside the iterative loop. This will be the subject of further research.

REFERENCES

- [1] U. Wachsmann, R.F.H. Fischer, and J.B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1361-1391, July 1999.
- [2] A.K. Khandani and W. Tong, "Application of shaping technique to multi-level turbo-coded modulation", in *Proc. 21st Biennial Symposium on Communications*, Queens University, Kingston, Canada, June 2002, pp 20-24.
- [3] D. Raphaeli and A. Gurevitz, "An improved pragmatic turbo encoding scheme for high spectral efficiency using constellation shaping," in *Proc. IEEE ICC'03*, Anchorage, Alaska, May 2003, pp. 2738-2742.
- [4] A.R. Calderbank and L.H. Ozarow, "Nonequiprobable signaling on the Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 36, pp. 726-740, July 1990.
- [5] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927-946, May 1998.
- [6] S. Le Goff, "Signal constellations for bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 49, pp. 307-313, Jan. 2003.
- [7] S. Le Goff, A. Glavieux, and C. Berrou, "Turbo codes and high spectral efficiency modulation," in *Proc. IEEE ICC'94*, New Orleans, Louisiana, May 1994, pp. 645-649.
- [8] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error correcting coding and decoding: turbo codes," in *Proc. IEEE ICC'93*, Geneva, Switzerland, May 1993, pp. 1064-1070.

Constructing Finite Pseudo-Random Strings using a Codeset-Based Entropy Measure

Ulrich Speidel

Department of Computer Science
The University of Auckland, New Zealand
ulrich@cs.auckland.ac.nz

May 25, 2004

Abstract

Nicolescu and Titchener showed that every finite string could be used to generate a unique recursively constructed variable-length code set belonging to the family of the T-codes. This paper shows how this duality may be exploited in the construction of random strings.

1 Introduction

T-codes are variable-length codes that can be constructed recursively from S using an algorithm called *T-augmentation* [5, 6, 13]¹. Nicolescu and Titchener [4, 10] showed that, given a finite alphabet S , any finite string $\sigma \in S^+$ defines a unique T-code set $S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}$, in which any string of the form σy , where $y \in S$, is one of the longest codewords. The algorithm that is used to derive the T-prefixes p_1, \dots, p_n and the T-expansion indices k_1, \dots, k_n for the construction of $S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}$ from σ is called *T-decomposition*. A detailed explanation of T-decomposition may also be found in [13].

Titchener [12, 9, 11, 15] further used T-decomposition to derive a complexity measure called T-complexity, and from it a T-information and T-entropy measure, the latter of which was shown experimentally to be closely related to the Kolmogorov-Sinai/Pesin entropy of the logistic map [16]. A summary of T-complexity, T-information, and T-entropy may be found in [17]. Titchener, Fenwick, and Chen [14] used T-information and T-entropy in order to derive a set of calibrated strings with different average T-entropies for the testing of compression algorithms.

¹Publications prior to 2003 were published under the author's pre-marital name, Günther (Guenther)

However, this paper looks at a conventional Shannon-like entropy [1, 2] of the T-Code set resulting from a T-decomposition and explores how this entropy may be used in the construction of pseudo-random strings.

2 The H_1 Entropy

The T-decomposition algorithm allows us to map an arbitrary finite string σ uniquely to a state machine - the decoder for the T-code set $S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}$, for which σy is one of the longest codewords. In [18, 19], the author proposed to use this duality of finite strings and T-code sets for the definition of an entropy, called H_1 .

H_1 is the conventional Shannon entropy of the T-code set $S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}$ arising from σ , provided that the code set is used to decode a random source (e.g., "tossing of a fair coin" in the binary case). The only condition we place on this source is that for all $x \in S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}$, the probability of decoding x as the next codeword is given by

$$P(x) = \#S^{-|x|}. \quad (1)$$

H_1 is thus defined as

$$H_1(\sigma) = - \sum_{x \in S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}} P(x) \log_2 P(x). \quad (2)$$

H_1 may also be computed using the following recurrence relation from [18, 19]:

$$H_1(p_{n+1}^{k_{n+1}} \sigma) = H_1(\sigma) \sum_{k'_{n+1}=0}^{k_{n+1}} \#S^{-k'_{n+1}|p_{n+1}|}. \quad (3)$$

The right hand side of Equation (3) may be converted into a closed form by using the summation

formula for the geometric series, a result that we will require later in this paper:

$$H_1(p_{n+1}^{k_{n+1}}\sigma) = H_1(\sigma) \frac{\#S^{-(k_{n+1}+1)|p_{n+1}|} - 1}{\#S^{-|p_{n+1}|} - 1}. \quad (4)$$

Note that the H_1 entropy is different from Titchener's T-entropy, and no comparison shall be attempted here. The "physical" meaning of $H_1(\sigma)$ is merely the expected length of a codeword in a random string decoded with the T-code set $S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}$.

3 The significance of H_1 in T-decomposition

T-decomposition is the algorithm that extracts the T-prefixes p_1, \dots, p_n and the T-expansion indices k_1, \dots, k_n for the construction of $S_{(p_1, p_2, \dots, p_n)}^{(k_1, k_2, \dots, k_n)}$ from σ . In its simplest form, we may view T-decomposition as a series of n decoding passes. In these passes, σy is decoded first over S , then over $S_{(p_1)}^{(k_1)}$, then over $S_{(p_1, p_2)}^{(k_1, p_2)}$, and so forth. After each decoding pass, the second-to last codeword decoded in σy is the T-prefix for the next level. The T-expansion parameter for the next level is obtained as the total number of adjacent copies of the newly-found T-prefix immediately to the left of and including the T-prefix itself.

Example: Let $\sigma = 10101000$ and let $y = 1$. Then $\sigma y = 101010001$. The T-decomposition of σy may thus be accomplished as follows:

1. First decoding pass over S (the dots denote codeword boundaries inside the string at the respective level):
1.0.1.0.1.0.0.0.1. The $k_1 = 3$ copies of the T-prefix $p_1 = 0$ are underlined.
2. Second decoding pass over $S_{(0)}^{(3)}$:
1.01.01.0001. The $k_2 = 2$ copies of the T-prefix $p_2 = 01$ are underlined.
3. Third decoding pass over $S_{(0,01)}^{(3,2)}$:
1.01010001. The single ($k_3 = 1$) copy of the T-prefix $p_3 = 1$ is underlined.

In each step, consider the string to the right of the newly-picked T-prefixes. It decodes as a single codeword (the last codeword in σy). It is also the string that (under its own T-decomposition) defined the T-code set over which we just decoded σy . Denote this string as $\tau_i y$, where i indicates the T-augmentation level. In our example we thus have $\tau_0 = y = 1$, $\tau_1 = p_1^{k_1} y = 0001$, etc. More generally,

$$\tau_i = p_i^{k_i} p_{i-1}^{k_{i-1}} \dots p_1^{k_1}. \quad (5)$$

Furthermore, consider the string π_i , defined as:

$$\pi_i \tau_i = \sigma. \quad (6)$$

In the $i + 1$ -th decoding pass, π_i is decoded over $S_{(p_1, p_2, \dots, p_i)}^{(k_1, k_2, \dots, k_i)}$. Note that $H_1(\tau_i)$, as the classical Shannon entropy of $S_{(p_1, p_2, \dots, p_i)}^{(k_1, k_2, \dots, k_i)}$, denotes nothing else but the expected/average length of a decoded codeword if a random source is being decoded using $S_{(p_1, p_2, \dots, p_i)}^{(k_1, k_2, \dots, k_i)}$.

4 H_1 and the T-decomposition of random strings

Under the proviso that π_i is a random string and that $|\pi_i|$ is large enough compared to $H_1(\tau_i)$, $H_1(\tau_i)$ may be regarded as the (possibly non-integer) expectation value for the length of p_{i+1} . Moreover, k_{i+1} may in this case be estimated as $\overline{k_{i+1}}$, once again taking the sum over the geometric series:

$$\overline{k_{i+1}} = \sum_{j=0}^{\infty} \#S^{-jH_1(\tau_i)} = \frac{1}{1 - \#S^{-H_1(\tau_i)}}. \quad (7)$$

We may thus estimate the length of τ_{i+1} as:

$$|\tau_{i+1}| = |\tau_i| + \overline{k_{i+1}} H_1(\tau_i). \quad (8)$$

N.B.: Since the T-complexity $C_T(\tau_i)$ is given by

$$C_T(\tau_i) = \sum_{j=1}^i \log_2(k_j + 1), \quad (9)$$

we may compute the expected T-complexity of τ_{i+1} , $\overline{C_T(\tau_{i+1})}$ by adding $\log_2(\overline{k_{i+1}} + 1)$ to the sum, which gives rise to the recurrence relation:

$$\overline{C_T(\tau_{i+1})} = C_T(\tau_i) + \log_2(\overline{k_{i+1}} + 1). \quad (10)$$

We may now be tempted to reinsert $H_1(\tau_i)$ and $\overline{k_{i+1}}$ into the recurrence relation of Equation (3) or (4) in order to compute the expectation value for $\overline{H_1(\tau_{i+1})}$. As Equation (3) requires $\overline{k_{i+1}}$ to be an integer, we have to use the closed form expression in Equation (4). Hence:

$$\overline{H_1(\tau_{i+1})} = H_1(\tau_i) \frac{\#S^{-(k_{n+1}+1)H_1(\tau_i)} - 1}{\#S^{-H_1(\tau_i)} - 1}. \quad (11)$$

The considerations above are based on a few assumptions that may not necessarily hold in every practical context. If σ is of finite length $\gg 1$ and we demand that all π_i for $|\pi_i| \gg H_1(\tau_i)$ are to be random, then we implicitly demand that σ be random, since $\pi_0 = \sigma$ and $H_1(\tau_0) = 1$, trivially.

One obvious question that remains is: What if $|\pi_i|$ is not large enough compared to $H_1(\tau_i)$? This is the case typically towards the end of the T-decomposition process, as π_i gets shorter, while the expected T-prefix length increases as the shorter codewords in $S_{(p_1, p_2, \dots, p_i)}^{(k_1, k_2, \dots, k_i)}$ are depleted. In this case, the possible run of p_{i+1} and hence k_{i+1} are limited by the length of π_i . In particular, consider the last decoding pass, where $i = n - 1$. Here, a value of $H_1(\tau_{n-1}) \gg 1$ implies $k_n = 1$ in the overwhelming majority of cases and hence $H_1(\tau_{n-1}) \approx |\pi_i|$. This effect increases with the length of σ (and hence n). In any case, a T-prefix p_n that gets repeated twice or more either puts us back within the $|\pi_i| \gg H_1(\tau_i)$ regime or leaves us with a σ that is clearly non-random.

5 Generating pseudo-random strings

We now have the tools to consider how H_1 might be used in the generation of pseudo-random strings. If we regard τ_i as the part of the pseudo-random string that we have already created, and presume that we will extend it using codewords from $S_{(p_1, p_2, \dots, p_i)}^{(k_1, k_2, \dots, k_i)}$ as building blocks, then $H_1(\tau_i)$ gives us feedback on how best to extend it. In essence, $H_1(\tau_i)$ gives a recommendation both for the length of the next T-prefix, p_{i+1} , which we have to select from $S_{(p_1, p_2, \dots, p_i)}^{(k_1, k_2, \dots, k_i)}$, and – indirectly via the result in Equation (7) – for k_{i+1} .

Note that H_1 works like a regulator: If we pick $|p_{i+1}| > H_1(\tau_i)$, then $H_1(\tau_{i+1}) < \overline{H_1(\tau_{i+1})}$, thus giving a “stronger recommendation” for the next round. The same is true for the reverse. The actual selection of the T-prefixes and T-expansion parameters still needs to be done by a separate algorithm for which H_1 provides the regulatory feedback signal.

Another approach is to presume that σ is a semi-infinite random string. Instead of starting each decoding pass at the beginning of the string, we now start at a position far enough to the left in order to be able to synchronize the decoder fully to level i before decoding at least one codeword $x \neq p_{i+1}$ followed by the run $p_{i+1}^{k_{i+1}}$. The self-synchronization properties (see, e.g., [13]) and the fact that σ is random guarantee self-synchronization of the decoder and thus the existence of such a starting position. In this case, we may use the aforementioned recurrence relations to predict C_T for random strings of length $|\tau_i|$.

6 Conclusions

The H_1 entropy is a conventional Shannon entropy and may thus be regarded as a predictor of the expected codeword length of codewords emitted by an information source whose output is presumed to be random. The output from the source is read as codewords from a T-code set, which in turn is determined by the source’s previous output. If this source is replaced with a controllable non-random source that we wish to regulate towards random output, H_1 yields the feedback required to close the regulator loop. However, H_1 does not prescribe the exact T-prefix and T-expansion parameters that are to be used in the construction of the output string/code set. A pseudo-random generator on this basis would thus be entirely compatible with the conventional wisdom that semi-infinite random strings may have finite non-random prefixes (or suffixes, depending on the direction) of arbitrary length. Another open problem is the fact that the past output of the “source” is one of the longest codewords in the T-code set and hence taken as its least likely output, when in fact it should be its “typical” output. This suggests that regulation may be biased towards maximum entropy rather than maximal randomness. The relationship between H_1 and Titchener’s T-entropy also remains to be investigated.

7 Acknowledgements

My thanks go to Mark Titchener, Cris Calude, Monica Dumitrescu, and Ludwig Staiger for useful discussions, questions, hints and encouragement.

References

- [1] C. E. Shannon: *A Mathematical Theory of Communications*. Bell Systems Technical Journal, 27:379, July 1948
- [2] C. E. Shannon: *A Mathematical Theory of Communications*. Bell Systems Technical Journal, 27:623, October 1948
- [3] A. Lempel and J. Ziv: *On the Complexity of Finite Sequences*. IEEE Trans. Inform. Theory”, 22(1), January 1976, pp. 75-81.
- [4] R. Nicolescu: *Uniqueness Theorems for T-Codes*. Technical Report. Tamaki Report Series no.9, The University of Auckland, 1995.
- [5] M. R. Titchener: *Generalized T-Codes: an Extended Construction Algorithm for Self-Syn-*

- chronizing Variable-Length Codes*, IEE Proceedings – Computers and Digital Techniques, 143(3), June 1996, pp. 122-128.
- [6] U. Guenther: *Data Compression and Serial Communication with Generalized T-Codes*, Journal of Universal Computer Science, V. 2, N 11, 1996, pp. 769-795. http://www.iicm.edu/jucs_2_11
- [7] U. Guenther, P. Hertling, R. Nicolescu, and M. R. Titchener: *Representing Variable-Length Codes in Fixed-Length T-Depletion Format in Encoders and Decoders*, CDMTCS Research Report no.44, Centre of Discrete Mathematics and Theoretical Computer Science, The University of Auckland, August 1997. <http://www.cs.auckland.ac.nz/research/CDMTCS/docs/pubs.html>.
- [8] U. Guenther, P. Hertling, R. Nicolescu, and M. R. Titchener: *Representing Variable-Length Codes in Fixed-Length T-Depletion Format in Encoders and Decoders*, Journal of Universal Computer Science, 3(11), November 1997, pp. 1207–1225. http://www.iicm.edu/jucs_3_11.
- [9] M. R. Titchener: *A Deterministic Theory of Complexity, Information and Entropy*, *IEEE Information Theory Workshop*, February 1998, San Diego.
- [10] R. Nicolescu and M. R. Titchener, *Uniqueness Theorems for T-Codes*, Romanian Journal of Information Science and Technology, 1(3), March 1998, pp. 243–258.
- [11] M. R. Titchener, *A novel deterministic approach to evaluating the entropy of language texts*, *Third International Conference on Information Theoretic Approaches to Logic, Language and Computation*, June 16-19, 1998, Hsi-tou, Taiwan.
- [12] M. R. Titchener, *Deterministic computation of string complexity, information and entropy*, *IEEE International Symposium on Information Theory*, August 16-21, 1998, MIT, Boston.
- [13] U. Guenther: *Robust Source Coding with Generalized T-Codes*. PhD Thesis, The University of Auckland, 1998. <http://www.tcs.auckland.ac.nz/~ulrich/phd.ps.gz>
- [14] M. R. Titchener, P. M. Fenwick, and M. C. Chen: *Towards a Calibrated Corpus for Compression Testing*, Data Compression Conference, DCC-99, Snowbird, Utah, March 1999.
- [15] M. R. Titchener: *A measure of Information*, IEEE Data Compression Conference, Snowbird, Utah, March 2000.
- [16] W. Ebeling, R. Steuer, and M. R. Titchener: *Partition-Based Entropies of Deterministic and Stochastic Maps*, *Stochastics and Dynamics*, 1(1), p. 45.
- [17] U. Guenther: *T-Complexity and T-Information Theory - an Executive Summary*, CDMTCS Research Report no.149, Centre of Discrete Mathematics and Theoretical Computer Science, The University of Auckland, February 2001. <http://www.cs.auckland.ac.nz/CDMTCS/researchreports/149ulrich.pdf>.
- [18] U. Guenther: *An Entropy Measure for Finite Strings based on the Shannon Entropy of a Code Set*, CDMTCS Research Report no.204, Centre of Discrete Mathematics and Theoretical Computer Science, The University of Auckland, December 2002. <http://www.cs.auckland.ac.nz/CDMTCS/researchreports/204ulrich.pdf>.
- [19] U. Speidel: *An Entropy Measure for Finite Strings based on the Shannon Entropy of a Code Set*, *IEEE International Symposium on Information Theory*, June 29-July 4, 2003, Yokohama, p.25.

Construction of a Convolutional ElGamal Elliptic Curve Cryptosystem

B. Ontiveros, I. Soto, R. Carrasco

Abstract:

In this paper we proposed the construction efficient cryptographic system, based on the combination of the ElGamal Elliptic Curve Algorithm and convolutional codes using the Viterbi decoding algorithm over the Gaussian channel. The originality is based on the construction of the combination of encryption and coding at the channel level. A detailed study of encryption, coding and the expansion factor is presented.

1. Introduction

In 1976, Diffie Hellman [2] introduced the concept of the public key cryptosystem. The security of this scheme is based on the intractability of the discrete logarithm problem in the multiplicative group of a large finite field. In 1985, ElGamal [3] made use of the discrete logarithm problem to construct a practical public key cryptosystem with security equivalent to Diffie Hellman's scheme. In 1985, Koblitz [6] and Miller [4] independently proposed the implementation of a public key cryptosystem using elliptic curve groups over finite fields. The Elliptic curve discrete logarithm problem appear to be much more difficult than the discrete logarithm problem. [5].

The above cryptosystems are implemented at the upper layer of the OSI Model. The main contribution of this paper is to show a simple combination of a non-linear elliptic curve with convolutional codes to provide secure transmission and error correction over a Galois Field at the channel level with reasonable complexity.

In order to improve the convolutional_EC cryptosystem we constructed elliptic curves with a small class number using the Complex Multiplication (CM) method [1].

This paper is organized as follows: Section 2 presents the Mathematical Background

Section 3 describes the flow of information for this new infrastructure. Section 4 presents the results for a suitable method of construction. Finally, conclusions are given in the final section.

2. Mathematical background

When using elliptic curves and codes for cryptography it is necessary to construct elliptic curves with a given or known number of points over a given finite field, in order to represent the input alphabet.

The CM method generates an elliptic curve of a suitable order. The roots of Weber polynomials have to be computed. Each root of such a polynomial determines two possible elliptic curves. However, only one of them has the desired suitable order. Equation 1 defines an elliptic curve E over a finite field F_p , where $p > 3$, which can be given as

$$E(F_p): y^2 = x^3 + ax + b \quad (1)$$

$a, b \in F_p$. Associated with E , there are two important quantities: the discriminant $\Delta = -16(4a^3 + 27b^2)$ and the j -invariant $j = 1728(4a)^3 / \Delta$ where $\Delta \neq 0$.

Given $j_0 \in F_p$ there is an elliptic curve, E , defined over F_p such that $j(E) = j_0$. Let $k = j_0 / (1728 - j_0)$. Then the equation $E: y^2 = x^3 + 3kx + 2k$ is an elliptic curve with j -invariant $j(E) = j_0$.

Let E be defined over F_p and its order be $\#E(F_p) = p + 1 - t$. Let p be an odd prime such that $4p = t^2 + Ds^2$ for some $t, s \in Z$. Then there is an elliptic curve E defined over F_p such that $\#E(F_p) = p + 1 - t$.

3. System description

Figure 1 shows the Convolutional_ElGamal Elliptic Curve (Convolutional_EGEC)

cryptosystem composed of the encryption process (sender), in this case represented by Alice, and at the other end of the channel the Decryption process (receiver), represented by Bob.

The Convolutional_EGEC cryptosystem as the encryption scheme is described using elliptic curve groups with small class number. Let $E(F_p)$ be the small class number elliptic curve and $\mathbf{a} \in E(F_p)$ is a fixed point, a public known point on E , preferably a generator of E . User Bob chooses a random integer a , where a is his private key. With the private key he compute his public key as $\mathbf{b} = a\mathbf{a}$.

To transmit a message x sender Alice chooses a random integer k . In the EGEC module the sender uses a fixed point $\mathbf{a} \in E(F_p)$ from the Public Directory to compute $y_1 = k\mathbf{a}$ and $y_2 = x + k\mathbf{b}$, which are elliptic curve points.

The output information y_1, y_2 from the EGEC module is introduced into the MAPPING module, not shown in figure 1. In order to compress the data, these compressed symbols are denoted by (y_1^*, y_2^*) .

Then by using the public parameters of convolutional codes, (the convolutional code selected is G) in the Module Encoder, Alice encodes (y_1^*, y_2^*) . After the encoding process Alice obtains z and sends it to Bob. Using BPSK modulation the codeword is sent to the channel where it is distorted by the Gaussian Channel. This codeword corresponds to information that has security and channel parity.

To read the message, Bob decrypts using the convolutional code parameters from public directory and decodes using the Viterbi algorithm. Then Bob applies the DEMAPPING module, not shown in figure 1, in order to obtain (y_1, y_2) . In the EGEC⁻¹ module Bob uses his private key to compute $y_2 - a.y_1 = x$ and obtains the

message x' , which is a good approximation of x .

4. Construction method

In order to construct a suitable curve for the Convolutional_EGEC the method is summarized in the following steps:

1. Given a prime number p check whether $4p = t^2 + Ds^2$ has a solution (t, s) , with t, s integers.
2. Search for the smallest D in $4p = t^2 + Ds^2$ along with t (s is not needed in the computations).
3. There are only two possible orders for the curves which can be constructed, These are $\#E(F_p) = p + 1 \pm t$. Check if one of the orders has an admissible factorization (by admissible factorization we mean a prime or nearly prime number as defined in the standards). If not, find another D and corresponding t . Repeat until an order with admissible factorization is found.
4. Construct the Weber $W_D(x)$ polynomial using the formulas given in the standards, $W_D(x) = \prod \left(x - j \left(\frac{-b + \sqrt{-D}}{2a} \right) \right)$. (The class of polynomial for D is a fixed monic polynomial with integer coefficients. In particular, it is independent of p).
5. Compute the roots (modulo p) of either polynomial. For the purpose of constructing suitable elliptic curves over a finite field, we first need to compute the j -invariant. Find a root j_0 of $W_D(x) \pmod{p}$. This j_0 is the j -invariant of the curve to be constructed.
6. Given a root j_0 and $k = j_0 / (1728 - j_0) \pmod{p}$ the first curve will be $E: y^2 = x^3 + 3kx + 2k$. The second curve, called the twist of the first will be $y^2 = x^3 + 3kc^2x + 2kc^3$, where c is any quadratic non-residue in F_p .

7. Check the order of the curve. If it is not $p + 1 - t$ then construct the twist using a randomly selected non square $c \in F_p$.
8. Select a convolutional code $C(n, k)$.

Observe that the combination can use convolutional codes with very long constraint length in order to protect (y_1^*, y_2^*) [7].

Case study: Constructing an elliptic curve over F_p

Previous work has been reported using supersingular curves, but in this paper we explore small class numbers, which provide better security parameters. Let p be the 64-bit prime, $p = 12870883345109083757$. Then search for the smallest Discriminant, $D = -307$, and check if two possible orders for the curve ($\#E(F_p) = p + 1 \pm t$), which can be constructed, have an admissible factorization. If not, find another D and corresponding t . Repeat until an order with admissible factorization is found. Then we construct a Weber $W_D(x)$ polynomial, with ClassNumber = 3. This specifies the degree of the class polynomial computed as:

$$W_D(x) = x^3 + 2907045x^2 + 1610550x + 634500$$

Compute the roots (modulo p) of the Weber polynomial. In this case a root of Q over F_p is $z = 3365961130054963131$. For the purpose of constructing suitable elliptic curves over a finite field we first need to compute the j -invariants. Given a root j_0 and $k = j_0 / (1728 - j_0) \pmod{p}$ the first curve will be $E: y^2 = x^3 + 3kx + 2k$. The elliptic curve over $F_{12870883345109083757}$ is

$$E: y^2 = x^3 + ax + b, \text{ where:}$$

$$a = 8297695495300455506$$

$$b = 6272509432155594431$$

5. Results

In this section three sets of experiments will be considered. Firstly, those related with the

construction time. Secondly, the message expansion factor introduced by the convolutional EGEC Cryptosystem in comparison with RSA ECC. And thirdly, those related with the performance gain introduced by the combination.

Figure 2 represents the time to construct versus discriminant for different class numbers. Small class numbers take less time to construct than larger class numbers. Equation 3 shows the message expansion factor of the ElGamal Elliptic Curve scheme:

$$\frac{\#bits}{y} = \frac{4 \lceil \log_2 p \rceil}{2 \lceil \log_2 p \rceil} = 2 \quad (3)$$

With the convolutional_EGEC the message expansion factor is 1. Table 1 shows the effect introduced by the expansion factor in regards to RSA, ElGamal, ECC and convolutional_EGEC.

	Encrypted Message (bits)
RSA	1024
ElGamal	2048
ECC	321
convolutional_EGEC	100

Table 1: Size of encrypted 100-bit messages.

Figure 3 shows a comparison of the combination of convolutional_EGEC using $a = 8297695495300455506$, $b = 6272509432155594431$ and $K = 7$ with AWGN. Observe that the gain introduced by the designed system is 0.4 dB for a $BER = 10^{-4}$ when compared with a system without encryption.

6. Conclusions

In this paper an efficient cryptographic system has been presented using ElGamal elliptic curves over finite fields and convolutional codes.

The reduction of the expansion factor depends on the mapping selected to perform the compression, from the source data to the cursor size.

Observe that the gain of the designed system is 2.4 dB for a $BER = 10^{-4}$ when compared with an uncoded system and the

gain is 0.4 dB for BER=10⁻⁴ when compared with a system without encryption.

7. Acknowledgements

The authors would like to thank Fondecyt-Chile, project 1030149/2003, for their financial support.

8. References

[1] A. Atkin, F. Morain. *Elliptic Curves and primality proving*, Math. Comp. 61, N° 203, pages 29-68, 1993.

[2] W. Diffie and M. E. Hellman. New directions in cryptography. IEEE Transactions on Information Theory, IT-22:644-654, 1976.

[3] ElGamal, T., A public key cryptosystem and signature scheme based on discrete logarithms, IEEE

Trans. on Information Theory IT-31, 473-481, 1985.

[4] V. S. Miller. Use of elliptic curves in cryptography. In CRYPTO '85, pages 417-426, 1986.

[5] A.J. Menezes. *Elliptic Curve Public Key Cryptosystems*. Kluwer Academic Publishers, 1993.

[6] N. Koblitz. *Elliptic curve cryptosystems*. Mathematics of Computation, 48:203-209, 1987.

[7] Sklar, Bernard. *Digital Communications Fundamentals and Applications*. ISBN 0130847887. Second edition, 2001. Prentice Hall, Inc

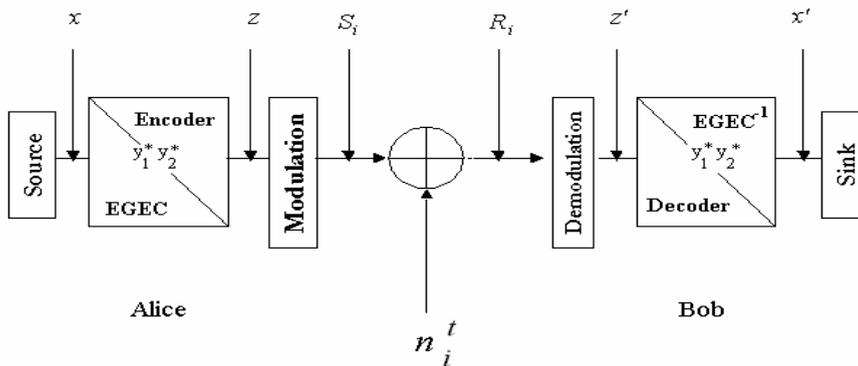


Figure 1: System diagram of the Convolutional ElGamal Elliptic curve Cryptosystem with AWGN, with public parameters $E(F_p)$, $C(n, k)$ and $\mathbf{a} \in E(F_p)$.

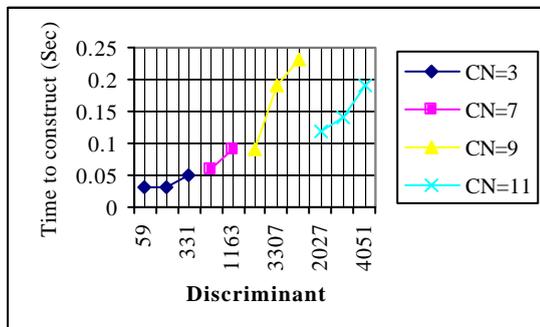


Figure 2: Time to construct versus Discriminant for different Class numbers.

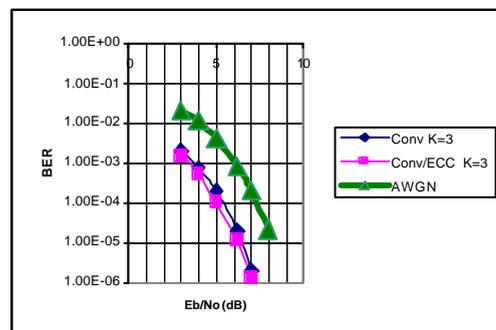


Figure 3: Comparison of the combination of convolutional_EGEC using $a = 8297695495300455506$, $b = 6272509432155594431$ and $K=7$.

Decoding with Arbitrary Pulse-Shape in Optical PPM Communications

Chuanyun Zou

Department of Telecommunication and Information Engineering

Guilin University of Electronic Technology

Guilin, Guangxi 541004, P. R. China

E-mail: hmzchy@gliet.edu.cn

Abstract—Two Maximum-Likelihood (ML) decoding algorithms and sub-optimal one are presented for any pulse-shape pulse position modulation (PPM) in direct-detection optical channel. They are based on the multi-samples per chip, which can give joint estimation of the symbol and the synchronization clock shift. Their performances are compared through computer simulations.

I. INTRODUCTION

Pulse-position modulation (PPM) is a technique that achieves very good average-power efficiency and is widely used in many applications of optical wireless links [1], [2]. Most of results in the published papers [3]-[5] about the decoding to optical PPM signal are based on the single sample per chip with ideal (rectangular) pulses. In practice, however, the optical communication link is often constrained by bandwidth of the channel and the electronic components. As a result, the received pulses are often highly filtered during transmission, and appear as non-rectangular shape. This is especially true when the channel is non-directed indoor infrared radiation, or involves space propagation through clouds, or rain. Although a few works [6] [7] have focused on decoding in optical PPM communication systems with non-rectangular pulse-shape, their analyses are mainly based on the single sample per chip. In this paper, we present the maximum likelihood (ML) decoding algorithms with arbitrary pulse-shape based multi-samples per chip, by which the data and the

This work was supported by the National Natural Science Foundation of China under Grant 69982004 and the Education Dept. of Guangxi Zhuang Autonomous Region in China,.

synchronization clock shift could be obtained at the same time.

This paper is organized as follows. In Section II, we present three schemes of optimal decoding: Maximum Likelihood Chip (MLC), Maximum Likelihood Sample (MLS), and Maximum Accumulating of Samples (MAS). In Section III, the computer simulations for comparison of performance are illustrated. Finally, the concluding remarks can be found in Section VI.

II. SCHEMES OF DECODING

Under Q -ary PPM format, an optical signal pulse is transmitted in one of Q positions, or time chips, in a frame of duration T as shown in Fig.1. The PPM frame consists of Q τ -second chips in modulation time T_S followed by a guard time T_D in which there are D τ -second empty chips. Thus, there are M chips in a frame, the frame length (in seconds) is $T = T_S + T_D = (Q + D)\tau = M\tau$. A single optical pulse transmitted at the i th chip ($i=0,1,2,\dots,Q-1$) of modulation time per frame conveys the $\log_2 Q$ bits of information.

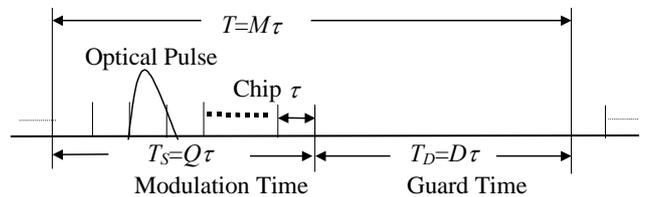


Fig.1. The Digital PPM Symbol Structure

The analysis can be directly based on the photoelectrons at the output of the photodiode detector. In the Q -ary PPM signaling, if we sampled the output of the photodiode P times every chip, the pho-

ton-counting vector of one symbol can be described as,

$$K = [k(0), k(1), \dots, k(PM - 1)]$$

where $k(j)$ ($j=0, 1, \dots, PM - 1$) is the number of the photoelectrons at the output of the optical detector over interval Δ , which is the integrating time of the demodulator following the optical detector. If a PIN photodiode is used, every $k(j)$ will be a Poisson distributed random variable with probability

$$P(k(j) / \lambda(j)) = (\lambda(j)\Delta)^{k(j)} e^{-\lambda(j)\Delta} / k(j)! \quad (1)$$

where $\lambda(j) = \lambda_s(j) + \lambda_b$ is the light intensity of the received signal (photons/s) in interval Δ at the sampling time j , which is composed of signal light $\lambda_s(j)$ and background noise light λ_b . The $\lambda_s(j)$ can be described as $\lambda_s f(j - N)$, N is the start position of the optical PPM pulse, which is a random variable modulated by the source information, λ_s is the maximum intensity of received signal, $f(j)$ is the normalized optical pulse waveform with duration w , that is

$$f(j) = \begin{cases} \neq 0 & 0 \leq j \leq w - 1 \\ = 0 & j < 0, j \geq w \end{cases} \quad (2)$$

When $\Delta < \tau/P$ (independent samples) and $w < DP$ (no interframe interference), we get the likelihood function

$$P(K / N = n) = \prod_{j=0}^{PM-1} P(k(j) / \lambda(j)) \quad (3)$$

$$= c \prod_{j=0}^{PM-1} \left(1 + \frac{\lambda_s}{\lambda_b} f(j - n) \right)^{k(j)}$$

where

$$c = \prod_{j=0}^{PM-1} (\lambda_b \Delta)^{k(j)} \exp(-(\lambda_s f(j - n) + \lambda_b) \Delta) / k(j)!$$

is independent of n , and hence cannot contribute to the decision. Thus, we have log likelihood function

$$l(n) = \sum_{j=0}^{PM-1} k(j) \ln(1 + x f(j - n)) \quad (4)$$

where $x = \lambda_s / \lambda_b$ is peak value signal-noise ratio (SNR). The maximum likelihood PPM receiver is computing $l(n)$ and choosing n according to the largest one, that is

$$\hat{n} = \arg(\max(l(n))). \quad (5)$$

If we assume that there is accurate system synchronization clock, the pulse could only occur in one of the Q positions $\{0, P, 2P, \dots, (Q-1)P\}$, so the pulse position n can be searched in these Q chip points. This decoding algorithm is called Maximum Likelihood Chip (MLC) and the estimation of the symbol I is

$$\hat{I} = \hat{n} / P \quad I \in \{0, 1, \dots, Q - 1\}. \quad (6)$$

In the practical optical wireless PPM communications, because there is always no accurate clock recovery, there may be some shift between the receiver sampling clock and the transmitter one. The start of the pulse will go beyond the Q positions $\{0, P, 2P, \dots, (Q-1)P\}$, so the pulse position must be searched in QP samples $\{0, 1, 2, \dots, QP-1\}$. Thus the estimation of the symbol I and the synchronization shift Δn are

$$\hat{I} = \lfloor \hat{n} / P + 0.5 \rfloor \quad I \in \{0, 1, \dots, Q - 1\} \quad (7)$$

$$\Delta n = \hat{n} - \hat{I} P. \quad (8)$$

In (7) and sequel of this paper, $\lfloor y \rfloor$ denotes the largest integer less than y . 0.5 means $-P/2 < \Delta n < P/2$.

This decoding algorithm is called Maximum Likelihood Sample (MLS), which can demodulate the received symbol and estimate the synchronization clock shift at the same time.

If the received signal is ideal rectangular optical pulse with chip width P , we have

$$l(n) = \ln(1 + x) \sum_{j=0}^{P-1} k(n + j) \quad (9)$$

since the $\ln(1+x)$ is independent of n and cannot contribute to the estimation, so the detection statistic can be chosen as

$$l(n) = \sum_{j=0}^{P-1} k(n + j). \quad (10)$$

(10) is the accumulation of P samples after position n , and the \hat{n} is chosen to have the largest $l(n)$. The detection method based on (10) is called Maximum Accumulation of Samples (MAS). When the optical pulse waveform is unknown, MAS is the sub-optimal algorithm, but its calculation is simpler than MLS

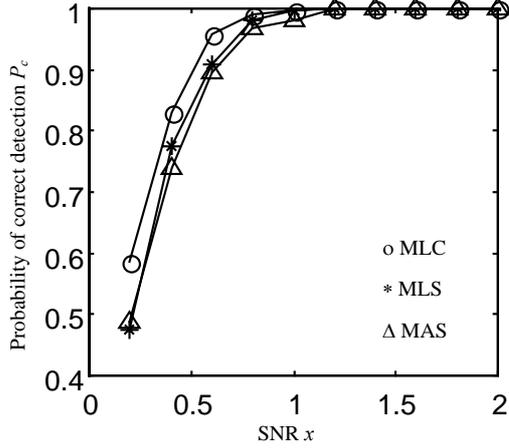


Fig. 2. Probability of correct detection P_c versus SNR x where $Q=4, D=3, P=3, b = 20, \beta=1$

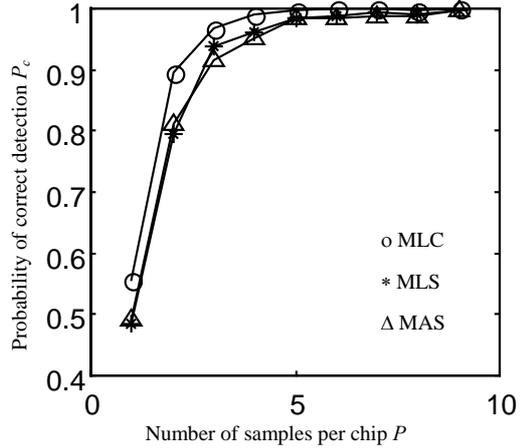


Fig. 4. Probability of correct detection P_c versus number of samples per chip P where $b = 20, Q=4, D=3, x=0.6, \beta=P/3$.

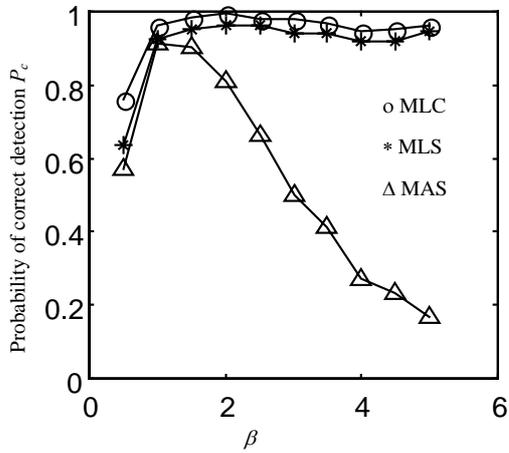


Fig. 3. Probability of correct detection P_c versus pulse-width w or β where $b = 20, Q=4, P=3, x=0.6, D=12$.

and MLS, it does not need knowledge of the pulse shape and can be used in practice.

III. COMPARISON

We will compare the performances of three decoding methods described by MLC, MLS and MAS through computer simulations. In the sequel, we will denote the probability of correct detection by P_c . The simulation results are shown in Figs.2 to 4, which plot the probability of correct detection P_c as a function of the various system parameters, such as SNR x , pulse width w , and number of samples per chip P . The simulation results based on the exponential pulse, which is expressed as

$$f(j) = \frac{j}{\beta} e^{1-\frac{j}{\beta}} \quad (11)$$

where β is the time of pulse waveform peak value. The pulse duration w is taken as $[6\beta]$.

In each plot, the data points of MLC, MLS and MAS are denoted by o, * and Δ , respectively. It can be seen from the graphs that the MLC has the best decoding performance, the MLS is a little inferior to MLC, and the MAS has the worst performance. From Fig. 3, the MAS is much more sensitive to the pulse width than the MLC and MLS. When the pulse spreads into adjacent chip (inchip interference) severely ($w \geq [1.5P]$), the performance of MAS will be significantly degraded, but the two ML receivers still perform well. What is same important, however, is that the performance of MAS, which is much easier to implement than two ML receivers, is comparable to the performance of the latter when no severe in-chip interference. In this case, the MAS receiver will be preferred for its simplicity in practice.

Other observations from Fig. 2 and Fig. 4 are that all receivers improve monotonically as a function of SNR x and number of samples per chip P . For large values of x and P , the performance of all receivers is about the same.

IV. CONCLUSIONS

Several methods of decoding are presented and analyzed in optical PPM systems with any pulse

shape. The methods are called MLC, MLS and MAS, respectively, here. The MLC is an optimal decoding in assumed perfect synchronization. The MLS is an optimal decoding on multi-samples per chip and can give joint estimation of symbol and the synchronization clock shift. Both MLC and MLS have good performance of combating pulse spreading, but they require knowledge of pulse shape. The MAS is a sub-optimal decoding, its performance will be significantly degraded when the pulse spreading into adjacent chip. The MAS, however, is much easier to implement than two ML receivers and it is comparable to the performance of the latter when no severe in-chip interference. In this case, the MAS receiver will be preferred in practice for its simplicity. All conclusions are based on Poisson photodetected observables.

REFERENCES

- [1] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, pp. 265–298, Feb. 1997.
- [2] M. Fitzmaurice and R. Bruno, "NASA/GSFC program in direct detection optical communications for intersatellite links," in *Proc. Optical Space Communications Meeting*, Paris, France, Apr. 26–29, 1989, pp. 10–23.
- [3] S. Patarasen and C. N. Georghiades, "Maximum-likelihood symbol synchronization and detection of OPPM sequences," *IEEE Trans. Commun.*, vol. 42, pp. 2282–2290, June 1994.
- [4] C. N. Georghiades, "Optimum joint slot and symbol synchronization for the optical PPM channel," *IEEE Trans. Commun.*, vol. COM-35, pp. 632–636, June 1987.
- [5] M. P. Advani and C. N. Georghiades, "Jointly optimal receivers for the optical pulse-position modulation channel," *IEEE Trans. Commun.*, vol. 38, pp. 179–186, Feb. 1990.
- [6] G. Prati, "Joint pulse spreading estimation and decoding in stretched pulse PPM optical channels" *IEEE Trans. Commun.*, vol. 33, pp. 760–765, Aug. 1985.
- [7] G. Prati and R. M. Gagliardi, "Decoding with stretched pulses in laser PPM communications" *IEEE Trans. Commun.*, vol. 31, pp. 1037–1045, Sep. 1983.

Design of a digital BPSK demodulator using Field Programmable Gate Arrays

Faruque Ahamed, PhD student and Dr. Frank A. Scarpino, Professor, ECE Dept, University of Dayton,
300College Park, OH 45469,USA.E-mail:ahamedfz@notes.udayton.edu

Abstract

Recent progress of hardware description languages (e.g., Verilog, VHDL) and digital implementation in Field Programmable Gate Arrays (FPGA), substantial academic digital design projects become practicable and challenging. That is, the time and effort to implement significant design projects may be undertaken without sacrificing the broad educational demands placed upon the modern engineering student. In the present paper, the design of an all-digital BPSK demodulator is described. The paper details the design components (e.g., Booth multipliers, PN-generators) as well as the simulation of the entire system. The entire system is designed using the Verilog hardware description language and implemented on an Altera 10k FPGA device. The current paper described is the design and simulation of a digital BPSK demodulator system (correlation detector).

I. Introduction

Digital modulation is the process by which digital symbols are transmitted into waveforms that are compatible with the characteristics of the channel [1]. The modulation process converts a baseband signal into a bandpass signal compatible with available transmission facilities. At the receiver end, demodulation must be accomplished to recognize the signals. The process of deciding which symbol was transmitted is referred to as a demodulation process. The receiver generates a signal that is phase locked to the carrier. Binary Phase Shift keying (BPSK) does not require but may use a coherent receiver. The coherent receiver is called the correlation receiver because it correlates the received signal composed of the transmitted signal plus noise with a sinusoidal signal that is phase locked to the transmitted carrier. The purpose of the correlation receiver is to reduce the received symbol to a single point or statistic that is used by the decision circuit to determine which symbol was transmitted (either 0 or 1). In practice, this single point is a fixed voltage. Since the system described in this paper is digital, the decision point here is a number, obtained by an S/H (sample and hold) device. The decision circuit is a voltage comparator (digital number comparator) that is set such that if the input

voltage (number) is above a threshold level, the comparator indicates a “1” is received; if the input voltage (number) is below that level, a received “0” is indicated [1], [2], [3], [4].

II. Description of the coherent demodulation system of BPSK

With reference to Fig. 1 which is a coherent BPSK system, consider a signal either $s_1(t)$ or $s_2(t)$ is transmitted representing a binary symbols of 1 and 0 respectively and defined by

$$s_1(t) = \sqrt{\frac{2E}{T}} \cos(\omega_0 t + \phi) \quad (1a)$$

$$\text{and} \quad s_2(t) = \sqrt{\frac{2E}{T}} \cos(\omega_0 t + \phi + \pi) \quad (1b)$$

Equation (1b) may be written as

$$s_2(t) = -\sqrt{\frac{2E}{T}} \cos(\omega_0 t + \phi) \quad (2)$$

where $0 \leq t \leq T$, the phase term ϕ is an arbitrary constant, so that the analysis is unaffected by setting $\phi = 0$. The parameter T is the bit duration and E is the transmitted signal energy per bit. Since equations (1) and (2) differ only by a relative phase shift of the 180 degrees, they are referred to as antipodal signals. For the antipodal case, only a single basis function is needed and it can be expressed is as follows:

$$\psi(t) = \sqrt{\frac{2E}{T}} \cos(\omega_0 t) \quad \text{For } 0 \leq t \leq T \quad (3a)$$

The function of $\psi(t)$ of equation (3a) is called the reference signal. The input signal either $s_1(t)$ or $s_2(t)$ is correlated with the reference signal of $\psi(t)$. This correlation gives the desired result. With respect to Fig. 1, if there is no noise then, we can express the correlator input signal $r(t)$ is as follows,

$$r(t) = s_i(t) \quad (3b)$$

where, $i = 1$ or 2 . So, the correlation function is generated

$$z(t) = \int_0^T r(t)\psi(t)dt \quad (3c)$$

where, $z(t)$ is correlator output. From equation (3c), we can write,

$$z(t) = \int_0^T \psi(t)s_1(t)dt \quad (4a)$$

or

$$z(t) = \int_0^T \psi(t)s_2(t)dt \quad (4b)$$

By using equation (1a), (2) and (3a), equation (4a) and (4b) reduces to

$$z(t) = E \quad (5a)$$

and

$$z(t) = -E \quad (5b)$$

The value of $z(t)$ from equation (5a) and (5b), a decision is made that a transmitted signal was either $s_1(t)$ or $s_2(t)$. So, therefore, if the value of $z(t)$ yields E , a decision is made that $s_1(t)$ was transmitted and if it is $-E$, a decision is made that $s_2(t)$ was transmitted.

III. Description of the digital system design and simulation process

Fig. 2 depicts BPSK transmitted carrier signal and PN source data. Fig. 3 describes the modulation process that access at the transmitter location and the detection process that takes place at the receiver. In this project, the simulation system (see Fig. 3) contains two (2) Read-only memory (ROM) elements. The memory element named ROMA is used to generate a sampled data stream that represents the receiver phase-locked reference carrier signal. The ROMA is equivalent to $\psi(t)$ (reference to Fig. 1). Memory element ROMB of Fig. 2 which is equivalent to input signal either $s_1(t)$ or $s_2(t)$ (reference to Fig. 1) provides a sampled data stream that represents the carrier that would normally be generated at the transmitter. The accumulator

output is called $z(t)$ (see Fig. 3) which is equivalent to the output from the integrator (see Fig. 1). The PN generator simulates input data that comes from the information source at the transmitter end of the system. A 2's complement circuit generates a sampled data stream that represents the BPSK modulated carrier that comes from the sending end of the system.

The PN circuit is used to generate a random data pattern that would be available at the transmitter site. The sample carrier signal from ROMB and the pseudo-random data pattern are fed to a two's complement circuit which generates sampled data representing a received, modulated carrier signal. In this project, eight (8) samples have been taken each cycle of the reference and transmitted signals (but in Fig. 2, sixteen samples have been taken to make the smooth curve). The input data is 1 or 0 according to the BPSK modulation scheme. If the input data is 1, the phase of the XTMR signal is unchanged but if it is 0, the 2's complement circuit yields a 180 degree's phase shift of the XMTR signal. The demodulator includes a signed multiplier that takes 16 bit sample data as an input from the reference signal and transmitted signal and produces an output (product) which is thirty two (32) bits long. The demodulator also includes a 32 bit accumulator that stores the product samples from the multiplier for eight sample times. The final bit decision is made depending upon the final numerical value (S) of the accumulator (see Fig. 4). The accumulation represents the integral

$$S \approx \int_0^T \phi(t)s_i(t)dt \quad (6)$$

where, S = accumulator value, $i = 1$ or 2 , and T = bit time. Depending on the accumulator value of the integral as represented by (6), a decision is made that either a "1" or "0" was transmitted. If the accumulator value is positive over the integration period (see Fig. 4), a decision is made that one (1) is transmitted. Otherwise a decision is made that zero (0) is transmitted. Here, we made a decision every eight (8) sample times, which is one bit time in duration and in this case is one cycle of the transmitted carrier. In the real world the poor signal quality is asserted when $T_2 < S < T_1$ (reference to Fig. 4) and indicates a low confidence in the data coming from the demodulator output.

IV. Discussion on simulation result

In this project, both the memory of ROMA and ROMB are used as a source data of reference and transmitted signal respectively. The Booth multiplier takes data from ROMA and ROMB as a multiplicand and multiplier respectively. ROMA is the source of a simulated carrier signal that would normally be generated at the transmitter. ROMB is the source of a carrier signal that is locked to the transmitted carrier signal (as received at the receiver location). The ROMB signal is further processed in the simulation by inverting the phase as necessary to conform to the simulated data signal (as generated by the PN generator).

From the simulation result of Fig. 5, we see that that the BPSK demodulator produces output data (decision), which exactly matches the input data as expected and it comes one (1) bit time later. In this project, the clock frequency is taken 1 MHz, sampling rate around 30 KHz and the bit rate was approximately 3720 b/s. As with any synchronous digital system, the system operating speed scales directly with the clock frequency.

V. Applications of BPSK Demodulator

This type of demodulator has practical application in military and commercial communication systems.

VI. Conclusion

The digital design, simulation and FPGA implementation of BPSK demodulation system using Altera design tools and simulation software have been discussed details in this paper. The Altera maxplus II and Quartus simulation software were utilized to produce the simulation results presented here (only maxplus II). The paper is also shown the digital architecture used to implement the system. In this project, the different components (1) read only memory (ROM) module, (2) a signed digital (Booth) multiplier, (3) an m-sequence (PN) generator and (4) an accumulator used to perform the role of an integrator have been designed and simulated. Then integrating the distinct parts into an operating system, culminating in the simulation of the integrated components. It can be exposed to a typical system development project as well as the details involved in the particular system under consideration.

References

- [1] Bernard Sklar, Digital Communications – Fundamentals and Applications, 2nd ed., New Jersey : Prentice Hall PTR, 2001.
- [2] Floyd M. Gardner, “BPSK/QPSK Timing-Error Detector for Sampled Receivers,” IEEE Transactions on communications, Vol. Com-34, No. 5, pp. 423-429, May 1986.
- [3] Ron Sorace, “Probability of Error for BPSK Modulation from Electrostatic Interference,” IEEE Transactions on Aerospace and Electronic systems, Vol.35, No.4, pp. 1383-1387, October 1999.
- [4] Douglas A. Hill and John B. Bodie, “Carrier Detection of PSK Signals,” IEEE Transactions on communications, Vol.49, No. 3, pp. 487-496, March 2001.

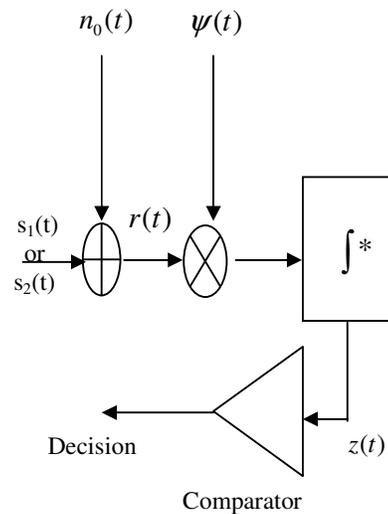


Fig. 1. An idealized single channel system diagram

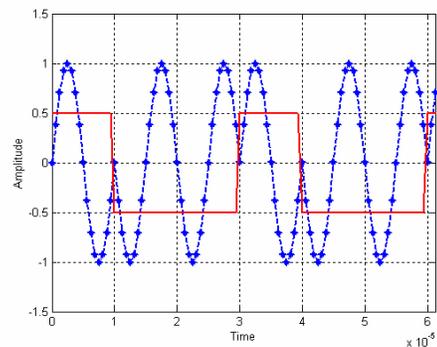


Fig. 2. BPSK transmitted carrier signal (dotted line) and PN source Data (solid line)

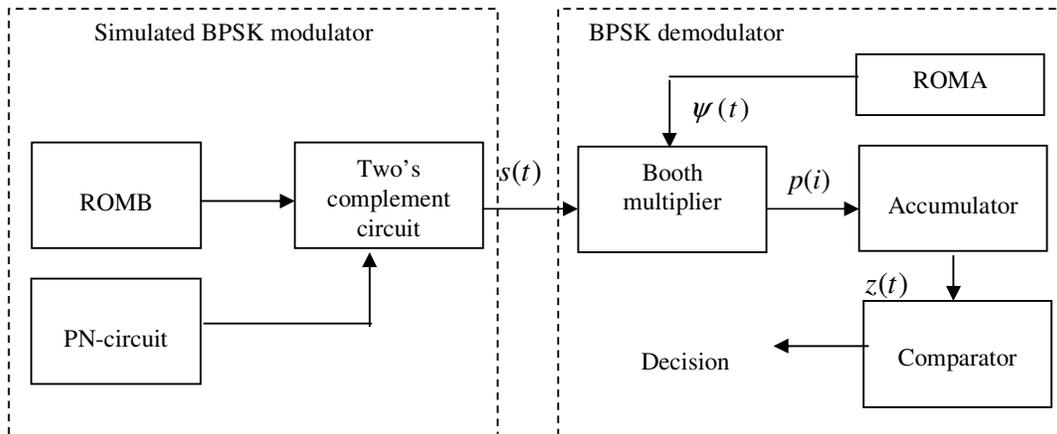


Fig. 3. Architecture of digital implementation and simulation system

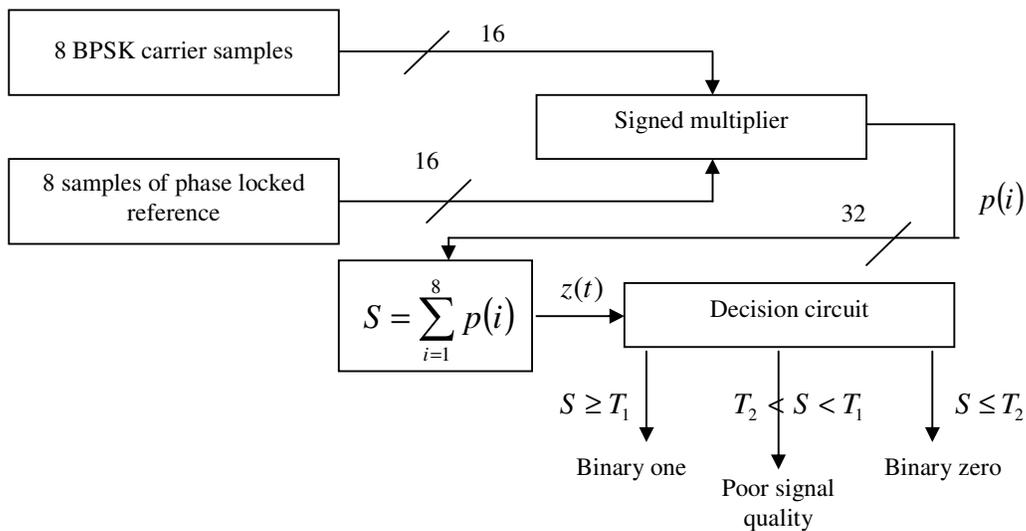


Fig. 4. Architecture of computational section of digital BPSK demodulator.

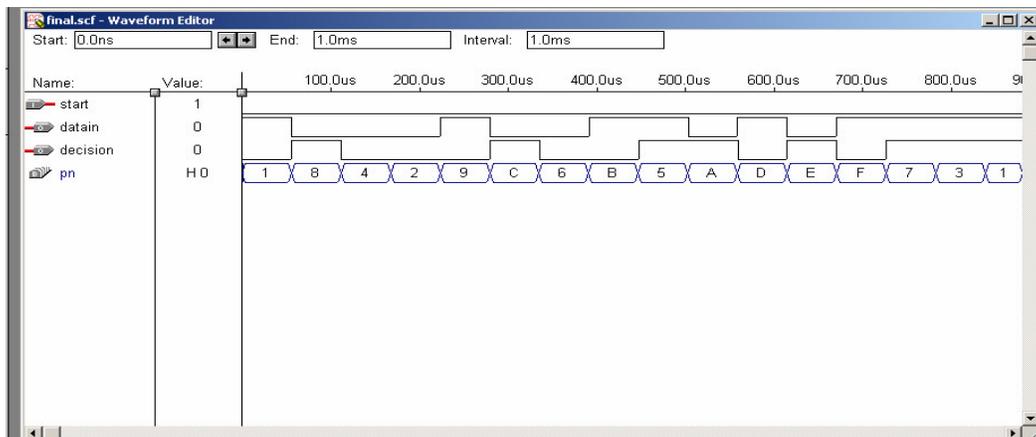


Fig. 5. Simulation result using maxplus II

Dual semiconductor optical amplifier polarisation independent wavelength conversion using four-wave mixing

David I. Forsyth and Michael J. Connelly.

*Optical Communications Research Group, Dept. of Electronic and Computer Engineering,
University of Limerick, Limerick, Ireland.
e-mail: david.forsyth@ul.ie*

Abstract: Semiconductor optical amplifiers (SOAs) have been arranged with the aim of obtaining polarisation-free wavelength conversion, using four-wave mixing (FWM). Two SOAs were used in a polarisation diversity arrangement, with an optoelectronic feedback loop added. Results at bit rates up to 2.5 Gb/s indicate the technique achieves a polarisation dependence of < 0.5 dB in the FWM-generated conjugate.

Introduction

A key function in future high capacity DWDM optical networks will be the ability to convert data from one wavelength channel to another [1]. This makes it re-configurable and helps reduce blocking probability in such wavelength-routed networks [2]. It also increases the flexibility and capacity of limited WDM networks, centralises network management and allows easier protection switching [1, 3]. In addition, it can help resolve packet contention and reduce optical buffering requirements [1].

Previous wavelength converter types include difference-frequency generation (DFG) in electro-optic crystals and passive semiconductor waveguides, and four-wave mixing (FWM) in silica fibres, SOAs, FP lasers, DFB lasers and passive semiconductor waveguides [4]. SOAs have numerous advantages at present [3]: they are more compact, have low power consumption, have all

optical ultra-fast response, are readily integrated and mass produced, and they utilise non-linear effects to produce high speed wavelength conversion with continuous wide-range tuning (up to 65 nm downshift [5]).

Four-wave mixing (FWM) is a coherent, non-linear process involving the interaction of two or more optical fields in an optical semiconductor to create a new optical field at a different wavelength [6]. It is an attractive generation of a signal replica (plus sideband) from two injected optical fields (i.e a pump + modulated/cw signal), and this new conjugate signal is a spectrally inverted replica of input probe signal [7]. Efficient FWM is achieved when the pump and signal are co-polarised and is theoretically characterised by several ultrafast and competing dynamics occurring within the SOA converter [6, 7]. FWM SOA wavelength conversion is easily tuneable to convert up or down over a large range (~ 65 nm) [5].

A key drawback of SOA FWM wavelength conversion, however, is that it is very *polarisation* sensitive [6]. Most FWM schemes are single-pump using only one SOA and, since the polarisation state (SOP) of the incoming signal is in reality an uncontrolled variable, this situation can be heavily polarisation dependent (in some cases the newly-generated conjugate can change by > 10 dBm or even completely disappear with changing signal input polarisation state) and therefore potentially a

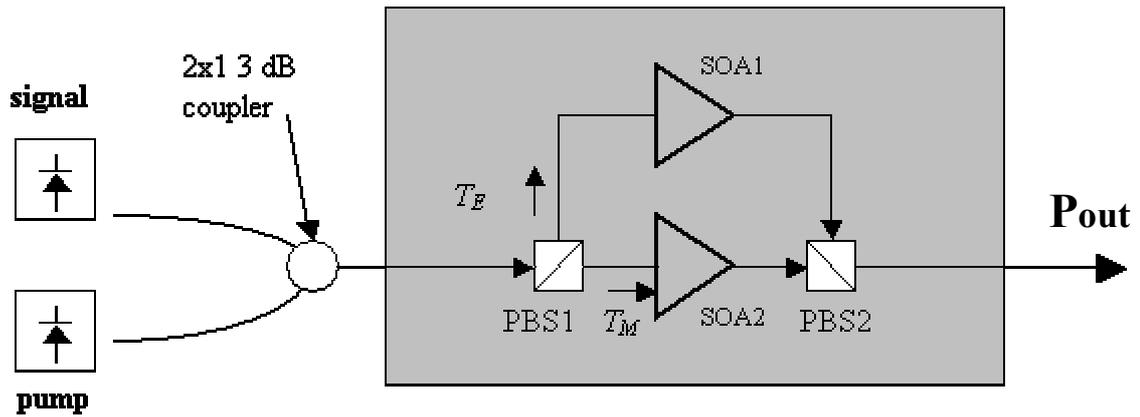


Figure 1: Schematic polarisation-diversity SOA FWM layout.

disaster for real communication links. Hence, the goal is to be as diverse as practically possible to polarisation effects using certain SOA FWM configurations.

Many types of SOA FWM polarisation-diversity schemes have been used previously - these are usually with two pumps and one SOA [4, 9, 10, 11], one pump and two SOAs [12, 13] or with two pumps and two SOAs [14]. Figure 1 shows a schematic layout of a one pump and two SOA-type scheme, which will be the method employed in this work [12]. A pump and signal probe are coupled into a system consisting of two polarising beamsplitters/combiners and two SOAs. The pump is polarised at 45° to a polarisation beam splitter's (PBS1) principal axes, so that half of its power goes to each SOA. The input to each SOA is a pump field and the co-polarised component of the signal field. These mix in each SOA to produce a conjugate signal with the same polarisation. These orthogonally polarised conjugates are recombined by PBS2. So, if each SOA has the same gain and conversion efficiency, then the scheme is polarisation independent. The aim of our work is to create a practical wavelength/frequency converter system, using a similar polarisation-diversity SOA FWM scheme, to achieve efficient and reasonably polarisation-insensitive wavelength conversion at high bit rates.

Experimental Arrangement

A polarisation-diversity scheme was created, as shown in Figure 2, with all three SOAs computer controlled for bias voltage and temperature. A pseudorandom-bit-sequence (PRBS) of $2^{15} - 1$ modulated at 2.488 Gb/s was used for the input stream of data (i.e the probe input to be wavelength-converted) and, with pump/probe wavelength selection, the desired new wavelength could easily be chosen. A down conversion of around 4 nm was used, the pump at 1550 nm and probe at 1552 nm. With the probe switched off, the input pump power to each SOA was equalized using polarisation controller 1. The probe was then activated and adjusted to a value just less than the pump power. Polarising beam splitter PBS1 created the necessary orthogonal SOPs through each SOA, and recombination was achieved by PBS2. A 0.1 nm FBG-based filter then selected the wavelength of interest. By adjusting the bias voltage to one SOA, any slight difference in SOA gains could be compensated for. Also the path lengths from PBS1 to PBS2 through each SOA were equalized to put the conjugates in phase. A further SOA was used as a preamp to improve receiver sensitivity of FWM product. The EDFA before PBS1 ensured saturation operation of SOA FWM [6]. Additionally, an optoelectronic feedback loop was applied to one of the SOAs built-in APC

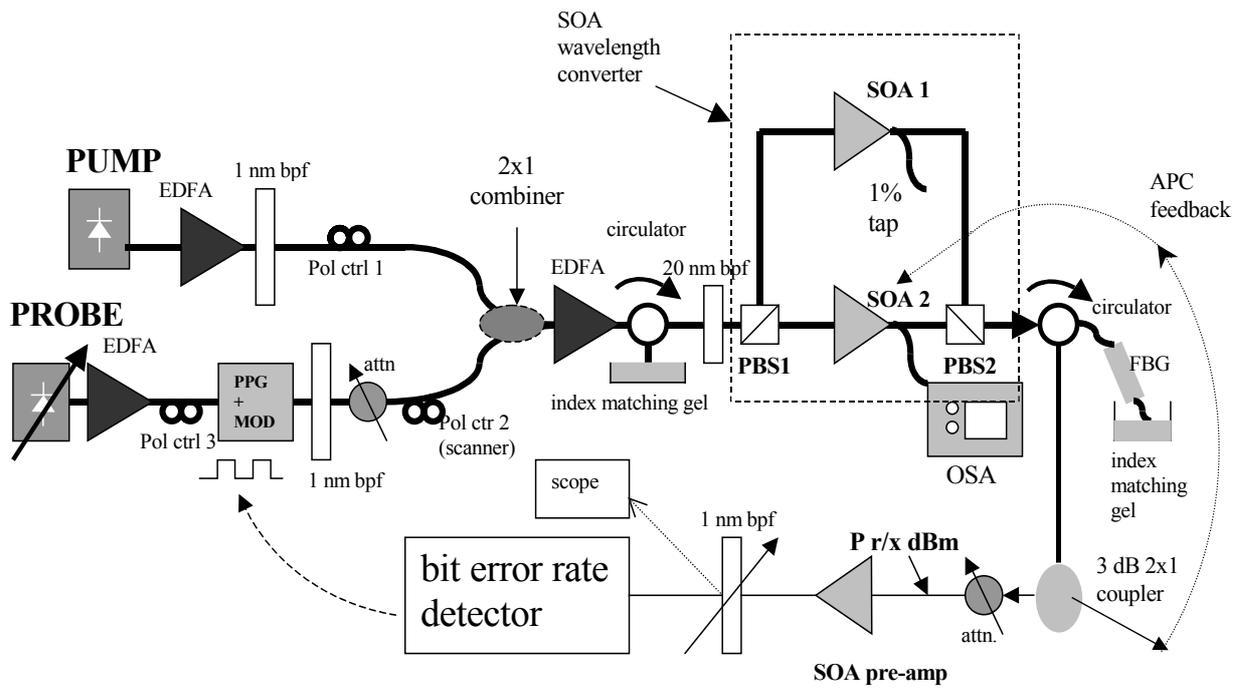


Figure 2: Experimental set-up

(automatic power control) facility. Using the above system, the quality and polarisation sensitivity of the new FWM wavelength-converted signal was monitored over periods of up to 12 hrs.

Results

Figure 3 shows the eye diagram received on the scope at $P_{r/x}$ (mean) = -30.5 dBm (no attenuation):

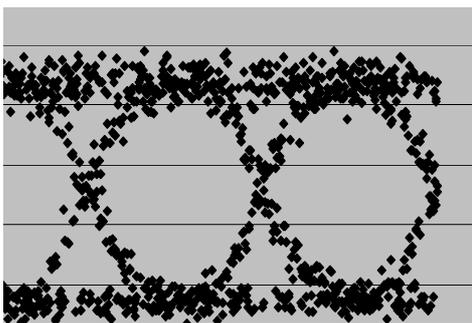


Figure 3: Eye diagram of FWM signal

Measured Q factor was around 8, jitter ~18 ps and the extinction ratio was ~7.4 dB. The pump and probe input powers were 4.9 dBm and -7.5 dBm respectively at the wavelength converter input. At each SOA input, these were 0.5 dBm (when split

50:50) and from -19 dBm to -10 dBm (depending on probe SOP) respectively. Both SOA gains were equalised prior to the measurements. Figure 4 shows BER test results performed on the FWM signal, revealing an improvement of ~1 dB in receiver sensitivity at 10^{-9} BER with the feedback:

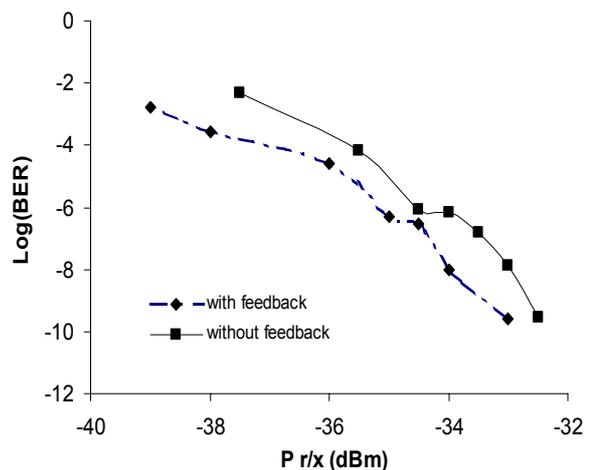


Figure 4: BER performance

As the polarisation of the input signal was changed over all possible states (by scanning polarisation controller 2), the output power of the corresponding wavelength-converted signal changed by less than

1.0 dB without feedback, and < 0.5 dB with feedback applied.

Conclusions

This work has shown that the dependence of the FWM-generated conjugate on the SOP of the input probe in a SOA FWM polarisation-diversity wavelength conversion scheme can be improved to better than 0.5 dB with feedback.

Acknowledgements

The authors are pleased to acknowledge the support of the Tellabs Foundation in Ireland for this work.

References

- [1] I. White, R. Penty, Matthew Webster, Yew Jun Chai and Adrian Wonfor, Wavelength Switching Components for Future Photonic Networks, *IEEE Communications Magazine*, pp. 74 - 81, September 2002.
- [2] D. Nasset, T. Kelly and D. Marcenac, All-Optical Wavelength Conversion Using SOA Nonlinearities, *IEEE Communications Magazine*, pp. 56-61, December 1998.
- [3] "Semiconductor Optical Amplifiers", Michael J. Connelly, Kluwer Academic Press (2002).
- [4] J.P.R. Lacey, M.A. Summerfield and S.J. Madden, Tunability of Polarization-Insensitive Wavelength Converters Based on Four-Wave Mixing in Semiconductor Optical Amplifiers, *Journal of Lightwave Technology*, Vol. 16, Issue 12, pp. 2419 - 2427, December 1998.
- [5] J. Zhou, N. Park, K.J. Vahala, M.A. Newkirk and B.I. Miller, Four-wave mixing wavelength conversion in semiconductor travelling-wave amplifiers measured to 65 nm of wavelength shift, *IEEE Photonics Technology Letters*, Vol. 6, pp. 984 - 987, 1994.
- [6] D.F. Geraghty, R.B. Lee, M. Verdiell, M. Ziari, A. Mathur and K.J. Vahala, Wavelength conversion for WDM communication systems using four-wave mixing in semiconductor optical amplifiers, *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 3, Issue 5, pp. 1146 - 1155, October 1997.
- [7] S. Diez, C. Schmidt, R. Ludwig, H.G. Weber, K. Obermann, S. Kindt, I. Koltchanov and K. Petermann, Four-wave mixing in semiconductor optical amplifiers for frequency conversion and fast optical switching, *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 3, Issue 5, pp. 1131 - 1145, October 1997.
- [8] A. Mecozzi, S. Scotti, A. D'Ottavi, E. Iannone and P. Spano, Four-wave mixing in traveling-wave semiconductor amplifiers, *IEEE Journal of Quantum Electronics*, Vol. 31, Issue 4, pp. 689 - 699, April 1995.
- [9] Polarisation-independent phase conjugation of lightwave signals, R.M. Jopson and R.E. Tench, *Electronics Letters*, Vol. 29, Issue 25, pp. 2216 - 2217, 9 December 1993.
- [10] R. Schnabel, U. Hilbk, T. Hermes, P. Meissner, C. Helmolt, K. Magari, F. Raub, W. Pieper, F.J. Westphal, R. Ludwig, L. Kuller and H.G. Weber, Polarization insensitive frequency conversion of a 10-channel OFDM signal using four-wave-mixing in a semiconductor laser amplifier, *IEEE Photonics Technology Letters*, Vol. 6, Issue 1, pp. 56 - 58, January 1994.
- [11] K.K. Chow, C. Shu, M.W.K. Mak and H.K. Tsang, Polarisation-insensitive widely tunable wavelength converter using polarisation-diversified SOA fibre-ring laser, *Electronics Letters*, Vol. 39, Issue 1, pp. 96 - 97, 9 January 2003.
- [12] J.P.R. Lacey, S.J. Madden and M.A. Summerfield, Four-Channel Polarization-Insensitive Optically Transparent Wavelength Converter, *IEEE Photonics Technology Letters*, Vol. 9, No. 10, October 1997.
- [13] U. Feiste, R. Ludwig, E. Dietrich, S. Diez, H.J. Ehrke, D. Razic and H.G. Weber, 40 Gbit/s transmission over 200 km of standard-fiber using polarisation independent mid-span spectral inversion, *24th European Conference on Optical Communication*, Vol. 1, pp. 531 - 532, 20-24 September 1998.
- [14] A. Mecozzi, Applications of four-wave mixing in semiconductor optical amplifiers, *24th European Conference on Optical Communication*, Vol. 1, pp. 647 - 648, 20-24 September 1998.

Effects of the Prefilter Bandwidth on Symbol Synchronizers

António D. Reis^{1,2}, José F. Rocha¹, Atilio S. Gameiro¹, José P. Carvalho²

¹Dep. de Electrónica e Telecomunicações / Instituto de Telecomunicações, Universidade de Aveiro, 3810 Aveiro, Portugal

²Dep. de Física, Universidade da Beira Interior Covilhã, 6200 Covilhã, Portugal

Abstract

In this paper we present four types of synchronizers, namely the analog, the hybrid, the combinational and the sequential.

Then, we will consider four prefilter bandwidth PF(s), namely $fc=\infty$, $fc=5.tx$, $fc=2.tx$ and $fc=1.tx$, where tx is the transmission rate.

The main objective is to study as the prefilter bandwidth affects the performance of each synchronizer.

Key words: Synchronism in Digital Communications

I. INTRODUCTION

In this work we present four types of symbol synchronizers namely, the analog, the hybrid, the combinational and the sequential.

The difference between them is inside of the signal comparator.

Fig.1 shows the general configuration of the symbol synchronizers with prefilter.

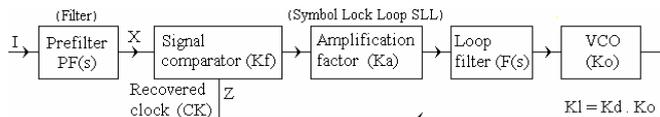


Fig.1 General symbol synchronizer with prefilter

The signal comparator detects the phase of the positive clock transition relatively to the symbol center of the information data.

The others blocks are identical in all types, thus the amplification factor is a controlled gain that gives the same condition to all synchronizers.

The loop filter select the low frequencies and eliminates the high frequencies.

The VCO is the voltage controlled oscillator that correspond to the recovered clock.

After we go present the prefilter considering four versions with different bandwidth ($PF(s)=\infty$, $PF(s)=5*tx$, $PF(s)=2*tx$, $PF(s)=1*tx$).

Next, we present the behavior of each synchronizer (analog, hybrid, combinational and sequential) tested with the four prefilter bandwidth.

After, we show the design, tests and results with some comparisons [1, 6, 7].

Finally, we present the main conclusions.

II. TYPES OF CLOSED LOOP SYNCHRONIZERS

The four topologies of closed loop symbol synchronizers are

obtained introducing the respective signal comparator [3].

A. Analog synchronizer

Fig.2 shows the analog topology, in which the signal comparator is constituted by two analog multipliers (for positive and negative transitions) and some delays.

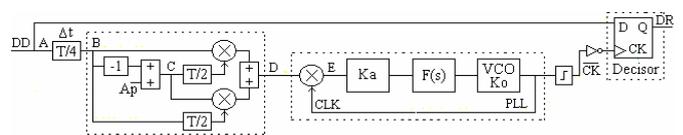


Fig.2 Analog synchronizer

The input signals (main input and VCO output) are both analog.

B. Hybrid synchronizer

Fig.3 shows the hybrid topology, in which the signal comparator is constituted by an exor that generates transition pulses followed by a switch that divides the pulse in two.

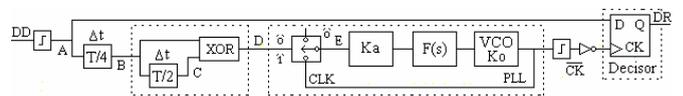


Fig.3 Hybrid synchronizer

The input signal is already digital but the VCO output is still analog.

C. Combinational synchronizer

Fig.4 shows the combinational topology, in which the signal comparator is constituted by an exor that detects the transitions, followed by a denux that divides the pulses.

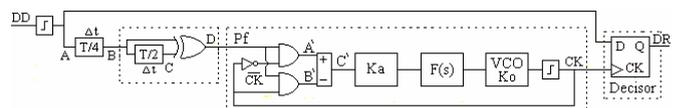


Fig.4 Combinational synchronizer

The two signal comparator inputs are both digital and the output is only function of the present inputs.

D. Sequential synchronizer

Fig.5 shows the sequential topology, in which the signal comparator is constituted by a flip flop with exor (variable pulse) and an exor with a delay (fixed pulse) [2].

^{1,2}UA-UBI

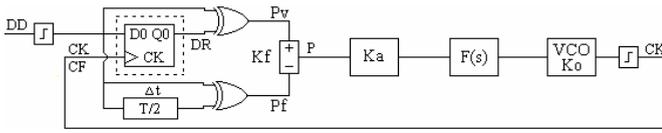


Fig.5 Sequential synchronizer

The two signal comparator inputs are both digital but now its output depends simultaneously on the two entries and also on the phase comparator state (memory).

III. PREFILTERS PF(s)

The prefilter that we use is the Butterworth filter in which we consider four versions with different bandwidths.

Fig.6 shows the four prefilters bandwidths. In (a) the filter bandwidth is $PF(s)=\infty$, in (b) the bandwidth is $PF(s)=5*tx$, in (c) the bandwidth is $PF(s)=2*tx$ and in (d) the bandwidth is $PF(s)=1*tx$ [4].

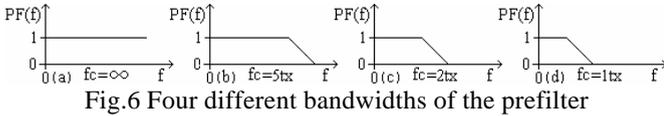


Fig.6 Four different bandwidths of the prefilter

We opted by the Butterworth filter of order one, that has less performance but much more speed.

A. Prefilter bandwidth ($fc=\infty$)

In this case the Butterworth filter has a bandwidth $PF(s)$ of infinite value ∞ .

B. Prefilter bandwidth ($fc=5tx$)

In this case the Butterworth filter has a bandwidth $PF(s)$ equal to five times the data rate $5*tx$.

C. Prefilter bandwidth ($fc=2tx$)

In this case the Butterworth filter has a bandwidth $PF(s)$ equal twice the data rate $2*tx$.

D. Prefilter bandwidth ($fc=1tx$)

In this case the Butterworth filter has a bandwidth $PF(s)$ equal to the data rate $1*tx$.

IV. DESIGN, TESTS AND RESULTS

A. Test setup

Fig.7 shows the setup that we used to get the jitter-noise curves of the various synchronizers [5].

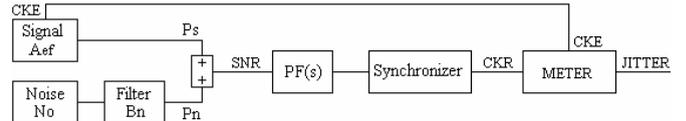


Fig.7 Block diagram of the test setup

The signal to noise ratio SNR is given by P_s/P_n , where P_s is the signal power and P_n is the noise power. They are defined as $P_s=A_{ef}^2$ and $P_n=N_o.B_n=2\sigma_n^2\Delta\tau B_n$. A_{ef} is the RMS amplitude, B_n is the external noise bandwidth, N_o is the noise power spectral density, σ_n is the noise standard deviation and $\Delta\tau$ is the sampling period (inverse of samples per unit time).

The prefilter is not used here, but can be useful in system with high noise quantities ($PF(s)=1$).

B. Jitter measurer

Fig.8 shows the jitter measurer (METER) that consists of a RS flip-flop which detects the variable phase of the recovered clock (VCO) relatively to the fixed phase of the emitter clock.

This relative phase variation is the recovered clock jitter.

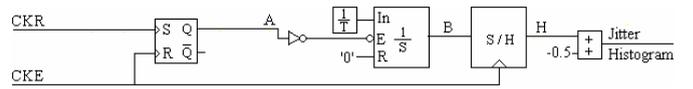


Fig.8 The jitter measurer

The others blocks convert this phase variation into an amplitude variation, which is the jitter histogram.

Fig.9 shows the waveforms that illustrate the operation mode of the jitter measurer.

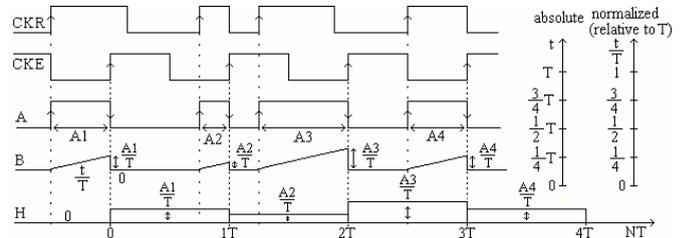


Fig.9 Waveforms at the jitter measurer

The jitter histogram is then sampled and processed by an appropriated program giving the average m , jitter variance in squared radians σ_n^2 , jitter standard deviation in unit intervals root mean squared UIRMS and jitter standard deviation in unit intervals peak to peak UIPP.

We have used also others jitter measurers with similar results.

C. Loop parameters design

To establish guaranteed comparisons it is necessary to test all the synchronizers in equal conditions.

We use a normalized transmission rate $tx=1$ baud ($f_o=1$ Hz) what facilitates the analyses and allows one more easy extrapolation for other rhythms of transmission. We use an equivalent external noise bandwidth $B_n=5$ Hz for all SLL. For

the closed loop symbol synchronizers SLL, we use a loop noise bandwidth $B_l=0.02\text{Hz}$.

For analog SLL, the relation between signal to noise ratio SNR and jitter variance σ_n is $\text{SNR}=\text{Aef}^2/\text{No.Bn}=\text{Aef}^2/(2\sigma_n^2 \cdot \Delta\tau \cdot \text{Bn})=(0.5)^2/(2\sigma_n^2 * 10^{-3} * 5)=25/\sigma_n^2$. This relation is more complicated for the others symbol synchronizers.

We will now present the loop parameters design for the various PLLs considering the first (1st) and the second order loop (2nd).

- 1st order loop:

In the 1st order loop, the filter $F(s)=0.5\text{Hz}$ eliminates only the high frequency, but maintain the loop characteristics. This cutoff frequency $F(s)=0.5\text{Hz}$ is 25 times higher than $B_l=0.02\text{Hz}$. Then the transfer function of the 1st order is

$$H(s)=\frac{G(s)}{1+G(s)}=\frac{KdKo}{s+KdKo} \quad (1)$$

the loop noise bandwidth for the SLLs is

$$B_l=\frac{KdKo}{4}=\text{Ka}\frac{KfKo}{4}=0.02\text{Hz} \quad (2)$$

so for the analog SLL with $K_m=1$, $A=1/2$, $B=0.45$ we have

$$\text{Ka}\frac{K_mABKo}{4}=0.02\text{Hz} \rightarrow \text{Ka}=0.08\frac{2.2}{\pi} \quad (3)$$

for the hybrid SLL, with $K_m=1$, $A=1/2$ and $B=0.45$ we have

$$\text{Ka}\frac{K_mABKo}{4}=0.02\text{Hz} \rightarrow \text{Ka}=0.08\frac{2.2}{\pi} \quad (4)$$

for the combinational SLL ($K_f=1/\pi$) we have

$$\text{Ka}\frac{KfKo}{4}=\text{Ka}\frac{(1/\pi)2\pi}{4}=0.02\text{Hz} \rightarrow \text{Ka}=0.04 \quad (5)$$

and for the sequential SLL ($K_f=1/2\pi$) we have

$$\text{Ka}\frac{KfKo}{4}=\text{Ka}\frac{(1/2\pi)2\pi}{4}=0.02\text{Hz} \rightarrow \text{Ka}=0.08 \quad (6)$$

This formulas are useful in synchronizers design

- 2nd order loop:

The transfer function with $F(s)=\frac{1+sT_2}{sT_1}$ is

$$H(s)=\frac{sKdKo(T_2/T_1)+KdKo/T_1}{s+sKdKo(T_2/T_1)+KdKo/T_1} \quad (7)$$

$$=\frac{sA+B}{s^2+s2\mathbf{x}Wn+Wn^2} \quad (8)$$

and the loop noise bandwidth is

$$B_l=\frac{\mathbf{x}Wn}{2}\left(1+\frac{1}{4\mathbf{x}^2}\right) \quad (9)$$

Taking ($\xi=1$ and $B_l=0.02$) and solving the above equations we obtain for $F(s)$

$$F(s)=\frac{1+s63}{s977} \quad (10)$$

so for the analog SLL we have

$$Kd=KaKf=Ka(1/2)(1/2)=\frac{1}{2\mathbf{p}} \rightarrow \text{Ka}=\frac{2.2}{\mathbf{p}} \quad (11)$$

for the hybrid SLL we have

$$Kd=KaKf=Ka(1/2)(0.45)=\frac{1}{2\mathbf{p}} \rightarrow \text{Ka}=\frac{2.2}{\mathbf{p}} \quad (12)$$

for the combinational SLL we have

$$Kd=KaKf=Ka\frac{1}{\mathbf{p}}=\frac{1}{2\mathbf{p}} \rightarrow \text{Ka}=0.5 \quad (13)$$

and for the sequential SLL we have

$$Kd=KaKf=Ka\frac{1}{2\mathbf{p}}=\frac{1}{2\mathbf{p}} \rightarrow \text{Ka}=1 \quad (14)$$

This formulas can be used in others synchronizers.

D. Results

We studied the jitter-noise behavior of the analog (ana), hybrid (hib), combinational (cmb) and sequential (seq) synchronizers. Each synchronizer was tested with four prefilter bandwidths ($f_c=\infty$, $f_c=5\text{tx}$, $f_c=2\text{tx}$ and $f_c=1\text{tx}$).

Fig.10 shows the jitter-noise curves of the analog synchronizer with the four prefilter bandwidths.

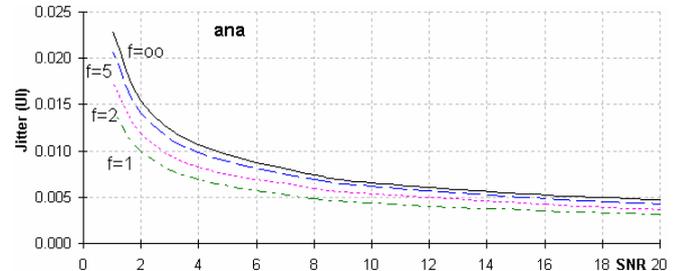


Fig.10 Analog synchronizer with the four prefilters $PF(s)$

We verify that generically (for low and high SNR) the jitter diminishes when the prefilter bandwidth decreases from $f_c=\infty$ until $f_c=1\text{tx}$.

Fig.11 shows the jitter-noise curves of the hybrid synchronizer with the four prefilter bandwidths.

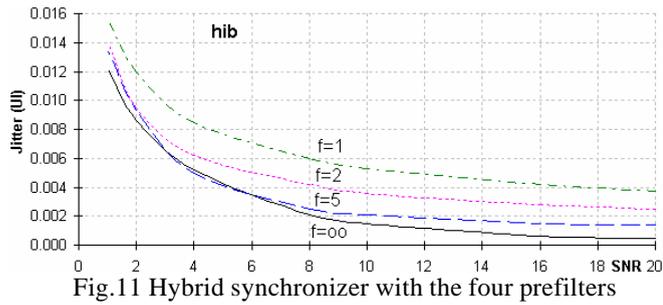


Fig.11 Hybrid synchronizer with the four prefilters

We verify that, for low SNR ($SNR < 4$), the prefilters are advantageous. So in this case the jitter diminishes when the filter bandwidth decreases from $f_c = \infty$ until $f_c = 1 \cdot tx$.

However for high SNR ($SNR > 4$) the prefilter is disadvantageous. So in this case the jitter increases when the filter bandwidth decreases from $f_c = \infty$ until $f_c = 1 \cdot tx$.

Fig.12 shows the jitter-noise curves of the combinational synchronizer with the four prefilter bandwidths.

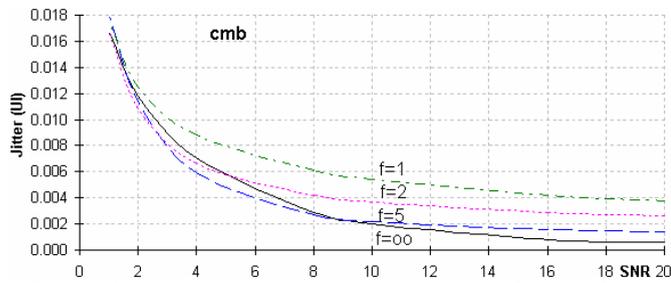


Fig.12 Combinational synchronizer with the four prefilters

We verify that, for low SNR ($SNR < 4$), the prefilters are advantageous. So in this case the jitter diminishes when the filter bandwidth decreases from $f_c = \infty$ until $f_c = 1 \cdot tx$.

However for high SNR ($SNR > 4$) the prefilter is disadvantageous. So in this case the jitter increases when the filter bandwidth decreases from $f_c = \infty$ until $f_c = 1 \cdot tx$.

Fig.13 shows the jitter-noise curves of the sequential synchronizer with the four prefilter bandwidths.

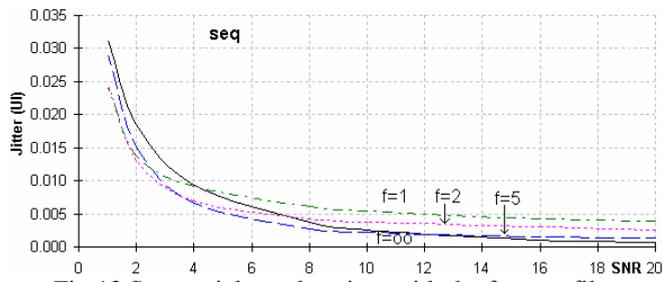


Fig.13 Sequential synchronizer with the four prefilters

We verify that, for low SNR ($SNR < 4$), the prefilters are advantageous. So in this case the jitter diminishes when the filter bandwidth decreases from $f_c = \infty$ until $f_c = 1 \cdot tx$.

However for high SNR ($SNR > 4$) the prefilter is disadvantageous. So in this case the jitter increases when the filter bandwidth decreases from $f_c = \infty$ until $f_c = 1 \cdot tx$.

V. CONCLUSIONS

We studied each one of four synchronizers namely the analog (ana), hybrid (hib), combinational (cmb) and

sequential (seq) with a prefilter of four different bandwidths ($PF(s) = \infty$, $PF(s) = 5 \cdot tx$, $PF(s) = 2 \cdot tx$, $PF(s) = 1 \cdot tx$).

Then we tested each setup in order to obtain its output clock jitter (UI) as function of the input signal to noise ratio SNR.

We noted that, for the synchronizers without input limiter (analog) the prefilter is always advantageous (for low and high SNR), so the jitter diminishes when the filter bandwidth decreases.

However for the synchronizers with input limiter (hybrid, combinational and sequential) the jitter depends on the SNR value. So for low SNR ($SNR < 4$) the prefilter is advantageous, then the jitter diminishes when the prefilter bandwidth decreases on the other hand for high SNR ($SNR > 4$) the prefilter is disadvantageous then the jitter increases when the prefilter bandwidth decreases.

In general, we concluded that, when the prefilter bandwidth diminishes the distortion of the signal increases and then its immunity to the noise diminishes. This phenomenon is more accentuated in the synchronizers with input limiter that ignore low noise margin.

This is comprehensible because the limiter has noise margin and therefore low noise is ignored, but high noise obliges the gates to switch on-off provoking output spikes.

VI. ACKNOWLEDGMENTS

The authors are thankful to the program FCT (Foundation for Science and Technology).

VII. REFERENCES

- [1] – Werner Rosenkranz, “Phase Locked Loops with limiter phase detectors in the presence of noise”, IEEE Transactions on Communications com-30, October 1982.
- [2] - Charles R. Hogge, “A Self Correcting Clock Recovery Circuit”, Journal of Lightwave Technology, Dec. 1985.
- [3] – A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho “Phase and Data Synchronizers in the Presence of Noise”, 5th International Conference on Telecommunications p.59, Zagreb 17-18 June 1999.
- [4] – A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho “Effects of the Input Signal Shape on the PLL Jitter”, 7th International Symposium on Recent Advances in Microwave Thecnology 177, Málaga 13-17 December 1999.
- [5] – A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho “A New Technique to Measure the Jitter”, III Conference on Telecommunications p.64, F. Foz-PT 23-24 April 2001.
- [6] - A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho, “Open Loop Symbol Synchronizers”, 9th International Conference on Telecommunications p.286, Beijing-CN 23-26 June 2002.
- [7] - A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho, “Mixed Loop Symbol Synchronizers”, 3th International Symp. on Communication Systems Networks and Digital Signal Processing p.324, Stafford-UK 15-17 July 2002.

Estimating the Position and Velocity of Mobiles With Using GRNN Neural Networks

Mohammad Reza Sabagh , Nasser Sadati

*Intelligent Systems Laboratory, Electrical Engineering Department,
Sharif University of Technology, P.O.Box 11365-9363, Tehran - Iran
email: sabagh_54@yahoo.com*

Abstract-Determining the position and velocity of mobile phone is an important issue for handover control and hierarchical cellular networks are concerned, since the efficient allocation of moblies to macro or micro cells depends on its present velocity. In this paper a new approach for tracing a mobile by evaluating subsequent signal-strength measurements to different base stations. The basic idea resembles multi-dimensional scaling (MDS), a well-recognized method in statistical data analysis. The characterisation of velocity and position in GSM investigation is based on General Regression Neural Network (GRNN), which is known as a network with one-pass learning algorithm with a highly parallel structure. Even with sparse data in a multidimensional measurement space, the algorithm provides smooth transitions from observed value to another.

I. INTRODUCTION

A Rapid growth is expected in the load of cellular radio networks in the near future. Since the number of channels, either physical, Time Division Multiple Access(TDMA) or Code Division Multiple Access(CDMA) is a restricted resource for such systems, increasing demand can only be satisfied by diminishing the transmission power and coverage area of the cells. This, however, increases the number of handovers and the corresponding administration overhead. Hierarchical cell structures seem to be a good compromise between an efficient use of available channels while simultaneously keeping the number of handovers small. Fast moving mobiles are assigned to larger cells and stationary or slow moving stations are allocated to microcells, which are used in areas of high utilization. Of course physical channels must be segregated in overlapping cell areas.

For an efficient channel assignment, information about the position and velocity of mobile stations is inevitable in hierarchical networks. Two quantities are needed to obtain distance and speed information; signal strengths of different base stations measured by a mobile, and corresponding propagation times. Both parameters are subject to changes caused by short term fading, shadowing, and reflections such as a sophisticated method is necessary to translate signal strength into distance information.

In this paper, a method called multi-dimensional scaling (MDS) is used [2]. The necessary data are available in the GSM system, where each 0.48 s the downlink signal levels of six neighbouring base stations are transmitted on a discrete scale from 0-63. for this purpose, the GRNN neural networks based method has been proposed which is capable of estimating the position and velocity of mobiles[9].

II. ESTIMATING THE POSITIONS VIA SIGNAL STRENGTH MEASUREMENTS

To clarify the concept, we first introduce the proposed method under simplified assumptions. Let $Z_i \in \mathfrak{R}^2$ denote the position of base station i and $d_i(x) = \|x - Z_i\|$, the euclidean distance of $x \in \mathfrak{R}^2$ from Z_i . If the average signal strength follows a propagation law of the type $c d^{-\alpha}$, d denoting the distance from the transmitter, c is constant and $\alpha \geq 2$, the attenuation exponent, then the average signal strength $S_i(x)$ of base station i is measured at position x is obtained as :

$$S_i(x) = c [d_i(x)]^{-\alpha} \quad (1)$$

Let $\gamma_i(t)$ denote the measured signal strength of base station i ; $i=1, \dots, n$ at a certain position.

Obviously, $\gamma_i(t)$ is subject to random fluctuations due to short-term Rayleigh and Rice Fading. The transformed values $\gamma_i^\alpha(t)/c = \delta_i(t)$ correspond to the distance from base station i , such that the solution $\hat{\mathbf{x}}(t)$ of

$$\text{minimized } f(\mathbf{x}) = \sum_{i=1}^n [d_i(\mathbf{x}) - \delta_i(t)]^2 \text{ over } \mathbf{x} \in \mathfrak{R}^2 \quad (2)$$

is a least squares estimator of the actual position at time t . positions of the base stations are a priori fixed, and only the position of the mobile has to be fitted to the transformed signal data. A local minimum of (2) can be calculated numerically by a Newton-type iteration

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}_f^{-1}(\mathbf{x}_k) \nabla f(\mathbf{x}_k) \quad k \in \mathbb{N}_0 \quad (3)$$

Where the $\nabla f(\mathbf{x})$ denote the gradient at differentiable point \mathbf{x} given by:

$$\nabla f(\mathbf{x}) = 2 \sum_{i=1}^n \frac{d_i(\mathbf{x}) - \delta_i(t)}{d_i(\mathbf{x})} (\mathbf{x} - \mathbf{Z}_i) \quad (4)$$

and with $\mathbf{x} = (x_1, x_2)'$ and $\mathbf{Z}_i = (u_i, v_i)'$, the Hessian is determined as

$$\mathbf{H}_f(\mathbf{x}) = \begin{bmatrix} f_{11}(\mathbf{x}) & f_{12}(\mathbf{x}) \\ f_{12}(\mathbf{x}) & f_{22}(\mathbf{x}) \end{bmatrix} \quad (5)$$

where

$$f_{11}(\mathbf{x}) = 2 \sum_{i=1}^n \frac{\delta_i(t)(x_1 - u_i)^2}{d_i^3(\mathbf{x})} + \frac{d_i(\mathbf{x}) - \delta_i(t)}{d_i(\mathbf{x})} \quad (6)$$

$$f_{12}(\mathbf{x}) = 2 \sum_{i=1}^n \frac{\delta_i(t)}{d_i^3(\mathbf{x})} (x_1 - u_i)(x_2 - v_i) \quad (7)$$

$$f_{22}(\mathbf{x}) = 2 \sum_{i=1}^n \frac{\delta_i(t)(x_2 - v_i)^2}{d_i^3(\mathbf{x})} + \frac{d_i(\mathbf{x}) - \delta_i(t)}{d_i(\mathbf{x})} \quad (8)$$

Iteration (3) solves $\nabla f(\mathbf{x}) = 0$. It is well known that this algorithm usually get stuck at local minima and does not necessarily find the global minimum. However, a reliable estimation $\hat{\mathbf{x}}(t_i)$ of a mobile position at time t_i is a reasonable starting point for finding the next solution $\hat{\mathbf{x}}(t_{i+1})$ by (3). Applying the newton iteration (3) to each of the 748 data sets, $[\delta_1(t_i), \dots, \delta_6(t_i)]$; $i=1, \dots, 748$, this iteration is stopped at a relative accuracy of $\varepsilon=10^{-4}$.

III. ESTIMATING VELOCITY

Estimated locations at subsequent time points t_i is shown as $\hat{\mathbf{x}}(t_i) = [\hat{x}_1(t_i), \hat{x}_2(t_i)]'$;

$i=1, \dots, m$. For the GSM system, we have $t_i = 0.48i$ second, when normalizing $t_0 = 0$. We restrict our attention to k preceding estimated positions which are labeled as $\hat{\mathbf{x}}(t_1), \dots, \hat{\mathbf{x}}(t_k) \in \mathfrak{R}^2$ for notational convenience. If the movement of the mobile is linear with constant speed vector \mathbf{a} and velocity $\|\mathbf{a}\|$, then the true position at time t_i is given by

$$\mathbf{X}(t_i) = t_i \mathbf{a} + \mathbf{b} \quad (9)$$

where \mathbf{b} is the position at time $t_0 = 0$.

The parameters \mathbf{a} and \mathbf{b} in this linear regression setup are estimated from the observed values $\hat{\mathbf{x}}(t_1), \dots, \hat{\mathbf{x}}(t_k)$ by the solutions $\hat{\mathbf{a}}, \hat{\mathbf{b}}$ of the least square criterion

$$\text{minimize } \sum_{j=1}^k \|\hat{\mathbf{x}}(t_j) - (t_j \mathbf{a} + \mathbf{b})\|^2 \quad \mathbf{a}, \mathbf{b} \in \mathfrak{R}^2 \quad (10)$$

minimization of equation (10) decomposes into

$$\begin{aligned} \min \sum_{j=1}^k [\hat{x}_\lambda(t_j) - (t_j a_\lambda + b_\lambda)]^2 + \\ \min \sum_{j=1}^k [\hat{x}_2(t_j) - (t_j a_2 + b_2)]^2 \\ a_1, b_1 \in \mathfrak{R} \quad a_2, b_2 \in \mathfrak{R} \end{aligned} \quad (11)$$

From standard linear regression in statistics, the solution of either sum is well known to be

$$\hat{\mathbf{a}}_\lambda = \frac{\sum_{j=1}^k (t_j - \bar{t}) [\hat{x}_\lambda(t_j) - \bar{x}_\lambda]}{\sum_{j=1}^k (t_j - \bar{t})^2} \quad (12)$$

$$\hat{\mathbf{b}}_\lambda = \bar{x}_\lambda - \hat{\mathbf{a}}_\lambda \bar{t} \quad , \quad \lambda = 1, 2 \quad (13)$$

where

$$\bar{t} = (1/k) \sum_{j=1}^k t_j \quad , \quad \bar{x}_\lambda = (1/k) \sum_{j=1}^k \hat{x}_\lambda(t_j) \quad (14)$$

k is usually small with values between 5-20. Hence, a solution of (10) is given by (12), (13), (14) through $\hat{\mathbf{a}} = (\hat{a}_1, \hat{a}_2)'$ $\hat{\mathbf{b}} = (\hat{b}_1, \hat{b}_2)'$, the actual position at time t_i is then estimated by the value of the regression line at $t=t_i$, i.e.

$$\hat{\mathbf{x}}(t_i) = t_i \hat{\mathbf{a}}(t_i) + \hat{\mathbf{b}}(t_i) \quad (15)$$

Differentiating $\hat{\mathbf{x}}$ at $t=t_i$ simultaneously yields a smoothed estimation of the instantaneous speed vector at time t_i , namely

$$\hat{S}(t_i) = \hat{a}(t_i) \quad i=1, \dots, m \quad (16)$$

The corresponding velocity is obtained as :

$$\hat{V}(t_i) = \|\hat{a}(t_i)\| = [\hat{a}_1^2(t_i) + \hat{a}_2^2(t_i)]^{1/2} \quad (17)$$

This equations discussed completely in [2],[3],[4],[5],[6],[7],[8].

IV. GENERAL REGRESSION NEURAL NETWORK

GRNN is an artificial neural network that is usually defined as a network composed of a large number of simple processors (neurons) that are massively interconnected. They operate in parallel, and take past measurement for estimation of continuous variables and converges to the underlying (linear or nonlinear) regression surface in to concern [9],[1].

Assume that $f(x,y)$ represents the known continuous joint probability density function of a random vector variable x , and a scalar random variable y . Let x be a particular measured value of the random variable x , the conditional mean of y that gives x (also called the regression of y on x) is given by

$$E[y|x] = \frac{\int_{-\infty}^{\infty} yf(x,y)dy}{\int_{-\infty}^{\infty} f(x,y)dy} = y_e(x) \quad (18)$$

in [9] $\hat{y}(x)$ is obtained as

$$y_e = \hat{y}(x) = \frac{\sum_{i=1}^m A^i \exp(-\frac{C_i}{\delta})}{\sum_{i=1}^m B^i \exp(-\frac{C_i}{\delta})} \quad (19)$$

where

$$\begin{aligned} A^i(k) &= A^i(k-1) + y^i \\ B^i(k) &= B^i(k-1) + 1 \\ C_i &= \sum_{j=1}^p |x_j - x_j^i| \end{aligned}$$

$A^i(k)$ is the sum of the y values, and $B^i(k)$ is the number of samples assigned to cluster i . the value of δ in (19) gives the smallest error that should be used in the final network.

Fig.1 shows the overall block diagram of the GRNN in its adaptive form represented by (19):

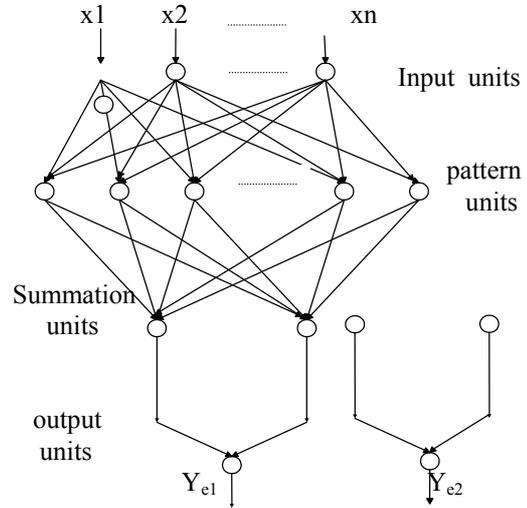


Fig. 1. GRNN block diagram

V. SIMULATION

Simulation is done in two parts, a communication and a neural network part. Communication part is simulated by using the algorithms derived in [2],[3]. The neural network part uses the GRNN model shown in Fig.2.

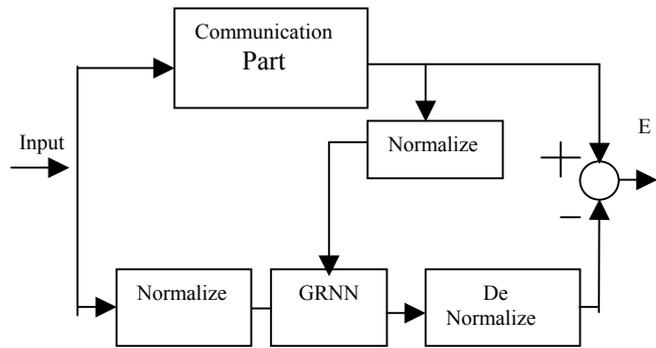


Fig. 2. Model of GRNN for simulation

As a preprocessing step, it is usually necessary to scale all input and output variables in Fig.2 x_n and y_n are normalization values of x and y . For training the GRNN the values of x_n and y_n are used. The Purpose of the training is to minimize the output error (E) as seen in Fig 2. Relationship between y_n and y is given by:

$$y_n = \frac{y_n^{\max} - y_n^{\min}}{y^{\max} - y^{\min}} (y - y^{\min}) + y_n^{\min} \quad (20)$$

$$y = \frac{y_n^{\max} - y_n^{\min}}{y_n^{\max} - y_n^{\min}} (y_n - y_n^{\min}) + y_n^{\min} \quad (21)$$

where

y_n^{\max} , y_n^{\min} , y^{\max} , y^{\min} are the minimum and maximum of Denormalization and Normalization variables. In this paper there are 6 inputs and 3 outputs. Inputs are the received power of the Mobile from six BTS and outputs

are the position and velocity. The position contains the values of X and Y in cartesian coordinates of (X,Y).

VI. SIMULATION RESULTS

Figures 3-6 show the comparison between the results obtained from GRNN and real status. The simulation results are done with different number of training data and neurons. As shown, the GRNN has a very good estimation ability. Even with a little number of training and neurons, it estimates the position and velocity of mobile. It can be seen that the estimation results of GRNN outperform those obtained from the real status.

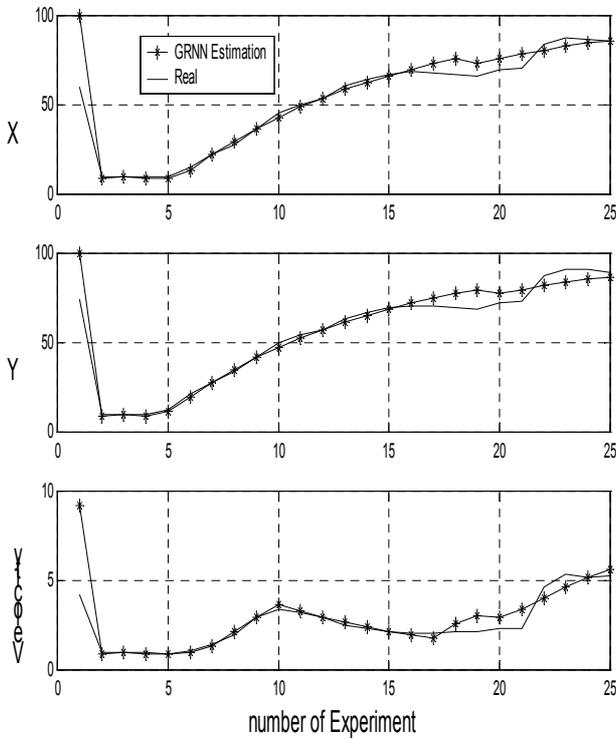


Fig. 3. GRNN Estimation with Number of Training data = 300 and Neurons = 20

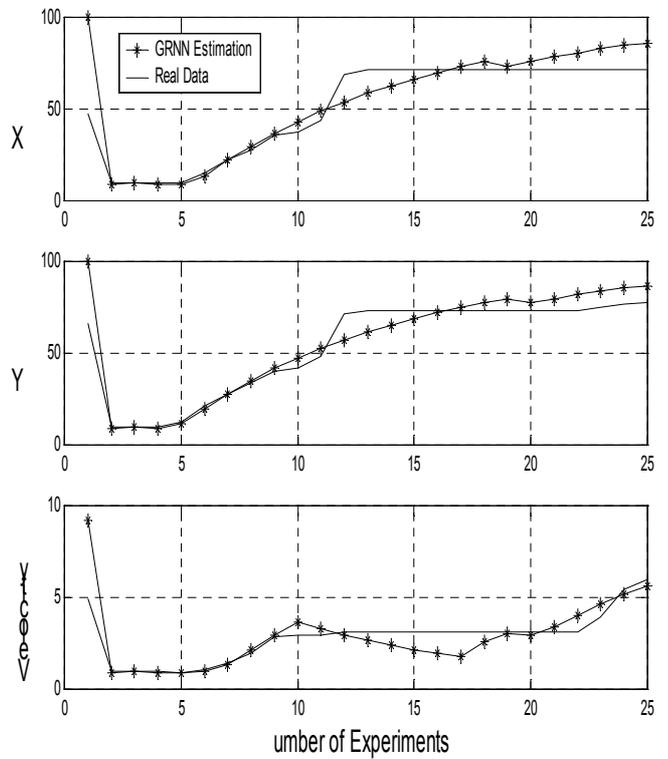


Fig. 4. GRNN Estimation with Number of Training data = 300 and Neurons = 10

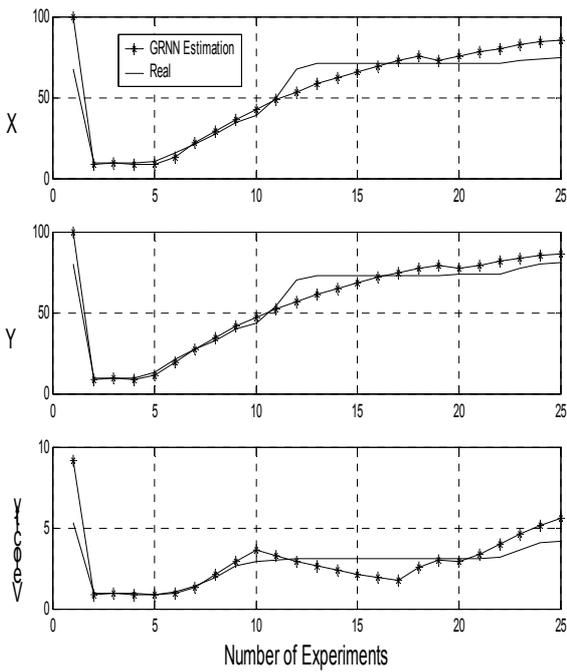


Fig 5. GRNN Estimation with Number of Training data = 200 and Neurons = 20

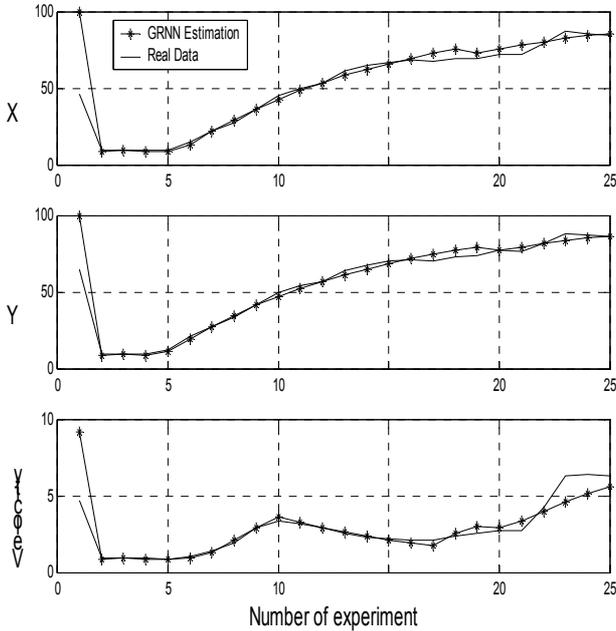


Fig 6. GRNN Estimation with Number of Training data = 200 and Neurons = 10

References:

- [1] P.G. Balis and O.R. Hinton "Characterisation of mobile velocity in GSM using a simple feed-forward neural network" Electronics Letters 17 th september 1998 vol.34 no.19
- [2] HELLEBRANDT,M.,MATHAR,R. and SCHEIBENBOGEN,M. "Estimating position and velocity of mobiles in a cellular radio network" IEEE Trans.,1997,VT-46,(1),pp.65-71
- [3] JAKES,W.C. "Microwave mobile communications" wiley,Newyork,1997
- [4] W.G.Figel, N. H. Shepherd,and W. F. Trammell, "Vehicle location by a signal attenuation method " ,IEEE Trans. Veh.Technol,vol. VT-18,pp. 104-109,nov. 1969
- [5] M.Hata, "Empirical formula for propagation loss in land mobile radio services," IEEE Trans. Veh. Technol., vol. VT-29, pp.317-325,Aug.1980
- [6] M. Hata and T. Nagatsu, "Mobile location using signal strength measurements in a cellular system", IEEE Trans. Veh. Technol., vol. VT-29,pp. 245-252,may 1980
- [7] C.W. Sung and W. S. Wong, "User speed estimation and dynamic channel allocation in hierarchical cellular systems",IEEE 44 th veh. Technol. Conf,Stockholm,sweden,1994,pp.91-95
- [8] K.Kawabata,T.Nakamura,and E.Fukuda, "Estimating velocity using diversity reception", IEEE 44 th veh. Technol.conf. Stockholm,sweden,1994, pp. 371-374
- [9] Donald F. Specht , "A General Regression Neural Network" , IEEE Transaction on Neural Networks vol. 2, No.6 ,november 1991

Expurgated Tangential Bound of Low Density Parity Check Codes

Vu-Duc Ngo, Hae-Wook Choi and Sin-Chong Park

Systems VLSI Lab, SITI Research Center, Information and Communications University, Taejon City, Korea
 Phone: +82-42-866-6910, Fax:+82-42-866-6921
 Email: {duc75,hwchoi,scpark}@icu.ac.kr

Abstract—In this paper, the tightened union bound and tightened tangential bound, let say expurgated tangential bound, with respect to the block error probability for Low Density Parity Check Codes are derived. These work are carried out by eliminating the useless codewords out of the union bound. The elimination of these codewords is done by applying Verdu's theorem and calculating the Hamming weight of the codewords that belong to the irreducible set. These new bounds not only are simpler than old bound techniques in the sense of the complexity of computation but also extend the reliable region of E_b/N_o in which the bounds yield meaningful results.

I. INTRODUCTION

Low density parity check codes (LDPC) were first introduced in 1961 by Gallager [1] and rediscovered by Mackay and Neal [2]. A lot of work were done and the results of experiments showed that like Turbo codes, they demonstrate low BER at low E_b/N_o . The *union bound* is the conventional method that we can use to theoretically evaluate the performance of any codes, unfortunately, it is just tight for the high area of E_b/N_o . The conventional *tangential bound* or *conventional tangential bound* [8] was derived based on the *conventional union bound*, therefore it also presented loose bound for the low area of E_b/N_o .

In this paper, we derive the new tight bounds for LDPC codes based on Verdu's theorem [3], this theorem introduces the method to refine the *union bound* as well as the *union tangential bound* [8,9] to make them tighter. Moreover, the computations of the new tight bounds are done over the *irreducible set* (will be introduced later) that is the subset of the codeset, our proposed bounds can be archived with less complexity of computation in comparison with the conventional bounds. The new bounds yield the meaningful results in the area of low E_b/N_o as well.

The paper is organized as follows: in Section 2, we reintroduce the ensemble of LDPC codes briefly adhering to Gallager in [1], and briefly present the upper bound on the ensemble distance spectrum of these LDPC codes. In Section 3, we introduce the way to tighten the *union bound* for Linear Block Code especially for an ensemble of LDPC codes based on Verdu's theorem [3]. In Section 4, the tightened Union Tangential Bound is presented. Our conclusions are presented in section 5.

II. DISTANCE SPECTRUM OF LDPC

The ensembles of (n, j, k) LDPC codes are defined by Gallager in [1] and built by a sparse parity check matrix, containing mostly 0's and only a relatively small number of 1's. Considering a (n, j, k) parity check matrix as a matrix of n columns that has the number of 1's as j ($j \geq 3$) and k ($k > j$) in each column and each row, respectively, otherwise they are 0's. Therefore, the parity check matrix has (nj/k) rows.

Let $\overline{N(l)}$ be the ensemble average of binary codeword of Hamming weight l ($0 \leq l \leq n$). Gallager [1] derived an upper bound on $\overline{N(l)}$ as follows:

$$\overline{N(l)} \leq \binom{n}{l}^{1-j} \times \overline{N_1(l)}, \quad (1)$$

where $\overline{N_1(l)}$ is the average number of codewords weight l that satisfies any of the j subblocks of n/j parity checks. By applying the Stirling approximation, the final equation can be represented as

$$\overline{N(l)} \leq C(\lambda, n) \exp(-nB(\lambda)), \quad (2)$$

where $\lambda = l/n$ and

$$C(\lambda, n) = \left[2n\pi\lambda(1-\lambda) \right]^{\left(\frac{j-1}{2}\right)} \times \exp\left[\frac{j-1}{12n\lambda(1-\lambda)} \right], \quad (3)$$

$$B(\lambda) = (j-1)H(\lambda) - \frac{j}{k} \left[\mu(s) + (k-1) \ln 2 \right] + js\lambda. \quad (4)$$

To optimize the upper bound on $\overline{N_1(l)}$ in equation (2), s need to be an arbitrary variable that satisfies the following equation

$$\lambda = \frac{\partial \left[\ln \left[2^{-k} \left((1+e^s)^k + (1-e^s)^k \right) \right] \right]}{k\partial s}. \quad (5)$$

The function $H(\lambda)$ is the entropy function defined by

$$H(\lambda) = -\lambda \ln \lambda - (1-\lambda) \ln (1-\lambda). \quad (6)$$

The upper bound in equation (2) can be considered as an upper bound on ensemble average weight distribution of codes, belonging to the Gallager ensemble. The weight

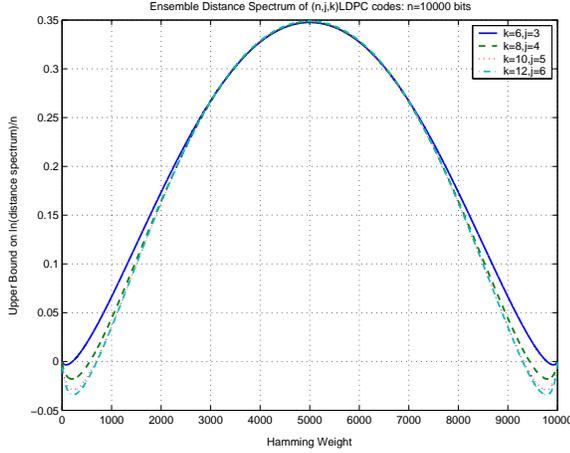


Fig. 1. Distance spectrums.

distribution $\overline{N(l)}$, so called the distance spectrum of (n, j, k) LDPC codes, is shown in Fig. 1 for some ensembles with rate near 1/2 and a block length of 10000 coded bits (We use similar parameters with Sason and et al. [5] for further comparisons). This figure also shows the improvement in the ensemble performance of (n, j, k) LDPC codes for fixed block length, archived by increasing the values of k and j while keeping the rate of j/k constant.

III. TIGHT UNION BOUND APPROACH

A. Union bound

Let us define the codeset of one linear code as $\mathbf{C} = \{c_0, c_1, \dots, c_{n-1}\}$, and suppose that the all zero codeword is transmitted over the AWGN channel. Carrying out the ML decoding method, the union bound of block error probability can be defined as the following upper bound

$$Pe = \bigcup_{j=1}^{n-1} Pr(c_j|c_0) \leq \sum_{c_i \in \mathbf{C}} A_j Q\left(\frac{\delta_j^E}{2\sigma}\right), \quad (7)$$

where A_j are the number of codewords at Euclidean distance δ_j^E , $j = 1, 2, \dots, n-1$, from the transmitted (all zero) codeword.

B. Verdu's theorem

In the inequality (7), the union bound is bounded by using the sum of probabilities of all the individual error events instead of a union of all the probability of error events.

If one of the pairwise error region (PER) - let say - $\chi_j = \{y : |y - c_j| < |y - c_0|\}$, where y denotes the received vector, is included in the union of the remaining PERs then it can be excluded from the union as well as the sum in the right hand side of the inequality (7). Therefore, in principle, we can exclude all the terms that do not correspond to the hyperplanes that form the faces of the Voronoi region of the transmitted codeword c_0 . The method of eliminating these

terms from the union bound was derived by Verdu [3]. In this paper, Verdu showed how these terms can be excluded from the union bound in the case of binary antipodal transmission over the Gaussian channel with intersymbol interference.

A generalization of the Verdu's theorem was presented by Biglieri et al. [6], it shows a sufficient way for excluding a given error codeword c_j from the union bound. The generalization can be summarized as follow.

Given codeset $\mathbf{C} = \{c_0, c_1, \dots, c_{n-1}\}$ of a linear code corresponds to the set of input messages $\mathbf{U} = \{u_0, u_1, \dots, u_{n-1}\}$, if there exists the irreducible set belongs to \mathbf{C} , denoted by $\Omega = \{\xi_0, \xi_1, \dots, \xi_{s-1}\}$ and $\Omega \subset \mathbf{C}$, where s stands for the size of this set, and any $c_k \in \mathbf{C}$ that satisfies the conditions

$$\begin{cases} c_k \notin \mathbf{C} \\ c_k = c_0 + \alpha_j \sum_{j=1}^{s-1} \xi_j, \quad \alpha_j = 0, 1; \quad \xi_j \in \Omega \\ c_0 : \text{transmitted codeword (all zero codeword)} \end{cases} \quad (8)$$

This means that any codeword c_k can be represented as the sum of some elements of Ω . Moreover, since Ω is defined as the irreducible set, its all elements can not be decomposed further. Therefore, the pairwise error probability $Pr(c_k \rightarrow c_0)$ can be expurgated from the union bound defined in the equation (7). Consequently, it can be tightened as

$$Pe \leq \sum_{j=1}^{s-1} Pr(c_0 \rightarrow \xi_j) = \sum_{\xi_j \in \Omega} B_j Q\left(\frac{\delta_j^0}{2\sigma}\right), \quad (9)$$

where B_j is the number of codewords of Ω that has the Euclidean distance equal δ_j^0 from all zero codeword. Since Ω is the subset of \mathbf{C} then it is clear that $\sum_{c_i \in \mathbf{C}} A_j Q\left(\frac{\delta_j^E}{2\sigma}\right) > \sum_{\xi_j \in \Omega} B_j Q\left(\frac{\delta_j^0}{2\sigma}\right)$. The question is how to work out the irreducible set $\Omega = \{\xi_0, \xi_1, \dots, \xi_{s-1}\}$ in terms of size as well as its cardinality. The following part will address this problem.

C. Tighten union bound for LDPC

In [10], we proved that for a LDPC code with parity check matrix H defined by (n, j, k) , there exists the irreducible set $\Omega = \{\xi_0, \xi_1, \dots, \xi_{s-1}\}$. Furthermore, for any $\xi_j \in \Omega$, we have its Hamming weight defined by $w(\xi_j)$ and $w(\xi_j) \leq (n \times j + k)/k$. Consequently, if BPSK is used as modulation scheme for transmitted signal over the AWGN channel, we can tighten the union bound with respect to block error probability in equation (7) of LDPC as the following inequality

$$Pe \leq \sum_{l=1}^{(n \times j + k)/k} \overline{N(l)} \times Q\left(\sqrt{l \frac{2E_b}{N_0}}\right), \quad (10)$$

where $\overline{N(l)}$ is calculated in (2).

Using the results of upper bound on the ensemble distance spectrum of (n, j, k) LDPC codes for some cases with rate near 1/2 and a block length of 1000 coded bits calculated in Section 1, we can calculate the tight bound based on equation

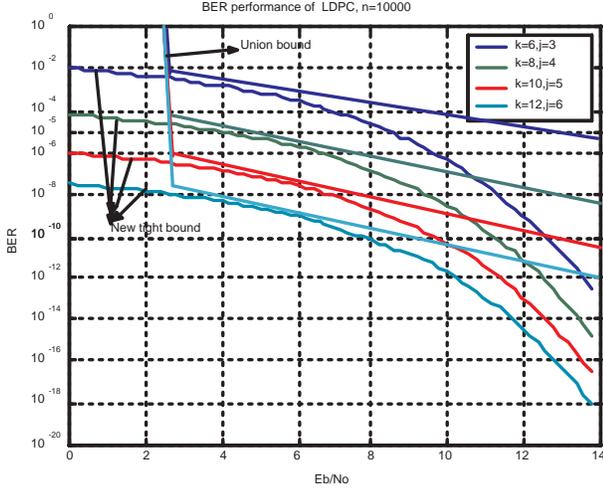


Fig. 2. Union bound and Tightened union bound.

(10) as the Fig. 2.

The above figure shows that the improvement of the BER performance comes along with the improvement of the ensemble performance when we increase the values of j and k while keep the rate of j/k constant. It is clear that our archived bound on LDPC is tighter than the union bound calculated in [5], and in the area of low E_b/N_0 this bound has meaningful results due to the heavy weight codewords are excluded. We name this tight bound as *tightened union bound* (Tightened UB).

IV. TIGHTEN TANGENTIAL BOUND OF LDPC

The union tangential bound was firstly proposed by Berlekamp [8] and lately represented by Herzberg and Poltyrev [9]. These authors proved that the union tangential bound is tighter than the union bound. Since the union tangential bound is still calculated with full cardinality of the codeset or full distance spectrum of the code without any refined work, we can apply the similar manner which applied in the previous approach of tightening the union bound to tighten it for LDPC codes.

First of all, according to Herzberg and Poltyrev [10], the union tangential is illustrated by the Fig. 3. Suppose the noise is decomposed into two components z_1 and z_2 , they are independent and have the normal distribution as

$$f_z(z_i) = \frac{1}{\sqrt{2\pi}\sigma} \times \exp(-z_i^2/2\sigma^2), \quad i = 1, 2. \quad (11)$$

Let us define C_i is the transmitted codeword with the power equal n and its neighbor is C_j , the Euclidean distance between them is δ_{ij}^E . Therefore, the radius of the sphere of the transmitted codeword is \sqrt{n} (normalized). Let the event of error

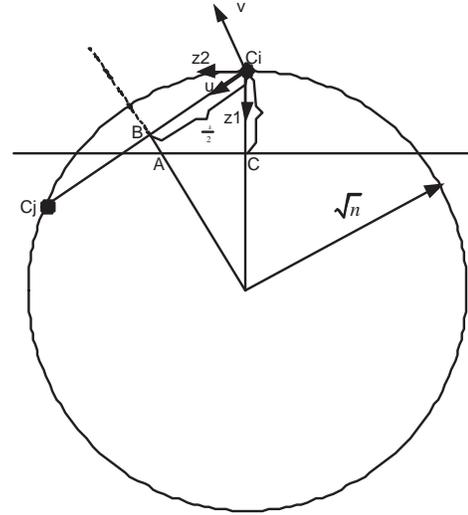


Fig. 3. Tangential bound illustration.

at the output of the decoder is denoted by \mathbf{E} . The block error probability, Pe , can be written as follows

$$\begin{aligned} Pe &\leq \min_j \left(Pr(\mathbf{E}|z_1 \leq \gamma) Pr(z_1 \leq \gamma) + Pr(z_1 > \gamma) \right) \\ &= \min_j \left(Pr(\mathbf{E}, z_1 \leq \gamma) + Pr(z_1 > \gamma) \right) \end{aligned} \quad (12)$$

We define the \mathbf{E}_{ij} is the error event that the received vector is closer to the codeword C_j than the transmitted C_i caused by AWGN. By geometrical calculation, we obtain $\overline{AC} = \frac{\sqrt{n}-z_1}{\sqrt{n-(\delta_{ij}^E/2)^2}}$ and $\overline{BA} = \frac{\sqrt{n}\gamma - (\delta_{ij}^E/2)^2}{\sqrt{n-(\delta_{ij}^E/2)^2}}$, it follows that

$$Pr(\mathbf{E}_{ij}, z_1 \leq \gamma) = Pr\left(u \geq \delta_{ij}^E/2, v \geq -\frac{\sqrt{n}\gamma - (\delta_{ij}^E/2)^2}{\sqrt{n - (\delta_{ij}^E/2)^2}}\right), \quad (13)$$

where u and v are two other directions shown in Fig. 3. By taking projection of the noise vector on two directions, the union tangential bound of block error probability can be simplified as

$$Pe \leq \sum_{k=1}^n A_k Q\left(\frac{\delta_k^E}{2\sigma}\right) Q\left(-\frac{\sqrt{n}\gamma_0 - (\delta_k^E/2)^2}{\sigma\sqrt{n - (\delta_k^E/2)^2}}\right) + Q\left(\frac{\gamma_0}{2\sigma}\right) \quad (14)$$

where A_k is the average number of codewords at distance δ_k^E from a transmitted codeword (all zero codeword), n is the dimension of code's ensemble. The optimum value $\gamma = \gamma_0$ that makes the bound tightest can be found by solving the following equation

$$f(\gamma) = \sum_{k=1}^n A_k \times Q\left(\frac{(\sqrt{n}-\gamma)\delta_k^E}{2\sigma\sqrt{n - (\delta_k^E/2)^2}}\right) - 1 = 0. \quad (15)$$

Now for LDPC codes, we can apply the similar way that was used in equation (10) to tighten the union bound to refine the union tangential bound to make it tighter. Assume that the all

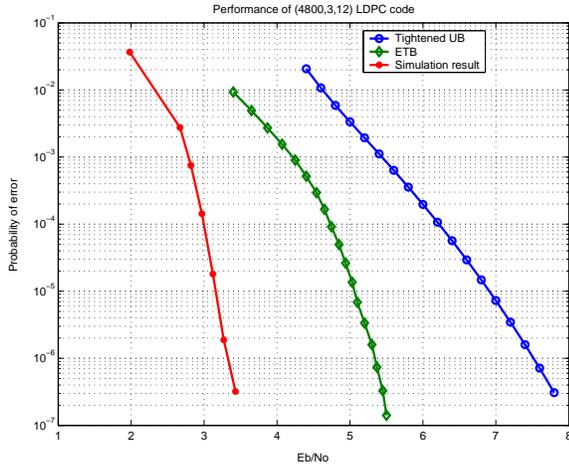


Fig. 4. Simulation results and theoretical bounds.

zero codeword of an (n, j, k) LDPC code is transmitted, the union tangential bound can be tightened as the final following inequality

$$Pe \leq \sum_{l=1}^{(n \times j + k)/k} \frac{N(l)}{N} Q\left(\frac{\delta_l^E}{2\sigma}\right) Q\left(-\frac{\sqrt{n}\gamma_0 - (\delta_l^E/2)^2}{\sigma\sqrt{n - (\delta_l^E/2)^2}}\right) + Q\left(\frac{\gamma_0}{2\sigma}\right) \quad (16)$$

The above bound can be called as *expurgated tangential bound* (ETB). The numerical results of *tightened union bound* (Tightened UB) and ETB bound of a $(4800, 3, 12)$ LDPC code are depicted on the following Fig. 4. This figure also shows the simulation results of the above code with 10 iterations using sum product algorithm.

Since the *expurgated tangential bound* is closer to the simulation result, it follows that we can use this bound to evaluate the performance of LDPC codes instead of the conventional union bound.

V. CONCLUSION

In this paper, we studied the BER performance of ML decoded LDPC codes using union bound for binary input AWGN. We also studied the Gallger's upper bound on the ensemble distance spectrum of LDPC codes with their influent parameters [1, 4, 5] (see Fig. 1). The significant step is that we based on the theorem derived by Verdu [3] as well as figured out the irreducible set of LDPC codes to calculate the new tight bounds. These bounds are tighter than union bound and union tangential bound respectively as well as simpler in calculation in terms of the complexity.

REFERENCES

[1] R. G. Gallager, Low density parity check codes, M.I.T Press, Cambridge, Massachusetts, 1963.

[2] D.J.C. Mackay and R.M. Neal, "Near Shannon limit performance of low density parity check codes," *Electron. Letter*, Vol. 32, pp. 1645-1646, August 1996.

[3] S. Verdu, "Maximum likelihood sequence detection for intersymbol interference channel : A new upper bound on error probability," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 1, pp. 62-68, January. 1987.

[4] I. Sason and S. Shamai, "On improved upper bounds on the decoding error probability of block codes over interleaved fading channels, with applications to turbo-like codes" , *IEEE Trans. Information Theory*, vol. 47, no. 6, pp. 2275 - 2299, September 2000.

[5] I. Sason and S. Shamai, "Tangential sphere bounds on the ensemble performance of ML decoded low density parity check codes" , Proceedings of the Third ITG Conference on Source and Channel Coding, pp. 161 - 166, Munich, Germany, January 17 - 19, 2000.

[6] E. Biglieri, G. Caire, and G. Taricco, "Expurgating the union bound to error probability: A generalization of the Verdu-Shields theorem". ISIT 97, Ulm, Germany, p. 373, June 1997

[7] T.-Y. Hwang, "Decoding linear block codes for minimizing word error rate," *IEEE Trans. Information Theory*, vol. IT-25, pp. 733-737, Nov. 1979.

[8] E. R. Berkelamp, "The technology of error-correction codes", Proc. IEEE, vol. 68, pp. 546-593, May 1994.

[9] Herzberg. H, Poltyrev. G, "The error probability of M-ary PSK block coded modulation schemes" *Communications, IEEE Transactions on* , Volume: 44 Issue: 4 , April 1996 , Page(s): 427 -433

[10] Vu-Duc Ngo and Sin-Chong Park, "Tightening Union Bound by Applying Verdu Theorem for LDPC," PIMRC'2003 Beijing, China, The 14th IEEE International Symposium on, Vol. 1, pp 420-423, Sept. 2003.

Iterative turbo decoder design with convolutional interleavers

Sina Vafi, Tadeusz Wysocki

University of Wollongong, Northfields Ave, 2500, Wollongong, Australia
{sv39,wysocki}@uow.edu.au

Abstract

This paper verifies performance of iterative turbo decoding with a convolutional interleaver instead of a block interleaver. Two interleaver structures are proposed that differ with the number of inserted stuff bits introduced at the end of each block. For each of these structures, the relevant modifications to the convolutional iterative decoder operation are presented. Based on the conducted simulations and comparison with the previously introduced block interleavers, the suitable interleaver for arbitrary data stream can be selected.

1 Introduction

Interleaving is known as an essential factor influencing performance of turbo codes. In most of designs, a turbo code is implemented as a block code when one block interleaver is used. Another interleaver family, are non-block interleavers, such as convolutional interleavers that have comparable delay with block interleavers, and a simplified implementation [1]. In the case of convolutional interleavers, unlike block interleavers, one input data bit affects the interleaver more than once. Therefore, it is necessary to use continuous decoding methods for such a turbo code.[2] [3]

In order to consider a turbo code with the convolutional interleaver as a block code, it is vital to return the interleaver memories to the known state by inserting stuff bits at the end of each block. Block code-word is then created, which allows us to implement a conventional iterative decoder that is known as a sub-optimum turbo decoder. Considering the interleaver specifications and length of input data stream, the number of inserted stuff bits is altered, which affects the performance of turbo codes.

In this paper, the performance of turbo codes with the convolutional interleaver of different length is verified, when the conventional iterative decoding is applied. Furthermore, because of the conducted optimisation of the interleaver,[4], that reduces stuff bits numbers, modification to the iterative decoder is pro-

posed. For different interleaver lengths, the simulation results of the proposed interleavers have been compared with the previously published results for square block interleavers, which are known as the suitable interleavers for a block data length of less than 1000 bits and then the suitable convolutional interleaver characteristics is selected. The results indicate that the new interleaver with inserting acceptable stuff bits to the input data produces better performance of turbo codes.

The paper is organized as follows: Section 2 briefly explains the structure of convolutional interleavers as block interleavers. In section 3, the structure of iterative turbo decoder using the convolutional interleavers is introduced. In section 4, based on the obtained simulation results from different interleavers, the best structure in terms of turbo codes performances is selected. Finally, section 5 gives summary and conclusions.

2 Convolutional Interleaver Structure

A convolutional interleaver consists of T parallel lines of delay elements. In general, each successive interleaver line has M more delay elements than the previous line. Figure 1 shows this structure with $M=1$. A zero stuff bits insertion to the interleaver memories at the end of each data block clears the memory contents of the interleaver and the interleaved data block is obtained. Depending on the length of input data stream, the interleaved data will be terminated at one of the interleaver lines, which can be determined by the result of $\text{Rem}(L,T)$, where $\text{Rem}(L,T)$ gives the remainder of L/T operation. For input data stream $\{x_0, x_1, x_2, \dots, x_{L-1}\}$ with length L , different interleaved data blocks for the interleaver with $T=3, M=1$ value would be :

$$\begin{aligned} \text{Rem}(L,T)=0: & \{x_0, 0, 0, x_3, x_1, 0, x_6, \dots, x_{L-7}, 0, x_{L-2}, x_{L-4}, 0, 0, x_{L-1}\} \\ \text{Rem}(L,T)=1: & \{x_0, 0, 0, x_3, x_1, 0, x_6, x_4, \dots, x_{L-3}, x_{L-5}, 0, 0, x_{L-2}\} \end{aligned}$$

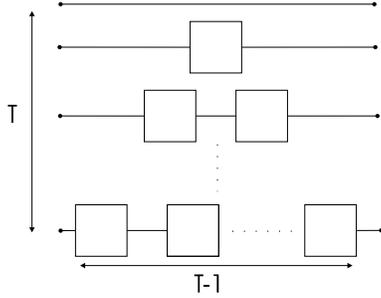


Figure 1: General Convolutional interleaver structure with $M=1$.

$\text{Rem}(L,T)=2:$

$$\{x_0, 0, 0, x_3, x_1, 0, x_6, \dots, x_{L-4}, x_{L-6}, 0, x_{L-1}, x_{L-3}\}$$

Since inserting stuff bits reduces usage of channel bandwidth, an optimisation is performed to control number of stuff bits. It can be achieved by adding a zero deletion block after interleaving which deletes extra zero stuff bits that are inserted at the end of each block. Considering the interleaver with the mentioned parameters, in this example, the interleaved data blocks after zero bit deletion would be :

$\text{Rem}(L,T)=0:$

$$\{x_0, 0, 0, x_3, x_1, 0, x_6, \dots, x_{L-2}, x_{L-4}, x_{L-1}\}$$

$\text{Rem}(L,T)=1:$

$$\{x_0, 0, 0, x_3, x_1, 0, x_6, x_4, \dots, x_{L-3}, x_{L-5}, x_{L-2}\}$$

$\text{Rem}(L,T)=2:$

$$\{x_0, 0, 0, x_3, x_1, 0, x_6, \dots, x_{L-4}, x_{L-6}, x_{L-1}, x_{L-3}\}$$

In this paper, trellis termination and truncation are applied for the first and the second Recursive Systematic Convolutional (RSC) encoders of a turbo encoder, respectively. Therefore, when stuff bits are inserted to the convolutional interleaver after the trellis termination, they have no effect on the systematic and the first parity data. Hence, these bits can be eliminated from the end part of the mentioned data. Based on the outlined characteristics, we can apply the conventional iterative turbo decoder using the new convolutional interleavers.

3 Iterative turbo decoder structure for Convolutional Interleavers

In turbo encoder with the convolutional interleaver, the second encoded parity data have a different length from the encoded systematic and the first parity data, due to stuff bits inserted into the interleaver. Hence, on the basis of the presented encoding method, the relevant decoder is developed. Figure 2 shows the structure of the conventional iterative de-

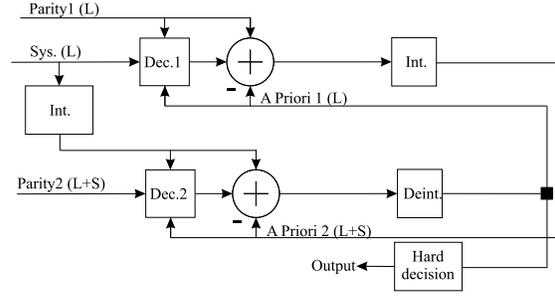


Figure 2: Iterative turbo decoder structure with non stuff bits deletion for convolutional interleavers.

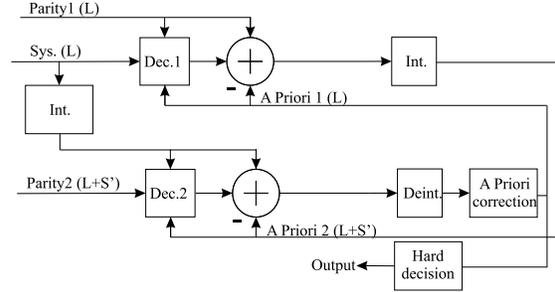


Figure 3: Iterative turbo decoder structure for optimised convolutional interleavers.

coder for convolutional interleaver with non stuff bits deletion together with information length in different positions of the decoder, which are shown in parentheses. In the first component decoder where the systematic and the first parity information are used, the extrinsic information length is L . When the systematic and the extrinsic information pass through the interleaver, their lengths equal to the total number of the interleaver memories, i.e. S bits, will be increased. Therefore, the a priori information length for the second decoder would be $L+S$, which is equal to the length of the second encoded parity information. By passing the obtained extrinsic information through the deinterleaver, which inverts the action of the interleaver, the a priori information length for the first component decoder will be changed to L .

The proposed decoder can be used for the turbo codes with the optimised convolutional interleaver. In this case, interleaver adds S' stuff bits ($S' < S$) equal to $\frac{T(T-1)}{2}$ value to the systematic and extrinsic information length of the first decoder, whose length will be equal to the length of the second parity. Furthermore, a priori correction should be performed after deinterleaving in order to remove the added stuff bits from the extrinsic information and generate corrected information with length of L , compatible with the conducted optimisation at the interleaver. Figure 3 shows the structure of the modified decoder.

4 Simulation results

In the simulations, turbo codes with the specifications of $(m=2, 1, 7/5)$, $(m=4, 1, 35/23)$, where m represents number of the RSC encoder memories, are employed and only the first RSC encoder is terminated to the zero states. The performance of each turbo code has been verified for the optimised and non stuff bits deletion interleavers with different $\text{Rem}(L, T)$ values. We implemented Soft Output Viterbi (SOVA) Algorithm, compatible with the structure presented in [5] as the iterative turbo decoding method, which is of more interest than other methods due to lower complexity in practical design. In this paper, we did not follow any improvement on this algorithm to make similar performance as the Maximum A Posteriori (MAP) algorithm or its modifications. Hence, in order to evaluate performance of turbo codes with the new interleavers, the results have been compared with the previously published square block interleaver results where SOVA was employed [5] [6]. 8 iterations and Additive White Gaussian Noise (AWGN) have been considered in each case.

Under these assumptions, simulation has been performed for the full and half rates turbo codes with the interleaver length $L=169$, and period $T=8$, which results have been illustrated in figures 4 and 5 respectively.

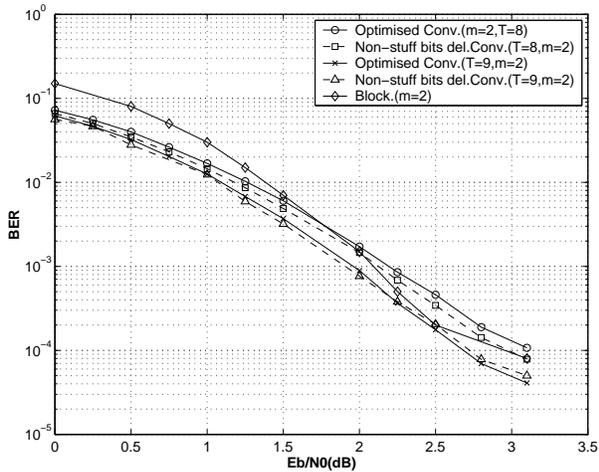


Figure 4: Performance of full rate turbo codes with interleaver periods $T=8, T=9$ and length $L=169$.

For lower signal to noise ratios, application of this interleaver has improved turbo codes performance in comparison with the block interleaver, while in higher signal to noise ratios, where error floor occurs, it shows undesirable behaviour. In order to remove this drawback, convolutional interleavers with higher periods $T=9$ and $T=10$ have been employed. The relevant simulations have been shown in Figures 4 to 7. In the considered example, the obtained results from interleaver with period $T=10$ have the closest perfor-

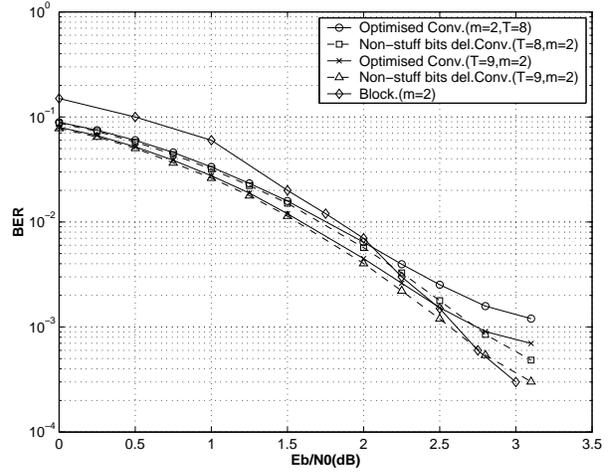


Figure 5: Performance of half rate turbo codes with interleaver periods $T=8, T=9$ and length $L=169$.

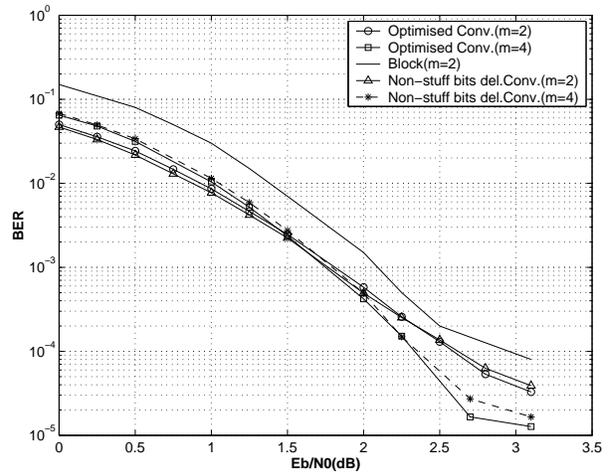


Figure 6: Performance of full rate turbo codes with interleaver $T=10$ and length $L=169$.

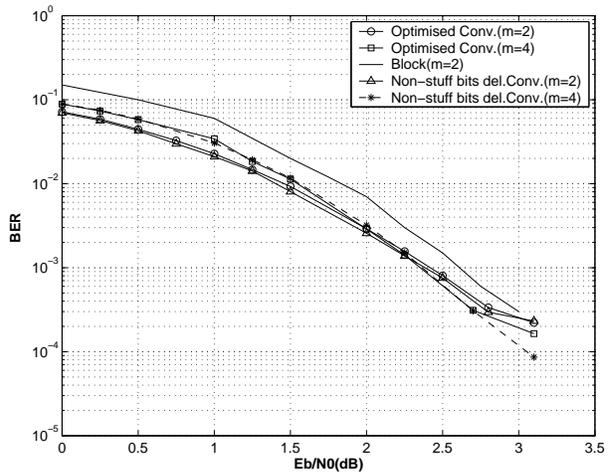


Figure 7: Performance of half rate turbo codes with interleaver $T=10$ and length $L=169$.

mance to the square interleaver (13×13). It shows that inserting the redundancy equal to 9 percent of the whole data block improves the code performance for the full and half rate of about 0.25 dB and 0.3-0.5 dB, respectively. Furthermore, this interleaver applies 45 delay elements which is less delay than 169 in a compared block interleaver and results in better performance of the turbo code.

For the full rate turbo codes with convolutional interleaver, as seen in Figures 4 and 6, with increasing the interleaver period, their performance is better when the optimised interleaver is applied. This is especially evident in the error floor region and is an expected result, since our calculation of a weight distribution for the codes with optimised interleaver [7] indicate that it produces lower multiplicities for free distance value and other low weights than in the case of other interleavers.

More verifications have been achieved in Figure 8 by implementation of the convolutional interleaver periods $T=15$ and $T=20$ with the length $L=1024$. Again, both convolutional interleavers have close performance. However, for the half rate turbo codes with the interleaver $T=15$ similar to obtained graphs from Figures 5 and 7, we can observe different behavior, with the turbo code using the optimised interleaver exhibiting slightly worth performance. It is most likely due to the presence of free distance value with higher multiplicities for the optimised interleaver in this case. The given results show that increasing the number of stuff bits from 105 to 190 bits for the optimised interleaver has improved the turbo codes performance for about 0.3-0.4 dB. In this example, application of the optimised interleaver with the period $T=20$, gives better performance than the non stuff bits deletion interleaver, while achieves an acceptable stuff bits ratio equal to 6 percent of the whole block of data.

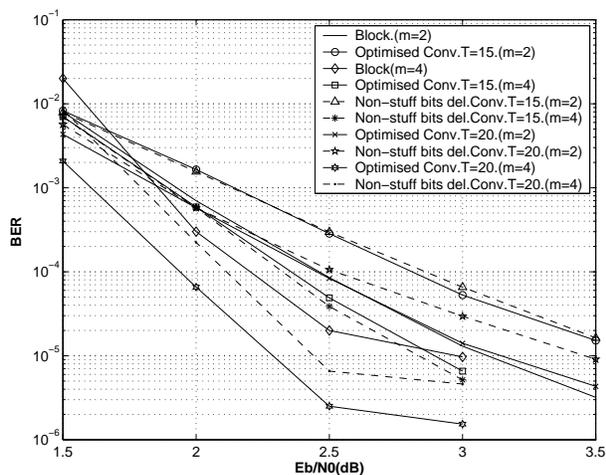


Figure 8: Performance of half rate turbo codes with the interleaver length $L=1024$.

In comparison with the obtained results for the interleaver of length $L=169$, we can conclude that the optimised interleaver is more reliable in long data blocks, which use less stuff bits. However, as mentioned before, inserting stuff bits reduces usage of channel bandwidth and can deteriorate performance of turbo codes for some applications that are sensitive to noise. Therefore, compromise should be performed in determining the code performance and the stuff bits number, which will be followed in the future works.

5 Conclusions

In this paper, the new structure of the convolutional interleaver as a block interleaver has been presented. Based on the conducted simulations, performance of this interleaver and its optimised structure in different turbo codes has been verified and compared with the previously published results. The results show that using the optimised interleaver creates better performance with less delay than when block interleavers are applied. More optimization will be followed in the future work to obtain the optimum interleaver in case of stuff bits reduction.

References

- [1] G.D.Forney, "Burst-correcting codes for the classic bursty channel," *IEEE Trans. Commun.*, vol. COM-19, pp. 772-781, Oct. 1971.
- [2] E.K.Hall and G.Wilson, "Stream-oriented turbo codes," *IEEE Trans.Inform.Theory.*, vol. 47, no. 5, pp. 1813-1831, July. 2001.
- [3] B.Bai, X.Ma, and X.Wang, "Novel algorithm for continuous decoding of turbo codes," *IEE proc. Commun.*, vol. 146, no. 5, pp. 271-274, Oct. 1999.
- [4] S. Vafi, T. Wysocki, and I. Burnett, "Convolutional interleaver for unequal error protection of turbo codes," *7th international DSPCS and 2nd WITSP, Australia.*, pp. 485-491, Dec. 2003.
- [5] J. Hagenauer, P.Robertson, and L.Papke, "Iterative(turbo) decoding of systematic convolutional codes with the map and sova algorithms," *1994 ITG conference on source and channel coding, Munich, Germany*, pp. 21-29, Oct. 1994.
- [6] L. Hanzo, T.H.Liew, and B.L.Yeap, *Turbo coding, turbo equalisation, and space-time coding : for transmission over fading channels*, J. Wiley, 2002.
- [7] S.Vafi and T.Wysocki, "Computation of the free distance and low weight distribution of turbo codes with convolutional interleavers," *Accepted for the 15th PIMRC conference*, 2004.

Low Complexity Residual Phase Tracking Algorithm for OFDM-based WLAN Systems

Suvra S. Das[†], Ratnam V. Rajakumar[§], Muhammad I. Rahman[†], Arpan Pal[‡],
Frank H.P. Fitzek[†], Ole Olsen[†], Ramjee Prasad[†]

[†] Department of Communications Technology, Aalborg University, e-mail: ssd@kom.auc.dk

[‡] Tata Consultancy Services, [§] IIT Kharagpur, India

Abstract— This paper presents the design of an efficient low complexity residual phase tracking algorithm for Orthogonal Frequency Division Multiplexing (OFDM) based Wireless Local Area Network receivers. In this paper the focus is on mitigation of the residual carrier frequency synchronization offset and sampling frequency offset. We propose a novel algorithm that uses piecewise linear approximation to estimate the complex exponential of the phase angle at pilot locations instead of estimating the actual phase angle by highly complex (costly) angle computation and search functions. This helps in reducing the implementation cost of an OFDM receiver. By means of analysis and simulation we show that our design combines both high performance and low complexity.

I. INTRODUCTION

Wireless Local Area Networks (WLAN) are becoming part of omnipresent communication infrastructures. WLANs are being applied in hotels, airports, cafes and various other locations. Different types of terminals are already (or planned to be) equipped with WLAN such as laptops, PDAs, and mobile phones. After the first phase where WLANs were penetrating the consumer market, now we are in the next phase supporting enhanced quality of services (QoS). First systems were based on direct sequence or frequency hopper spread spectrum technology. Currently Orthogonal Frequency Division Multiplexing (OFDM) is the chosen technology for enhanced and future high data rate WLAN systems.

OFDM is a key technology to mitigate the multi-path effect of the wireless channel. It has been already used in DVB-T, IEEE 802.11a and 16a. It is being considered for IEEE 802.20 and IEEE 802.11n as well. OFDM will remain as the key enabling technology for achieving higher data rates in wireless packet based communication in next few years to come [1]. Targeting the mass market of wireless modules low cost solution has to be found, having in mind the tradeoff between efficiency and price. Cost of an OFDM receiver largely depends on the implementation complexity of the synchronization and channel estimation algorithms. High accuracy is needed in synchronization since coherent demodulation of OFDM is extremely sensitive to such errors. Carrier and sampling frequency acquisition and maintenance with high accuracy is vital for successful transmission of long packets. Residual Carrier Frequency Offset (CFO) and Sampling Frequency Offset (SFO) tracking (*phase tracking Figure 1*) are thus very critical part of OFDM receivers. But, the tracking module is highly complex [2], [3], and thus has significant potential for cost optimization of OFDM receivers.

The synchronization impairments that an OFDM receiver has to mitigate are Frame Timing Offset (FTO) or Symbol Timing Offset (STO), taken care of in the *Time Synch*, CFO compensated by the *Frequency Synch* block of Figure 1. Such algorithms are very well discussed in several literatures [4], [5] and many more. In this article, we deal with the residual CFO and SFO errors, jointly termed as *Residual Phase Errors*. *Residual Phase Error* is the combined error due to non-exact carrier frequency offset correction and existing sampling frequency offset. In a real environment the synchronization

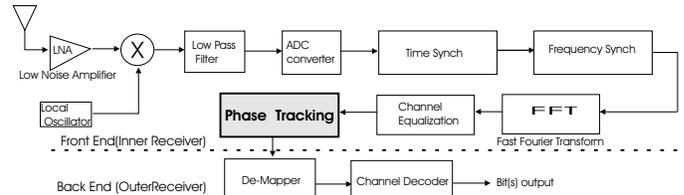


Fig. 1. OFDM receiver front end architecture

blocks placed at the receiver front end are not able to estimate the exact carrier frequency offset due to circuitry noise and fixed word length effects. Moreover due to sampling frequency offset there is a slowly increasing timing offset. The receiver has to continuously track and compensate for these effects in order to improve the efficiency of the system. Literatures describing correction algorithms use the search function *argmax* [2], [3], [6], [7] after complex-conjugate-multiply-add operations. They also need to compute the *inverse tangent* [3], [5], [8], [9], [10] to find the phase angles. The implementation complexity is very high for all these necessary function blocks [3].

In this paper we propose a novel algorithm for residual phase tracking without using either of the afore mentioned complex (costly) arithmetic functions. It computes the complex exponential of the phase angle at the pilot-tone locations instead of the phase angles to minimize the implementation complexity. The design of our algorithm is such that it can be very easily applied to any coherent OFDM demodulation scheme for WLAN type of networks.

The remainder of the article is organized as follows. In Section II we describe the system under investigation. Then we briefly present the well known mathematical analysis describing residual synchronization errors in OFDM systems. Our proposed novel algorithm is described in Section IV. A detailed discussion on the performance evaluation of our scheme is given in section V.

II. SYSTEM UNDER INVESTIGATION

A. Frame Format

OFDM systems vary greatly in their implementation. It is thus important that we describe the frame format referred in this article. We consider the IEEE802.11a [11] frame format as described in Figure 2. A transmitted packet has an all pilot (known data at Transmitter (Tx) and Receiver (Rx)) training sequence known as *PREAMBLE* at its beginning. It is used for packet start identification, automatic gain control system, symbol timing synchronization, *initial carrier frequency synchronization* and channel estimation. A Guard Interval (*GI*) follows the *PREAMBLE* which in turn is followed by the *SIGNAL* field. The *SIGNAL* field has information about the packet length and modulation format used in the frame. It is a Binary Phase

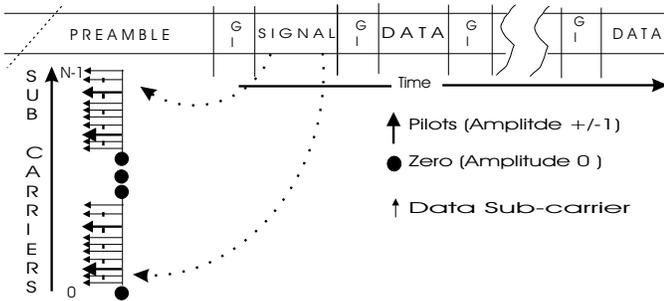


Fig. 2. OFDM receiver front end architecture

Shift Keying (BKSP) modulated OFDM symbol. Then follows a sequence of *DATA* fields separated by *GIs*, i.e the OFDM symbols carrying information. There are 64 subcarriers used in the Fast Fourier Transform (FFT) block of the OFDM system under consideration. Not all subcarriers are used to carry information. Some subcarriers such as the zero frequency component (to avoid carrier transmission) and the higher frequency subcarriers are made zero in order to avoid the filtering effect on the subcarriers due to analogue components. Hence only 52 subcarriers carry non-zero power, among which, pilots tones (values ± 1) are present at four distinct locations (subcarriers -21,-7,7 and 21). Thus only 48 subcarriers carry data information.

B. OFDM Transmission Signal Model

An OFDM symbol consists of a sum of subcarriers that are modulated by using any linear modulation method, such as *Binary Phase Shift Keying* (BPSK) or *Quadrature Amplitude Modulation* (QAM). The transmitted baseband signal for l^{th} OFDM symbol, $s_l(t)$ can be expressed as [2]:

$$s_l(t) = \frac{1}{\sqrt{T_d}} \sum_{k=-N/2}^{N/2-1} X_{l,k} e^{j \frac{2\pi}{T_d} k [t - lT_{sym} - T_{CP}]} \quad (1)$$

Where $X_{l,k}$ are constellation points to IDFT input at k^{th} subcarrier of l^{th} OFDM symbol; T_{sym} , T_d , T_{CP} and T_s are duration of complete OFDM symbol, data part, Cyclic Prefix (CP or GI) and sampling period respectively. Similarly, N_{sym} , N_d , N_{CP} defines samples for complete OFDM symbol, data part and CP respectively. $T_{sym} = T_d + T_{CP}$, $N_{sym} = N_d + N_{CP}$ and $N = N_d$. Further on we shall omit the scaling factor for simplification of representation. The signal in Equation 1 is transmitted over frequency-selective fading channel, which is characterized by its low-pass-equivalent impulse response $h(\tau, t)$ plus AWGN $n(t)$. The channel is considered to be quasi-static during the transmission of a complete packet, thus $h(\tau, t)$ simplifies to $h(\tau)$ [5]. It is further assumed that the effect of channel response $h(\tau)$ is restricted to the interval $t \in [0, T_{CP}]$, in another words, the length of CP is chosen to be longer than the maximum possible delay spread, τ_{max} . In this way the guard interval (the cyclic prefix) is able to completely absorb the tail of the pulse of the previous symbol. Received baseband signal at the receiver antenna can be written from [5]

$$r(t) = \int_0^{\tau_{max}} s_l(t - \tau) h(\tau) d\tau \quad (2)$$

$$= \sum_{k=-N/2}^{N/2-1} H_k X_{l,k} e^{j \frac{2\pi}{T_d} k [t - lT_{sym} - T_{CP}]} + n(t) \quad (3)$$

where $n(t)$ is additive white gaussian noise. Here, H_k is the channel transfer function (CTF) for k^{th} subcarrier.

III. RESIDUAL PHASE ERROR IN THE RECEIVER

Residual Phase Error has already been defined in Section I. FTO and CFO corrected signal after FFT can be expressed from [2], [5]

$$R_{l,k} = H_k X_{l,k} e^{j \frac{2\pi}{N} (N_{CP} + lN_{sym}) \phi_k} e^{j (\pi \frac{N_d - 1}{N} \phi_k + \theta)} \frac{\sin(\pi \phi_k)}{\sin\left(\frac{\pi \phi_k}{N_d}\right)} + N_{l,k} \quad (4)$$

Where $\phi_k \approx k\zeta + \xi T_d$; θ the carrier phase offset; $\xi = \delta(F_{rx} - F_{tx})$; F_{tx} and F_{rx} are the local oscillator frequencies at the transmitter and the receiver respectively and δ implies residual error after initial carrier frequency offset correction. ζ is the receiver sampling frequency offset defined through $T'_s = T_s(1 + \zeta)$; where T'_s is the receiver sampling period and T_s the transmitter sampling period. Residual carrier frequency offset in the signal even after CFO correction together with the sampling frequency offset cause phase rotation in each subcarrier to increase with OFDM symbol index l as can be seen from Equation 4. Cumulative phase increment severely limits the number of OFDM symbols that can be transmitted in one packet. The receiver thus has to continuously track and compensate for the effect (*Phase Tracking* block of Figure 1).

Since the time invariant terms are inseparable from the channel transfer function, we can write

$$R_{l,k} = X_{l,k} H'_k e^{j l \phi_k C} + N_{l,k} \quad (5)$$

where $H'_k = H_k e^{j \frac{2\pi}{N} N_{CP} \phi_k} e^{j (\pi \frac{N_d - 1}{N} \phi_k + \theta)} \frac{\sin(\pi \phi_k)}{\sin\left(\frac{\pi \phi_k}{N_d}\right)}$

and $C = 2\pi \frac{N_{sym}}{N}$.

So, we can call H'_k as the equivalent channel transfer function. $\sigma_N^2 = E\{|N_{l,k}|^2\}$ is the additive noise power [2], where $E\{\dots\}$ denotes expected value. There is an ICI term present in the received signal which can be represented as additional noise term. This leads to degradation in the available SNR. The power of the ICI term of the k^{th} sub-carrier of the l^{th} OFDM symbol is given by [6]

$$\sigma_{l,k-ICI}^2 \approx \frac{\pi^2}{3} (k\zeta)^2 \quad (6)$$

In OFDM-WLAN environment of 64 sub-carriers OFDM symbol, this effect is very small at values of $\zeta \sim 10^{-5}$ and thus $\sigma_{l,k-ICI}^2$ can be omitted now for the algorithm under discussion. We need to mention here that if the slowly increasing sampling timing drift due to SFO reaches one sampling period, then we either miss a sample or oversample it. This leads to irreducible ISI. Rob-stuff [12] method is used to combat such an effect. This particular situation is not addressed by our algorithm.

IV. OFFSET CORRECTION ALGORITHM

In this section we elaborate on the proposed residual phase correction algorithm. First, after channel equalization the estimate of the exponential part ($e^{j l \phi_k C}$) of Equation 5 is computed at the pilot locations, instead of estimating the phase ($l \phi_k C$). Then a running time-averaging is done to increase the SNR of the estimates. Finally using the estimates at pilot locations we piece-wise-linearly interpolate (Figure 3) the compensating complex exponential at all the data subcarriers. These are then multiplied with $R_{l,k}$ after channel equalization.

The effect of noise in the channel estimate can be reduced to as low as 0.41dB by using a channel estimator Gain of 10dB [2]. Thus for now we assume ideal channel compensation. If the estimated

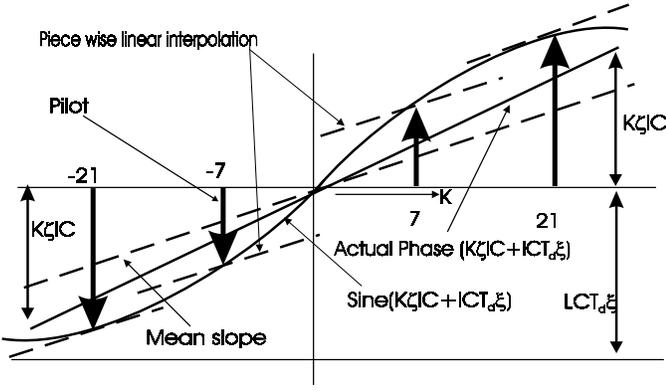


Fig. 3. Piece wise linear interpolation

channel transfer function is \hat{H}_k we assume $\hat{H}_k \approx H'_k$. It is to be noted for IEEE 802.11a type wireless networks a basic assumption is that the channel quasi-static in the whole packet duration. The *Long Training Sequence* [11] present in the *PREAMBLE* is used to estimate the channel once every packet. This estimate is valid for equalizing all OFDM symbols in entire packet since channel variability is much slower than the packet duration. We can write the received subcarriers after FTO, initial CFO correction and channel compensation as

$$R'_{l,k} = R_{l,k} H'_k / \hat{H}_k = X_{l,k} (\alpha_{l,k})^l + N''_{l,k} \quad (7)$$

where $R'_{l,k}$ is the received subcarrier after timing, frequency and channel compensation and $N''_{l,k}$ as the new noise term; $\alpha_{l,k} = e^{j\phi_k C}$ and we define $\beta_{l,k} = e^{j\phi_k C}$. If $P_{l,p}$ (values ± 1) is the pilot tone at p^{th} subcarrier index of the l^{th} OFDM symbol. For all further computation we have p at pilot indexes only. The algorithm is stated as

$$\hat{\alpha}_{l,p} = \nu \hat{\alpha}_{l-1,p} + (1 - \nu) R'_{l,p} \cdot R'^*_{l-1,p} \quad (8)$$

$$\hat{\alpha}_{l,p} = \hat{\alpha}_{l,p} / |\hat{\alpha}_{l,p}| \quad (9)$$

$$\hat{\beta}_{l,p} = \nu \hat{\beta}_{l-1,p} \hat{\alpha}_{l,p} + (1 - \nu) R'_{l,p} \cdot P^*_{l,p} \quad (10)$$

$$\hat{\beta}_{l,p} = \hat{\beta}_{l,p} / |\hat{\beta}_{l,p}| \quad (11)$$

with initial conditions $\hat{\alpha}_{0,p} = \hat{\beta}_{0,p} = 1$ where \hat{x} indicates estimate of x ; x^* denotes complex conjugate of x ; ν is the memory factor used for averaging. Equation 8 estimates the increment from $(l-1)^{th}$ OFDM symbol to l^{th} OFDM symbol. It updates previous estimates. Equation 10 estimates the compensating complex exponential for the l^{th} OFDM symbol. Equation 8 and 10 uses averaging to increase the SNR of the estimate. A detailed derivation and analysis of the algorithm is given in [13]. It has to be noted that we are not computing the phase angle, rather the complex exponential of the phase angle. Then we interpolate the real and imaginary parts separately. The straight solid line in Figure 3 is the *actual phase* that needs to be estimate, and the curve is the sinusoid of the phase. We are estimating this sinusoid at the pilot locations. There will be a sine and a cosine term, only one component is shown to reduce the complexity of the figure. We approximate them to be piece-wise-linear for small angles. Then the mean slope of each of them is estimated as given below in Equation 12. Note that we are estimating the slope as a complex entity in one single equation since the real and imaginary parts do not interact in the equation stated below.

$$m_l = \frac{1}{2} \left(\frac{\hat{\beta}_{l,p_{21}} - \hat{\beta}_{l,p_{-7}}}{21 - (-7)} + \frac{\hat{\beta}_{l,p_{-7}} - \hat{\beta}_{l,p_{-21}}}{7 - (-21)} \right) \quad (12)$$

Piece-wise-linear-interpolation is done to find the complex multiplication factor for each subcarrier for compensating residual phase error as states below.

$$y_{l,k} = \hat{\beta}_{l,p} - m_l(p - k) \quad (13)$$

Where m indicates the nearest pilot index to k^{th} subcarrier. The for compensation we use $R''_{l,k} = R'_{l,k} \cdot y_{l,k}^*$. The maximum errors that may occur will be at the farthest subcarriers where the difference is largest because of larger sub-carrier index (see Equation 4 & 5).

V. PERFORMANCE OF ALGORITHM

In [13] received signal after compensation by the proposed algorithm for small angles is given by

$$\begin{aligned} R''_{l,k} &= X_{l,k} + X_{l,k} \{LB(p - k)\}^2 \\ &\approx X_{l,k} \end{aligned} \quad (14)$$

where $B = \zeta C$. Since ζ is in the order of 10^{-5} the degradation due to the approximation used in the proposed algorithm described in section IV is almost negligible. The noise power, for small angles, has been shown in [13] to be almost equal to the noise power present in the system before applying the algorithm. If E_b is the energy per bit and $N_o/2$ the noise power density and since for BPSK, symbol energy and bit energy are the same we can say that the probability of error for BPSK modulation is given by $P_e = Q\{\sqrt{\frac{2E_b}{N_o}}\}$ [14] does not degrade significantly for low residual phase errors. This is, can be understood in the light of Equation 14.

A. Simulation Parameters

Simulations were performed with parameters from IEEE 802.11a WLAN standard: $N = N_d = 64$, $N_{CP} = 16$, BPSK modulation with 1/2 code rate convolution code, but constraint length was kept as 3; *Carrier frequency* = 5.4 GHz; *Sampling Frequency Offset* = 50 ppm. We have taken the largest SFO as per the standard's requirement since, very small SFO does not have significant effect on the performance of the receiver. Residual carrier frequency error after CFO correction was varied from $\sim 30Hz$ to $\sim 4kHz$ to test the algorithm's performance for small (800bits) and larger (4000bits) packet lengths.

B. Performance Comparison

For comparison in correspondence to our objective, an algorithm that estimates phases angles $l\phi_k C$ at the pilot locations using inverse tangent function is taken with reference to [3], [8]. A linear interpolation of the phase is done for all data sub-carriers. The compensation is done by multiplying the received signal of Equation 5 by $e^{-j\widehat{l\phi_k C}}$, where $\widehat{l\phi_k C}$ is the estimated phase angle. In Figure 4 and 5, we denote our algorithm as *alg-1* (dotted line) and the algorithm used as reference for comparison as *alg-2* (solid line). A reference Bit Error Rate (BER) level of 10^{-5} was chosen from IEEE802.11a [11] standard for performance comparison.

1) *Low Residual Phase Error*: Figure 4 shows the BER vs SNR curve for both algorithms for low residual phase error. It can be observed that both algorithms perform very closely for residual carrier frequency error in the range of $\sim 30Hz$ (0.005ppm) to $\sim 300Hz$ (0.05ppm) with sampling frequency offset of 50 ppm for both small and medium sized packet of length 800 and 4000 bits. BER of $\sim 10^{-5}$ is achieved by both algorithms at almost same SNR of about (7dB). This confirms our analysis for small angles as stated in the beginning of section V. It may be noted that the estimation of the phase angle by angle functions is not perfect because of corruption

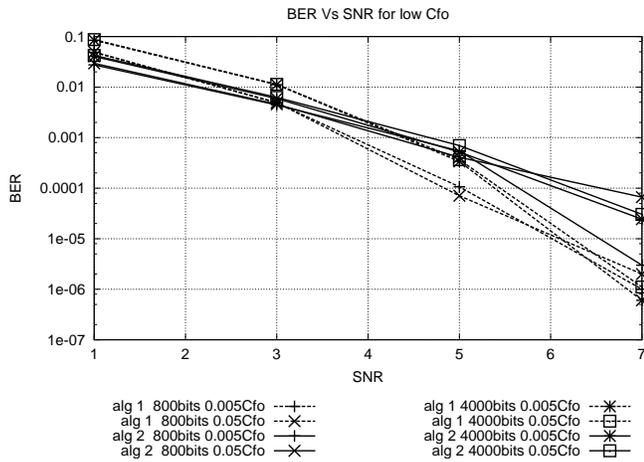


Fig. 4. BER Vs SNR for low residual phase

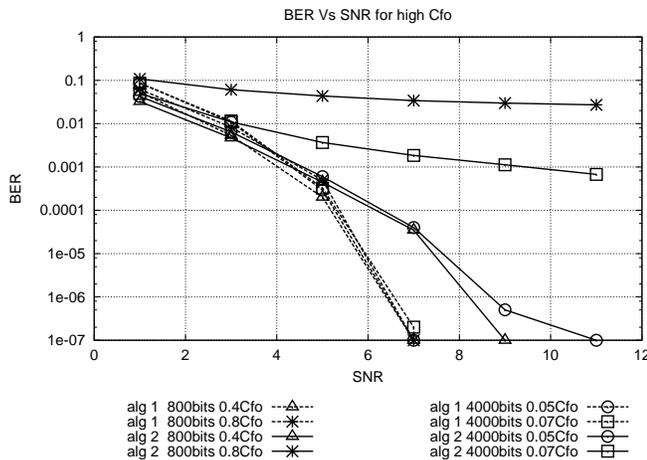


Fig. 5. BER Vs SNR for high residual phase

of the complex number by noise. At low angles this does not cause much deviation though.

2) *High Residual Phase Error*: Figure 5 shows graphs of BER vs SNR for both algorithms from simulation results for comparatively higher residual phase errors. Residual carrier frequency offset in the range of $\sim 300\text{Hz}$ (0.05ppm) to $\sim 4\text{kHz}$ (0.8ppm), sampling frequency offset of 50ppm for packet lengths of 800bits and 4000bits were considered. For 800bits packet length, residual carrier frequency offset of $\sim 2\text{kHz}$ (0.4ppm) *alg-2* requires only about 2dB more SNR in comparison to *alg-1* to meet the required performance level. But for residual carrier frequency offset of $\sim 4\text{kHz}$ (0.8ppm) for same previous packet length of 800bits *alg-2* does not meet the required BER, where as *algo-1* does meet the requirement. For larger packet length of 4000bits it can be seen that at residual carrier frequency offset of $\sim 300\text{Hz}$ (0.05ppm) *alg-2* requires about 4dB more SNR to reach the BER of 10^{-5} as compared to *alg-1*. For larger residual carrier frequencies $\sim 400\text{Hz}$ (0.07ppm) *alg-2* never reaches the desired performance criteria stated above, whereas *alg-1* does so without significant change in SNR requirement.

VI. CONCLUSION

We have presented an efficient low complexity residual phase correction algorithm, using piece-wise-linear approximation technique,

for OFDM based WLAN receivers to replace the complex (costly) inverse tangent and argmax function implementation.

It has been observed that for low residual phase errors the proposed algorithm performs almost identical to the algorithm using phase angle estimation by inverse tangent functions. For higher residual phase errors the proposed algorithm's performance is still stable when the other scheme almost fails. By being robust to larger residual errors it reduces stringent performance requirement of the *Freq Synch* block. This creates provision for lower performance requirement of the *initial freq synch* block followed by our tracking algorithm as given in Figure 1. Higher accuracy performance requirements of the *Freq Synch* block increases complexity [3] and implementation cost of the receiver. Further, our algorithm does not use the *argmax* or *inverse tangent* functions which are very costly in terms of hardware implementation. It may be mentioned here that low complexity implementation of *inverse tangent* function may be done using table lookup or *cordic* structure. With increased required resolution, table look up uses more space and *cordic* suffers from higher latency. In contrast, our algorithm does not have such limitations. We have thus seen for WLAN type of packet based wireless network using OFDM scheme the proposed algorithm for residual phase tracking can prove highly effective in reducing cost of receivers without compromising on performance.

REFERENCES

- [1] Richard Van Nee & Ramjee Prasad, *OFDM for Wireless Multimedia Communications*. Artech House Publishers, 2000.
- [2] M. Speth, S.A. Fechtel, G. Fock & H. Meyr, "Optimum Receiver Design for Wireless Broad-Band Systems Using OFDM - Part I," *IEEE Transactions on Communications*, vol. 47, no. 11, November 1999.
- [3] —, "Optimum Receiver Design for Wireless Broad-Band Systems Using OFDM - Part II: A case study," *IEEE Transactions on Communications*, vol. 49, no. 4, April 2001.
- [4] A. Dey, R.V. Rajakumar, S S Das, A. Pal, Balamurali., "Synchronization Algorithms For The IEEE 802.11 a/g Wireless Lan, <http://kom.auc.dk/~ssd/reports/>," *10th National Conference on Communications, Indian Institute of Science, Bangalore*, Jan 30 - Feb 1 2004.
- [5] Klaus Witrisal, "OFDM Air Interface Design for Multimedia Communications," Ph.D. dissertation, Delft University of Technology, The Netherlands, April 2002.
- [6] B. Yang, K.B. Letaief, R.S. Cheng & Z. Cao., "Timing Recovery for OFDM Transmission," *IEEE Journal ON Selected Areas in Communications*, vol. 18, no. 11, November 2000.
- [7] Baoguo Yang; Zhengxin Ma; Zhigang Cao, "ML-oriented DA sampling clock synchronization for OFDM systems," *WCC - ICCT 2000*, vol. 1, pp. 781 - 784, 21-25 Aug 2000.
- [8] Juha Heiskala & John Terry, *OFDM Wireless LANs: A Theoretical and Practical Guide*, 2nd ed. Sams Publishing, July 2001.
- [9] I. Abhayawardhana, V.S.; Wassell, "Residual frequency offset correction for coherently modulated ofdm systems in wireless communication," in *Vehicular Technology Conference*, vol. 2, VTC Spring 2002 2002, pp. 777-781.
- [10] Miaoudakis;et.al, "An all-digital feed-forward CFO cancellation scheme for HIPERLAN/2 in multipath environment," in *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2002*, 4, Ed., Sept 2002, pp. 15-18.
- [11] IEEE Std 802.11a-1999, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications High-speed Physical Layer in the 5 GHz Band," IEEE, Tech. Rep., 1999.
- [12] P. M. M. Pollet, T.; Spruyt, "The ber performance of ofdm systems using non-synchronized sampling," in *GLOBECOM '94*, vol. 1. IEEE, 28 Nov.-2 Dec. 1994 1994, pp. 253 - 257.
- [13] S.S.Das et. al, "Design of Low-Complexity OFDM-WLAN Systems Part I, <http://kom.auc.dk/~ssd/reports/>," Aalborg University, Department of Communication Technology , Technical Report R-04-1004, ISSN 0908-1224, ISBN 87-90834-45-3, February 2004.
- [14] J.G.Proakis, *Digital Communications*, 3rd ed. Mc Graw Hill, 1995.

LOWER ORDER BASED ADAPTIVE RECEIVER FOR DS-CDMA COMMUNICATION SYSTEM

S. A. Jimaa¹, M. E. Jadah¹, B. S. Sharif²

¹ Etisalat University, College of Engineering & Information Sciences, PO Box 980, Sharjah, UAE

² School of Electrical, Electronic, and Computer Engineering, University of Newcastle upon Tyne, Newcastle upon Tyne, UK

ABSTRACT

This paper investigates the performance of using the lower order non-quadratic cost function adaptive algorithm (L_p) in the adaptation of a non-linear receiver, coupled with a second-order phase tracking subsystem, for asynchronous DS-CDMA communication system impaired by double-spread multipath channel and Additive White Gaussian Noise (AWGN). The non-linear receiver comprises feedforward filter (FFF), feedback filter (FBF), and an equalizer/second order phase locked loop (PLL). The investigation studies the performance of the mean-square error (MSE) and the bit-error-rate (BER) using various cost functions and step-sizes. Computer simulation results indicate that the proposed receiver's algorithm with the cost function of power 1.9 gives the best performance. Furthermore the results show that the best value of the algorithm's step-size is 0.0005. Finally it is demonstrated that non-linear receivers adapted by the proposed algorithm will have faster convergence rate and similar BER performance, in comparison with the NLMS adaptive receiver.

1. INTRODUCTION

Various adaptive MMSE receivers have been proposed for the detection of DS-CDMA systems. For AWGN channels, adaptive MMSE receivers were developed based on the standard MSE cost function [1], [2]. Demodulation of DS-CDMA signals is conventionally achieved with a matched-filter receiver, which exploits the low cross-correlation between signatures of different users. In most of DS-CDMA systems, transmitters send information independently resulting in users arriving asynchronously at the receiver. In such a system, signatures are unable to maintain their orthogonality resulting in multiple-access interference (MAI). Another major drawback associated with DS-CDMA is the near-far problem whereby a weak signal from distant user suffers from interference from a strong signal from a nearby user. The linear and non-linear MMSE detectors considered in [3] are single-user detectors in the sense

that they demodulate the bit stream of one user at a time. For high speed phase coherent communications as required in 3G mobile systems, it is shown in [4] that the adaptive MMSE detector is not able to track the phase of the incoming signal especially when the received signal experiences a deep fade. On the other hand, the proposed receiver in [3] performs optimal phase tracking and channel equalization jointly. So far, the Least-Mean-Square (LMS) algorithm, based upon $E[e^2(t)]$ has proved popular for many applications because of its simplicity and ease of implementation. However, many alternatives based upon non-mean-square-error cost functions can also be defined to improve the adaptation performance in specific statistical environments [5]-[7]. Hence the proposed receiver in this paper is based on using the lower order cost function, $E[e^p(t)]$, algorithm to update the tap coefficients of the feedforward and feedback filters. A second-order digital phase-locked loop (DPLL) is used here to track the phase of the incoming signal. Extensive tests have been carried out to determine the best value of the proposed algorithm cost function's power and its step size.

2. SYSTEM MODEL

In this work, asynchronous DS-CDMA system with K active users has been considered. QPSK with symbol duration T_s is assumed while the chips of the spreading sequence have duration T_c . The unit-energy signature waveform of the k^{th} user is given by:

$$s_k(t) = \frac{1}{\sqrt{T_s}} \sum_{j=1}^N \delta_k[j] \Psi(t - (j-1)T_c) \quad (1)$$

where $\delta_k[n] \in \{-1, +1\}$ is the n^{th} chip of the k^{th} user and $N = T_s/T_c$ is the processing gain of the system. The chip waveform $\Psi(t)$ is zero for $t \notin (0, T_c)$. Each user's transmitted signal is assumed to pass through a frequency-selective Rayleigh fast-fading (Doppler frequency = 129 Hz) channel. AWGN, resulting from receiver thermal noise, is also considered. The received signal from K users is represented as:

$$r(t) = \sum_{k=1}^K r_k(t) + n(t) \quad (2)$$

where $r_k(t)$ is represented as:

$$r_k(t) = \sum_{i=-\infty}^{\infty} d_k[i] \sum_{l=1}^L c_k[l] s_k(t - iT_s - t_k[l]) \quad (3)$$

$d_k[i] = a_k[i] + jb_k[i]$ is the i^{th} QPSK symbol of the k^{th} user, and $a_k[i], b_k[i] \in \{-1, +1\}$, L is the number of paths of user k , the complex quantity $c_k[l]$ is the amplitude and phase variation of the l^{th} path of user k and $t_k[l]$ is the reception time of the l^{th} path of user k . The amplitude and phase of the $c_k(l)$ in (3) are Rayleigh and uniformly distributed, respectively. The delay of each user is defined by the arrival time of the first ray of that user, $t_k[l]$ where $t_k[l] \in [0, T_s)$.

3. RECEIVER STRUCTURE

The proposed receiver model is shown in Fig. 2. Joint operation of equalization and phase tracking is considered here [3]. Assuming perfect estimation of the transmission time of each user, the signal to the input of the k^{th} user chip-match filter (CMF) is delayed by $T_s - \tau_k$, and then sampled at $\Delta = T_c$. Without loss of generality, the user of interest is assumed as user number 1. The N taps of the FFF are arranged in a row vector α_1^T , and the input signal samples currently stored in the FFF are given by:

$$r_1(n) = [r(nT_s + T_c + \tau_1), \dots, r(nT_s + NT_c + \tau_1)]^T \quad (4)$$

The FBF has tap weights β_1^T and operates on M previous detected symbols.

$$d_1(n) = [d_1^{\wedge}[n-1], \dots, d_1^{\wedge}[n-M]]^T \quad (5)$$

If we define the coefficients vector u_1 and data vector x_1 as:

$$\mathbf{u}_1 \equiv [\alpha_1^T, \beta_1^T]^T, \quad \mathbf{x}_1[n] \equiv [r_1^T, d_1^T]^T \quad (6)$$

then the soft-symbol estimate of the n^{th} QPSK symbol of the 1st user is $d_1^{\wedge}[n] = u_1 \cdot x_1[n] e^{-j\theta_1^{\wedge}[n]}$ which in turn will be fed to the hard symbol decision to produce

$$d_1^{\wedge}[n] = \text{sgn}[\text{Re}\{d_1^{\wedge}[n]\}] + j \text{sgn}[\text{Im}\{d_1^{\wedge}[n]\}] \quad (7)$$

The MSE of the receiver at instant time n is given by:

$$\boldsymbol{\varepsilon}_1 = E\{|\boldsymbol{\varepsilon}_1[n]|^2\} = E\{|d_1[n] - d_1^{\wedge}[n]|^2\} \quad (8)$$

While the mean lower-order error (ML_pE) of the receiver at time n is given by:

$$\boldsymbol{\varepsilon}_1 = E\{|\boldsymbol{\varepsilon}_1[n]|^p\} = E\{|d_1[n] - d_1^{\wedge}[n]|^p\} \quad (9)$$

where $p=1.9$. In decision-directed mode, $d_1[n]$ should be substituted by $d_1^{\wedge}[n]$. Due to the time varying nature of the channel response, optimal values of the receiver parameters are also time varying. Therefore, the equalizer must be designed to adaptively compensate for time variations in the channel characteristics. The NLMS adaptive algorithm with an optimum step-size of 0.001 is used to update the equalizer as follows:

$$u_1[n+1] = u_1[n] + \frac{\mu_{NLMS}}{\|x_1[n]\|^2 + \gamma} x_1[n] \boldsymbol{\varepsilon}_1^*[n] e^{-j\theta_1^{\wedge}[n]} \quad (10)$$

The proposed L_p adaptive algorithm is used to update the equalizer as follows [7]:

$$u_1[n+1] = u_1[n] + p\mu_{Lp} x_1[n] \boldsymbol{\varepsilon}_1^{*(p-1)}[n] e^{-j\theta_1^{\wedge}[n]} \quad (11)$$

$$\theta_1[n+1] = \theta_1[n] + K_1 \phi_1[n] + K_2 \sum_{i=0}^n \phi_1[n] \quad (12)$$

$$\phi_1[n] = \text{Im}\left\{d_1^{\wedge}[n] d_1^{\wedge*}[n]\right\} \quad (13)$$

where $|\bullet|^2$ denotes the squared Euclidean norm, $\phi_1[n]$ is the phase detector output, $K_1 \gg K_2$ are proportional and integral tracking constants, μ_{NLMS} is the NLMS step size, μ_{Lp} is the L_p algorithm step size, and γ is a small positive constant used to ensure stability if the input signal power is low. It should be noted that the structure shown in Fig. 2 is applicable to all other users.

4. L_p ADAPTIVE ALGORITHM

The choice of a cost function is central to the design of adaptive algorithms using the method of gradient descent. Different cost functions lead to different algorithms [7]. The traditional quadratic cost function used in the NLMS algorithm is defined in Eqn. (8). A lower order algorithm L_p is derived from the LMS by using $p=1.9$ giving rise to a non-quadratic cost function. Justification of using this value of p will be given in section 5. Such a lower order cost function can be described as in Eqn. (9), and it is applied to the method of gradient descent using the iterative algorithm of Eqn. (11). It should be noted that a normalized version of the proposed L_p has been tested and it offers the same performance. Hence the focus is on the L_p algorithm.

5. RESULTS AND DISCUSSION

The proposed receiver is tested by means of simulations in an asynchronous system where the arrival time of the first ray of each user satisfies $\tau_k[1] \sim U(0, N)$. 31-chip Gold sequences are used and the modulation scheme is QPSK. The multipath channel considered here is Rayleigh multipath frequency selective fading implies that $T_m > T_s$, the delay time is greater than the symbol

time. The channel is fast fading with a Doppler frequency of 129 Hz. In this paper it is assumed that we have 5 equal power users. All the performances shown in Figs. 2-7 are obtained using the channel specified as follows: The first path is delayed by 0.0 μ sec (0 chips) with an amplitude of 0.7, the second path is delayed by 1.5 μ sec (46 chips) with an amplitude of 0.5, and the third path is delayed by 1.875 μ sec (58 chips) with an amplitude of 0.332. Firstly, we demonstrate in Fig. 2 the MSE performance of the proposed lower order receiver for different values of p , where p is the power of the lower-order cost function. Several values of p were tested (1.2, 1.4, 1.7, and 1.9) and Fig. 2 clearly demonstrates that the L_p algorithm with $p=1.9$ achieves the fastest convergence. In Fig. 3 the BER performance of the proposed receiver is examined. It can be observed that the BER performance of the propose receiver adapted by the L_p algorithm with $p=1.9$ achieves the best performance. Moreover extensive tests have been carried out to examine the performance of the proposed receiver using powers of $2.0 > p > 1.9$. These tests revealed that $p=1.9$ gives better performance in both MSE and BER. However, the detailed results of this investigation have been omitted here due to space limitations. In both figures the value of μ_{L_p} equals to 0.0005. In Figs. 4&5 the MSE and BER performances of the L_p adaptive receiver for various step-sizes are examined. It is clear that the L_p performance with $\mu=0.0005$ performs better which confirms the earlier observation. To highlight the advantage of using the lower-order algorithm in updating the tap weights of both the FFF and the FBF, simulations results plotted in Fig. 6 show that lower-order algorithm exhibits faster convergence in comparison with the traditional NLMS. Furthermore, the BER plots in Fig. 7 show that similar performance to the NLMS can be achieved by using the L_p algorithm.

6. CONCLUSIONS

This paper considers the use of a lower order algorithm as an alternative to the NLMS and shows that the MSE performance is improved. Extensive computer

simulation tests show that the best performance is obtained when the power (p) of the cost function is 1.9 and the stepsize is 0.0005. The proposed receiver structure using L_p adaptive algorithm provides a fast convergence compared to the NLMS adaptive algorithm, albeit at similar computational complexity.

7. REFERENCES

- [1] U. Madhow and M. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Comms.*, vol. 42, pp. 3178-3188, December 1994.
- [2] S. L. Miller, "An adaptive direct-sequence code-division-multiple-access receiver for multiuser interference rejection," *IEEE Trans. Comms.*, vol. 43, pp. 1746-1755, April 1995.
- [3] S. S. Petsis, B. S. Sharif, C. C. Tsimenidis, and O. R. Hinton, "MMSE-based receivers for asynchronous DS-CDMA communication systems in the presence of double-spread multipath channels," *The 10th IEEE international conference ICECS2003*, Sharjah, UAE, 14-17 December 2003.
- [4] A. N. Barbosa and S. L. Miller, "Adaptive detection of DS/CDMA signals in fading channels," *IEEE Trans. Comms.*, vol. 46, pp. 115-124, January 1998.
- [5] S. A. Shah and C. F. N. Cowan, "Modified stochastic gradient algorithm using nonquadratic cost functions for data echo cancellation," *IEE Proc., Vis. Imag Sig. Proc.*, vol. 142, no. 3, pp. 187-191, June 1995.
- [6] C. Rusu and C. F. N. Cowan, "Adaptive data echo cancellation using cost function adaptation," *Journal of Signal Processing*, 80 (2000), pp. 2457-2473, Nov. 2000.
- [7] S. A. Jimaa, "Adaptive Equalization Using Lower-Order-Cost-Function In The Presence of Co-Channel Interference", *Proceedings of the IASTED International Conference on Modeling and Simulation - MS'99*, pp. 254-258, 5-8 May 1999, Philadelphia, Pennsylvania-USA.

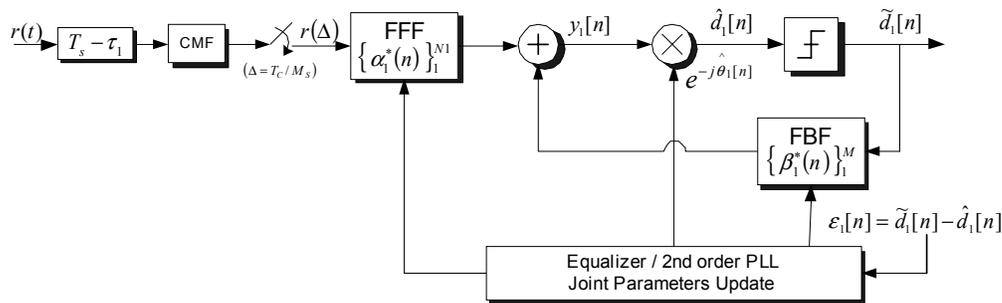


Figure 1 Receiver structure

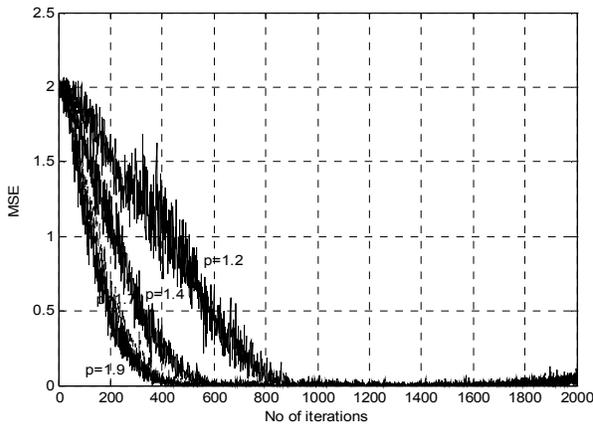


Fig. 2 Learning curves of the L_p algorithm for various cost functions, step-size=0.0005

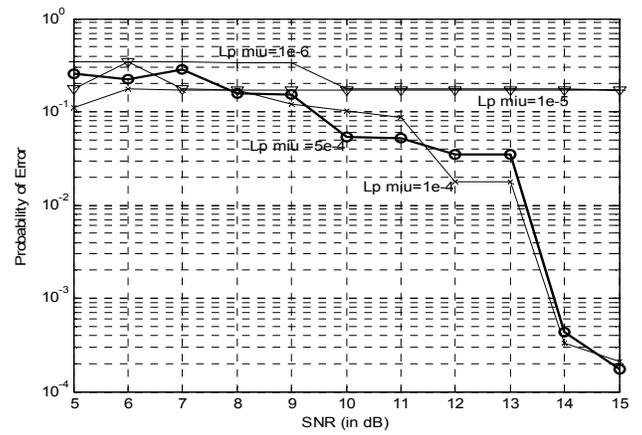


Fig. 5 BER performance of the L_p algorithm for various step-sizes, $p=1.9$

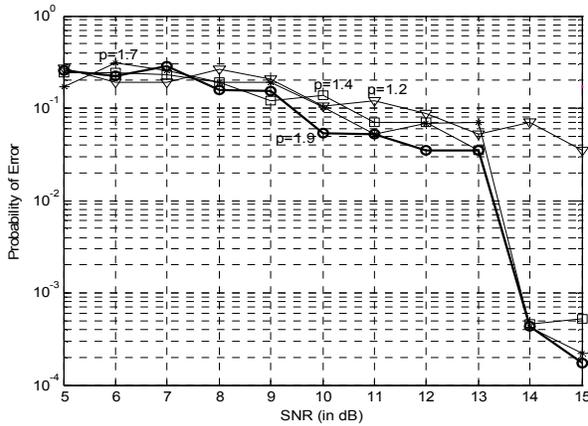


Fig. 3 BER performance of the L_p algorithm for various cost functions, step-size=0.0005

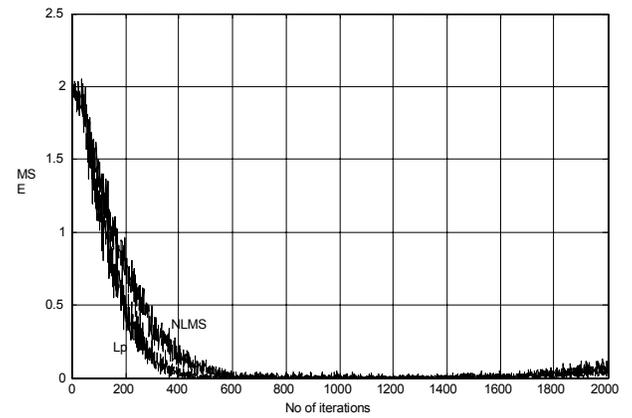


Fig. 6 Learning curves of the proposed algorithm (L_p) and the NLMS algorithm

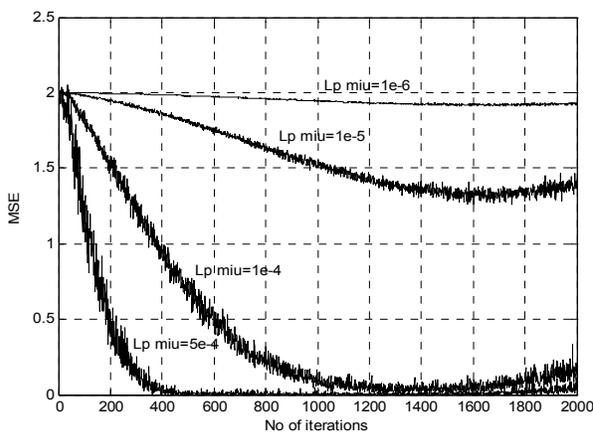


Fig. 4 Learning curves of the L_p algorithm for various step-sizes, $p=1.9$

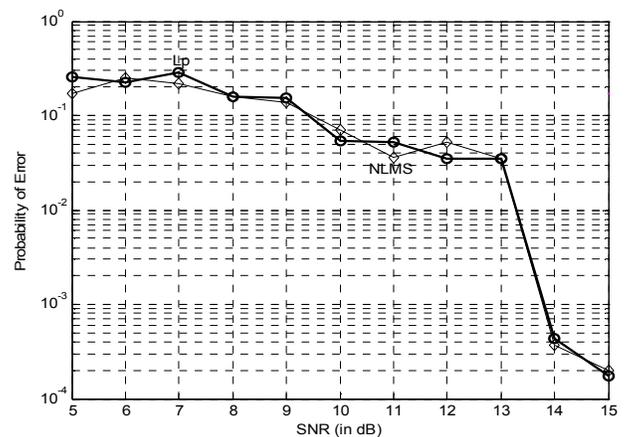


Fig. 7 BER performance of the proposed algorithm (L_p) and the NLMS algorithm

MODULATION AND PROTOCOL RECOGNITION IN COMMUNICATION SYSTEMS

Ali R. Attar, Abbas Sheikhi, Habibollah Abiri, Ali Zamani

Department of Electrical and Electronic Engineering, Shiraz University
P.O.BOX : 71345-1457, Shiraz, Iran
e-mail: ar_attar@hotmail.com, sheikhi@shirazu.ac.ir
web: www.shirazu.ac.ir

ABSTRACT

Automatic Modulation Recognition (AMR) has been a major research area for the past decade. In this paper we introduce a new method based on time-domain and spectral features of the received signal. We have used neural network as the classifier. Some analog and digital modulations including AM, LSSB, USSB, FM, ASK2, ASK4, ASK8, PSK2, PSK4, PSK8, FSK2, FSK4, FSK8 and MSK are considered. Then using information from the received signal like baud rate, carrier frequency and modulating scheme the protocol used for signal transmission is detected.

1. INTRODUCTION

Automatic Modulation Recognition (AMR) can be used in many applications. Electronic warfare [1], Spectrum surveillance and management [2], Electronic support measure [3] and Universal demodulator [4] are some of its applications to mention a few.

Different methods of AMR can be categorized in two broad fields : Pattern recognition and Decision theoretic. In the past, decision making was the main method used by researchers like [1],[5], and [6], but in the previous years pattern recognition methods are dominant especially using neural networks. Some of the papers dealing with this situation are [2],[3]and[4].

The aim of this paper is to introduce a proper method in order to automatically recognize the modulating scheme and data communication protocol. Some authors consider a three step method for decoding data from an unknown received signal and the input/output data relationship like Fig. 1.[7].

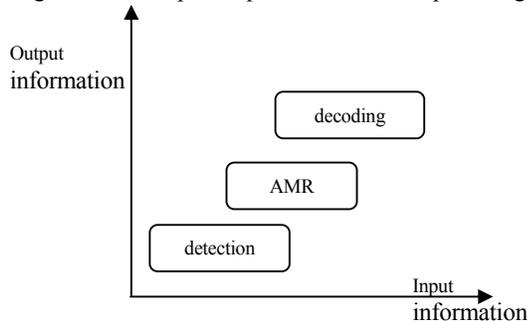


Figure 1: The input/output information relationship

We claim that by using an extra step named protocol recog

nition like Fig. 2, data decoding is done more easily.

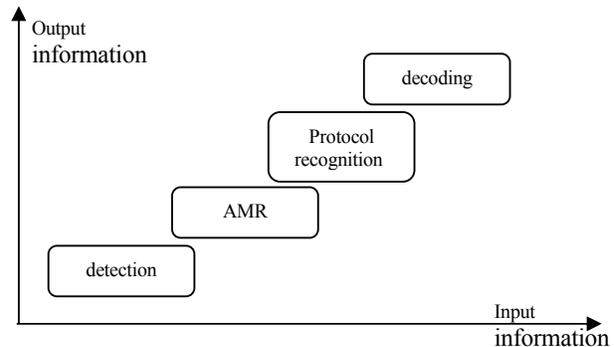


Figure 2: The input/output information relationship in our method

In this paper we will introduce a new modulation classifier using features of received signal in both time and spectral domain and then using this information plus baud rate and carrier frequency estimation the protocol used can be deduced using a database of protocol features.

There are a number of different modulation recognition methods introduced by researchers of the field. Most of these methods are proper for a limited number of modulations. Each method uses its own assumption about the known parameters of signal like carrier frequency, SNR, Baud rate, simulation method and so on. So, combining two method together can be a difficult job.

Also, there are many features of the signal proper for recognition of specific characteristics of the modulated signal. We have use some of the features previously introduced by other authors for some specific applications to develop a new method proper for classification of a broad category of signals.

The modulation set we choose includes AM, LSSB, USSB, FM, MASK, MPSK, MFSK all for M=2,4,8 and MSK (minimum shift keying). There are a number of different methods to classify subsets of the above mentioned modulation schemes but there is no unique way to recognize all of them. The above modulation schemes are selected form nine communication systems that we try to recognize them. These systems are ACARS, ALE, ATIS, FMS-BOS, PACTOR-II, PSK31, DGPS, GOLAY and ERMES. We have used neural networks with backpropagation training method as modulation classifier.

In part 2 the problem definition and introduction of the features will be presented. Part 3 contains description of the

neural network structure. The carrier and baud rate estimation method is introduced in Part 4. Part 5 is the result of our simulations and finally the conclusions will be presented in part 6.

2. PROBLEM DEFINITION

Each type of modulation changes a characteristics of carrier signal according to the message to be sent. The main characteristics of the carrier signal are frequency, phase and amplitude. So if we desire to recognize different modulation schemes, we should find some features that shows the variation of these characteristics. Analoge modulations are created according to the following formula [8] :

$$y(t)=A[1+m\cos(2\pi f_c t+k_d \int x(t)dt)]$$

m is the AM modulation depth, k_d is FM modulation index, x(t) is the modulating signal and f_c is the carrier frequency.

The digital modulations are created as follows[9]:

$$\text{MASK} : S_m(t) = \text{Re}[A_m U(t)e^{j2\pi f_c t}] \quad m=1,2,\dots,M$$

$$\text{MPSK} : S_m(t) = \text{Re}[AU(t)e^{j[2\pi f_c t + \frac{2\pi}{M}(m-1)]}]$$

$$m=1,2,\dots,M$$

$$\text{MFSK} : S_m(t) = \text{Re}[AU(t)e^{j2\pi(f_c+f_m)t}]$$

$$m=1,2,\dots,M$$

For the minimum shift keying, we have used the concept of continuous phase modulations :

$$\text{MSK} : S(t) = A \cos[2\pi(f_c + \frac{1}{4T} I_n)t - \frac{n\pi}{2} I_n + \theta_n]$$

$$nT \leq t \leq (n+1)T$$

Where

$$\theta_n = \pi h \sum_{k=-\infty}^{n-1} I_k \quad \text{and } I_n \text{ are data amplitudes.}$$

And h is called modulation index.

Now we need some features to show the amount of variation in amplitude, phase and frequency of the signal. As we have used neural networks it is possible to use hierarchical method for classification [3]. More details will be given in part 4. So we grouped the following modulation schemes as metagroup1:AM, LSSB, USSB, FM, ASK2, ASK4, PSK2, PSK4, FSK2 and FSK4 and the remaining ones as metagroup2: ASK8, PSK8, FSK8 and MSK.

For the first metagroup we can use the nine features presented by Nandi and Azzouz [10]. They are as follows :

$$1- \gamma_{\max} = \max |DFT(a_{cn}(i))|^2 / N_s$$

N_s : No. of samples per block

a_{cn} : Normalized-centred instantaneous amplitude

$$2- \sigma_{ap} = \sqrt{1/c(\sum_{an(i)>at} \phi_{NL}^2(i)) - (1/c \sum_{an(i)>at} |\phi_{NL}(i)|)^2}$$

ϕ_{NL} : Centred-nonlinear component of instantaneous phase

C : No. of samples in ϕ_{NL} for which $a_n(i) > a_t$, the

threshold value of $a_n(i) = \frac{a(i)}{\text{mean}(a(i))}$

And a(i) is the instantaneous amplitude.

$$3- \sigma_{dp} = \sqrt{1/c(\sum_{an(i)>at} \phi_{NL}^2(i)) - (1/c \sum_{an(i)>at} \phi_{NL}(i))^2}$$

$$4- P = \frac{P_L - P_U}{P_L + P_U} \quad \text{where}$$

$$P_L = \sum_{i=1}^{f_{cn}} |X_c(i)|^2, \quad P_U = \sum_{i=1}^{f_{cn}} |X_c(i + f_{cn} + 1)|^2$$

$X_c(i)$: Fourier transform of RF signal

$$5- \sigma_{aa} = \sqrt{1/N_s(\sum_{i=1}^{N_s} a_{cn}^2(i)) - (1/N_s \sum_{i=1}^{N_s} |a_{cn}(i)|)^2}$$

$$6- \sigma_{af} = \sqrt{1/c(\sum_{an(i)>at} f_N^2(i)) - (1/c \sum_{an(i)>at} |f_N(i)|)^2}$$

$f_N(i)$: Normalized-centred instantaneous frequency

$$7- \sigma_a = \sqrt{1/c(\sum_{an(i)>at} a_{cn}^2(i)) - (1/c \sum_{an(i)>at} a_{cn}(i))^2}$$

$$8- \mu_{42}^a = \frac{E\{a_{cn}^4(i)\}}{\{E\{a_{cn}^2(i)\}\}^2}$$

where E{ } means Expected value.

$$9- \mu_{42}^f = \frac{E\{f_N^4(i)\}}{\{E\{f_N^2(i)\}\}^2}$$

f_N : Normalised-centred instantaneous frequency

For the second group we use the spectrum of the signal as the feature as proposed in reference [2]. In this neural network Welch periodogram of signal is used as feature for classification. To reduce the dimension of input data, The main lobe of periodogram containing most of information is used and the remaining parts are not considered. Choosing the proper interval of periodogram is done after checking all the modulation set spectrums. The neural network structure and its characteristics are introduced in the next section.

Finally we need methods to estimate baud rate and carrier frequency of signal. For carrier recovery we have used the zero crossing method[11]. The baud rate is estimated using the method of [12] and a brief explanation of the two methods are given in part 4

3. NEURAL NETWORK STRUCTURE

We have used the concept of hierarchical neural networks described by [3]. In this method classification can be done in successive steps. The outputs can be classified in groups called metagroup and the neural network classifies these metagroups first. Then classification can be done within each metagroup in the same manner till getting the final result. Our neural network classifies two metagroups mentioned in part 2. The input data is classified as one of the

first metagroup members or just as metagroup2. So we do not need another neural network for classification of first metagroup. Then signals belonging to the second metagroup are classified using another network structure to the final output result. This method is shown in the Fig. 3.

Net1 is a feedforward network with two hidden layers. The structure is chosen after an extensive simulation tests. The number of nodes are 9 in the input layer, 75 in the first hidden layer, 75 in the second hidden layer and 11 in the output layer. The number of nodes is chosen to get the best performance results. The activation function used is log-sigmoid.

The network is trained using backpropagation method. For better result we have used a variable rate backpropagation algorithm. This training scheme converges faster and avoids falling in a shallow minimum and gets better results. We have used 240 block of data each containing 2048 samples of the signal in SNRs varying from 0 to 55 dB for training the networks and another set of the same size but different from those used for training as the test set.

The second neural network named Net2 in the Fig. 3 is a one hidden layer feedforward neural network that is trained using the same method as Net1. It has only one input node.

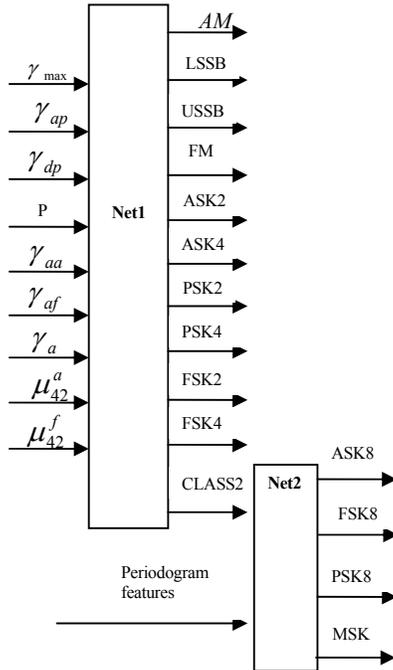


Figure 3: The neural network structure used in our method

4. CARRIER AND BAUD RATE ESTIMATION

Hsue and Soliman[11] introduced a method based on zero crossing to estimate the carrier frequency. The f_c is estimated by averaging $y_a(i)$ samples as :

$$f_c = \frac{N_y}{2 \sum_{i=1}^{N_y} y_a(i)}$$

where y_a is created as follows:

The received waveform is modelled as :

$$r(t)=s(t)+v(t) \quad 0 \leq t \leq T_s$$

If the signal is sampled using zero-crossing sampler, the time tags of zero-crossing points are recorded as a sequence $\{x(i), i=1,2,\dots,N\}$. Then the zero-crossing interval sequence $\{y(i)\}$ is defined as

$$y(i)=x(i+1)-x(i) \quad i=1,2,\dots,N-1$$

and zero-crossing interval difference sequence $\{z(i)\}$ is

$$z(i)=y(i+1)-y(i) \quad i=1,2,\dots,N-2$$

Finally variance σ_z^2 is estimated and $y(i)$ samples whose

$|z(i+1)|$ exceed $3.05 \sigma_z$ is discarded to obtain $\{y_a(i)\}$.

The length of the resultant sequence is N_y .

For baud rate estimation we have used Wegener[12] method.

The algorithm used is as follows :

sum=0

for each L_j in the sorted list : #bauds=round(L_j / est_{init})

sum=sum+($L_j / \#bauds$)

est_{final} =sum/N

In this algorithm L_j represents the number of samples between baud transitions $j-1$ and j (baud length) and N is the number of transitions. Before using the algorithm an initial estimation of baud length should be achieved.

5. SIMULATION RESULTS

We have used voice signal as the modulating signal instead of binary random data and we used 495120 samples of voice in total. The stop margin of training was 300 epochs or LMS rate less than 0.01. Then the maximum value of output nodes is considered for classification result.

In test period we used another set of 491520 to check the performance of the network. The results are presented in table 1 and 2 at the end of paper. In table 1 the performance result of the first neural network is included. In this case, ASK8, PSK8, FSK8 and MSK signals are classified all as class 2 data. We included them separately and the reader is aware that this does not mean the ability of recognition of them by the first neural network. In the second table the performance result of both the first and second neural network for the class 2 signals are introduced.

Although the modulation subset considered is different from reference [2] and [10], we compare the performance of our method with Nandi and Azzouz method in table 3 at SNR equal to 15 dB. The performance of Ghani method versus SNR is not presented in his paper, so we could not compare its performance with our method. As it is clear from table 3, the developed method presented here is in the same order of performance compared to references given above while recognizing new modulations that were not previously considered. Then using the modulation scheme information and estimated baud rate and carrier frequency and comparing these information with the known counterparts of them in the nine system mentioned earlier we can deduce the protocol. The data decoding then is a simple matter.

SNR	0 dB	5 dB	15 dB	25 dB	35 dB	45 dB	55 dB
AM	87.51	90	90	100	100	100	100
LSSB	15	55.5	90	100	100	100	100
USSB	17	85.7	95	100	100	100	100
FM	90	85	90	100	100	100	100
ASK2	90	100	100	100	100	100	100
ASK4	80	100	100	100	100	100	100
ASK8	75	95	100	100	100	100	100
PSK2	10	35	85	100	100	100	100
PSK4	15	65	95	100	100	100	100
PSK8	90	95	100	100	100	100	100
FSK2	85	75	100	100	100	100	100
FSK4	75	100	100	100	100	100	100
FSK8	95	100	100	100	100	100	100
MSK	70	85	95	100	100	100	100

Table 1: The percent of correct decision probability of the first neural network

SNR	0 dB	5 dB	15 dB	25 dB	35 dB	45 dB	55 dB
ASK8	15	60	80	100	100	100	100
PSK8	45	70	90	95	100	100	100
FSK8	30	75	87.5	97	100	100	100
MSK	20	60	95	100	100	100	100

Table 2: The correct decision probability of the second class

	AM	LSSB	USSB	FM	ASK2	ASK4	PSK2	PSK4	FSK2	FSK4
NANDI	88.5	99.8	98.5	90.1	96.8	86.5	99.5	96.8	99	99.5
OUR	90	90	95	90	100	100	85	95	100	100

Table 3: Comparison of our method with Nandi and Azzouz at 15 dB SNR

6. COCLUSIONS

We have developed a new method based on some of previously used features to recognize a new set of modulation schemes. The classification procedure is done by the hierarchical neural network method. The backpropagation training method with a variable learning rate is used. As it is clear from the table1, the overall performance of the method used in this paper is above 75% even in SNR as low as 15 dB. For SNR above 35 dB the performance reaches 100%. Although we considered nine communication systems in our project, expanding the number of protocols is straight forward using the principals given here.

REFERENCES

[1] S. S. Soliman and S. Hsue, "Signal classification using statistical moments," *IEEE Trans. On Communication*, vol. 40, Issue 5, pp. 908–916, May. 1992.

[2] N. Ghani and R. Lamontagne, "Neural networks applied to the classification of spectral features for automatic modulation recognition," *MILCOM '93*, Oct. 1993, vol. 1, pp.111-115.

[3] C. Louis and P. Sehier, "Automatic modulation recognition with hierarchical neural networks," *MILCOM '94*, Oct. 1994, vol. 3, pp.713-717.

[4] S. Taira, "Automatic classification of QAM signals by neural networks," in *Proc. ICASSP 2001*, May 2001, vol. 2, pp. 1309–1312.

[5] A. Polydoros and K. Kim, "On detection and classification of quadrature digital modulations in broad-band noise," *IEEE Trans. On Communication*, vol. 38, No. 8, pp. 1199–1211, Aug. 1990.

[6] P. C. Sapiano and J. D. Martin, "Maximum likelihood PSK classifier," *MILCOM '96*, Oct. 1996, vol. 3, pp.1010-1014.

[7] F. Delgosh, "Digital modulation recognition", Master's Thesis, Sharif Industrial university, 1998

[8] A. B. Carlson, P. B. Crilly and J. C. Rutledge, *communication systems: McGraw Hill*, Fourth edition, 2001.

[9] J. G. Proakis, *Digital communication: McGraw Hill international edition*, Second edition, 1989.

[10] E. G. Azzouz and A. K. Nandi, *Automatic modulation recognition of communication signals: Kluwer Academic Publishers*, Boston, 1996.

[11] S. Z. Hsue and S. S. Soliman, "Automatic modulation recognition of digitally modulated signals," *MILCOM '89*, Oct. 1989, vol. 3, pp. 645–649.

[12] A. W. Wegener, "Practical techniques for baud rate estimation," *ICASSP-92*, Mar. 1992, vol. 4, pp.681-684.

Multifunctional unit for Rayleigh Fading Channels combining Huffman and Convolutional decoding, MPSK demodulation, Synchronisation and a Statistical Channel Evaluation Technique

K.M.S. Soyjaudah, B. Rajkumarsingh

Electrical and Electronic Engineering Department, University of Mauritius
Reduit, Mauritius

Fax: (230) 465 7144

Tel: (230) 454 1041

Email: s.soyjaudah@uom.ac.mu

Technical Subject Category : Communication theory

Abstract

With the advent of modern high speed Digital Signal Processing (DSP) devices which support parallel processing, it is possible to combine many features of a digital communications receiver into a multifunctional unit. This paper describes a system which combines the functions of demodulation, error control, source decoding, synchronisation and channel estimation into a single functional unit. A framework employing the “multifunctional coding” concept has been implemented which will be adaptively controlled by making use of Statistical Channel Evaluation (SCE) data. Results of the simulation over a slow fading channel are presented.

1. Introduction

The adaptive system described in this paper is an application of multifunctional coding [1] in that it involves a combination of source decoding (Huffman code), error control (convolutional codes), demodulation (Multiple Phase Shift Keying (MPSK) in Rayleigh Fading channels adaptively controlled by making use of SCE data [2,3,4]. Such SCE information can be derived from the elements of the communication system itself, example the demodulator, decoder, synchroniser, etc. In this work, channel estimation and synchronisation is obtained through the use of soft maximum likelihood trellis decoding (SMLD) of the received signal. Synchronisation can be achieved intrinsically by using confidence metrics derived from the decoding of error codes. One such method is the Code Assisted Bit Synchronisation (CABS) algorithm [5] for convolutional codes based on trellis decoding. Based on the paper [6], decoders are designed that combine the processes of source and channel decoding. The latter, besides giving the same performance as the separate counterparts, has the advantage of reducing hardware complexity in implementation. A fixed rate code that can correct a fixed number of errors in a block of bits is inflexible and limits a system to a fixed throughput under all conditions [7-8]. Due to the time varying nature of the Rayleigh channel fading, the parity bits are not optimally used. As

channel conditions improve the fixed number of parity bits sent are greater than is required to correct the errors found. The excess parity bits are therefore not serving any useful purpose and they lower the data rate or throughput [9]. Prediction of the “deep fades” which occur in the channel can be used to obtain a more flexible and efficient system through the use of an adaptive scheme. The adaptive scheme selects the appropriate modulation scheme which also leads to an optimisation of channel bandwidth.

2. Combined Huffman, Convolutional codes and MPSK modulation

In this proposed scheme, when a character is read its equivalent Huffman codeword is obtained from its state diagram [10] and the latter encoded using the convolutional encoder. After encoding, depending on the MPSK scheme, the corresponding number of codeword bits are mapped onto a signal and transmitted as a symbol. The vectorial coefficients for each transmitted signal are calculated. As an example for QPSK modulation scheme, two bits in the codeword are taken at a time to form a symbol. The four possible encoded bits, mappings and vectorial coefficients are shown in Table 1.

Encoded Bits	Mapping	Vectorial Coefficients
00	S0	(1,0)
01	S1	(0,1)
10	S2	(-1,0)
11	S3	(0,-1)

Table 1: Coded bits, mapping and vectorial coefficients used in QPSK modulation scheme

2.1 Trellis design of Combined Huffman and rate $\frac{1}{2}$ convolutional codes employing QPSK modulation.

To design the combined Huffman, convolutional decoder and demodulator, a state diagram of the

Huffman code is first applied to a text file and the corresponding codes obtained as shown in Table 2.

Character	Huffman Code	Convolutional code
a	0	00
b	111	110110
c	101	111000
d	1101	11011000
e	1100	11011011
f	100	111011

Table 2. Huffman code for a data source

These data are then used to construct a state diagram given in Figure 1. Designing a decoder that performs source, channel decoding and demodulation, is actually the design of a new trellis. The new trellis has the properties of the Huffman tree, the trellis used for convolutional decoding and MPSK modulation. A trellis for a (2,1,2) convolutional code, with the following generator polynomial: $g^{(1)} = (1 \ 1 \ 1)$, $g^{(2)} = (1 \ 0 \ 1)$ is illustrated as an example.

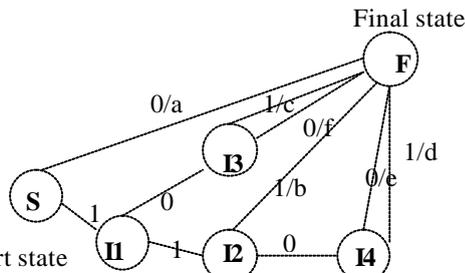


Figure 1: State diagram for the Huffman code in Table 2.

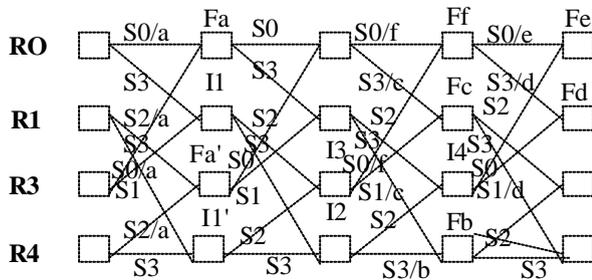


Figure 2 : Combined Huffman state and QPSK trellis diagram

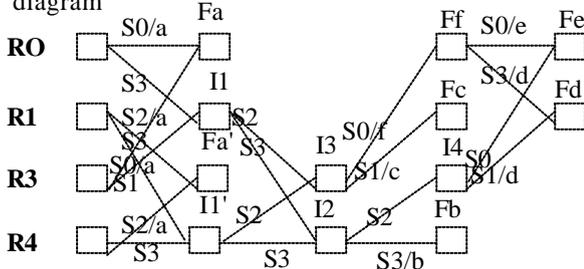


Figure 3. Combined Huffman and trellis diagram in reduced form

In order to realise this new trellis, we first superimpose the state diagram of the Huffman code on the trellis diagram as shown in Figure 2. The start state of the Huffman code can be associated with any

one of the states ($R0$ to $R4$) of the trellis diagram.

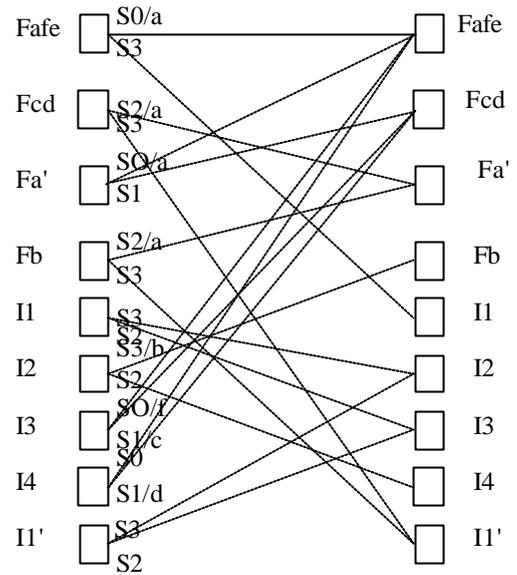


Figure 4 : Final trellis diagram for combined decoder

If the first symbol is a '00', a transition occurs from $R0$ to Fa leaving character 'a' as the output. The QPSK mapping in Table 1 shows that this symbol is transmitted by signal $S0$. In Figure 2, it is observed that not all the paths are being used. Hence this trellis can be reduced by eliminating these unused paths to get a reduced state diagram shown in Figure 3. This state diagram can further be reduced by the fact that a start state and a final can be considered to be the same state for a stream of data. Thus one start state $S0$ and three final state Fa , Ff and Fe can be considered to be a single state and is labelled $Fafe$. The final trellis diagram for the combined decoder is shown in Figure 4. The combined decoding and demodulation is just less complex to be implemented due to the fact that some states are deleted. It is also less complex in terms of mathematical computations.

3. Code Assisted bit Synchronisation (CABS) algorithm

The CABS technique, shown in Figure 5, is a form of intrinsic synchronisation in which the process of demodulation and decoding are combined, and symbol timing is derived from the viterbi decoder which also performs error correction of the received data. The starting point of the first symbol is decided in an arbitrary manner. The demodulator outputs, corresponding to one symbol interval, are passed to the decoder, which updates a buffer containing $L \times N$ trellis sections, where L is the search length of the decoder and N is the number of samples taken per symbol interval. The decoder operates on L trellis sections separated in time by N samples as shown in Figure 5. At each decoding step, the minimum path metric through the trellis is found and this is stored

along with the corresponding symbol, When all N decoding steps have been performed the minimum metric value among the N possible values indicates the synchronisation point and its corresponding decoded bit is the output from the decoder. N has been assigned a value of 4 in our simulations. Symbol timing information is then derived from the time shift associated with the chosen path.

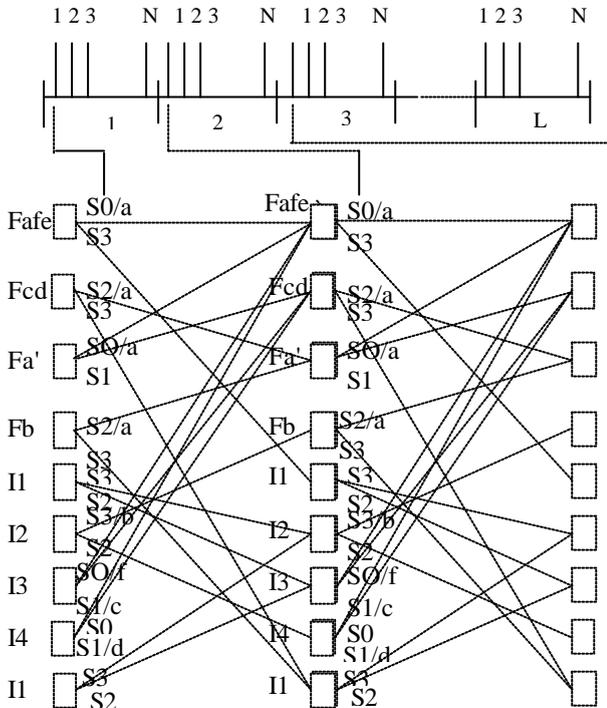


Figure 5 : CABS decoding procedure

4. SCE Techniques

Distortion imposed by the channel on the data stream causes errors at the receiver of a digital communication system. It is important to minimise the number of these errors and to maximise the throughput of the system. In order to optimise the system performance adaptively in response to channel conditions, an estimate of the receiver's error rate is required to initiate control actions. Statistical estimation techniques are useful tools for obtaining an on-line estimate of the channel state. As a by-product of the Viterbi decoding algorithm, the cumulative metric of the most likely path (first path) through the decoder trellis is available as an additional information besides the decoded output symbol. The SCE (error limiter) [2] retrieves this information and causes the decoder to make an additional pass through the trellis in search of the greatest likelihood error path (second path). The difference in the euclidean distance between these two paths is interpreted as a measure for the signal-to-noise ratio (SNR) in the transmission channel. As an example, the decoder metrics of rate 1/3 code is shown in Figure 6. Consequently the error probability of the decoded sequence could be

estimated. The error limiter averages the error detection metrics over a number of frames of data.

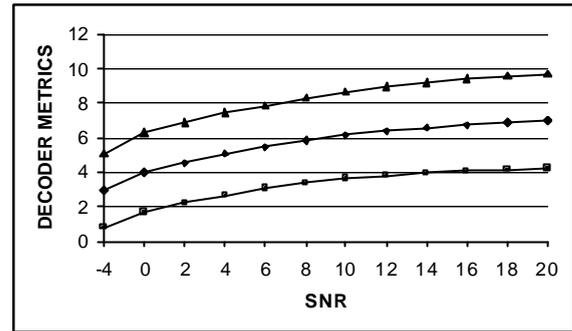


Figure 6. Mean and standard deviation of the error detection metrics vs SNR

The average is then compared with a set of previously obtained thresholds. The thresholds are arranged such that the overall output probability of error of the system is kept below a specified maximum value.

5.The Adaptive CABS Algorithm

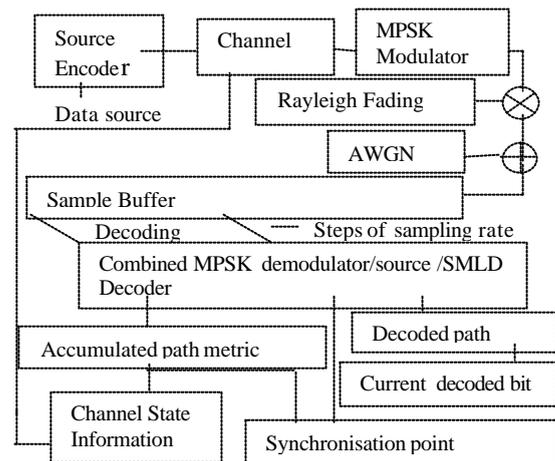


Figure 7. Block diagram of the proposed scheme

A simplified model of the proposed scheme is shown in Figure 7. The Viterbi decoder operates with a CABS algorithm to produce simultaneous demodulation, source decoding, error decoding and timing estimates. Metrics of the decoding process are then used to obtain Channel State Information (CSI). The transmitter receives the CSI and chooses a suitable modulation scheme. The modulation scheme is chosen according to a given set of thresholds in order to keep the BER below a certain level. The thresholds have been calculated for a BER below 1.0×10^{-4} [9]. The modulated baseband signals are then sent over the channel. The sources of distortion are assumed to be the AWGN and attenuation due to fading. With the CSI information the demodulator then recovers the components of the baseband signals [11,12]. Finally the combined CABS, source, channel decoder and demodulator restores the data sent. There are three possible schemes from which to

choose: code 1/3 with BPSK modulation, code 1/2 with QPSK modulation, code 2/3 with 8PSK modulation. Rate 1/3 with BPSK is used during the worst channel conditions and code 2/3 with 8PSK during the best channel conditions.

6. Simulations and results

Figures 9 and 10 compare the BER and throughput performance of the adaptive CABS scheme with the individual fixed rate error correcting codes and the ideal decoder. As can be seen the overall performance of the adaptive CABS system is an improvement over all the of the individual codes tested. For low SNR the adaptive scheme closely follow the rate 1/3 code in error performance as expected. This is the code with the highest error correcting ability of the set of codes available in this adaptive scheme. It can also be seen from this plot that the adaptive CABS system performs closely to the ideal decoder that it is almost indistinguishable.

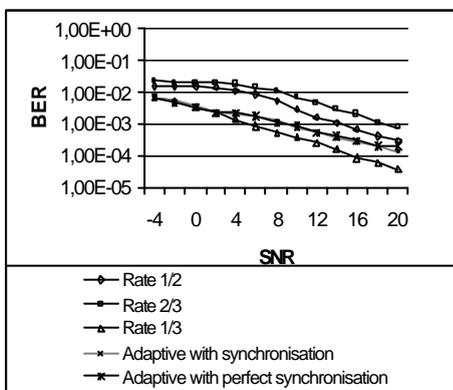


Figure 9. Comparison of performance of adaptive and individual

Although the adaptive algorithm more frequently selects the rate 1/3 code when the channel is bad, it switches to a higher rate code when the channel momentarily enters a good state. This leads to an overall improved throughput, as can be seen in Figure 10.

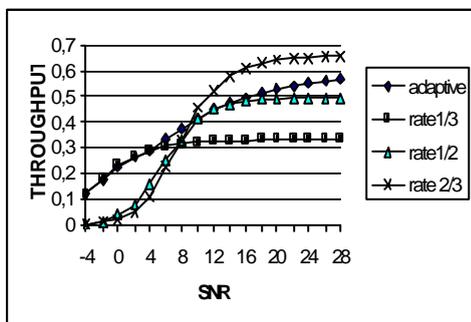


Figure 10. Comparison of throughput of the adaptive and the individual schemes

Thus a higher throughput is attained during very bad channel conditions than with a fixed rate code.

7. Conclusions

In this paper, an adaptive CABS system, controlled by SCE data, employing combined source, channel decoding, demodulation and synchronisation based on the multifunctional concept has been designed. The proposed synchronisation scheme provides good synchronisation close to the perfect case. Hardware and transmission overhead are also reduced. The adaptive scheme explores the time varying nature of Rayleigh fading channels and the required error performance characteristics can be specified and attained with significant improvement in the throughput. Moreover, since the different modulation schemes are adopted according to the modulation scheme, channel bandwidth is optimised.

References

- [1] Darnell M., Honary B., "Multifunctional coding schemes applicable to secure communications. Proc. IEE International Conference on Secure Communications, Conf. Proc. No. 269, 1986.
- [2] Zolghadr F., "Embedded coding algorithms applicable to time varying channels", Phd. Thesis, University of Warwick, Nov. 1989.
- [3] Arani F., Coulton P., Honary B., "Real-Time channel evaluation combined with decoding and synchronisation", IEEE Globecom '95 Telecommunications Conference, Singapore, Nov, 1995, vol 2, pp1343.
- [4] Shaw M., Honary B., Darnell M., "An RTCE assisted ARQ transmission Scheme: Design and Implementation", IEE Conf. Proc. On HF Communication, Vol. 284, pp. 43-53, 1988.
- [5] Honary B., Zolghadr F., and Darnell M., "Code Assisted Bit Synchronisation (CABS) Scheme", Electronic letters, Vol. 25, No. 2, January 1989.
- [6] Honary B. and Shaw M., "Combined Huffman and convolutional codes: Concatenated form without feedback", Electronic Letters 15th March 1990 Vol.26 No.6.
- [7] Bate D.B., "Adaptive coding algorithms for data transmission" P.h.D thesis Coventry Polytechnic, February 1992.
- [8] Honary B., Darnell., Vongas G., "Adaptive error control schemes for 2-200 MHz multiple-mechanism propagation paths", AGARD meeting October 1993.
- [9] Soyjaudah S., Rajkumarsingh B., "Adaptive Coding and Modulation Using Reed Solomon Codes for Rayleigh Fading Channels", Proceedings Eurocon 2001, Bratislava, Slovakia.
- [10] Huffman D.A., "A method for the construction of minimum redundancy codes", Proc. IRE, Vol. 40, pp 1098-1101, Sept. 1952.
- [11] Jakes W., "Microwave Mobile Communications", New York: Wiley, 1974, chapt. 1.
- [12] Alamouti S.M. and Kallel S., "Adaptive Trellis-Coded Multiple-Phase-Shift-Keying For Rayleigh Fading Channels", IEEE trans. Commun., vol. 42, No. 6, pp. 2305-2314, June 1994.

MULTILEVEL HARQ SCHEMES OVER WIRELESS CHANNELS

Qinglin Luo¹, Peter Sweeney²

¹ University of Surrey, GU2 7XH, Guildford, United Kingdom, q.luo@surrey.ac.uk

² University of Surrey, GU2 7XH, Guildford, United Kingdom, p.sweeney@surrey.ac.uk

Abstract –We propose three bandwidth efficient HARQ schemes based on multilevel coded modulation in this paper. These schemes combine bandwidth efficiency of coded modulation with power efficiency of forward error controlling to provide high efficient transmission. Simulations are carried out over wireless channels. Both throughput and reliability are found to be improved by employing HARQ schemes based on multilevel coded modulation instead of conventional trellis coded modulation. **Keywords** – Multilevel Coded Modulation, Hybrid ARQ, TCM

I. INTRODUCTION

HYBRID automatic-repeat-request (HARQ) that combine ARQ protocols with forward-error-correcting (FEC) schemes is generally considered to be the best error-control techniques for wireless applications[1]. Compared with standard ARQ, HARQ can improve the throughput of a system with the error correcting ability of FEC codes. Conventional codes such as block codes or convolutional codes can be used to construct HARQ schemes.

In HARQ, coded modulation is a standard method to prevent bandwidth expansion and achieve relatively high coding gains. Trellis coded modulation (TCM) [2] is the best known coded modulation schemes. Multilevel coded modulation (MLC) [3][4] is usually separately treated from standard TCM although they are proved to be same in principle [4] because of many significant features such as multistage decoding, unequal error protecting, etc.

HARQ schemes based on TCM have been investigated by many authors [5][6][7]. In this paper, we investigate selective-repeat type HARQ based on MLC and show that HARQ based on MLC with proper design can outperform that based on TCM. Various HARQ schemes based on MLC are constructed to suit for different requirements owing to the high flexibility of multilevel structure. Code combining is also found to be differently helpful to multilevel HARQ schemes.

This paper is organised as follows. Section II contains some brief description on overall system. HARQ schemes termed Scheme I, Scheme II and Scheme III are proposed in Section III. Section IV is devoted to the performance evaluation of all schemes. Some comparisons and analysis based on simulation results are presented in Section V. Section VI concludes this paper.

II. SYSTEM DESCRIPTION

A. HARQ Assumptions

We assume a selective repeat ARQ (SR-ARQ) scheme with suitably large buffers at the transmitter and receiver. Furthermore, we assume an error free feedback channel over which negative acknowledgements (NACK) can be sent.

B. Multilevel encoder and multistage decoder

The multilevel encoder and multistage decoder are sketched in Fig1 and Fig.2, respectively. First, a block of L binary source data bits $\mathbf{q} = (q_1, \dots, q_L)$, $q \in \{0,1\}$, is partitioned into M blocks $\mathbf{q}_i = (q_{i,1}, \dots, q_{i,l_i})$, $i = 0, \dots, M-1$, of length l_i , where

$$l_i = L \frac{R_i}{R}, \quad R = \sum_{i=0}^{M-1} R_i \quad (1)$$

and R_i is the code rate on level i , R is the overall code rate. \mathbf{q}_i is encoded by component binary encoder C_i generating words $\mathbf{x}_i = (x_{i,0}, \dots, x_{i,N-1})$, $x_{i,n} \in \{0,1\}$, $N = l_i / R_i$. At one time instant n , M bits from encoder outputs on all levels form a binary label $S_n = (x_{0,n}, \dots, x_{M-1,n})$, which is mapped onto a point \mathbf{s}_n represented by a pair of real-valued symbols $(s_{n,I}, s_{n,Q})$ on constellation size of $2^M - 1$, according to given mapping scheme and then sent to channels.

In the multistage decoder, at level i , a soft detector (8PSK demodulator) first calculates likelihood ratios of \mathbf{x}_i based on received symbol sequence \mathbf{y} and side information from lower levels $(\mathbf{x}'_0, \dots, \mathbf{x}'_{i-1})$,

$$L(x_{i,n}) = \ln \left[\frac{\sum_{k=0}^{2^M-1} P(\mathbf{s}_k, \text{where } S_{k,i} = 1 | \mathbf{x}'_0, \dots, \mathbf{x}'_{i-1}, \mathbf{y}_n)}{\sum_{k=0}^{2^M-1} P(\mathbf{s}_k, \text{where } S_{k,i} = 0 | \mathbf{x}'_0, \dots, \mathbf{x}'_{i-1}, \mathbf{y}_n)} \right] \quad (2)$$

And then send the result into a binary decoder corresponding to the encoder C_i . The decoding result is buffered as part of the overall output, and sent to the detectors of all higher levels simultaneously. At the end of

the decoding process, hard decision results based on $\hat{\mathbf{q}}_i, i=0, \dots, M-1$ from all levels are multiplexed to form the overall result $\hat{\mathbf{q}}$.

C. Wireless Channel Model

A coherent receiver model is considered. The in-phase and quadrature demodulator outputs are,

$$\begin{aligned} \mathbf{y}_{I,n} &= a_n s_{I,n} + G_{I,n} \\ \mathbf{y}_{Q,n} &= a_n s_{Q,n} + G_{Q,n} \end{aligned} \quad (3)$$

where a_n is a Rayleigh random variable representing the fading statistics with $E[a_n^2]=1$, and $G_{I,n}, G_{Q,n}$ are two uncorrelated Gaussian noise variables with zero mean and variance \mathbf{s}^2 , which are independent of the fading variable a_n .

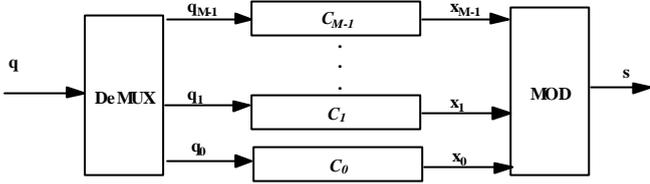


Fig.1. Multilevel Encoder

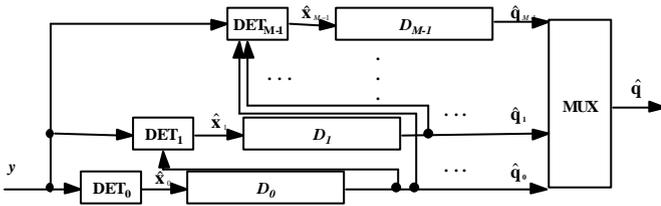


Fig.2 Multistage Decoder

III. MULTILEVEL HARQ SCHEMES

A. Scheme I

In this scheme, all levels share one CRC. If the CRC bits are attached at the end of a packet, the newly formed sequence including the CRC is $(q_{0,1}, \dots, q_{M-1,l_{M-1}-r}, c_1, \dots, c_r)$, where r is the CRC length. In the receiver, according to CRC checking result, one quality acknowledgement bit ($NACK$) is fed back to inform the transmitter about the receiving state of the whole packet. The sequence $(q_{0,1}, \dots, q_{M-1,l_{M-1}-r}, c_1, \dots, c_r)$ buffered in the transmitter will be retransmitted if $NACK$ is nonzero and the number of transmission attempts is less than the pre-set maximum value.

B. Scheme II

In this scheme, every level has its own CRC bits. The newly formed sequence after CRC attaching on level i is $(q_{i,1}, \dots, q_{i,l_i-r_i}, c_{i,1}, \dots, c_{i,r_i})$, ($i=0, \dots, M-1$), where r_i is the CRC bits length on level i . In the receiver, M bits of acknowledgement bits ($NACK_0, \dots, NACK_{M-1}$) are fed back to inform the transmitter about the receiving state of every level after CRC checking. For any single level i , the buffered sequence \mathbf{q}_i will be retransmitted if $NACK_i$ is nonzero and the number of transmission attempts is less than the pre-set maximum value.

C. Scheme III

The basic idea of this scheme is to make use of the unequal error protecting ability of different levels. Based on Scheme II, if retransmission of \mathbf{q}_i is requested, the priority of \mathbf{q}_i to be retransmitted on a better error-protected level will be increased. $NACK_i$ is used as the priority indicator of \mathbf{q}_i . In the receiver, after decoding on all levels is finished, the frames on all levels are restored into their original order, CRC on \mathbf{q}_i is checked and then operation on $NACK_i$ is carried out: if retransmission is requested, $NACK_i$ is increased by 1, otherwise it is reset to 0.

The packet structures of a three-level HARQ system after CRC attaching in Scheme I, II, and III are shown as Fig.3.

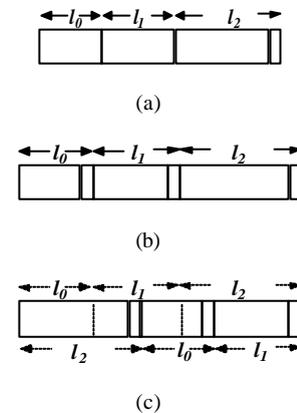


Fig. 3. Frame structures before encoding. a) Scheme I, b) Scheme II, c) Scheme III with $NACK_3 > NACK_1 > NACK_2$

IV. PERFORMANCE EVALUATION

Throughput and reliability are two basic performance parameters for an ARQ system. In this paper, throughput is defined as the average number of encoded data packets **accepted** by the receiver in the time that it takes the transmitter to send a single information packet.

Let T_r denote the average number of times a packet must be transmitted before it is accepted. Under the assumption of infinite buffering, $T_r = 1/(1 - P_d)$, where P_d is the probability of the events that a packet contains detectable errors, also the probability of the events that a retransmission is requested. Then the throughput of Scheme I can be written as,

$$\mathbf{h} = \frac{L-r}{L} \cdot \frac{R}{T_r} = \frac{L-r}{L} \cdot R \cdot (1 - P_d) \quad (4)$$

In Scheme II and III, each level can be looked as an independent ARQ channel, whose throughput can be expressed in the same form as equation (4). The overall throughput

$$\mathbf{h} = \sum_{i=0}^{M-1} \mathbf{h}_i = \sum_{i=0}^{M-1} \frac{l_i - r_i}{l_i} \cdot R_i \cdot (1 - P_{d,i}) \quad (5)$$

where variables with subscript i means the variable is for level i .

Reliability of an ARQ system is usually expressed in terms of the rate at which packets containing errors are accepted, i.e., the packet error rate (PER). In many cases, it is also expressed in terms of the bit error rate (BER) of the accepted packets. Here, we adopt BER terms.

For Scheme I, the BER is

$$P_b = \frac{n_u}{L} \quad (6)$$

where n_u is the number of undetectable erroneous bits in the packet accepted by the receiver.

For Scheme II and III, similar to the derivation of the throughput, the BER is

$$P_b = \sum_{i=0}^{M-1} P_{b,i} \cdot R_i = \frac{\sum_{i=0}^{M-1} n_{u,i}}{L} \quad (7)$$

The above equations give some insight into the evaluation of the performance of multilevel HARQ schemes. In the following section, some evaluation results based on simulations on an example system are presented.

V. EXAMPLE AND SIMULATION RESULTS

Rate compatible convolutional codes have been successfully applied into both multilevel coded modulation and HARQ in many previous works [8][9]. Since FEC in an HARQ system usually works at relatively small E_b/N_0 , and, at small E_b/N_0 the error performance of a systematic code is better than that of a non-systematic code having the same memory [10], we employ systematic convolutional codes as the component codes in our multilevel system.

The bandwidth efficiency of the three-level code is set to be 2 bits/symbol, which is distributed among three levels as 0.4, 0.75, 0.85. This distribution is decided first by calculating following the capacity rule proposed in [4] and then adjusting slightly according to simulation results so as to minimize the overall BER. Corresponding to this distribution, the mother code rates on level 0, 1, 2 are chosen to be 1/3, 1/2, 1/2, and generators are $(37, 25, 21)_{\text{oct}}$, $(37, 25)_{\text{oct}}$, $(37, 25)_{\text{oct}}$. Desired rates are obtained by puncturing [11].

Ungerboeck's set partitioning is used when the 8PSK modulator maps encoded data onto constellation. A CRC with generator $x^{16} + x^{12} + x^5 + 1$ is employed wherever required. The block size L is set to be 1024. All simulations are run over independent Rayleigh channels as described in Section II.

As a comparison, we also set up a convolutional TCM HARQ system based on a 16-state convolutional encoder and 8PSK modulation with Gray mapping. Same block size and same CRC generator is used as in multilevel HARQ. The throughput and BER curves of all systems are plotted in Fig. 4 and Fig.5.

From Fig.4, it can be observed that TCM HARQ can match multilevel HARQ Scheme I only for very low throughput. E_s/N_0 gain of at most 2 dB is observed for medium to high throughput by using multilevel HARQ Scheme I instead of TCM HARQ. Scheme II and III have the same throughput performance, which outperforms TCM HARQ by about 0.8 dB more than Scheme I due to the asynchronous retransmission strategy. For very high E_s/N_0 , the throughput curves of Scheme I and TCM HARQ approach the same value, which is higher than the value approached by Scheme II and III because fewer CRC bits are used.

In Fig.5, for the BER of 10^{-5} , gains of 4 dB, 1.5 dB, 4.5 dB are observed for multilevel HARQ Scheme I, II and III over TCM HARQ, respectively. Multilevel HARQ Scheme I outperforms TCM HARQ significantly for medium BER ($10^{-3} - 10^{-4}$). Scheme II outperforms Scheme I for low to medium BER because all levels are more closely 'watched' by their own CRCs in Scheme II, and fewer errors escape the CRC detection [12]. But it is outperformed for high BER because retransmissions of data on levels with undetectable errors in Scheme I help those errors to be detected in the next attempt at transmission. Scheme III performs equally with Scheme I for low to medium BER, but outperforms Scheme I significantly for high BER because the re-allocation to more powerful levels of some retransmitted data provides higher opportunity for the errors in them to be corrected.

REFERENCES

- [1] E. Ayanoglu, K. Y. Eng, and M. J. Karol, "Wireless ATM: Limits, challenges, and proposals," *IEEE Pers. Commun.*, pp. 18-33, Aug. 1996.
- [2] G. Ungerboeck, "Channel coding with multilevel/phase signaling," *IEEE Trans. Inform Theory*, vol. 25, no. 1, Jan. 1982, pp. 55-67
- [3] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inf. Theory*, vol. 23, pp.371-377, May 1977
- [4] U. Wachsmann, R. F. H. Fischer, and J. B. Hubber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1361-1391, July 1999
- [5] R. H. Deng, "Hybrid ARQ schemes employing coded modulation and sequence combining," *IEEE Trans. Commun.*, vol. 42, pp. 2239-2245, 1994
- [6] L. K. Rasmussen and S. B. Wicker, "The performance of type-I trellis coded hybrid-ARQ protocols over AWGN and slowly fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 418-428, 1994
- [7] A. Banerjee, D. J. Costello, Jr. and T. E. Fujia, "Performance of hybrid ARQ schemes using turbo trellis coded modulation for wireless channels", *Proceedings of the IEEE Wireless Communications and Networking Conference*, Chicago, Illinois, September 2000.
- [8] J. Hagenauer, "Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and their Applications," *IEEE Trans. Commun.*, COM-36, No. 4: 389-400, Apr. 1988.
- [9] S. Lin, D. J. Costello, Jr., and M. J. Miller, "Automatic-repeat-request error control schemes," *IEEE Commun. Mag.*, vol. 22, pp. 5-16, Dec. 1984.
- [10] C. Berrou, A. Glavieux, and P. Thitimajhima, "Near Shannon limit error correcting coding and decoding: Turbo-codes," in *Proc. Of IEEE Int. Conf. On Commun.*, pp. 1064-1070, May 1993.
- [11] 3GPP TSG-RAN "Multiplexing and channel coding (FDD)" 3GPP TS 25.212 v5.0.0
- [12] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

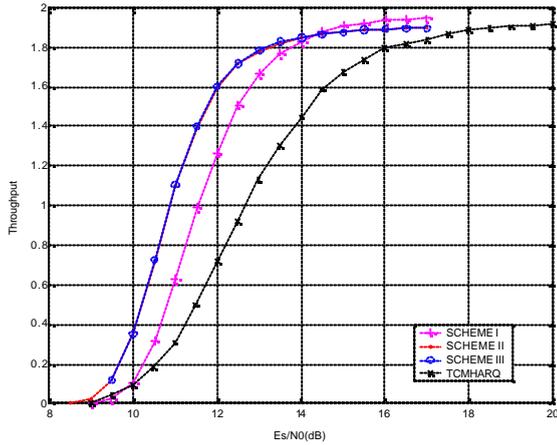


Fig.4 Throughput comparison among various HARQ schemes over independent Rayleigh fading channel

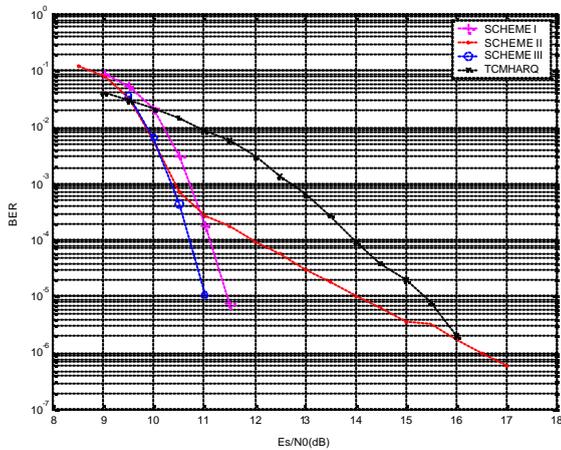


Fig.5 BER comparison among various HARQ schemes over independent Rayleigh fading channel

VI. CONCLUSION

Hybrid ARQ schemes termed Scheme I, Scheme II, Scheme III based on multilevel coded modulation are proposed in this paper. A three-level convolutional coded hybrid ARQ system over independent Rayleigh channel is explored as an example. Simulations on the system and on a convolutional trellis coded HARQ system with comparable configuration show that HARQ system based on multilevel coded modulation can perform better than that based on trellis coded modulation over wireless channels. Among multilevel HARQ schemes, Scheme III is the best choice for systems over wireless channels, either in sense of throughput or in sense of reliability.

Future work will be concentrated on applying code combining into all Schemes, and analyzing the influence of code combining on rate design rules in multilevel HARQ.

New Accurate Upperbounds for Maximum Likelihood Decoder of STBC

G.R. Mohammad-Khani, V. Meghdadi, J.P. Cances and G. Ferré

University of Limoges, ENSIL-GESTE, Parc ESTER Technopole, 87068 Limoges, France

E-mail : {reza_meghdadi, cances, ferre} @ensil.unilim.fr

ABSTRACT

In this paper we propose new decoding rules for STBC (Space-Time Block Codes) using Maximum Likelihood (ML) principles. Using algebraic tools we are able to derive a simple matrix model which is very easy to implement in simulation. After giving our particular detection rules we determine a new accurate upperbound for the BER performances of STBC transmitting schemes. Simulation results are given which show the accuracy of the derived upperbound.

I. INTRODUCTION

Transmit diversity has been studied extensively as a method of combating impairments in wireless fading channels [1-4]. STTC (Space-Time Trellis Coding) was at first introduced to combine signal processing at the receiver with coding techniques appropriate to multiple transmit antennas. Specific STTC designed for 2-4 transmit antennas perform extremely well in slow-fading environments and come close to the outage capacity [5-6]. However, the main drawback of such encoding schemes is the complexity of the decoding Viterbi based algorithm. To overcome the prohibitive complexity of STTC, simple diversity transmission schemes based on the construction of orthogonal matrices have been proposed [7-9]. The construction of orthogonal matrices leads to STBC techniques.

In this paper, we propose at first a generalized framework for the ML decoding of STBC. We then derive a matrix model particularly well suited to the description of MIMO (Multi-Input Multi-Output) systems. The obtained model has remarkably simple ML decoding algorithms based only on linear processing at the receiver. Using this matrix modeling, we find a new accurate upperbound for the BER performances of STBC transmitting schemes. Simulation results are given which show the accuracy of the derived upperbound.

II. SYSTEM DESCRIPTION

Fig. 1 illustrates the studied MIMO system, in which the mobile station is equipped with n transmit antennas and the base station is equipped with m receive antennas. The STBC coding rate is assumed equal to k/l (each packet (x_1, \dots, x_k) of k symbols corresponds to l STBC transmitted symbols (c_1^i, \dots, c_l^i) at the i th transmit antenna).

The coefficient $\alpha_{i,j}$ is the path gain from transmit antenna i to receive antenna j . They are modeled as samples of independent complex Gaussian random variables with variance 0.5 per complex dimension. At time t the received signal r_t^j at antenna j is given by

$$r_t^j = \sum_{i=1}^n \alpha_{i,j} c_t^i + n_t^j \quad (1)$$

where n_t^j are independent samples of a zero-mean complex Gaussian noise with variance $\sigma^2/2 = 1/2SNR$ per complex dimension. The average energy of the symbols transmitted from each antenna is normalized to be $1/n$. In order to derive a matrix form for equation (1), we define at first the vector $\mathbf{r}^j = [r_1^j \dots r_l^j]$

$$\mathbf{r}^j = \begin{bmatrix} \alpha_{1,j} & \dots & \alpha_{n,j} \end{bmatrix} \begin{bmatrix} c_1^1 & \dots & c_l^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \dots & c_l^n \end{bmatrix} + \mathbf{n}^j = \mathbf{h}^j \mathbf{C} + \mathbf{n}^j \quad (2)$$

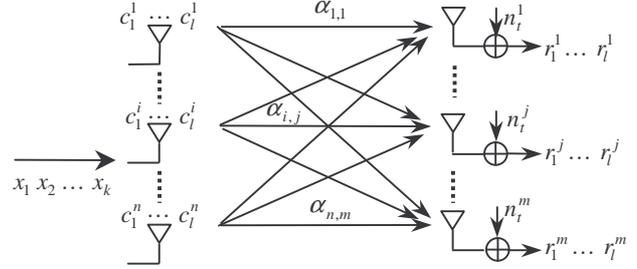


Figure 1: System Model of MIMO-STBC

\mathbf{h}^j contains the multipath gains between any transmit antenna and receive antenna j . \mathbf{C} is the block code matrix and $\mathbf{n}^j = [n_1^j \dots n_l^j]$ is a noise vector with auto-correlation matrix

$$E\{\mathbf{n}^j \mathbf{n}^{j\dagger}\} = \sigma^2 \mathbf{I}_{l \times l}$$

Considering the concatenation of all the vectors \mathbf{r}^j we have:

$$\mathbf{r} = \mathbf{H} \mathbf{C} + \mathbf{n} \quad (3)$$

where $\mathbf{r} = \begin{bmatrix} (\mathbf{r}^1)^T & \dots & (\mathbf{r}^m)^T \end{bmatrix}^T$, $\mathbf{H} = \begin{bmatrix} (\mathbf{h}^1)^T & \dots & (\mathbf{h}^m)^T \end{bmatrix}^T$ and

$\mathbf{n} = \begin{bmatrix} (\mathbf{n}^1)^T & \dots & (\mathbf{n}^m)^T \end{bmatrix}^T$. We suppose here that the channel parameters

are all known at the receiver. In this case the ML detection rule at the receiver can be summarized by equation (4).

$$\text{ML: } \underset{\mathbf{C}}{\text{Min}} \Omega = \sum_{j=1}^m \sum_{l=1}^l |r_l^j - \sum_{i=1}^n \alpha_{i,j} c_l^i|^2 \quad (4)$$

The metric Ω in detection rule (4) can be rewritten using (2)

$$\Omega = - \underbrace{\sum_{j=1}^m (\mathbf{r}^j \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger)}_{\Omega 1} + \underbrace{\sum_{j=1}^m \mathbf{h}^j \mathbf{C} \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger}_{\Omega 2} \quad (5)$$

It is not possible to further simplify (5) without the knowledge of the STBC structure. However, (5) implies that the metrics have to be computed for all possible codewords that means for all the possible combinations of k symbols. For example, in the case of a M -ary modulation, implies that the decoder has to check for M^k symbols.

III. COMPLEX STBC ORTHOGONAL DESIGN

Now, we show that using STBC orthogonal design enables to greater reduce the complexity of detection rule (5). We consider the most general case of complex orthogonal designs but the real case can be easily adapted to our description.

Encoding scheme :

The complex matrix $\mathbf{C}_{n \times l}$ represents a complex STBC orthogonal design of size n if it satisfies the following conditions :

1-Its elements are linear combinations of the k symbols (x_1, x_2, \dots, x_k) or their conjugates.

2- $\mathbf{C} \mathbf{C}^\dagger = \sum_{i=1}^k |x_i|^2 \mathbf{I}_{n \times n}$ These conditions imply that $n \leq l$.

We impose no condition on the elements of \mathbf{C} and this definition of orthogonality contains the examples well known in the literature [8], even the Few Sporadic Codes. According to the definition of orthogonal code, it is possible to write it in the following form:

$$\mathbf{C}_{n \times l} = \sum_{i=1}^k (x_i \mathbf{A}_i + x_i^* \mathbf{B}_i) \quad (6)$$

where \mathbf{A}_i and \mathbf{B}_i are constant matrices $n \times l$. The orthogonality of \mathbf{C} implies the following conditions

$$\begin{cases} \mathbf{A}_i \mathbf{A}_i^\dagger + \mathbf{B}_i \mathbf{B}_i^\dagger = \mathbf{I}_{n \times n} & \text{for } i, j = 1, \dots, k \text{ and } i \neq j \\ \mathbf{A}_i \mathbf{A}_j^\dagger + \mathbf{B}_j \mathbf{B}_i^\dagger = \mathbf{A}_i \mathbf{B}_i^\dagger + \mathbf{A}_j \mathbf{B}_j^\dagger = \mathbf{O}_{n \times n} \end{cases} \quad (7)$$

We will use this demonstration (equation (6)) in further compute.

IV. MAXIMUM LIKELIHOOD DECODING OF STBC

We start from equation (5), in which Ω_1 depend linearly of the transmitted symbols and Ω_2 is a quadratic form, containing terms $|x_i|^2, x_i x_j^*, x_i x_j, x_i^* x_j^*$. If it is possible to keep only the terms $|x_i|^2$

and to suppress the others, minimizing Ω is equivalent to optimize k metrics separately. Hence the computational task is considerably reduced. This is the case with a STBC orthogonal design as we described just before. Then for a complex orthogonal design matrix \mathbf{C} , the term Ω_2 can be further simplified as

$$\Omega_2 = \sum_{j=1}^m \mathbf{h}^j \mathbf{C} \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger = \sum_{q=1}^k |x_q|^2 \sum_{ij} |\alpha_{i,j}|^2 \quad (8)$$

For the first term Ω_1 , which is a real scalar, considering the linear dependence in variables x_1, \dots, x_k (and their conjugates) it is always possible to write it in the following form.

$$\Omega_1 = \sum_{j=1}^m -(\mathbf{r}^j \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger) = \sum_{q=1}^k -(p_q^* x_q + p_q x_q^*) \quad (9)$$

The computation of p_q will be further detailed. Using equation (8) and (9), we obtain (10)

$$\Omega = \sum_{q=1}^k \Omega_q \quad ; \quad \Omega_q = \left(\sum_{ij} |\alpha_{i,j}|^2 \right) |x_q|^2 - (p_q^* x_q + p_q x_q^*) \quad (10)$$

Since Ω_q only depends on x_q , minimization of Ω involves the optimization of each quantity Ω_q : $\text{Min}_{(x_1, \dots, x_k)} \Omega \equiv \text{Min}_{x_q} \Omega_q ; \forall q$

In this case, the decoding complexity is linear in the number of symbols (complexity with a M-ary modulation is $k.M$ instead of M^k). Furthermore, it is possible to simplify Ω_q and we will have.

$$\text{Min}_{x_q} \Omega_q \equiv \text{Min}_{x_q} \left| x_q - \frac{p_q}{A} \right|^2 = |x_q - p'_q|^2 ; \forall q = 1, \dots, k \quad (11)$$

where $A = \sum_{ij} |\alpha_{i,j}|^2$ is a real scalar and we call $p'_q = p_q/A$ the

decision parameter of symbol q . In fact, the ML decoder selects the closest constellation from the symbol decision parameter. To compute the term p_q , using equation (6), we can write :

$$\mathbf{h}^j \mathbf{C} = \sum_{i=1}^k x_i \mathbf{h}^j \mathbf{A}_i + \sum_{i=1}^k x_i^* \mathbf{h}^j \mathbf{B}_i = \mathbf{x} \mathbf{H}_R^j + \mathbf{x}^* \mathbf{H}_I^j \quad (12)$$

where $\mathbf{x} = (x_1 \dots x_k)$, $\mathbf{H}_R^j = \begin{pmatrix} \mathbf{h}^j \mathbf{A}_1 \\ \vdots \\ \mathbf{h}^j \mathbf{A}_k \end{pmatrix}$ and $\mathbf{H}_I^j = \begin{pmatrix} \mathbf{h}^j \mathbf{B}_1 \\ \vdots \\ \mathbf{h}^j \mathbf{B}_k \end{pmatrix}$

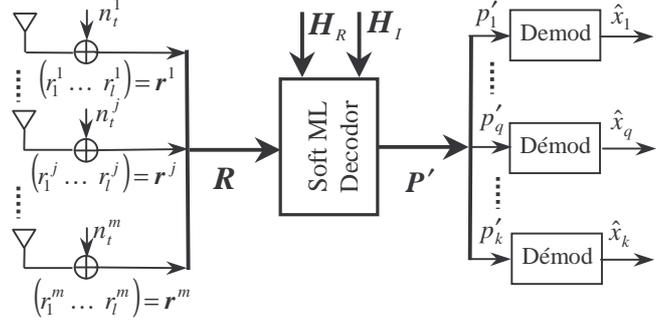


Figure 2: ML decoding for complex STBC orthogonal designs

Then using this expression we replace the term $\mathbf{h}^j \mathbf{C}$ in equation (9). The term $(\mathbf{r}^j \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger)$ is clearly the conjugate of $(\mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger)$.

$$\sum_{j=1}^m \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger = \mathbf{x} \mathbf{H}_R \mathbf{R}^\dagger + \mathbf{x}^* \mathbf{H}_I \mathbf{R}^\dagger \quad (13)$$

where $\mathbf{H}_R = (\mathbf{H}_R^1 \dots \mathbf{H}_R^m)_{k \times ml}$, $\mathbf{H}_I = (\mathbf{H}_I^1 \dots \mathbf{H}_I^m)_{k \times ml}$ and $\mathbf{R} = (\mathbf{r}^1 \dots \mathbf{r}^m)_{l \times ml}$. Using (13), (9) takes the following form :

$$\Omega_1 = -(\mathbf{x} \mathbf{P}^\dagger + \mathbf{x}^* \mathbf{P}^T) \quad (14)$$

where

$$\mathbf{P} = \mathbf{R} \mathbf{H}_R^\dagger + \mathbf{R}^* \mathbf{H}_I^T \quad (15)$$

When comparing equation (14) with (9), it is clear that the vector $\mathbf{P}' = (p_1 \dots p_k)/A$ contains all the sufficient statistics for ML Detection. The obtained matrix representation results in a simplified system implementation. Fig. 2 illustrates the ML decoding rule for complex STBC orthogonal designs, then we present the decoding rule as a complete computer algorithm.

Decoding algorithm :

Entry parameters :

- The STBC code with matrices \mathbf{A}_i et \mathbf{B}_i for $i = 1, \dots, k$.
- The received signals r_t^j for $j = 1, \dots, m ; t = 1, \dots, l$ (vector \mathbf{R}).
- The channel parameters $\alpha_{i,j} \forall i, j$ (the matrices \mathbf{h}^j and \mathbf{H}).

Programming steps:

- 1- Find the equivalent matrices \mathbf{H}_R^j and \mathbf{H}_I^j for $j = 1, \dots, m$ (from (12)) and matrices \mathbf{H}_R and \mathbf{H}_I (equation (13)).
- 2- Compute vector \mathbf{P} according to equation (15) and obtain the decision parameters $p'_q = p_q/A$ for all the symbols of a transmitted packet ($q = 1, \dots, k$).
- 3- Take the final decision for all transmitted symbols from (11).

V. DERIVATION OF ACCURATE UPPERBOUNDS

In this section we derive an accurate analytical bound to quantify the BER performances of the STBC ML decoding algorithm we described before. At first, using equation (13) for the definition of matrix \mathbf{R} , we write the received signals for all the set of receiving antennas in the following form :

$$\mathbf{R} = \mathbf{x} \mathbf{H}_R + \mathbf{x}^* \mathbf{H}_I + \mathbf{N} \quad (16)$$

where $\mathbf{N} = (\mathbf{n}^1 \dots \mathbf{n}^m)$. Before developing the mathematical derivation of the BER it is important to keep in mind some important characteristics of matrices \mathbf{H}_R and \mathbf{H}_I which contain channel informations and STBC structure. Using the STBC complex orthogonal design properties (equation (7)) and the definition of matrices \mathbf{H}_R and \mathbf{H}_I , it is easy to demonstrate that

$$\mathbf{H}_R \mathbf{H}_I^\dagger + (\mathbf{H}_R \mathbf{H}_I^\dagger)^T = \mathbf{0}_{k \times k} \quad (17-a)$$

$$\mathbf{H}_R \mathbf{H}_R^\dagger + \mathbf{H}_I \mathbf{H}_I^\dagger = \sum_{i,j=1}^n \sum_{k=1}^m |\alpha_{i,j}|^2 \mathbf{I}_{k \times k} \quad (17-b)$$

We have now to compute the noise statistics, which are present in the decision vector \mathbf{P}' . To do this we replace the vector \mathbf{R} in equation (15) by those of equation (16) and, using equation (17), this leads to

$$\mathbf{P}' = \mathbf{x} + \boldsymbol{\eta} ; \boldsymbol{\eta} = \frac{1}{\sum_{i,j} |\alpha_{i,j}|^2} [\mathbf{N} \mathbf{H}_R^\dagger + \mathbf{N}^* \mathbf{H}_I^T] \quad (18)$$

where $\boldsymbol{\eta}$ is a complex noise vector (Gaussian random vector if the channel gains are constant). In the output of our soft ML decoder, \mathbf{P}' give exactly the information symbols in addition to noise. Then using (18), the considered MIMO system is completely equivalent to a single-transmit single-receive system having an additive noise vector $\boldsymbol{\eta}$. For following we consider first the case of constant channel coefficients and next random channel coefficients.

a- The channel fading coefficients $\alpha_{i,j}$ are constant

Using statistics of noise and (17-b) it is now possible to compute the auto-correlation matrix of the noise vector $\boldsymbol{\eta}$.

$$E\{\boldsymbol{\eta}^\dagger \boldsymbol{\eta}\} = \frac{\sigma^2}{\sum_{i,j} |\alpha_{i,j}|^2} \mathbf{I}_{k \times k} \quad (19)$$

It is amazing to remark that the energy of each sample of $\boldsymbol{\eta}$ is equal to the energy of each sample of \mathbf{N} divided by a scale factor A . Since $\sigma^2 = 1/\text{SNR}$ and considering the energy of each transmitted symbol ($1/n$), the signal to noise ratio at the output of the ML decoder is

$$\text{SNR}_q = \frac{1/n}{\sigma^2 / \sum_{i,j} |\alpha_{i,j}|^2} = \left(\sum_{i,j} |\alpha_{i,j}|^2 / n \right) \text{SNR} \quad (20)$$

The parameter decisions in vector \mathbf{P} are fed to the demodulator (Fig. 2), using the well known error function Q , we find in the open literature [10] some useful expressions for the BER which can be applied to our system. Its amazing to remark that the signal to noise ratio at output of the soft ML decoder and BER depends only to absolute channel gains

b- The $\alpha_{i,j}$ are independent complex Gaussian random

In this case, the proposed method is to compute at first the BER, considering the case when $\alpha_{i,j}$ are constant and then to average over the probability density function of variables $\alpha_{i,j}$.

$$\bar{P}_e = E_\alpha \{P(e|\alpha)\} \quad (21)$$

Since the channel gains are Gaussian and independent, each variable $|\alpha_{i,j}|$ has a Rayleigh distribution and the joint density probability is the multiplication of the density probability of each random variable.

$$\alpha_{i,j} = (\alpha_{i,j}^R, \alpha_{i,j}^I) \sim \mathcal{N}(0, 1/2) \Rightarrow P(|\alpha|) = \prod_{i,j} 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \quad (22)$$

The computation of P_e leads to :

$$\bar{P}_e = E_\alpha \{P(e|\alpha)\} = \int_\alpha P(e|\alpha) \prod_{i,j} P(\alpha_{i,j}) \prod_{i,j} d|\alpha_{i,j}| \quad (23)$$

Then, to further compute the probability \bar{P}_e , one has to replace $P(e|\alpha)$ by the expressions already present for some kinds of modulation schemes. To approximate the values of $Q(x)$ we use the following rules [10]

$$Q(x) < \frac{1}{2} \exp(-x^2/2) \quad (24-a)$$

$$Q(x) \cong \frac{1}{\sqrt{2\pi}x} \exp(-x^2/2) \quad (24-b)$$

(24-a) is less precise than (24-b) but simpler to compute, (24-b) is a good approximation of $Q(x)$ particularly at high SNR's. For the incoming formulas we use the following notation

$$P_e = P(e|\alpha) \cong G_1 Q(G_2 \sqrt{\text{SNR}_q}) \quad (25)$$

Using (20) and (25), it is now possible to complete the computation of \bar{P}_e (23), the result in the case of formula (24-a) is the following

$$\begin{aligned} \bar{P}_e &< \frac{G_1}{2} \int_{|\alpha|} e^{-G_2^2 \text{SNR} (\sum |\alpha_{i,j}|^2 / 2n)} \prod_{i,j} \left(2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \right) d|\alpha_{i,j}| \quad (26) \\ &< \frac{G_1}{2} \left(\int_0^\infty 2\alpha e^{-\alpha^2 (1+G_2^2 \frac{\text{SNR}}{2n})} d\alpha \right)^{mn} = \frac{G_1}{2 \left(1+G_2^2 \frac{\text{SNR}}{2n} \right)^{mn}} \end{aligned}$$

It is important to remark that \bar{P}_e decreases exponentially with the number of antennas. It is then possible to obtain a more accurate bound at high SNR's, using (24-b) instead of (24-a).

$$\begin{aligned} \bar{P}_e &\cong \frac{G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{\text{SNR}}} \int_{|\alpha|} e^{-G_2^2 \text{SNR} (\sum |\alpha_{i,j}|^2 / 2n)} \frac{\prod_{i,j} 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} d|\alpha_{i,j}|}{\sqrt{\sum_{i,j} |\alpha_{i,j}|^2}} \\ &= \frac{G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{\text{SNR}}} \int_{|\alpha|} e^{-\left(1+G_2^2 \frac{\text{SNR}}{2n} \sum |\alpha_{i,j}|^2\right)} \frac{\prod_{i,j} 2|\alpha_{i,j}| d|\alpha_{i,j}|}{\sqrt{\sum_{i,j} |\alpha_{i,j}|^2}} \quad (27) \end{aligned}$$

The difficulty to compute (27) lies in the presence of the terms $\sqrt{\sum |\alpha_{i,j}|^2}$. To solve (27) it is possible to use a generalized polar transformation in a subspace of dimension mn . The polar transformation can be resumed with the following equations

$$\begin{aligned} M &= mn ; \alpha_d = |\alpha_{i,j}| ; d-1 = (i-1)m + (j-1) ; d = 1, \dots, mn \\ \alpha_M &= \rho \prod_{j=1}^{M-1} \sin \varphi_j , \alpha_i = \rho \left(\prod_{j=1}^{i-1} \sin \varphi_j \right) \cos \varphi_i , i = 1, \dots, M-1 \quad (28) \end{aligned}$$

With this polar transformation, we have $\varphi_i \in [0, \frac{\pi}{2}]$, $\sum \alpha_d^2 = \rho^2$ and computing the Jacobian, we obtain

$$d\alpha_1 d\alpha_2 \dots d\alpha_M = \rho^{M-1} \left(\prod_{i=1}^{M-1} (\sin \varphi_i)^{M-1-i} d\varphi_i \right) d\rho \quad (29)$$

Using this polar transformation and after some manipulation the integral (27) is transformed into

$$\bar{P}_e \cong \frac{2^{mn} G_1 \sqrt{n} \prod_{i=1}^{mn-1} \left[(\sin \varphi_i)^{2mn-2i} / (2mn-2i) \right]_0^{\frac{\pi}{2}}}{\sqrt{2\pi} G_2 \sqrt{\text{SNR}} \sqrt{\left(1+G_2^2 \frac{\text{SNR}}{2n} \right)^{2M-1}}} \int_{\rho=0}^{\rho} \rho^{2mn-2} e^{-\rho^2} d\rho \quad (30)$$

Using $\int_0^\infty \rho^{2K} e^{-\rho^2} d\rho = \frac{(2K)!}{K! 2^{2K+1}} \sqrt{\pi}$ we finally obtain

$$\bar{P}_e = \frac{G_1}{\left(1 + G_2^2 \frac{SNR}{2n}\right)^{mn}} \sqrt{\frac{1 + G_2^2 \frac{SNR}{2n}}{G_2^2 \frac{SNR}{2n}}} \left(\frac{(2mn-2)!}{((mn-1)!)^2 2^{2mn-1}} \right) \quad (31)$$

A other important fact is that the more the number of antennas (mn) increases, the more expression (31) is distant from (26) and gives better results.

VI. SIMULATION RESULTS

To check the approximations we use the two following STBC orthogonal designs with varying number of receive antennas and coding rate equal to 1 and 0.5 respectively. The results are shown on figures 3–5 for a QPSK modulation with perfect CSI.

$$G_1: \begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 & x_1^* & -x_2^* & -x_3^* & -x_4^* \\ x_2 & x_1 & x_4 & -x_3 & x_2^* & x_1^* & x_4^* & -x_3^* \\ x_3 & -x_4 & x_1 & x_2 & x_3^* & -x_4^* & x_1^* & x_2^* \\ x_4 & x_3 & -x_2 & x_1 & x_4^* & x_3^* & -x_2^* & x_1^* \end{pmatrix} G_2:$$

On figure 3, we present the performances of code G_1 comparing our bounds with simulation results for different number of receive antennas. The same work is done on figure 4 with G_2 . The accuracy of the improved upper bound (35) is such that the difference with simulation results is less than 0.2 dB. Figure 5 shows the symbol error probability (SER) as a function of the number of receive antennas at a given SNR (8 dB). The influence of the number of receive antennas is clearly demonstrated and the precision of our bounds is validated too.

VII. CONCLUSION

In this paper we have given a general framework, for ML decoding of STBC orthogonal, which is particularly easy to implement. Using this model we derive upper bounds for BER performances of such encoding schemes. Besides, using different STBC and varying number of receive antenna, simulation results confirm the accuracy of these bounds.

REFERENCES

- [1] G. J. Foschini Jr. and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas", *Wireless Personal Commun.*, Mar. 1998.
- [2] J. C. Guey, M. P. Fitz, M. R. Bell, and W. Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels", in *Proc. VTC'96*, 1996, pp. 136-140.
- [3] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity", *Int. J. Wireless Inform. Networks*, vol. 1, n° 1, 1994.
- [4] E. Telatar, "Capacity of multi-antenna Gaussian channels", AT&T-Bell Laboratories Internal Tevh. Memo., June 1995.
- [5] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communications : Performance analysis and code construction", *IEEE trans. Inform. Theory*, vol. 44, pp. 744-765, Mar. 1998.
- [6] V. Tarokh, A. Naguib, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communications : Performance criteria in the presence of channel estimation errors, mobility and multipaths", *IEEE trans. Commun.*, vol. 47, pp. 199-207, Feb. 1999.
- [7] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications", *IEEE J. Select. Areas. Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.
- [8] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "STBC from Orthogonal Designs", *IEEE trans. Inform Theory*, vol. 45, pp. 1456-1467, July 1999.

- [9] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "STBC for wireless communications : Performance results", *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451-460, Mar. 1999.
- [10] S. Benedetto and E. Biglieri, *Principles of Digital Transmission With Wireless Applications*, Plenum Publishers, 1999.

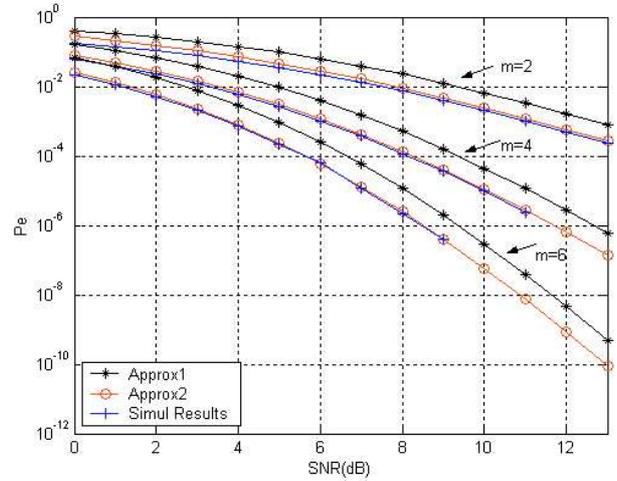


Figure 3 : SER as a function of SNR(E_s/N_0) for code G_1

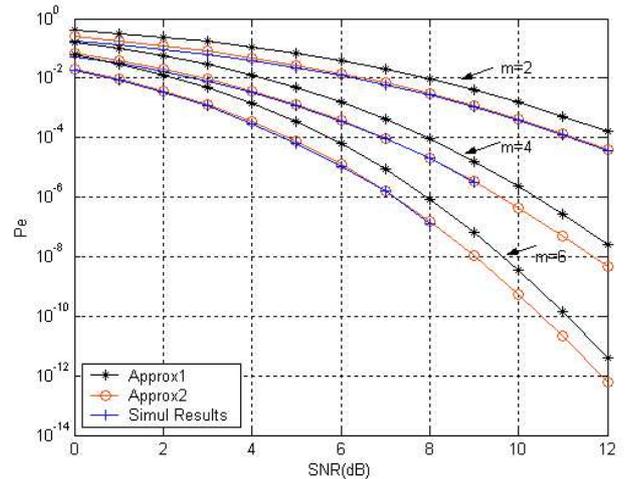


Figure 4: SER as a function of SNR(E_s/N_0) for code G_2

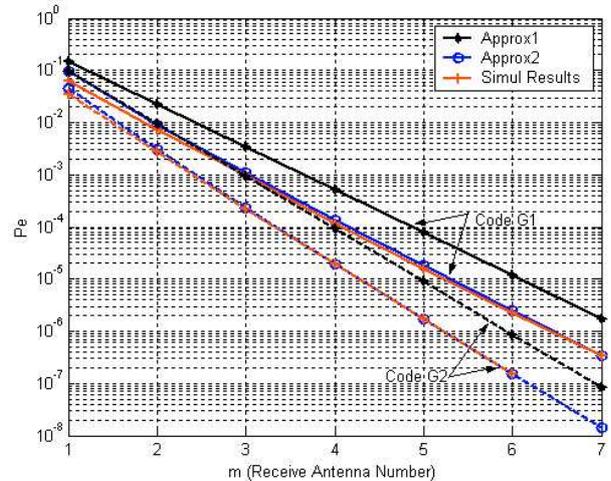


Figure 5: SER as a function of receive antenna number for SNR=8dB

New SLA Creation and Optimal Resource Management

Srećko Krile

University of Dubrovnik

Department of Electrical Engineering and Computing

Cira Carica 4, 20000 Dubrovnik, Croatia

Tel: 385 20 445-739, Fax: 385 20 435-590

srecko.krile@unidu.hr

Abstract: While DiffServ architecture solves the scalability problem of QoS provisioning, it fails to be the solution for end-to-end provisioning. A combination of IntServ/RSVP signaling with aggregate traffic handling mechanisms could solve such deficiencies. To obtain quantitative end-to-end guarantees in DiffServ architecture, based on traffic handling mechanisms with aggregate flows, some kind of congestion control through negotiation process (in new SLA creation) is necessary. For N quality-of-service levels (service classes) an efficient heuristic algorithm for end-to-end congestion control is being developed. The problem is seen as an expansion problem of link capacities in given limits from a common source. If the optimal expansion sequence has any expansion value that exceeds allowed limits (link capacity), a new SLA cannot be accepted or must be redefined through a negotiation process.

Index Terms: quality of service in DiffServ networks, constrained-based path selection, end-to-end QoS routing, SLA creation, traffic handling mechanisms.

1. Introduction

The classification of the aggregated flows (in DiffServ/MPLS cloud) is performed according to the SLA (Service Level Agreement) signed between a customer and the network operator (ISP). Each SLA contract specifies how much traffic may be sent (service class, bandwidth, delay etc.) and defines a time period for utilization of that service. Very important element for efficient end-to-end QoS routing is good prediction of traffic demands that is defined with limited number of SLA agreements. So, in the process of SLA creation the problem of new SLA acceptance for network operator exists.

SLA creation is in correlation with QoS routing, resource reservation mechanisms and admission control process (service invocation). It means that every new SLA acceptance directly influences on traffic handling mechanisms with other traffic flows (existing SLAs).

Necessity of some combination of IntServ (management per-flow) and DiffServ (management with aggregate flows) clearly represents a trade-off

between service granularity and scalability: as soon as flows are aggregated, they are not as isolated from each other as in IntServ architecture. In the moment of service invocation (explicit activation) the optimal routing sequence information for that traffic flow can be sent with RSVP (Resource Reservation Protocol) signaling protocol to MPLS routers, to ensure end-to-end guarantees. Sufficient resources must be available at any moment because congestion control in the SLA negotiation process are made former.

In section 2. The problems of new SLA creation and correlation with optimal resource management are investigating. Explanation of the mathematical model and heuristic approach for constraint-based path selection for new SLA creation is given in section 3.

2. New SLA Creation Problem

If network operator (ISP) wants to accept new SLA (between edge routers) it has to be checked with congestion control algorithm related on limited link resources and predicted traffic (caused with former accepted SLAs). For each communication link in the network given traffic demands (consist of number of SLAs) can be satisfied on different QoS levels (for example used bandwidth).

The optimal resource management can be seen as the optimal link capacity expansion problem with expansion values in allowed limits (capacity). If the optimal routing sequence has any link expansion with value that exceeds allowed limits (link capacity), it means that new SLA cannot be accepted or must be redefined through negotiation process.

Traffic demand (given in relative value as increment) on input of each edge router, represents the sum of all ingress and egress SLAs; see fig. 2. Congestion control must be done for traffic aggregated flows between edge routers (on each link), specially for definite period of time (critical moments); see fig 1. The optimal constraint-based path selection eliminates the possibility for traffic congestion. We need very effective tool to check such congestion possibility in the network. Some important papers about that problem are [1], [4] and [5]. In the paper [2] such algorithm is the part of service management architecture.

3. The Mathematical Model and Heuristic Approach

Let $G(A, E)$ denote a network topology, where A is the set of nodes and E the set of links. The source and destination nodes (edge routers in domain) are denoted by s and d , respectively; see fig 1. The number of QoS measures (e.g. bandwidth, delay) is denoted by z .

Consider a network $G(A, E)$ where each link is characterized by z -dimensional link weight vector, consisting of z nonnegative QoS weights $\{w_i(k, l), i = 1, \dots, z, (k, l) \in E\}$ as components. Given z constraints are denoted by $L_i, i = 1, \dots, z$. Definition of the multi-constrained (MCP) problem is to find a path P from s to d such that:

$$w_i(P) \stackrel{\text{def}}{=} \sum_{(k, l) \in P} w_i(k, l) \leq L_i \quad (3.1)$$

For $i = 1, \dots, z$.

In this paper we dealt about only one dimensional link weight vector, with only one constraint denoted with L . Given constraints are limited bandwidths for each link on the path: $L_m, m = 1, \dots, M$; see fig. 2. The link weight (cost) is the function of used capacity: lower used capacity (smaller bandwidth) gives lower weight (cost). The main condition is that given traffic demands must be satisfied. Nonlinear cost function is necessary if link weights are not positively correlated. The problem of

the optimal QoS routing can be seen as the minimum cost network flow problem in the multi-commodity single (common) source multiple destination network. Such problem can be solved as the capacity expansion problem (CEP) without shortages. Partially expansions for each link are made from common source in given limits (link capacity).

Transmission link capacities on the path between routers are capable to serve traffic demands for N different QoS levels (called *facilities*), for $i = 1, 2, \dots, N$. Facility i is used primarily to serve demands for QoS level i , but it can be used to satisfy traffic demands for QoS level j ($j > i$). Re-routing of traffic demands towards higher QoS level is the same thing as facility conversion toward lower QoS level; see fig. 2. In this model conversion of traffic demand is permitted only in the direction toward higher QoS level. The objective is to find optimal routing-policy that minimizes the total cost incurred over the whole path between edge routers (M interior routers and $M + 1$ transmission links) and to satisfy given traffic demands. An example of the optimal expansion solution for six interior routers can be seen in fig. 3. The flow theory enables separation of these extreme flows which can be a part of an optimal expansion solution from those which cannot be. With such heuristic approach we can obtain the optimal result with significant computational savings. Fig. 2 gives a network flow representation of CEP for three QoS

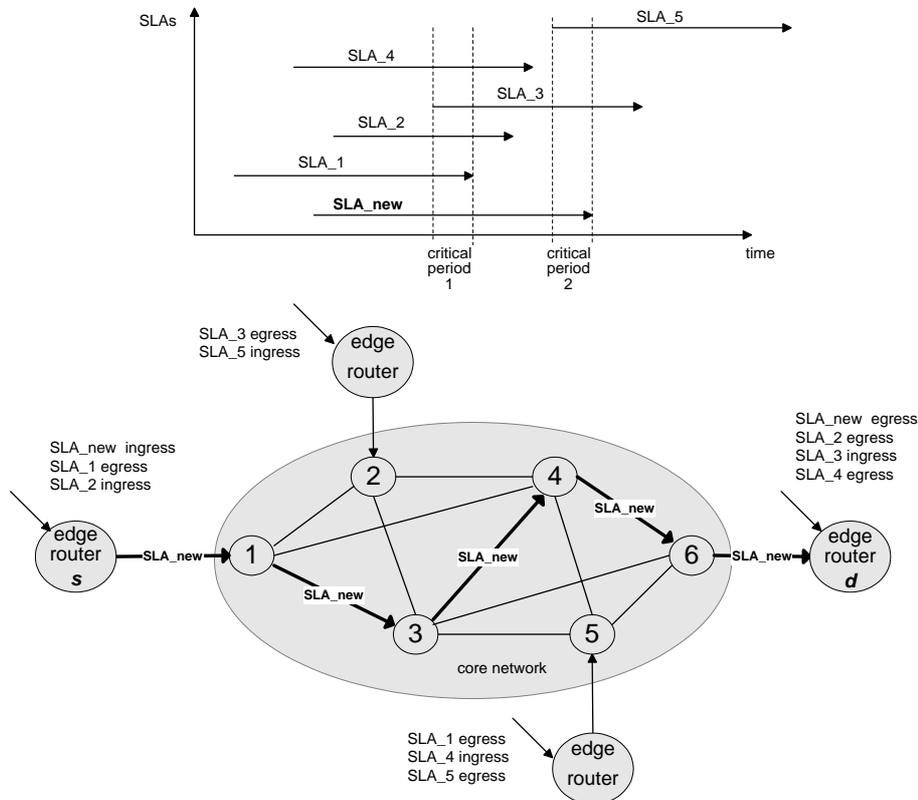


Figure 1. An example of number of SLAs in definite period of time. The optimal routing sequence for new SLA need not to be the shortest path solution.

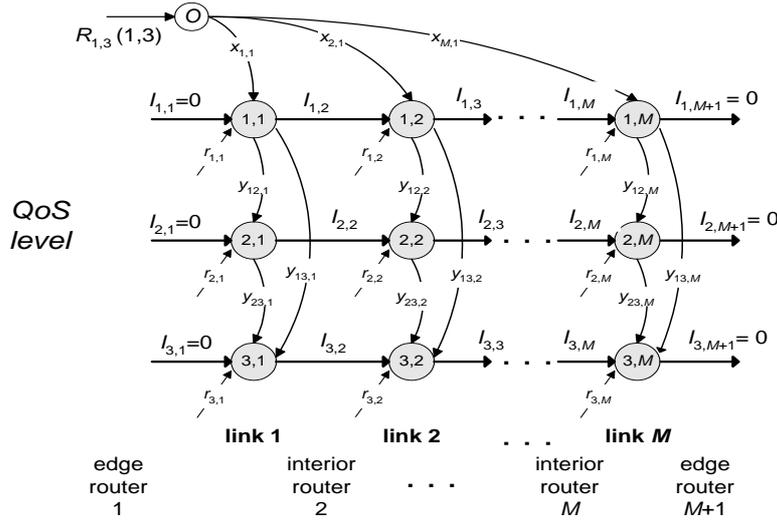


Figure 2. A numerical flow representation of the CEP problem

levels ($N = 3$) and M internal (core) routers included in the path.

On that diagram the m -th row of nodes represents a possible link capacity state of each transmission link between routers for i -th QoS level. Link capacity values are positive only (idle capacity), and shortages are not allowed. Horizontal links between them represent the traffic flow between routers. Common node "O" is the source for used capacity (expansions), introducing the new capacity on the link. Vertical links represent re-routing of traffic demands (equal to capacity conversion of facility).

In the mathematical model of CEP the following notation is used; see diagram on fig.2.:

i, j and $k =$ QoS level. The N levels are ranked from $1, 2, \dots, N$, and quality decreases with higher k .

$m =$ the order number of link on the path, connecting two successive routers. Path consists of M links ($m = 1, \dots, M$) between $M+1$ routers.

$u, v =$ the order number of capacity points in the sub-problem, $1 \leq u, \dots, v \leq M+1$.

$r_{im} =$ traffic demand increment for additional capacity of facility i (appropriate QoS level) on link m . For convenience, the r_{im} are assumed to be integer.

$I_{im} =$ the relative amount of idle capacity of facility i on the link m , related on the link before. Initially there is no capacity shortage between edge router and the interior router, $I_{i1} = 0, I_{i,M+1} = 0$.

$WI_{im} =$ upper limit for capacity of facility i on the link m .

$kl_i =$ the lowest step of possible facility change for QoS level i .

$x_{im} =$ the amount of used capacity for facility i on the transmission link m .

$Wx_{im} =$ upper limit for allowed expansion for facility i on the transmission link m .

$y_{ijm} =$ the amount of capacity of facility i on the link m , redirected to satisfy the traffic of lower level..

The CEP problem can be formulated as follows:

$$\min \left(\sum_{i=1}^M \left\{ \sum_{i=1}^N c_{im}(x_{im}) + h_{im}(I_{i,m+1}) + \sum_{j=i+1}^N g_{ijm}(y_{ijm}) \right\} \right) \quad (3.2)$$

So that we have:

$$I_{m+1} = I_{im} + x_{im} - \sum_{j=i+1}^N y_{ijm} - r_{im} \quad (3.3)$$

$$I_{im} = I_{i,M+1} = 0 \quad (3.4)$$

For $m = 1, 2, \dots, M; i = 1, 2, \dots, N; j = i + 1, \dots, N$.

The total cost on the path from edge to edge router includes some costs: the cost for capacity expansion $c_{im}(x_{im})$, the idle capacity cost $h_{im}(I_{i,m+1})$ as penalty cost to force the usage of the minimum link capacity (prevention of idle capacity), and the re-routing cost of traffic demands $g_{ijm}(y_{ijm})$. For expansion of link in allowed limits we can set the cost to zero. Costs are often represented by the fix-charge cost or with constant value. We assume that all cost functions are concave and non-decreasing, reflecting economies of scale, and they can change for appropriate link. With costs parameters we can influence on the optimization process, looking for the most appropriate routing solution.

3.1. Definition of the Capacity Point

Generalizing the concept of the capacity state for transmission link m in which the capacity state of each link is known within defined limits and which at least one capacity state satisfies $I_{im} = 0$, we define as a *capacity point*. In (3.1.1) α_m denotes the vector of capacities I_{im} for all QoS levels (facility types) on link m , and we call it capacity point.

$$\alpha_m = (I_{1m}, I_{2m}, \dots, I_{Nm}) \quad (3.1.1)$$

$$\alpha_0 = \alpha_{M+1} = (0, 0, \dots, 0) \quad (3.1.2)$$

Each column in the flow diagram from fig. 2 represents a capacity point, consisting of N capacity state values. (3.1.2) implies that idle capacities or capacity shortages are not allowed on the link between edge and interior router.

3.2. Sub-problem

Associated value between two capacity points, that represents minimum cost $d_{uv}(\alpha_u, \alpha_{v+1})$ we denoted as CES (*Capacity Expansion Sub-problem*). In CEP we have to find many cost values $d_{uv}(\alpha_u, \alpha_{v+1})$ that emanate two capacity points, from each node (u, α_u) to node $(v+1, \alpha_{v+1})$ for $v \geq u$. The approach described in [3] requires solving repeatedly a certain single location expansion problem (SLEP). Most of the computational effort is spent on computing the sub-problem values. Any of them, if it cannot be a part of the optimal sequence, is set to infinity.

Supposing that all d_{uv} values are known, the optimal solution for CEP can be found by searching for the optimal sequence of capacity points and their associated values. On that level the problem can be formulated as the shortest path problem between edge routers for an acyclic network, which nodes represent all possible capacity states.

Conclusions

In the process of new SLA creation possible congestion can be checked with proposed heuristic algorithm. Algorithm is based on mathematical model for the capacity expansion problem (CEP); see fig. 3. It means that such heuristic approach can be successfully applied for congestion control in the SLA creation process, that is in firm correlation with resource reservation mechanisms and admission control process. It will ensure end-to-end QoS routing guaranties, improving DiffServ granularity. The heuristic algorithm in all test-examples can achieve the best possible result (near-optimal expansion sequence), but requires the computation effort of $O(M^3 N^4 R_i^{2(N-1)})$. The required effort for one sub-problem is $O(N^2 M)$. The number of all possible d_{uv} values depends on the total number of capacity points. If there are no limitations (for WI_{im} and Wx_{im}) the complexity of such heuristic approach is pretty large and increases exponentially with N .

References

1. S. Giordano, S. Salsano, G. Ventre, "Advanced QoS Provisioning in IP Networks": The European Premium IP Projects, Communications, (2003), Vol.41 No 1., pp.30-36.
2. D. Kagklis, C Tsakiris, N. Liampotis, "Quality of Service: A Mechanism for explicit activation of IP Services Based on RSVP", Journal of Electrical Engineering, Vol. 54, No. 910, Bratislava, 2003, pp.250-254.

3. S. Krile, M. Kos, "A Heuristic Approach for Path Provisioning in Diff-Serv Networks", Proc. 7th ISSSTA (Int. Symp. On Spread-Spectrum Tech. & Application) - IEEE, Prag, 2002., pp.692-696.
4. K. Biswas, S. Ganguly and R. Izmailov, "Path provisioning for service level agreements in differentiated services networks, ICC 2002 - IEEE International Conference on Communications, vol. 25, no. 1, (2002), pp.1063-1068.
5. E. Bouillet, D. Mitra, K. G. Ramakrishnan, "The structure and management of service level agreements in networks", IEEE Journal on Selected Areas in Communications, vol. 20, no. 4, (2002), pp.691-699.

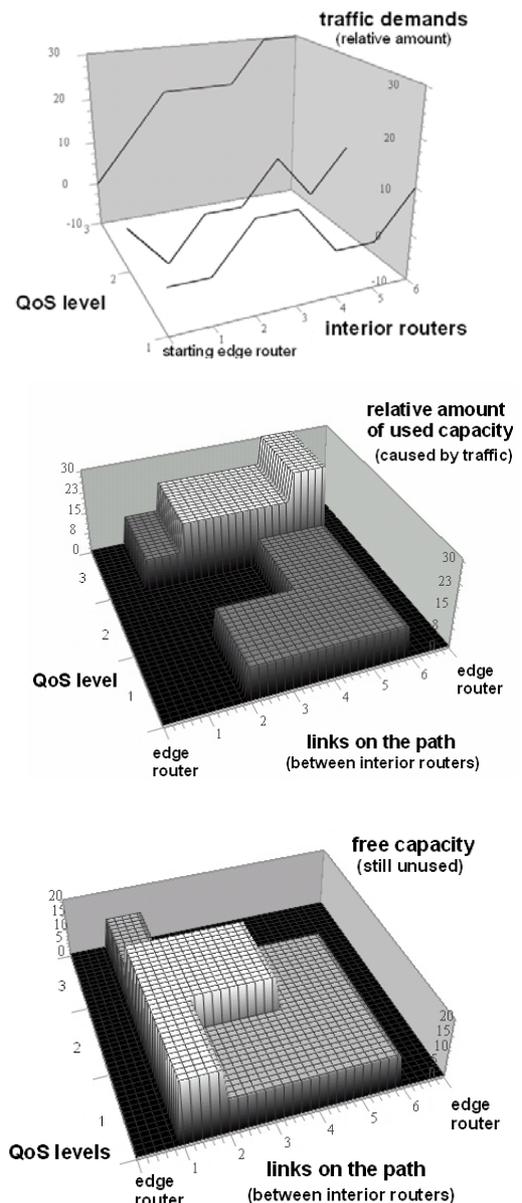


Figure 3. With such resource management tool we can analyse the capacity situation on the path for new SLA creation. In this example used capacities don't exceed given limits (max. 30 channels for each link), that means no congestion exists.

Noise and Crosstalk Analysis of SMZ Switches

R. Ngah, and Z. Ghassemlooy

Optical Communication Research Lab., School of Engineering & Technology, Northumbria University

Tel: +44 191 227 4902, Email: fary.ghassemlooy@unn.ac.uk

Abstract: Ultra-high speed OTDM networks require stable and compact all optical switches for demultiplexing and routing. In this paper we investigate the noise and crosstalk characteristics of all optical symmetric Mach-Zehnder (SMZ) switch using a numerical model. Two effects, the relative intensity noise (RIN) and channel crosstalk (CXT) that degrade the performance of the switch are investigated. Numerical simulation results show that RIN and CXT varies with different SOA length, data pulse width, RMS_{jitter} and the bit rate.

1. Introduction

All optical switches based on the cross-phase modulation (XPM) in conjunction with interferometric configurations such as Mach-Zehnder interferometers (MZIs), terahertz optical asymmetric demultiplexers (TOADs) and ultrafast nonlinear interferometers (UNIs) are widely studied [1]. Among these the monolithically integrated MZI switches are the most promising due to their compact size, thermal stability, and low-power operation. All optical demultiplexing employing MZI switches at 168 Gb/s has been reported [2]. Considering various MZI configurations, the symmetric Mach-Zehnder (SMZ) structure provides the highest flexibility and shortest switching window [3].

These switches are used in a number of applications such as optical routers. Recently, a 1x2 all optical router employing SMZ switches has been reported [4]. Therefore, understanding their noise and crosstalk characteristics is extremely important in order to reduce power penalty it may incur. This paper presents an analysis and subsequent simulations of the noise and crosstalk characteristics of a SMZ by mean of a computer simulation. The structure of this paper is as follows. The operation principles of the SMZ switch are explained in section 2. In Section 3, the noise and crosstalk characteristics of the switch and the sources introducing them are identified. Numerical simulation results for the noise and crosstalk performance of all optical SMZ switch are presented in Section 4. Finally, in Section 5, the concluding remarks are summarized.

2. Operation Principles

Figure 1 shows the block diagram of a typical SMZ switch composed of two semiconductor optical

amplifiers (SOAs), one in each arm of the interferometer, and a number of 3-dB couplers. SOAs are positioned in the same relative location within the interferometer. Control and data pulses are fed into switch via 3-dB couplers and co-propagate within the interferometer. With no control pulses, the SMZ is balanced in such a way that all the data signals emerge from the reflected output port (port 2). However, with the control signals present a differential phase shift is introduced between the two arms of the interferometer thus causing the data pulses to be switched to the transmitted port (port 1). Note that the temporal delay between the control pulses determines the nominal width of the switching window, $W(t)$ of SMZ. $W(t)$ at port 1 is given by [5]:

$$W(t) = 0.2 \left\{ G_1(t) + G_2(t) - 2\sqrt{G_1(t)G_2(t)} \cdot \cos(\Delta\phi(t)) \right\} \quad (1)$$

where G_1 and G_2 are the temporal gain profile of the data pulses, and $\Delta\phi$ is the phase difference between the data pulses which is related to the gain ratio and the linewidth enhancement factor κ and is given by [9]:

$$\Delta\phi = -0.5\kappa \ln(G_1 / G_2) \quad (2)$$

Solving (1) requires determination of G_1 and G_2 gains of the data signal at the output of the SOA1 and SOA2 at the temporal point, which are given as, respectively:

$$G_1(t) = \exp \left[\int_0^{L_{SOA}} \Gamma \cdot g \left(z, t + \frac{z}{V_g} \right) dz \right] \quad (3)$$

$$G_2(t) = \exp \left[\int_0^{L_{SOA}} \Gamma \cdot g \left(z, t + T_{delay} + \frac{z}{V_g} \right) dz \right] \quad (4)$$

where Γ is the confinement factor, g represents the differential gain of data and control pulses, t the time at which the temporal point of the data pulse enters the amplifier, T_{delay} the temporal delay between the control pulses, z/V_g the time increment in the z direction and V_g the group velocity of the control pulse

3. Noise and Crosstalk Analysis

The noise and crosstalk associated with all optical switches are the relative intensity noise (RIN) and channel crosstalk (CXT). The mathematical models reported in [5-6] are used in these analyses.

3.1 Relative intensity noise

A major concern in the ultra-high speed optical communications is the timing jitter that could be introduced by the optical transmitter, the spontaneous emission noise of lumped optical amplifier, and interaction between the data and control within the switching devices [6]. The arrival time of the jittered target pulses within the switching window is a random process of Gaussian distribution. The relative timing jitter between the control and the data pulses is transformed to the intensity fluctuations through the non-flat switching window profile, thus resulting in switching power penalties [5]. Using a square and symmetrical switching window profile it is possible to reduce the intensity fluctuation. Compared with TOAD, which has asymmetric switching window profile [7], the SMZ is less vulnerable to the timing jitter noise due to its symmetric switching window [4]. The output signal of the SMZ, $w(t)$ can be described by:

$$w(t) = \int_{-\infty}^{\infty} W(t) p(t - \tau) dt \quad (5)$$

where $W(t)$ is the switching window profile given by (1), $p(t)$ is the target data power profile, and τ is the relative pulse arrival time. The expected value of the output signal energy is then given as:

$$E[w(\tau)] = \int_{-\infty}^{\infty} w(t) p_t(t - \tau) dt \quad (6)$$

where p_t is the probability density function of the relative signal pulse arrival time given by:

$$p_t(t) = \frac{1}{\sqrt{2\pi} RMS_{jitter}} e^{-\frac{1}{2} \left(\frac{t}{RMS_{jitter}} \right)^2} \quad (7)$$

where RMS_{jitter} is the root mean square timing jitter

The variance of the target data signal energy as a function of τ is expressed as:

$$V[w(\tau)] = \int_{-\infty}^{\infty} w^2(t) p_t(t - \tau) dt - E^2[w(\tau)] \quad (8)$$

Assuming that the mean arrival time of the target channel is at the centre of the SMZ switching window, then the RIN induced by the timing jitter of the transmitted signal can be expressed by [5]

$$RIN(\tau) = \frac{V(\tau)}{E^2[w(\tau)]} \quad (9)$$

3.2 Channel crosstalk

The second effect that degrades the performance of the all optical switch is the channel crosstalk, which

is due to switching of non-target channels to the output port when the switching profile overlaps into adjacent signal pulses. Unlike NOLM and TOAD switches, SMZ optical switches with co-propagating signal characteristics eliminates the residual crosstalk, which is caused by a small XPM between the counter-propagating pulses within the switch [6]. The channel crosstalk defined as the ratio of the transmitted power of one non-target channel to that of a target channel is given by:

$$CXT = 10 \log \frac{E_{nt}}{E_t} \quad (10)$$

where E_{nt} and E_t are the output signal energy due to the non-target and target channel respectively, expressed by:

$$E_{nt} = \int_{t_c + T_{b/2}}^{t_c + T - T_{b/2}} W(t) p_p(t - t_c) dt \quad (11)$$

$$E_t = \int_{t_c - T_{b/2}}^{t_c + T_{b/2}} W(t) p(t - t_c) dt \quad (12)$$

where $p_p(t)$ is periodic train of data pulse, T_b is the bit slot duration, t_c is centre of the switching window and T is channel period.

4. Results

The model used for the SOA in the numerical model is the modified version of [7], and the parameters used in the simulation are listed in Table 1.

Table 1: Simulation Parameter

Parameter	Values
Amplifier length	0.3 mm
Amplifier cross section area	2.25e-13 m ²
Linewidth enhancement factor	4
Confinement factor	0.15
Differential gain	2.78e-20
Transparent carrier density	1e24
Spontaneous lifetime	350 ps
Group velocity	3.0e8/3.5
Data baseline bit rate	10 Gb/s
FWHM control pulse	1 ps
FWHM data pulse	1 ps
Wavelength of control and data	1550 nm
Control pulse peak power	2.5 W
Data pulse peak power	2.5 μW
RMS_{jitter}	1 ps

With $T_{delay} = 10$ ps, the gain profiles of data pulses, having propagated through the SOAs, calculated from (3) and (4) are shown in Fig. 2. In the absence of the

control pulse a data pulse passes through the SOAs and experience an initial gain of 20.1 dB. The gain profile drop rapidly to a value of 2.8 dB after a high power and short duration control pulse saturate the SOAs. $G_1(t)$ and $G_2(t)$ are then substituted into (1) and (2) to obtain the SMZ switching window profile. With identical gain profile, the SMZ symmetrical switching window with a FWHM of 10 ps is shown in the inset of Fig. 2

Equation 1 is then substituted into (9) and (10) for calculating the SMZ RIN and CXT respectively. Fig. 3(a) shows the SMZ RIN against the control signals separation, T_{delay} for different SOA lengths. As can be seen from the figure, for all values of SOA length the RIN decrease as the T_{delay} increases reaching a minimum value of ≈ -25 dB when T_{delay} is ≈ 7.5 ps. This is because, a larger T_{delay} results in a wide switching window profile, and consequently the RIN is less sensitive to timing jitter. For low values of T_{delay} , the RIN is more sensitive to the timing jitter, since the target signal tends to appear outside of the switching window due to the timing jitter. Beyond $T_{delay} \approx 7.5$ ps no further improvement in the RIN is achieved and the SOA with a longer length gives additional improvement of between 0.5 - 1.5 dB compared with the shorter length SOAs. Also from Figure 3(a), it can be seen that the shorter SOA length gives the highest RIN at larger T_{delay} .

The effect of data pulse width on the RIN is investigated as shown in Fig. 3(b). The figure shows that for $T_{delay} < 6$ ps the RIN is lower for the data pulse with width of 2 ps. Whereas for $T_{delay} > 6$ ps the data pulse with a small width gives a lower RIN reaching a minimum value at lower T_{delay} of 12 ps. Fig. 3(c) shows a plot of RIN versus T_{delay} for different values of the RMS_{jitter} . As expected, the RIN is higher for a larger value of RMS_{jitter} reaching a minimum value of ≈ -25 dB as in the case of a lower value of RMS_{jitter} , but at a higher threshold value of $T_{delay} = 14$ ps.

Using the same parameters as in Table 1 and for data pulses of FWHM of 1 ps, and for the total data rate of 100 Gb/s (10 channels at the base rate of 10 Gb/s) the simulated CXT against T_{delay} for different SOA length is shown in Fig. 4 (a). As shown in the figure, the channel crosstalk for all SOA lengths have a minimum value when T_{delay} is close to zero and increase as T_{delay} increases. It also shows that the shortest length SOA display the worst performance. This is due to lower gain of the SOA. Fig. 4(b) shows the CXT for SOA length of 0.3 mm at different total bit rates (the base rate is still 10 Gb/s). The plots illustrate the same characteristic for T_{delay} up to 10 ps. However, CXT for 100 Gb/s and 160Gb/s cases, increase rapidly beyond T_{delay} of 10 ps and 16 ps, respectively due to switching of non-target pulses from the adjacent channels. The demultiplexing of non-target signals from adjacent channels will occur unless the T_{delay} -induced

window width is comparable to or larger than the time slot of a OTDM channels, which is 25 ps, 10 ps and 6.25 ps for 40 Gb/s, 100Gb/s and 160 Gb/s demultiplexing, respectively.

5.0 Conclusion

A noise and crosstalk analysis has been presented in this paper. The noise and crosstalk associated with SMZ switches that degrade the performance of the all optical router are the relative intensity noise and the channel crosstalk. The relative intensity noise increases with a wider pulse width and a higher RMS_{jitter} . Meanwhile the degradation due to the channel crosstalk increases with a shorter SOA length and a higher total bit rate.

REFERENCES

- [1] R.P. Schrieck, M. H. Kwakernaak, and H. Melchior "All-Optical Switching at Multi-100-Gb/s Data Rates with Mach-Zehnder Interferometer Switches" *IEEE J. Quantum Electron.*, vol. 38, no. 8, August 2002, pp. 1053-1061.
- [2] S. Nakamura, Y. Ueno, K. Tajima, J. Sasaki, T. Sugimoto, T. Kato, T. Shimoda, M. Itoh, H. Hatakeyama, T. Tamanuki, and T. Sasaki, "Demultiplexing of 168-Gb/s data pulses with a hybrid-integrated symmetric Mach-Zehnder all optical switch," *IEEE Photon. Technol. Lett.*, vol. 12, Apr. 2000, pp. 425-427.
- [3] P. Toliver, R. J. Runser, I. Glesk, and P. R. Prucnal, "Comparison of three nonlinear optical switch geometries," *Proc. Con. Lasers and Electro-Optics (CLEO 2000)*, San Francisco, CA, May 2000, pp. 254-255.
- [4] R. Ngah, Z. Ghassemlooy, and G. Swift, "Simulation of an All Optical Time Division Multiplexing Router Employing Symmetric Mach-Zehnder (SMZ)," *HFPSC2002*, Sept. 2002, pp. 130-136.
- [5] K. Uchiyama, T. Morioka, S. Kawanishi, H. Takara, and M. Saruwatari, "Signal-to-Noise Ratio Analysis of 100 Gb/s Demultiplexing Using Nonlinear Optical Loop Mirror," *J. Lightwave Technol.*, vol. 20, n0. 2, Feb. 1997, pp. 618-624.
- [6] C. Y. Cheung, Z. Ghassemlooy, G. Swift, and A. Decker, "Crosstalk and Noise Characteristics of Non-linear Optical Loop Mirror Demultiplexer and Terahertz Optical Asymmetric Demultiplexer," *Proc. Inter. Sympo. On Commun. Sys., Networks and DSP*, July 2000, pp. 59-63.
- [7] G. Swift, Z. Ghassemlooy, A. K. Ray, and J. R. Travis, "Modelling of Semiconductor Laser Amplifier for the Terahertz Optical Asymmetric Demultiplexer," *IEE Proc. -Circuits Devices Syst.*, vol. 145, Apr. 1998, pp. 61-65.

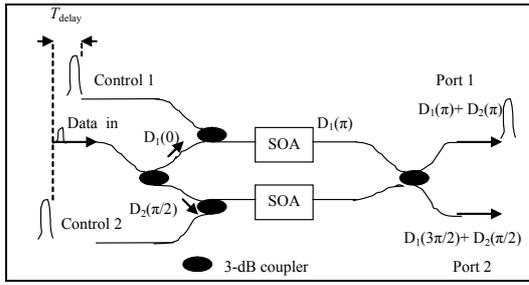


Fig. 1 SMZ configuration block diagram

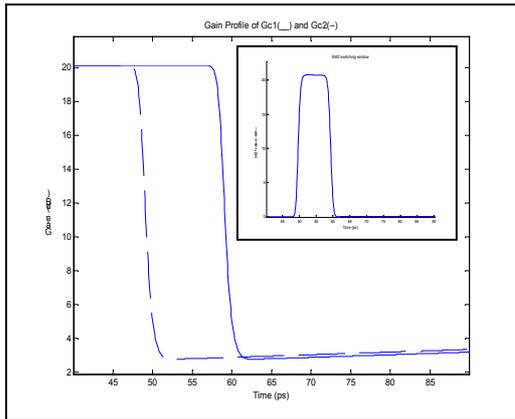
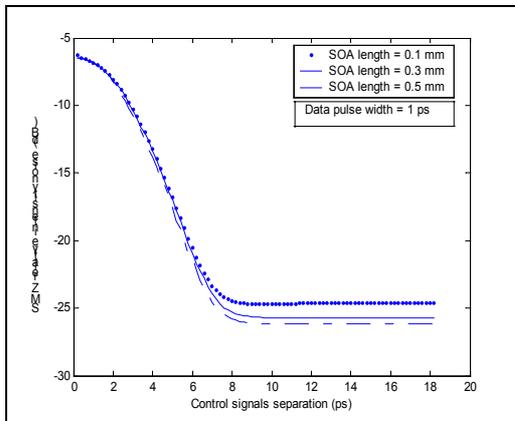
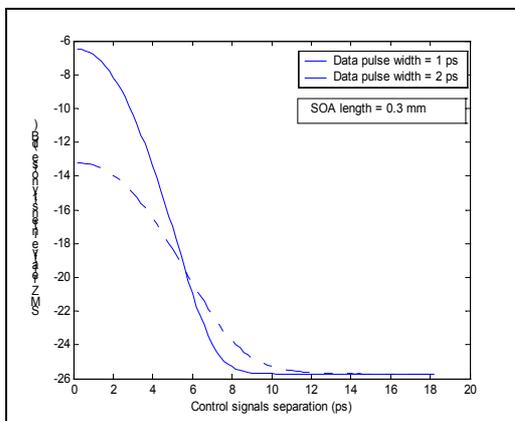


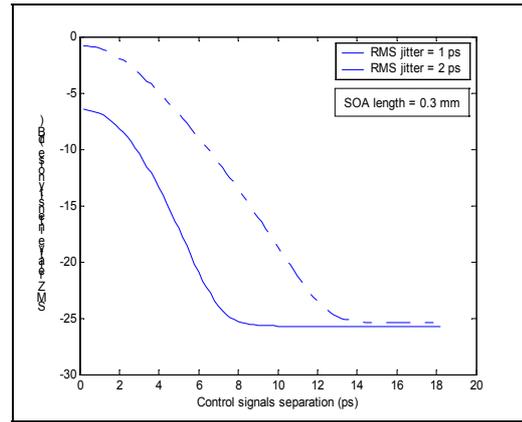
Fig. 2 Gain profile of the data signals in SMZ. Insert is the resulting switching window



(a)

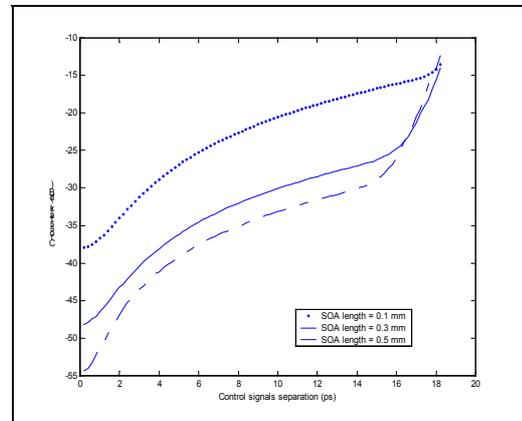


(b)

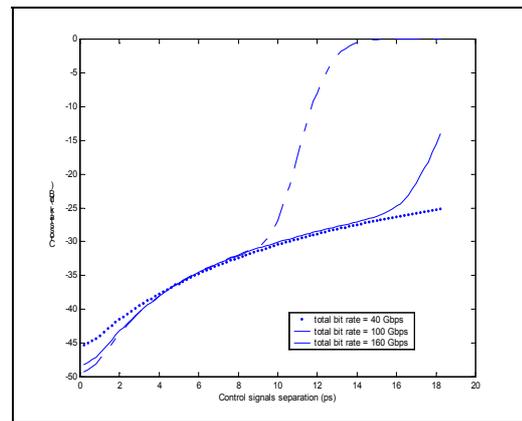


(c)

Fig. 3 Relative intensity noise vs. T_{delay} for different (a) SOA length, (b) data pulse width, and (c) RMS_{jitter}



(a)



(b)

Fig. 4 Channel crosstalk vs. T_{delay} for different (a) SOA lengths, and (b) total bit rates

OFDM Peak-to-Average Power Reduction Scheme with Spectral Masking

Himal A. Suraweera, Kusha R. Panta, Michael Feramez and Jean Armstrong
Department of Electronic Engineering, La Trobe University,
Melbourne Victoria 3086, Australia.
E-mail: j.armstrong@latrobe.edu.au

Abstract— Clipping and filtering can be a simple and effective method of reducing the peak-to-average power ratio (PAPR) for an orthogonal frequency division multiplexed (OFDM) signal. To be effective the clipping must be performed on an oversampled version of the time domain signal, but this results in out-of-band (OOB) power. It has previously been shown that the OOB signals can be filtered out using a discrete Fourier transform (DFT) based filter. However filtering causes peak regrowth, and an increase in PAPR. In this paper it is shown that if the OOB power is limited rather than completely eliminated much less peak regrowth occurs. Simulation results are presented for HIPERLAN2.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) technology is used in many broadband communications systems due to its robustness against frequency selective fading, impulse noise, high bandwidth efficiency and relatively simple receiver implementation.

One of the main disadvantages of OFDM is its high peak-to-average power ratio (PAPR). OFDM transmitters therefore require very linear output amplifiers with wide dynamic range. These are inefficient, expensive and in applications such as wireless local area networks (WLAN) can cause excessive drain on the batteries of portable equipment. Any amplifier nonlinearity causes intermodulation products resulting in unwanted out-of-band (OOB) power. Although PAPR is very large for OFDM, high magnitude peaks occur relatively rarely and most of the transmitted power is concentrated in signals of low amplitude.

The simplest approach to reducing the PAPR in OFDM signals is to clip the high amplitude signals. A variety of clipping techniques have been described in the literature [1], [2]. Some clip the outputs of the inverse discrete Fourier transform (IDFT) before interpolation. However the signal must be interpolated before analogue to digital conversion, and this will cause peak regrowth [3].

To avoid the problem of peak regrowth, the signal can be clipped after interpolation. However this causes significant OOB power. Some papers have described clipping of the interpolated signal followed by filtering [1], [4], [5], [6]. Two different filtering techniques are described in these papers. In [1] a conventional time invariant linear filter is used,

while [4] describes a DFT based filter. This filter passes all of the in-band components and removes the OOB components by nulling the associated discrete frequencies. This approach also has the advantage of causing no intersymbol interference (ISI).

Both forms of filtering result in some peak regrowth. If necessary, this peak regrowth can be reduced by repeating the clipping and filtering operation [4].

Clipping also results in some in-band distortion, but it has been shown that in practice the effect of this on overall bit error rate (BER) is negligible [7]. This is because the main effect of clipping is to shrink the overall signal constellation rather than to add clipping noise. Moreover, the clipping noise is added at the transmitter, rather than the receiver and is subject to frequency selective channel fading along with the wanted component signals.

In [6] the performance of these different filtering techniques are compared. These results suggest that the time-invariant linear filter results in less peak regrowth and lower PAPR than two other techniques including the DFT based filter. The simulation scheme in [6] makes use of a lattice wave-digital filter (LWDF). However this scheme results in greater OOB power than the DFT based filter. In this paper, we show that the improvement in the PAPR with the LWDF in [6] is mainly due to the difference in the amount of the OOB power that is allowed at the output of the filter. We present simulation results for the DFT based filter if some OOB power is allowed at the output of the filter. The amount of OOB power is set according to the spectral mask specified in the HIPERLAN2 standard. Simulation results show that for the same OOB power the DFT based filter causes less peak regrowth than the linear filter considered in [6].

This paper is organized as follows. In Section II we describe the OOB power in OFDM systems and PAPR reduction technique with spectral masking is introduced in Section III. We present simulation results on PAPR reduction and the BER performance of the system in Section IV and finally Section V contains some concluding remarks.

II. OUT-OF-BAND POWER IN OFDM SYSTEMS

The spectrum of the OFDM signal has some OOB power. The form of the OOB spectrum depends very strongly on the detailed design of the OFDM transmitter. However, the OOB

The work in this paper was supported by an Australian Research Council (ARC) Large Grant.

power is mainly due to two factors: the sidelobes of the spectra of the in-band subcarriers and intermodulation products due to nonlinearities in the transmitter output. In the case of an OFDM system with no windowing of each output symbol, then each subcarrier has a $\sin(x)/x$ spectrum. This results in relatively high sidelobes. The nonlinearities are due to the non-ideal amplifier characteristics, or due to deliberate clipping of the signal to eliminate signal peaks. It has been shown that the limiting factor becomes the amplifier nonlinearity when clipping is used [8].

Fig. 1 shows the spectrum of a HIPERLAN2 transmission without windowing and nonlinearities. The power has been averaged across a 1 MHz bandwidth as specified by the HIPERLAN2 standard. A simple way to reduce the level of these sidelobes is to apply some windowing to each symbol before transmission [8]. Filtering of the signal can also reduce the OOB power. If the signal is clipped (without filtering) the out-of-band power increases. Fig. 2 shows the spectrum when the signal is clipped at CR = 2dB and no filtering is applied. Clipping ratio is defined as the ratio of the clipping level to the root-mean-square power of the unclipped baseband signal,

$$CR = 20 \log_{10} \left(\frac{A}{\sigma} \right) \text{ dB.} \quad (1)$$

III. PAPR AND DFT BASED FILTERING WITH OUT-OF-BAND MASKING

A. PAPR

The PAPR of a continuous-time OFDM signal cannot be computed accurately by sampling the signal at Nyquist rate. Hence oversampling is essential to produce accurate PAPR estimates. The discrete time domain OFDM signal oversampled by a factor of I_1 can be expressed as,

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} a(k) \exp \left(\frac{j2\pi nk}{NI_1} \right) \quad (2)$$

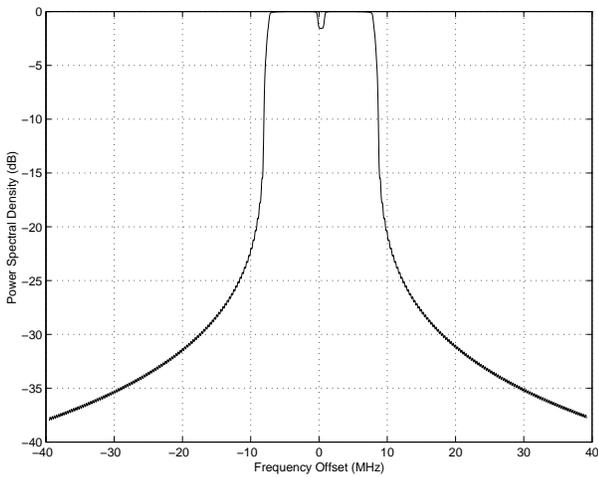


Fig. 1. Power spectral density of the HIPERLAN2 OFDM signal.

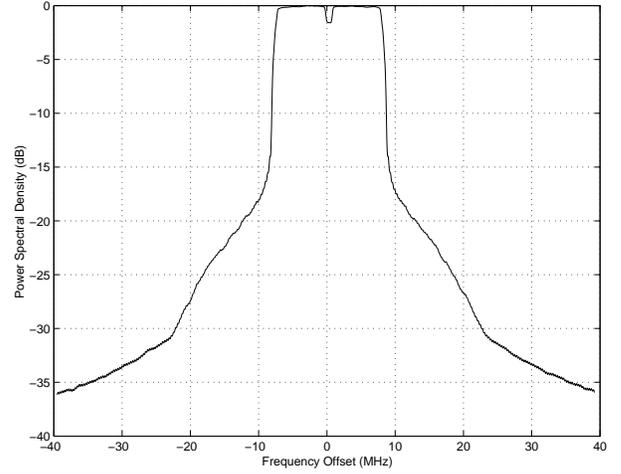


Fig. 2. Power spectra of the clipped HIPERLAN2 OFDM signal. CR = 2.0 dB.

for $n = 0, 1, \dots, NI_1 - 1$. The constructive superposition of complex subcarriers can contribute large peaks for the time domain OFDM signal resulting in a high PAPR. When there are enough subcarriers for the central limit theorem to apply, the real and imaginary components of the OFDM signal have Gaussian distributions. Thus the amplitude has a Rayleigh distribution. Clipping and filtering changes the probability distribution of the amplitude. In this paper we measure the effectiveness of the PAPR reduction techniques in terms of the cumulative distribution of the amplitude, as this is directly related to the OOB power which will result from any amplifier nonlinearities.

B. New Technique with Spectral Masking

Fig. 3 shows the block diagram of the new DFT based PAPR reduction scheme with spectral masking. Vector $\mathbf{A}_i = [a_0, \dots, a_{N-1}]$, which represents the data in each symbol i is converted from frequency to time domain using an oversize IDFT. N is the number of subcarriers in each OFDM symbol. For an oversampling factor of I_1 , the input vector is extended by adding $N(I_1 - 1)$ zeros in the middle of the vector. That is,

$$\tilde{\mathbf{A}}_i = [a_{0,i}, a_{1,i}, \underbrace{0, \dots, 0}_{N(I_1-1) \text{ zeros}}, \dots, a_{N-1,i}] \quad (3)$$

This results in trigonometric interpolation of the OFDM time domain signal [9]. The interpolated signal is then clipped. In this paper we use “amplitude clipping” and the nonlinearity acts on the envelope of the complex baseband signal.

Amplitude clipping :

$$f(r) = \begin{cases} x_n & |x_n| \leq A \\ Ae^{j\arg\{x_n\}} & |x_n| \geq A \end{cases} \quad (4)$$

Clipping is followed by frequency domain filtering to limit the OOB power. The filter consists of a DFT followed by an IDFT operation. The forward DFT transforms

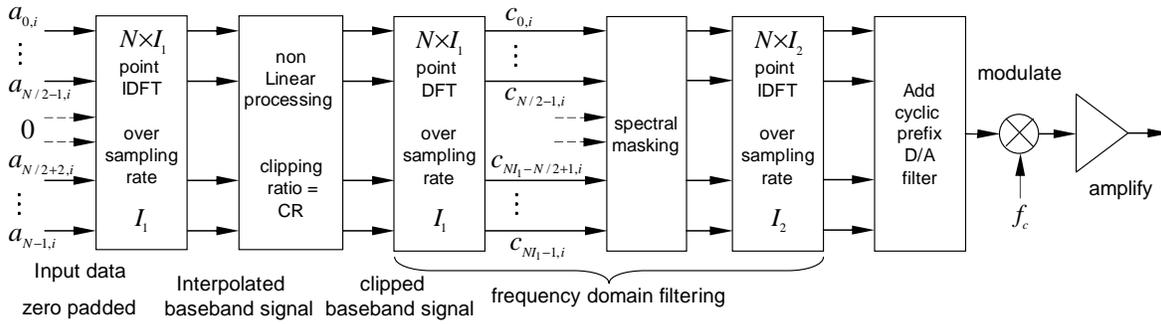


Fig. 3. Block diagram of the spectral masking PAPR reduced OFDM system.

the clipped signal back to the frequency domain. The in-band discrete frequency components of the clipped signal $c_0, \dots, c_{N/2-1}, c_{NI_1-N/2+1}, \dots, c_{NI_1-1}$ are passed unchanged while OOB spectral components are modified according to the spectral mask specifications given by the HIPERLAN2 standard [10]. This is different from the DFT filter described in [4] where all of the OOB components are nulled rather than limited.

The HIPERLAN2 standard [10] specifies the spectral mask which OFDM transmissions must meet. Fig. 4 shows this mask. Each OFDM symbol is composed of 64 subcarriers with a nominal bandwidth of 20 MHz. Of the 64 subcarriers, 47 are data carriers, 4 are pilots and 13 (the 0-th and some band-edge subcarriers) are not used and are set to zero. The band-edge subcarriers are not used as this reduces the analogue filtering requirements. The information bearing subcarriers have an effective bandwidth of 18 MHz while OOB power decays below 40 dBc at a frequency offset of 30 MHz away from the carrier frequency.

IV. PERFORMANCE OF PAPR REDUCTION SCHEME WITH SPECTRAL MASKING

In all simulations 4-QAM modulation was used. Figs. 5 and 6 show the complementary cumulative distribution function (CCDF) of PAPR for CR values of 4.5 and 2.5 dB respectively. Clearly if OOB masking is applied, it can be observed that negligible peak regrowth occurs. Fig. 7 shows the CCDF of PAPR by CCDF of PAPR by applying a different spectral

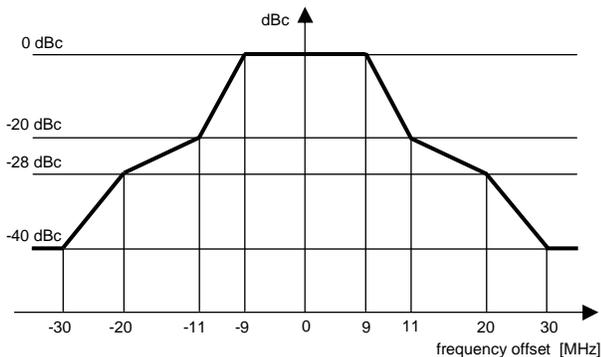


Fig. 4. Spectral power mask of OFDM based HIPERLAN2.

mask for CR = 4.5 dB. The applied mask is 3 dB below than that the mask of HIPERLAN2. In this case some peak growth is observed due to the distortion of the OOB components. However it is negligible compared to that of DFT based filter. Although the peak regrowth with spectral masking is negligible for the case of HIPERLAN2, this will not be true for applications such as digital video broadcasting where the OOB power requirements are much more stringent.

A. BER Performance

The BER performance of a clipped OFDM system depends on the in-band distortion. Hence the signal-to-clipping noise ratio (SCNR) which is defined as,

$$\text{SCNR}(k, i) = \frac{\alpha^2 E[|a(k, i)|^2]}{E[|N_d(k, i)|^2]} = \frac{\alpha^2 \sigma^2}{E[|N_d(k, i)|^2]} \quad (5)$$

and depends on the subcarrier index k is a good measure of the BER performance. In (5) $N_d(k, i)$ is the distortion noise component due to clipping for the k th subcarrier and α is a constant. It can be calculated by extending the Bussgang's theorem to the complex case [4] and is given by,

$$\alpha = (1 - e^{-\frac{A^2}{\sigma^2}}) + \sqrt{\frac{\pi A^2}{4\sigma^2}} \text{erfc}\left(\frac{A}{\sigma}\right) \quad (6)$$

where $\text{erfc}(x) = 2/\sqrt{\pi} \int_x^\infty e^{-y^2} dy$.

With the DFT based filter and nulling there is a trade off between the minimum PAPR and the in-band distortion. Using masking instead of nulling results in a lower PAPR with no increase in the in-band distortion, and so allows a given PAPR target to be achieved with less in-band distortion. Fig. 8 shows the BER performance results in a two-path fading channel where signal power of the echo path is half of that of the direct path. The delay of the echo signal equals 10 OFDM sample periods and perfect channel state information was assumed at the receiver. Cyclic prefix length equals 16. $I_1 = I_2 = 4$. The BER performance of the two OFDM systems, with no clipping and clipped at CR = 6 dB are almost identical having error rates of well beyond 10^{-6} at $E_b/N_0 = 18$ dB. The overall effect of channel fading is to minimize the effect of transmitter-added noise on the error performance of the OFDM system [7].

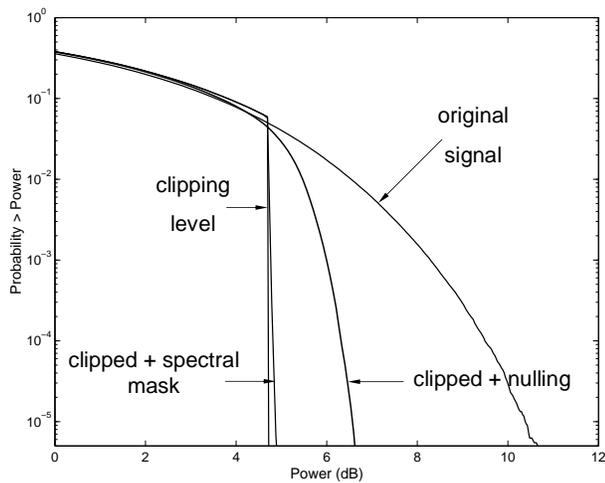


Fig. 5. CCDF for OFDM signals after clipping and spectral masking. CR = 4.5 dB.

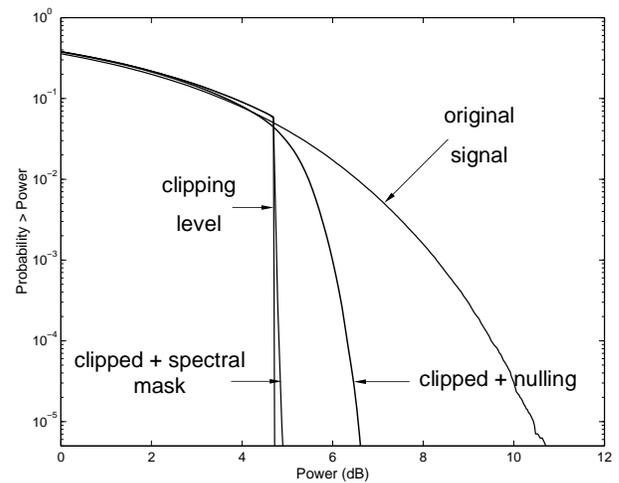


Fig. 7. CCDF for OFDM signals after clipping and spectral masking (3 dB below the HIPERLAN2 mask). CR = 4.5 dB.

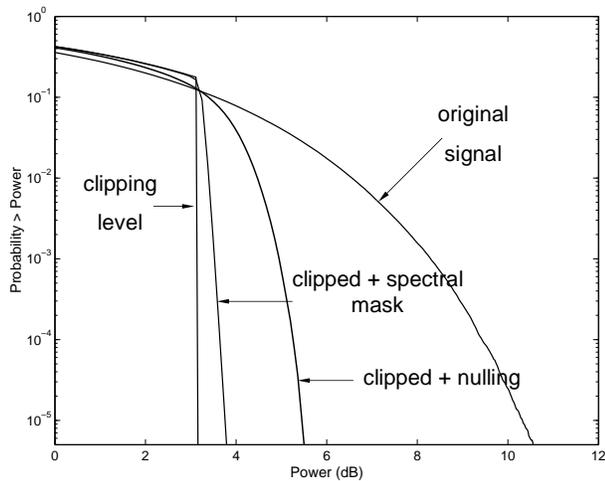


Fig. 6. CCDF for OFDM signals after clipping and spectral masking. CR = 2.5 dB.

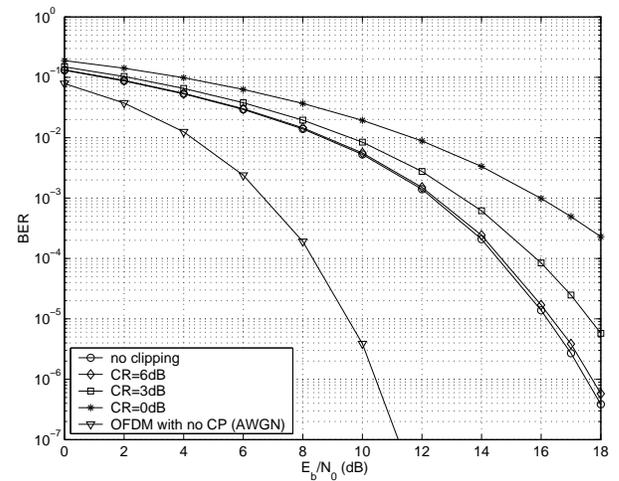


Fig. 8. BER versus E_b/N_0 in a multipath channel for varying CR values.

V. CONCLUSIONS

We have presented a detailed analysis of a DFT based filter with spectral masking to reduce the peak-to-average power ratio (PAPR) in OFDM signals. Spectral masking is used to limit the out-of-band (OOB) power and it is shown that the resulting peak-regrowth will be minimum. System degradation has been discussed in terms of in-band distortion and a bit error rate graph is presented in a frequency selective multipath channel. DFT based spectral masking filter introduces no in-band distortion and the effect of the clipping at the transmitter on the error performance of the OFDM system subject to multipath fading is minimal.

REFERENCES

- [1] X. Li and L. J. Cimini Jr., "Effects of clipping and filtering on the performance of OFDM", *IEEE Commun. Lett.*, vol. 2, pp. 131-133, May 1998.
- [2] R. O'Neill and L. N. Lopes, "Envelope variations and spectral splatter in clipped multicarrier signals," in *Proc. PMIRC '95*, Toronto, Canada, Sept. 1995, pp. 71-75.

- [3] C. Tellambura, "Phase optimization criterion for reducing peak-to-average power ratio for OFDM," *IEE Electron. Lett.*, vol. 34, pp. 169-170, Jan. 1998.
- [4] J. Armstrong, "Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering", *IEE Electron. Lett.*, vol. 38, pp. 246-247, Feb. 2002.
- [5] J. Armstrong, "New OFDM peak-to-average power reduction scheme," in *Proc. of IEEE Vehicular Technology Conf. (VTC' 2001 Spring)*, Rhodes, Greece, May 2001, pp. 756-760.
- [6] L. D. Kabulepa, T. Pionteck, A. Garcia and M. Glesner, "Design exploration for clipping and filtering PAPR reduction techniques in OFDM systems", in *Proc. 8th International OFDM Workshop (InOw'03)*, Hamburg, Germany, Sep. 2003, pp. 108-112.
- [7] K. Panta and J. Armstrong, "Effects of clipping on the error performance of OFDM in frequency selective fading channels," to appear in *IEEE Trans. Wireless Commun.*, 2004.
- [8] R. van Nee and R. Prasad, *OFDM for multimedia communications*, Boston, MA: Artech House, 2000.
- [9] D. Fraser, "Interpolation by the FFT revisited - an experimental investigation", in *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 37, pp. 665-675, May 1989.
- [10] ETSI TS 101 475: "Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; physical (PHY) layer".

Optical Fibre Refractive Index Reconstruction from Mode Near Field

X. Qian and A. C. Boucouvalas

Multimedia Communications Research Group,
School of Design, Engineering and Computing,
Bournemouth University, Fern Barrow, Poole, BH12 5BB, UK
{tboucouv, qxin}@bournemouth.ac.uk

ABSTRACT. A novel and accurate refractive index profile synthesis method for optical fibres is demonstrated using knowledge of the fundamental or higher order mode near field. This fundamental method is based on inverse transmission line principles. From Maxwell's equations, we derive a transmission line equivalent circuit for a circularly symmetric optical fibre with arbitrary refractive index. We demonstrate how to use this model to carry out the inverse problem of synthesis of the refractive index profile from mode near field data. We apply this method to construct waveguides supporting unusual near field patterns, and the accuracy of the reconstructed refractive index profile is examined numerically.

1. Introduction

The refractive index profile of an optical fibre plays an important role in characterizing the properties of the optical fibre. It allows the determination of the fibre's numerical aperture (NA) and of the number of modes propagating within the optical fibre, while defining intermodal and profile dispersion caused by the optical fibre itself. It is essential to establish an efficient and accurate method for measuring the refractive index profile. A number of techniques, [1], [2], [3], have been proposed for determining the refractive index distribution of optical fibres from the propagation mode near field. The most well known rely on the seminal theoretical work by Morishita [2], which relies on an inverse solution of the scalar wave equation for the refractive index profile. In [1], the measurement of the near field intensity is improved using a scanning optical microscopy technique rather than conventional optics. Improvements from [2], have been recently reported in [3], which is a robust method to noise and errors, and non-iterative, but reported for planar waveguides only. We have shown that our transmission line technique can be applied in optical fibres and can determine exactly the mode propagation constants [4]-[5], and cutoff wavelengths of waveguide modes [6]. In general from knowledge of the monomode optical fibre near electric field, we can even synthesize the exact refractive index profiles numerically using this powerful technique [7].

In this paper, first we are extending our results and demonstrate that the same refractive index profile can

be synthesised not only from the mode field of the HE_{1l} mode but also from knowledge of higher order mode fields, hence confirming the generality of this method. We then proceed with synthesizing refractive index profiles which support unusual mode fields such as linear, or sigmoid. Sigmoid fields are very interesting because they offer a mode fields distribution in the core which is very flat, unlike ordinary step index fibres which support Bessel function mode field distributions.

2. Transmission line Representation of Optical fibres

We divide a cylindrical optical waveguide into a large number of concentric homogeneous cylindrical layers of thickness δr , permittivity ϵ , permeability μ and conductivity σ in Fig.1.

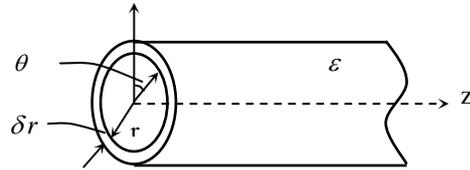


Figure 1. Homogeneous optical fiber thin cylindrical layer

The E and H components of Maxwell's equations for any such layer can be written as (1) and (2), [4]

$$\left. \begin{aligned} \beta r E_{\theta} - l E_z &= \omega \mu r H_r \\ l H_z - \beta r H_{\theta} &= (\omega \epsilon - j \sigma) r E_r \\ \frac{\partial(\omega \mu r H_r)}{\partial r} &= -j \omega \mu (l H_{\theta} + \beta r H_z) \end{aligned} \right\} \quad (1)$$

where $\gamma^2 = \beta^2 + \left(\frac{l}{r}\right)^2 - \omega^2 \mu \epsilon + j \omega \mu \sigma$, β is the propagation constant, l is the azimuthal mode number (integer), and ω is the mode frequency. We restrict our analysis to the case $\sigma = 0$, $\mu = \mu_0$, $\epsilon = n^2 \epsilon_0$, where n is the refractive index of the layer at distance r from the axis.

Equation (3) represents two independent transmission lines with voltages V_s , V_d and currents I_s , I_d .

$$\left. \begin{aligned} \frac{\partial[(\omega\varepsilon - j\sigma)rE_r]}{\partial r} &= -(\sigma + j\omega\varepsilon)(IE_\theta + \beta rE_z) \\ \frac{\partial(IH_\theta + \beta rH_z)}{\partial r} &= -\frac{\gamma^2}{j\omega\mu}\omega\mu rH_r + \beta H_z - \frac{l}{r}H_\theta \\ \frac{\partial(IE_\theta + \beta rE_z)}{\partial r} &= -\frac{\gamma^2}{\sigma + j\omega\varepsilon}(\omega\varepsilon - j\sigma)rE_r \\ &\quad + \beta E_z - \frac{l}{r}E_\theta \end{aligned} \right\} (2)$$

$$\left. \begin{aligned} Z_B &= \frac{1}{2}(\delta r)^2 \gamma_s^2 Z_P \\ Z_P &= \frac{Z_0}{jn r \delta r k_0 (\beta^2 + (\frac{l}{r})^2)} \end{aligned} \right\} (6)$$

After some algebra similarly to [7], (1) and (2) can be transformed into

$$\left. \begin{aligned} \frac{\partial V_s}{\partial r} &= \frac{-\gamma_s^2}{j\omega\varepsilon_0 n F} I_s, \quad \frac{\partial I_s}{\partial r} = -j\omega\varepsilon_0 n F V_s \\ \frac{\partial V_d}{\partial r} &= \frac{-\gamma_d^2}{j\omega\varepsilon_0 n F} I_d, \quad \frac{\partial I_d}{\partial r} = -j\omega\varepsilon_0 n F V_d \end{aligned} \right\} (3)$$

where $\gamma_s^2 = \beta^2 + (\frac{l}{r})^2 - n^2 k_0^2 \mp \frac{2nk_0\beta l}{(\beta r)^2 + l^2}$

(- for *HE* and + for *EH* modes).

The corresponding characteristic impedances are

$$\left. \begin{aligned} Z_s &= \frac{\gamma_s}{j\omega\varepsilon_0 n F} \\ Z_d &= \frac{\gamma_d}{j\omega\varepsilon_0 n F} \end{aligned} \right\} (4)$$

The above equations are recognized as the well known transmission line equations with the solution represented by the following electric circuit Fig.2.

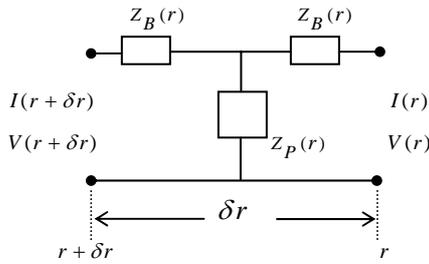


Figure2. Equivalent circuit of a dielectric waveguide

With impedances

$$\left. \begin{aligned} Z_B &= Z_s \tanh(\gamma_s \frac{\delta r}{2}) \\ Z_P &= \frac{Z_s}{\sinh(\gamma_s \delta r)} \end{aligned} \right\} (5)$$

where δr is the length of the transmission line.

Since δr is infinitesimal, $\frac{\delta r}{r} \ll 1$, we finally have

Normalizing (6) with respect to k_0 gives

$$\left. \begin{aligned} \bar{Z}_B &= \frac{1}{2}(\delta \bar{r})^2 \bar{\gamma}_s^2 \bar{Z}_P \\ \bar{Z}_P &= \frac{Z_0}{jn \bar{r} \delta \bar{r} (\bar{\beta}^2 + (\frac{l}{\bar{r}})^2)} \end{aligned} \right\} (7)$$

where $\bar{r} = rk_0$, $\delta \bar{r} = \delta rk_0$, $\bar{\beta} = \frac{\beta}{k_0}$

$$\bar{\gamma}_s^2 = \gamma_s^2 / k_0^2 = \bar{\beta}^2 + (\frac{l}{\bar{r}})^2 - n^2 \mp \frac{2n\bar{\beta}l}{(\bar{\beta}\bar{r})^2 + l^2}$$

$\bar{Z}_P = Z_P \times k_0$, $\bar{Z}_B = Z_B \times k_0$, $Z_0 = 120\pi$ is the free space impedance.

A waveguide can therefore be modelled as a cascade of T circuits, the impedances, voltages and currents of which depend on the waveguide physical and optical properties. For determining the refractive index profile from knowledge of E_r , we assume the following boundary condition: At $r = \infty$, we assume $Z_{prev} = 0$

and $n=n_2$ (silica refractive index). The equivalent circuit for a cylindrical thin layer, Fig.1, of constant refractive index n and thickness $\delta \bar{r}$ at distance \bar{r} from the core is represented as an electric circuit in Fig.3. From circuit theory, we may use the following recursive relation to determine the values of $Z_{p,n}$,

$$Z_{B,n} \text{ and } Z_{prev} = Z_n.$$

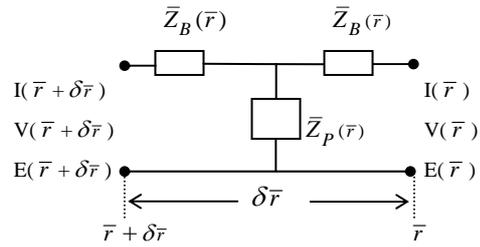


Figure3. Equivalent circuit for a cylindrical thin layer at constant refractive index n and thickness $\delta \bar{r}$ at distance \bar{r} from the core, where V, I, E are the electric voltages and currents in the circuit and E is the corresponding mode Electric field.

From (8), since we know that $\bar{\beta}$ is the effective refractive index and for typical waveguides lies between n_1 and n_2 . For the time being, we also assume

we have full knowledge of $\bar{\beta}$. We also know λ_0 , \bar{r} , $\delta\bar{r}$, $n(\infty) = n_2$.

$$\left. \begin{aligned} \bar{Z}_n &= \frac{(\bar{Z}_{n-1} + \bar{Z}_{B,n})\bar{Z}_{P,n} + \bar{Z}_{B,n}}{\bar{Z}_{n-1} + \bar{Z}_{B,n} + \bar{Z}_{P,n}} \\ \bar{Z}_{B,n} &= \frac{1}{2}\bar{\gamma}^2\delta\bar{r}^2\bar{Z}_P \\ \bar{Z}_{P,n} &= \frac{Z_0}{n(r)\bar{r}\delta\bar{r}(\bar{\beta}^2 + \frac{l^2}{\bar{r}^2})} \end{aligned} \right\} \quad (8)$$

Where $\bar{\gamma}^2 = \bar{\beta}^2 + \frac{l^2}{\bar{r}^2} - n^2(r) - \frac{2\bar{\beta} \ln(r)}{(\bar{\beta}^2 \bar{r}^2 + l^2)}$

Hence $I(\bar{r} + \delta\bar{r}), V(\bar{r} + \delta\bar{r})$ can be calculated for any radius

$$E(\bar{r}) = \frac{I_E(\bar{r})Z_0}{n^2\bar{r}}, \quad I_s = \frac{2I_E}{\sqrt{n}} \quad (9)$$

Since we know $E(\bar{r})$ and $I_s(\bar{r})$, hence we can calculate $n(\bar{r})$, and synthesize the refractive index profile..

$$n(\bar{r}) = \left[\frac{I_s(\bar{r})Z_0}{2E(\bar{r})\bar{r}} \right]^{2/3} \quad (10)$$

3. Numerical Results

Fig.4 shows the exact calculated fields of HE_{11} , HE_{12} and HE_{21} modes of a segmented core optical fibre. By using (10) with wave number $l=1$ for HE_{11} , HE_{12} and $l=2$ for HE_{21} , we can use the inverse method to reconstruct refractive index profile of the segmented optical fibre. Fig.5(a) shows the reconstructed refractive index profile of a segmented optical fibre which agrees with the original refractive indices of $n_1=1.51508$, $n_2=1.508$, and $n_3 = 1.512$ using the HE_{11} mode near field data. We compare its accuracy with the original (exact) refractive index in Fig.5(b), which shows the error (%) versus the normalised radius. The error in refractive index shows small oscillations about the exact value in the core. The error in the cladding is much smaller. In Fig.6 we show the effect of inaccuracies in $\bar{\beta}$ on the ripple in the reconstructed Δn for HE_{11} , HE_{12} and HE_{21} modes. We observe that for each mode the ripple increases with using the incorrect $\bar{\beta}$. Since the min. error occurs at the exact $\bar{\beta}$, this gives us a means of locating the unknown $\bar{\beta}$, since we can simply start the reconstruction with $\bar{\beta} = n_2$ and repeat the reconstruction process with a new $\bar{\beta}$ within $n_2 \leq \bar{\beta} \leq n_1$ until the observed ripple is minimised. At the observed minimum error ripple corresponding to the exact $\bar{\beta}$, the reconstructed refracted index is also exact. HE_{11} , HE_{12} and HE_{21}

modes have their own exact $\bar{\beta}$, the values as shown in Fig. 6. The HE_{11} mode has the maximum $\bar{\beta}$ among the three modes.

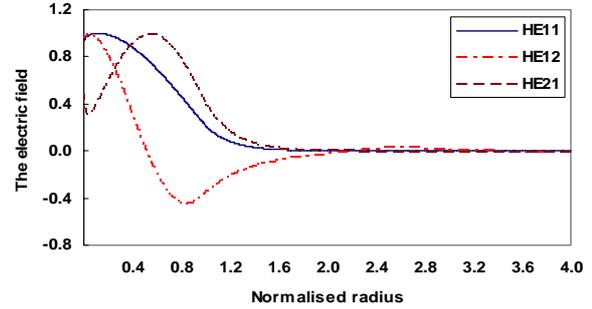
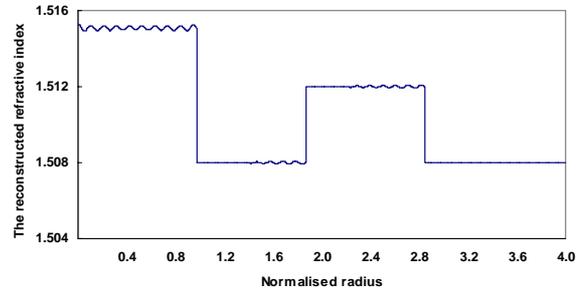
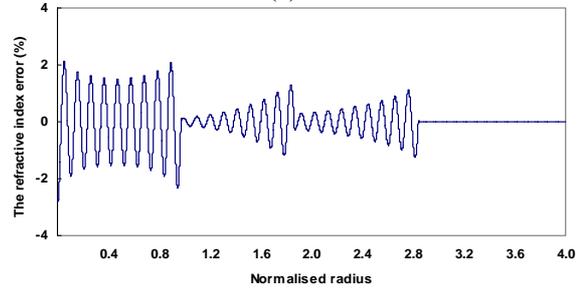


Figure4. The HE_{11} , HE_{12} and HE_{21} modes near electric fields of a segmented core fibre with $n_1=1.51508$, $n_2=1.508$, and $n_3 = 1.512$ and $\lambda = 0.87054 \mu m$.



(a)



(b)

Figure5. (a).The reconstructed refractive index using (10), and HE_{11} mode field data. (b).The % error in the reconstructed refractive index profile, (the difference between the exact and reconstructed refractive index profile) $(100(n_{exact} - n_{reconstructed}) / n_{exact})\%$

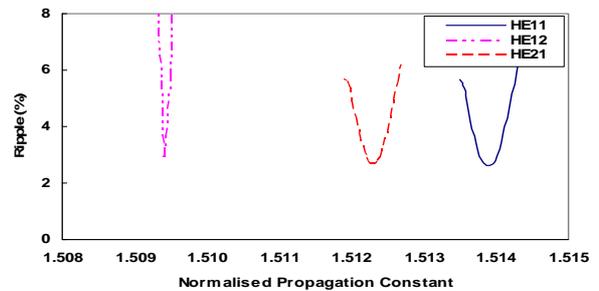


Figure6. The refractive index difference, ripple (%) of the synthesized refractive index versus values for $\bar{\beta}$ offset from exact

In order to illustrate further the power of our technique, let us now consider optical fibres with special near field profile. Fig.7(a) shows two examples of desired triangular, (linear) near electric field with different slopes. In Fig.7(b), we can see the refractive index reconstruction supporting the fields. We demonstrate that our inverse transmission line method reconstructs the refractive index profile successfully for this special near field profile. Fig.8(a) is another example this time of a sigmoid near field profile. Sigmoids show near constant core field intensities. The fibre refractive index reconstruction using sigmoid near field data is demonstrated in Fig.8(b). In this case, we have forced the cladding index to be constant. This is allowed with our technique and results in deeper gratings in the core since the technique is compensating for the enforcement in a shorter radius range (core only).

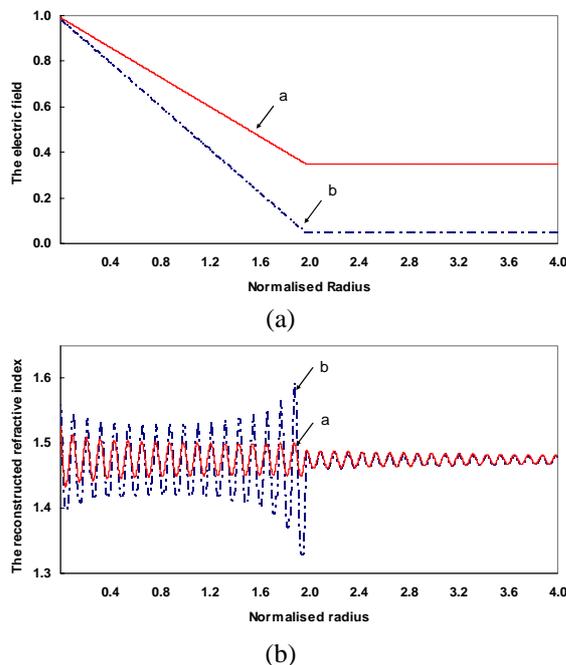


Figure7.(a).Examples of near electric field (triangular profile) with different slopes. (b).The reconstructed refractive index profiles from the triangular profile electric fields.

References

- 1 L. Dhar, H. J. Lee, E. J. Laskowski, S. K. Buratto, H. M. Presby, C. Narayanan, C. C. Bahr, P. J. Anthony, M. J. Cardillo "Refractive index profiling of optical waveguides using near-field scanning optical microscopy", OFC 1996, paper FB-5, pp.303-304.
- 2 K. Morishita "Refractive-Index-Profile Determination of Single Mode Optical Fibres by a Propagation-Mode Near Field Scanning Technique", IEEE J. Lightwave technology Vol LT-1, No 3, Sept. 1983, pp.445-449.
- 3 J. H. Lin and C. K. Chen "An inverse Algorithm to Calculate the refractive Index Profiles of Periodically Segmented Waveguides from the Measured Near-Field Intensities", IEEE J. Lightwave Technology, Vol.20, No 1, Jan 2002, pp.58-64.
- 4 C. D. Papageorgiou and A. C. Boucouvalas "Propagation constants of cylindrical dielectric waveguides with arbitrary

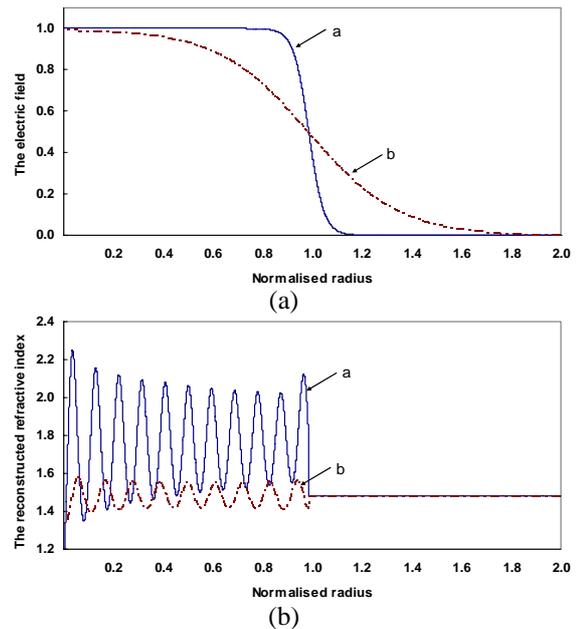


Figure8.(a).Examples of near electric field (sigmoid profile) with different slopes. (b).The reconstructed refractive index profiles from the sigmoid profile electric fields.

4. Conclusions

In this paper, a new, simple and accurate numerical index profile synthesis procedure has been developed and demonstrated with practical examples. The method uses inverse transmission line principles and relies on the modelling of a thin uniform cylindrical layer of an optical fibre to a transmission line circuit. The method requires knowledge of the near electric field of the optical fibre and the reconstruction is theoretically exact. It allows us to create waveguides supporting modes of nearly arbitrary field distribution, which can be useful in many applications in fibre optic sensing and other applications such as refractive index profiling instrumentation.

- refractive index profile, using the 'Resonance' technique", Electronics Letters, Vol 18, No 18, 2 September 1982, pp.768-788.
- 5 A. C. Boucouvalas and S. C. Robertson "Optical waveguide transverse transmission line equations and their use in determining mode properties", JOERS Advanced Fibre Measurement Symposium, National Physical Laboratories, London, 10 September 1985.
- 6 A. C. Boucouvalas and C. D. Papageorgiou "Cutoff frequencies in optical fibres of arbitrary refractive index profile using the 'resonance' technique", IEEE Journal of Quantum Electronics, Vol QE-18, No 12, Dec. 1982, pp. 2027-2031.
- 7 A. C. Boucouvalas and X. Qian "Optical Fibre Refractive Index Profile Synthesis from Near Field", IEEE Globecom 2003, OP6-5, San Francisco, USA, November 2003.

Optimum Noise Thresholds in Decision Directed Impulse Noise Mitigation for OFDM

Jean Armstrong, Michael Feramez and Himal A. Suraweera
Department of Electronic Engineering, La Trobe University,
Melbourne Victoria 3086, Australia.
Email: j.armstrong@latrobe.edu.au

Abstract—Impulse noise is a significant problem in some OFDM applications including digital television broadcasting. In this paper we study the optimum threshold values for a novel decision directed impulse noise mitigation algorithm. In this algorithm the noise component in each received input sample is estimated based on preliminary decisions on the transmitted data. When the noise estimate is above a given threshold, it indicates that impulse noise is present in the sample and the estimated noise component is subtracted from the input sample before final demodulation. In this paper the optimum threshold levels for varying impulse noise parameters are calculated.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) technology is used in many digital broadband communication systems. One of the advantages of OFDM compared to single carrier systems is that it is more resistant to the effects of impulse noise because of the spreading effect of the discrete Fourier transform (DFT). However impulse noise can still cause significant problems in OFDM systems. This is a major practical problem in digital video broadcasting (DVB) [1-5].

The theoretical effects of impulse noise in multicarrier systems have been analyzed [6], and a number of techniques for mitigating the effect of impulse noise have been described. One approach is to identify peaks in the received time domain signal and reduce these by either clipping or nulling the sample [5], [8], [9]. This is effective only for impulse noise with peaks larger than the wanted OFDM signal. This will be true only in very extreme cases. In high signal to noise environments such as broadcast television, the impulse noise can be well above the background Gaussian noise, yet well below the OFDM signal.

Several authors have used techniques that operate on the signal in the frequency domain [7], [10], [11]. Häring and Han Vinck [7] describe an iterative process in which information is exchanged between estimators operating in the time and frequency domains. The simulation results they present are for extreme cases with very large noise impulses. In [10], impulses are detected in the frequency domain by identifying subcarriers with extreme values. In [11] the positions of noise impulses are identified using pilot tones.

Very recently, decision directed impulse mitigation has been developed separately and independently by two groups [4], [12-14]. Some details of the techniques are slightly different, but the basic concept is the same. Preliminary decisions are made about the transmitted data and from these an estimate is made of the noise in the received signal. The estimated noise is subtracted from the original signal before final demodulation. When the input noise is impulsive, the technique substantially reduces the noise power. The technique depends on the fact that the signal appears random in the time domain and highly structured in the discrete frequency domain whereas for the impulse noise the converse is true.

In [13], [14], a more theoretical approach is taken, and an analysis of the decision and noise estimation processes is presented. Whereas [4] has a more practical emphasis with results being presented for noise captured from a real world interference source.

The decision directed estimation technique requires a non linear noise estimation function. In this paper threshold type non-linearities are investigated and the optimum thresholds found for varying impulse noise parameters.

II. IMPULSE MITIGATION TECHNIQUE

Fig. 1 shows the block diagram of a receiver with the new mitigation technique. The received OFDM baseband signal samples are given by

$$x(l) = r(l) + n_g(l) + n_i(l) = r(l) + n_t(l) \quad (1)$$

where $r(l)$ is the wanted OFDM signal, $n_g(l)$ is the Gaussian noise and $n_i(l)$ is the impulse noise. $n_t(l) = n_g(l) + n_i(l)$ is the total noise at the input. The samples $x(l)$ are optionally passed through a non-linearity that clips or nulls large samples. The samples at the output of the non-linearity, $z(l)$, are serial-to-parallel converted to form the vector of N complex samples that are input to the N -point DFT. The output of the DFT is the N -point vector Z . Preliminary decisions, $\hat{D}_p(k)$, about the transmitted data are made based on $Z(k)$. The noise component of $Z(k)$ is $N_t(k)$. The observed noise, $N_p(k)$, is calculated using

$$N_p(k) = Z(k) - \hat{D}_p(k) \quad (2)$$

This work was supported by an Australian Research Council (ARC) Large Grant.

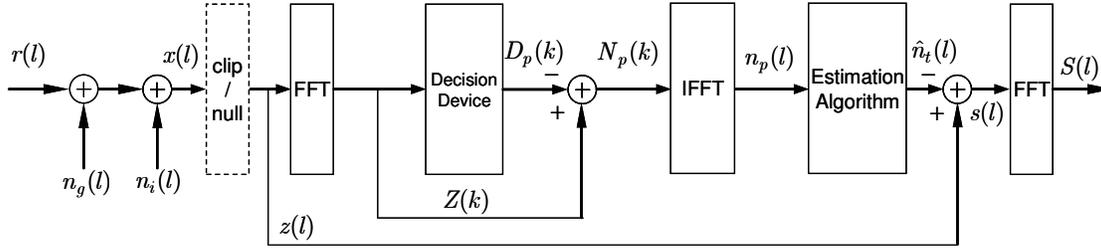


Fig. 1. Block diagram of receiver with impulse mitigation.

Except for extreme cases, most of the received subcarriers are correctly decoded and the observed noise is exactly equal to the received noise in that subcarrier. In the cases where the subcarrier is incorrectly decoded, ‘decision noise’ will be added to the observed value.

OFDM is more resistant to the effects of impulse noise than single carrier systems because of the spreading effect of the receiver DFT operation. The energy of each impulse is spread evenly across all of the subcarriers in that symbol. When there is more than one impulse in a received symbol period, T , the contributions combine linearly in each subcarrier. When there are enough impulses during T for the central limit theorem to apply, $N_i(k)$ has a Gaussian distribution.

The vector N_p is then converted back into the discrete time domain using an inverse FFT to give the vector \mathbf{n}_p . If there are no decision errors, $n_p(l) = n_t(l)$. However even in the presence of decision errors $n_p(l)$ contains some information about $n_t(l)$. $n_p(l)$ is then input to an estimation device to generate an estimate $\hat{n}_t(l)$ of the total input noise. This is subtracted from $z(l)$ to generate $s(l)$. The rest of the receiver is a standard OFDM receiver consisting of DFT etc.

The task of the estimation algorithm is to estimate the presence and size of noise impulses. A number of algorithms are possible. In this paper, we consider two estimation algorithms: a threshold operating on the real and imaginary components separately and an estimation algorithm operating on the amplitude of each sample. If the estimated noise component in a sample is above the threshold, then the estimated value is subtracted from the input signal before the second stage of demodulation if it is below the threshold the input sample is unchanged.

The operation of the amplitude non-linearity is described by:

$$\begin{aligned} \hat{n}_t(l) &= an_p(l) \quad \text{for } |n_p(l)| > \alpha \\ &= 0 \quad \text{for } |n_p(l)| < \alpha \end{aligned} \quad (3)$$

To describe the second non-linearity, represent $\Re(n_p(l))$ as r_p , $\Re(n_t(l))$ as r_t and $\Re(n_d(l))$ as r_d . Then the operation of the real non-linearity is described by

$$\begin{aligned} \hat{r}_t &= ar_p \quad \text{for } |r_p| > \alpha \\ &= 0 \quad \text{for } |r_p| < \alpha \end{aligned} \quad (4)$$

The operation of the imaginary non-linearity is identical.

For the technique to be effective in reducing the overall bit error rate (BER) of the system, $n_t(l)$ must be impulsive (not stationary Gaussian) and the estimation algorithm must be non-linear. If the estimation process is linear, it will appear to improve the received constellation as each point moves towards the value $\hat{D}_p(k)$, but the points will move closer to both incorrect and correct decision points. However, the technique is very effective if non-linear processing is used and the noise is impulsive. This depends on the fact that for large values of $n_t(l)$, $n_p(l) \approx n_t(l)$.

Fig. 1 shows an optional clipping or nulling function operating on the received baseband samples. This reduces the effect of very large noise impulses that are above the envelope of the OFDM signal. However simulations have shown that this improves the performance only in very extreme cases [13].

III. IMPULSE NOISE MODELS

A number of models for impulse noise have been presented in the literature [1], [2], [15], [16]. Some characterize only the probability density function of the amplitude of the noise, whereas others also consider the time correlation of impulse events. Very recent research by the BBC, which measured a variety of impulse noise sources, has shown that many of the impulse noise sources of practical importance in OFDM applications can be modeled as gated Gaussian noise [1], [2].

In this paper we use a particular form of gated Gaussian noise, where the noise is the sum of additive white Gaussian noise (AWGN) of variance σ_n^2 and a second higher variance Gaussian noise component which lasts for a fraction, μ , of the time duration of each OFDM symbol and which has variance σ_i^2 during this time. (i.e. the variance is calculated over only μT not over T). In general $\sigma_i^2 \gg \sigma_n^2$. The total noise power is then $\sigma^2 = \mu\sigma_i^2 + \sigma_n^2$. Each of these variances is for the real and imaginary components taken separately. The impulsive samples are spread randomly throughout each OFDM symbol.

The gated Gaussian model is used because it gives a good indication of the performance of OFDM systems. Here the critical factor is whether the BER for each symbol is above the threshold at which the error correcting coding will reduce the final BER to an acceptable level, rather than the BER averaged over the entire received signal. It also allows the length and power of the impulse noise to be varied in a way that makes clear the practical implications of the technique. For example, in the context of DVB, it indicates how the resistance to impulse noise can be improved by increasing the transmitter power or choosing the 8k rather than 2k mode.

IV. SIMULATION RESULTS

Matlab simulations were used to examine how the performance of decision directed noise mitigation technique depends on the non-linear estimation process. The performance is measured both in terms of symbol error rate (SER) and the normalized mean square error of the noise estimation. The normalized error is given by: $E\{|\hat{n}_t - n_t|^2\} / E\{|n_t|^2\}$.

The simulations are for a flat fading channel. For each simulation, the average power in each of the real and imaginary components of the wanted OFDM signal is unity. Figs. 3 and 4 show the resulting SER as a function of E_b/N_0 where N_0 is the single sided spectral density of the white Gaussian (non impulse component). The impulse noise parameters are $\sigma_i^2 = 0$ dB and $\mu = 0.01$. In other words, for each plot, the impulse noise is kept constant and the effect of varying E_b/N_0 is measured. 64QAM modulation and 2048 subcarriers were used. In the simulations it was assumed that all subcarriers were carrying data and no pilot tones were used. The thresholds are standardized in terms of the standard deviation of the wanted OFDM signal. Fig. 2 shows the results for amplitude thresholding, while Fig. 3 shows the results for real/imaginary thresholding. The weighting factor was set at $a = 1$. It is clear that for these parameters the use of an amplitude threshold give better performance than a real/imaginary threshold. Simulation for other parameters (not shown) indicate that this is true in general, so the rest of the simulation results are for the amplitude case. Amplitude thresholding gives better results because in the gated Gaussian model, both the real and imaginary components of a given sample are either impulsive or non-impulsive. This will also be true for real world impulse noise sources.

Fig. 4 shows how the normalized mean square error in the noise estimation process varies with threshold. This is of interest because calculation of mean square error rather than BER is more mathematically tractable. In [13] an analysis was presented in terms of minimum mean square error. A comparison of Figs. 2 and 4 indicates that the threshold which gives the minimum mean square error (MMSE) (0.4) is not the

same as the threshold which gives the minimum SER (0.5). This is also true for other parameters and is because the time domain noise is partially correlated.

Fig. 5 explores this relationship. From this it can be seen that the BER follows a typical curve (inverted in the x-axis because the plot is versus noise rather than SNR) for higher thresholds but not for lower thresholds.

Fig. 6 shows the normalized mean square error as a function of threshold for $\sigma_i^2 = -12$ dB and $\mu = 0.16$. In other words the impulsive noise is for a longer proportion of the symbol period but at a lower level. The values are chosen so that the total noise power is the same as in the earlier figures. For these values the impulse mitigation is still effective but Fig. 6 shows (and the SER simulations confirm) that the selection of the threshold level is much more critical in this case. This is because the impulsive noise is only slightly above the background white Gaussian noise and the 'decision noise'.

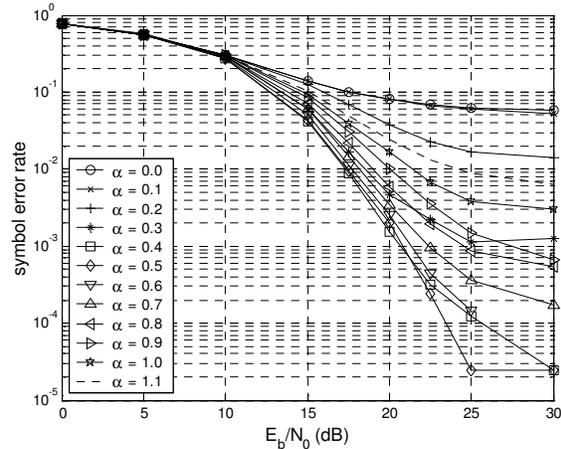


Fig. 2. SER versus E_b/N_0 for amplitude threshold and varying α , for $\sigma_i^2 = 0$ dB and $\mu = 0.01$.

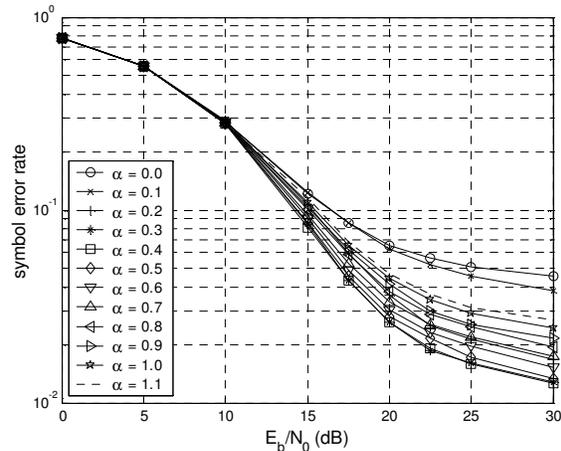


Fig. 3. SER versus E_b/N_0 for real and imaginary thresholding and varying α , for $\sigma_i^2 = 0$ dB and $\mu = 0.01$.

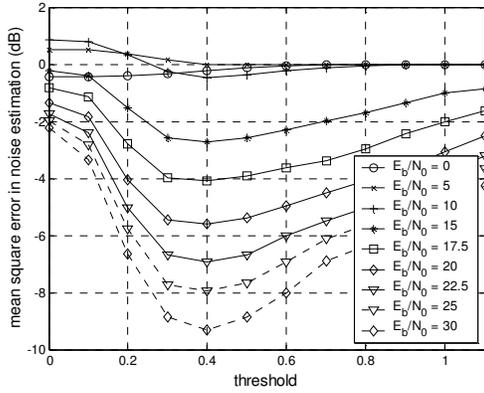


Fig. 4. Normalized mean square error versus amplitude threshold, α and varying E_b/N_0 , for $\sigma_i^2 = 0$ dB and $\mu = 0.01$.

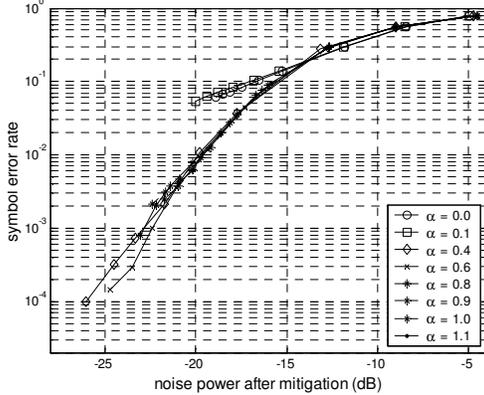


Fig. 5. SER versus normalized mean square error for varying amplitude threshold α and $\sigma_i^2 = 0$ dB and $\mu = 0.01$.

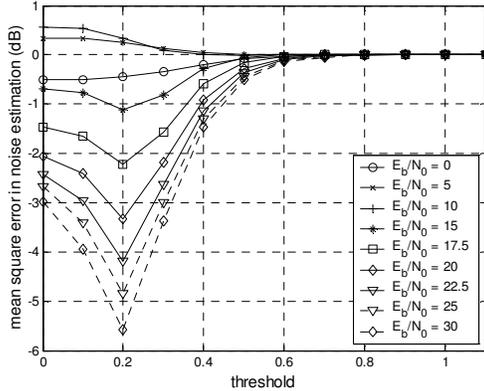


Fig. 6. Normalized mean square error versus threshold amplitude threshold, α and varying E_b/N_0 , for $\sigma_i^2 = -12$ dB and $\mu = 0.16$.

V. CONCLUSIONS

Simulation results have been presented to show the effect of using different non-linear estimation algorithms in decision directed impulse mitigation for OFDM. It is shown that better performance is obtained using thresholds based on the amplitude of the complex baseband signal rather than non-linearities which operate on the real and imaginary components separately.

The ultimate measure of performance is the overall SER, however mean square error is more mathematically tractable. It was found that for given simulation parameters, the threshold which results in the minimum mean square error is generally lower than the threshold which gives the minimum SER. This is because of correlation between the estimation errors in different samples. For given impulse noise parameters, the optimum threshold varies only slightly with the level of background white noise

The impulse noise mitigation gives the greatest improvement in performance when the impulsive noise energy is concentrated in a small proportion of a symbol period. For a given total impulse noise energy, the choice of threshold becomes more critical if the energy is spread over more signal samples.

REFERENCES

- [1] J. Salter and J. Lago-Fernández, "DTT comparison of 64-QAM with 16QAM – Co-channel interference from PAL, echoes and impulsive interference," EBU Technical Review, Oct. 2003
- [2] J. Lago-Fernández and J. Salter, "Modelling impulsive interference in DVB-T: Statistical analysis, test waveforms and receiver performance," BBC R&D White Paper, [Online]. Available: <http://www.bbc.co.uk/rd/pubs/whp/whp080.html>
- [3] Y. Wu, "Performance comparison of ATSC 8-VSB and DVB-T COFDM transmission systems for digital television terrestrial broadcasting," *IEEE Trans. Cons. Elect.*, pp. 916-924, Aug. 1999.
- [4] S. V. Zhidov, "Impulse noise suppression in OFDM based communication systems," *IEEE Trans. Cons. Elect.* vol. 49, pp. 944-948, Nov. 2003.
- [5] J. H. Stott, "Detection and removal of clipping in multi-carrier receivers," European Patent Office, EP 1043874, Oct. 2000
- [6] M. Ghosh, "Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems," *IEEE Trans. Commun.*, vol. 44, pp. 145-147, Feb. 1996.
- [7] J. Häring and A. J. Han Vinck, "Iterative decoding of codes over complex number for impulsive noise channels," *IEEE Trans. on Inform. Theory*, vol. 49, pp.1251-1260, May 2003.
- [8] T. N. Zogakis, P. S. Chow, J. T. Aslanis and J. M. Cioffi, "Impulse noise mitigation strategies for multicarrier modulation", in *Proc. IEEE ICC93*, May 1993, pp. 784-788.
- [9] H. A. Suraweera, C. Chai, J. Shentu and J. Armstrong, "Analysis of impulse noise mitigation techniques for digital television systems," in *Proc. 8th International OFDM Workshop*, Hamburg, Germany, Sept. 2003. pp. 172-176.
- [10] M. Slišković, "Signal processing algorithm for OFDM channel with impulse noise", in *Proc. IEEE ICECS 2000*, Lebanon, Dec. 2000, pp. 222-225.
- [11] F. Abdelkefi, P. Duhamel and F. Alberge, "On the use of pilot tones for impulse noise cancellation in Hiperlan2," in *Proc. ISSPA*, Kuala Lumpur, Malaysia, Aug. 2001, pp. 591-594.
- [12] J. Armstrong and H. A. Suraweera, "Method for multicarrier reception," Australian Prov. Patent 2003905104, Sept. 2003.
- [13] J. Armstrong and H. A. Suraweera, "Impulse noise mitigation for OFDM using decision directed noise estimation," *IEEE ISSSTA 2004*, in press.
- [14] J. Armstrong and H. A. Suraweera, "Decision directed impulse noise mitigation for OFDM in frequency selective channels," *IEEE GLOBECOM 2004*, submitted.
- [15] S. Y. Tan, Y. H. Chew, T. T. Tjhung and M. W. Wu, "A two-dimensional impulse noise model and error probability of QAM in DSL," *IEEE Trans. Commun.*, vol.51, pp. 129-134, Feb. 2003.
- [16] A. D. Spaulding and D. Middleton, "Optimum reception in an impulsive interference environment-part I: Coherent detection," *IEEE Trans. Commun.*, vol. 25, pp. 910-923, Sept. 1977.

Performance Evaluation of Medium Dependence of Turbo Codes

Seema Verma & Rekha Govil

Department of Computer Science & Electronics

Banasthali Vidyapith (INDIA) – 304022

Phone No.:91+1438+228647 ; Fax: 91+1438+228649

vseema@banasthali.ac.in ; seemaverma3@yahoo.com

ABSTRACT

Turbo codes, introduced in 1993 are error correcting codes, and are only recently being employed in some of the communication protocols. The performance of any error correcting scheme can be estimated by extensive modeling of the transmission media and such studies bear important decisive implications in the design of transmission systems. In the present paper the results of the simulation studies of their performance dependence on medium are reported. The studies have been made for the communication channels twisted pair, coaxial cable, microwaves & optical fibers.

KeyWords

Turbo Codes, Mathematical Modelling of AWGN Channel, Rayleigh Fading, Chi-square channel model.

1. Introduction

Information transferred within an electronic communication channel is always liable to corruption by noise within channel. Signals conveying information can be so contaminated by noise that information becomes erroneous. It may be possible to reduce the level of noise but its complete elimination is not possible. The codes referred to in the channel coding theorem do not prevent the occurrence of errors but rather allow their presence to be detected and corrected. As such the codes are known as error – detecting and error – correcting codes or for short **error – control codes**. Turbo codes, a new class of error correcting code were introduced in 1993 and are now considered among the most important developments in coding theory. Parallel concatenated coding schemes employing convolutional codes as constituent codes linked by an interleaver have been proposed in the literature as “**Turbo Code**” [1]. Turbo Codes are formed by two constituent convolutional encoders separated by an interleaver. The decoder consists of two Soft-Input Soft-Output (SISO) modules connected by an interleaver-deinterleaver, and an iterative algorithm that swaps from one SISO module to the other.

Performance near the ultimate Shannon limits (up to 0.5 dB) can be obtained with large interleavers, and uniformly good performance is possible for any interleaver size [2]. Applications areas of Turbo Codes include space communications (the proposal of a new standard for the telemetry link coding is in a very advanced stage by CCSDS), coding for ATM and wireless applications, fading channels, digital direct broadcast satellite services, CDMA, channel equalization, and combined carrier estimation and decoding [3-5]. Various studies on performance evaluation of Turbo Codes have been carried out in the literature. S. Benedetto and G. Montorsi derived an analytical upper bound to the average performance of such coding schemes, which is independent of the interleavers used, and shows the influence of the interleaver length on the code performance [6]. Oscar Y. Takeshita et al have investigated the performance of a class of asymmetric Turbo Codes [7]. He considered the BER performance characteristics of asymmetric Turbo Codes in both the “waterfall” (i.e. at smaller SNR’s) and the “error-floor” (i.e. at higher SNR’s) region. Cristian Schlegel looked at the performance of Turbo Codes with respect to various information theoretic error bounds [8]. Volker Kuhn worked on the comparison of turbo codes employing short interleavers with convolutional codes in terms of bit error rate performance and computational effort [9]. Eric K. Hall investigated the Design and Analysis of Turbo Codes on Rayleigh Fading Channels. Furthermore, the design issues relevant to Turbo Codes are examined for the correlated fading channel [10]. He examined stream-oriented Parallel Concatenated (Turbo) Codes. Anderson has worked on Turbo Coding for deep space applications [11]. Claude Berrou studied the encoding and decoding processes of Turbo Codes [12]. None of the studies deal with the dependence of performance of turbo

Codes on the communication medium. In the present paper the results obtained from a study on performance evaluation of medium dependence of Turbo Codes are presented. Communication through Twisted pair, Coaxial cable, Microwaves & Optical fibers have been simulated by modeling the channel and the performance of Turbo Codes have been studied with varying S/N ratio as applicable in practical situations.

2.1 Methodology

Each channel has been modeled by using mathematical modeling techniques as per the channel characteristics and through simulation, the performance of Turbo Code with respect to variation in frame size, signal to noise (SNR) ratio and code rate has been studied. Out of the various models proposed in the literature proposed in literature, the following have been chosen for this work:

- Additive White Gaussian Noise Channel Model for Twisted Pair & Coaxial Cable,
 - Rayleigh Slow fading Channel Model for Microwave media, and
 - Chi-square Channel Model for Optical Fiber
- Each of these is described below in brief.

A. AWGN Model: Twisted Pair & Coaxial Cable

Terrestrial wired channels are AWGN channels [13]. The white gaussian noise is a model for the thermal noise generated by random movement in the receiver. The voltage distribution of this noise follows a normal or a gaussian distribution i.e. a bell shaped curve. The mean of this noise distribution is Zero while it's variance is a function of the noise spectral density. Assuming binary shift keying (BPSK) modulation, the amplitude of the transmitted signal is $\pm\sqrt{E_s}$, depending on whether the bit is a 1 or a 0, where E_s is the energy per symbol that is equal to the energy per bit since BPSK modulation is used without coding. The AWGN channel is described by

$$Z_i = A_1 W_i + N_i, \quad i = 1, 2, 3, \dots \quad \text{---(1)}$$

where $W_i \in \{-1, +1\}$ is the BPSK signal of unit energy and N_i is a gaussian noise sequence with zero mean and variance $N_0/2$. For AWGN channel, $A_1 = 1$ [14]. The noise variance is given by

$$\sigma^2 = (N_0 f_s)/2 \quad \text{---(2)}$$

Therefore, the normalized signal to noise (SNR) ratio is

$$E_b/N_0 = [(\pm)2T_b f_s]/2\sigma^2 \quad \text{---(3)} \quad \text{where } T_b \text{ is the time period of a bit } T_b f_s \text{ is the number of samples per}$$

bit and Energy per bit is normalized to unity. If 1 sample per bit is used, the variance σ^2 can be represented as:

$$\sigma^2 = 1/2(E_b/N_0) \quad \text{---(4)}$$

This formula is used to compute the AWGN noise variance. The additive white Gaussian noise can be found by taking the product of the standard deviation of the noise and a normally distributed random number with unit variance.

B. Rayleigh Model: Microwave

The Rayleigh model arises from the combination at the receiver of many point scatterer contributions, each having only a small fraction of the received energy. The Rayleigh fading model can be justified from a Central Limit Theorem Thesis. Under the above mentioned assumptions, appropriate demodulation and sampling give our system model the discrete representation,

$$y_k = a_k \cdot x_k + n_k \quad \text{---(5)}$$

In this representation, x_k is a BPSK symbol amplitude having values of $\pm\sqrt{E_s}$, where E_s is the received energy per symbol, and n_k is an AWGN component with zero mean and variance $N_0/2$. The channel gain 'a' is described by a probability density function and an autocorrelation function. The channel gain is Rayleigh distributed with probability density function,

$$p_a(a) = 2ae^{-a^2}, \quad a \geq 0 \quad \text{---(6)}$$

The channel gain is scaled to have a mean-square value of $E_a[a^2] = 1$ indicating that the expected received power will be E_s .

C. Chi-Square Channel Model: Optical Fiber

We consider amplitude shift keying (ASK) where the signal is modulated to be either 0 or an optical pulse of duration T_s [15]. A practical receiver consists of an optical band pass filter of bandwidth B_0 , a photo detector and an electrical filter of bandwidth B_0 that integrates over the bit period T_s . For systems employing optical amplifiers, under low-power operation of the fiber channels (i.e. quasi-linear regime), amplified spontaneous emission noise (ASE) dominates all other sources of noise. $M = B_0/B_e > 1$ denotes the number of modes per

polarization state in the received optical spectrum, where B_0 and B_e are the optical bandwidth and the electrical bandwidth of the system at the detector, respectively. Noise n_i can be mathematically represented as a Fourier series expansion with Fourier coefficients that are assumed to be independent Gaussian random variables with zero mean and variance $N_0/2$. After passing through optical amplifier, the received signal (the integral of the output of the photo detector) is given by

$$x = \sum_{i=1}^{2M} (s_i + n_i) \mathbf{a}_i, \quad \text{---(7)}$$

where s_i and n_i denote the signal and the ASE noise projected to $2M$ orthonormal basis. Signal energy is

$$\sum_{i=1}^{2M} s_i^2 = 2E_s \quad \text{for transmitting 1, ---(8a)}$$

$$\sum_{i=1}^{2M} s_i^2 = 0 \quad \text{for transmitting 0 ---(8b)}$$

where E_s is the average energy of the transmitted signals (assuming equal probability).

2.2. Results

The study of performance of Turbo Codes in various media has been done by varying SNR, frame sizes and code rates through simulating the transmission in various media by suitable mathematical models for each. The results of the simulation studies are shown in figure -1, which gives a comprehensive picture in terms of Bit Error Rate (BER) versus Signal to Noise Ratio (E_b/N_0). The plot reveals interesting output as given below:

- (i) In Twisted Pair, the BER does not show significant variation over the entire range of SNR up to 5dB. Also the BER does not vary much with respect to variation in code rate and frame size and its value is around 0.2
- (ii) The results for transmission through Coaxial Cable are also same except that BER is an order of magnitude lower

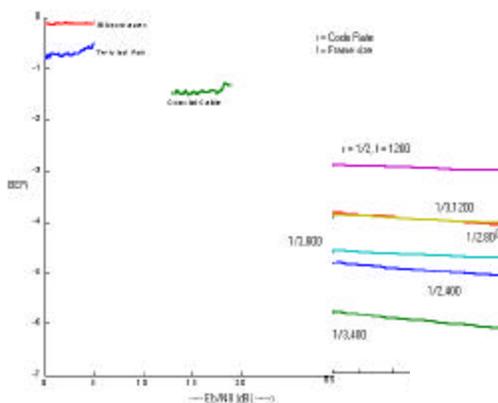


Figure 5.1 BER vs. E_b/N_0 (dB) for Various Transmission Media (Note: BER is on log)

with a value around 0.04. This is an expected result as the SNR is much higher in Coaxial cable as compared to Twisted Pair.

(iii) In Microwave transmission also, the results tell that BER does not vary with variation in SNR, code rate or frame size but the value of BER is greater than 0.7 which is very high. This makes Turbo Codes unsuitable for use in Microwave transmission. This result confirms earlier studies.

(iv) The performance evaluation results in Optical Fiber transmission are very encouraging. The results not only show a very low BER (in the range of $1E-05$ to $1E-03$) but also show variation in BER with respect to code rate and frame size. The studies bring out the obvious facts i.e. - the BER varies directly with the code rate, - the BER varies directly with the frame size. All these results are summarized in above figure which shows BER vs. SNR for all the transmission media studied in this work.

3. Conclusions

In the present paper we have presented simulation results on the studies of medium dependence on the performance of turbo Codes which are recently developed error correcting codes. The studies are made by appropriately modeling different communication channels such as Twisted pair, Coaxial cable, Microwaves and optical fiber and studying the performance of Turbo Codes with respect to Signal to Noise Ratio with varying code rate and frame size. The results show that Turbo Codes can be successfully employed with Coaxial Cables and optical fibers but not with Microwaves. In Optical fibres, the performance does not degrade even with larger frame sizes, a result very significant for practical applications.

4. Acknowledgement

We are thankful to AICTE for providing the financial support under the R&D/TAPTEC Scheme for the purchase of Matlab(version 6.1).

References

- [1] C. Berrou, A. Glavieux, P. Thitimajshima, "Near Shannon Limit error Correcting Coding and Decoding : Turbo Codes, " *Proc. 1993*

- IEEE Global Communication Conference*, pp. 1064-1070.
- [2] www.ee.virginia.edu/research/CCSP/turbo-codes
- [3] www.engelschall.com
- [4] www.ksu.edu
- [5] www.ece.arizona.edu/~ryan
- [6] S. Benedetto, G. Montorsi, "Performance Of Turbo Codes," *Electronic Letters*, pp. 163-165, 2nd Feb. 1995
- [7] Oscar Y. Takeshita, Oliver M. Collins, Peter C. Massey, Daniel J. Costello, "A Note On Asymmetric Turbo-Codes," *IEEE Comm. Letters*, Vol. 3, pp. 69-71, March 1999.
- [8] Christian Schlegel, Lance Perez, "On Error Bounds And Turbo Codes," *IEEE Communications Letters*, Vol. 3, pp. 205-207, July 1999.
- [9] Volker Kuhn, "Evaluating the Performance of Turbo Codes And Turbo-coded Modulation in a DS-CDMA Environment," *IEEE Journal On Selected Areas In Communications*, Vol. 17, pp. 2138-2147, December 1999.
- [10] Eric K. Hall, Stephen G. Wilson, "Design And Analysis Of Turbo Codes On Rayleigh Fading Channel," from www.ee.virginia.edu/research/CCSP/turbo-codes/
- [11] Stephen A. Fechtel, "A Novel Approach to Modelling and Efficient Simulation of Frequency-Selective Fading Radio Channels," *IEEE Journal On Selected Areas In Communications*, pp. 422-431, April 1993
- [12] Claude Berrou, Alain Glavieux, "Near Optimum Error Correcting Coding And Decoding : Turbo-Codes," *Transactions On Communications*, pp. 1261-1271, October 1996
- [13] Chandran Naveen, "Hybrid ARQ with parallel & serial concatenated Convolutional codes for next generation Wireless Communication", Thesis submitted to College of Engineering at West Virginia University, 2001.
- [14] Zhu GC & Fady I. Alazazi, "Soft decision COVQ for TURBO CODED AWGN & Rayleigh Fading channels", *IEEE Communication Letters*, vol-5, no. 6, pp. 257-259, June 2001
- [15] Li J., Cai Y., & Narayanan KR, "On the Bit error rate of product Accumulate Codes in Optical fiber Communication", Electrical Engineering Deptt. Texas A&M University, College station, 1999.

Performance Evaluation of MRC for MC-CDMA Communication Systems over Nakagami-m Channels

Joy Iong-Zong Chen and Roger K. S. Miao

Department of Communication Engineering Da Yeh University
 No. 112, Shan-Jiau Rd. Da-Tsuen, Chang-Hwa 51505 Taiwan R. O. C.
 Tel: 01-886-48-511888 ex: 2523
 Fax: 01-886-48-511245
 E-mail: jchen@mail.dyu.edu.tw

Abstract – The performance for a multi-carrier code-division multiple-access (MC-CDMA) system working in uncorrelated and correlated fading channels with Nakagami distribution is investigated in this paper. We adopt the BPSK modulation scheme with an alternative expression for Q-function to evaluate the BER (bit error rate) performance of the MC-CDMA system. The sum of Gamma variates is adopted to derive a closed-form solution for arbitrarily correlated channel in order to avoid the difficulty of explicitly obtaining the pdf (probability density function) for the SNR (signal-to-noise ratio) at the MRC (maximum ratio combining) output. The results with simple expression obviously show that the performance degradation of the MC-CDMA system is sensitive to the correlation of fading channels.

Keywords: MC-CDMA, MRC, uncorrelated channel, correlated channel.

I. Introduction

Due to the advantages of spectrum efficient, interference immune, high data rate, and insensitivity to frequency selective channel, etc. Such that multiple access system bases on direct sequence CDMA (code-division multiple-access) have drawn recent interest in the application of wireless radio systems [1]. Especially, multi-carrier CDMA (MC-CDMA) appears to be a considerable candidate for future mobile radio communication systems.

The BER (bit error rate) analysis of MC-CDMA based on considering different kinds of assumptions, so far, have been dedicated in numerous previously researches [1, 2, 3]. The performance evaluation of MC-CDMA over multipath fading channels was studied in [2]. The results presented in [3] are for uplink channel using MRC (maximal ratio combining) with the assumption frequency offsets condition in correlated fading. The performance of MC-CDMA in non-independent Rayleigh fading was studied in [4]. In [5], which is by use of the method of CF (characteristic function) and residue theorem to calculate the performance for downlink MC-CDMA systems. Both of the envelopes and phases correlation are considered in [6], evaluated the performance of a MC-CDMA system operates in Rayleigh fading channel. The literature in [7] illustrated the error probability for MC-CDMA systems assumed that the transmission channel is in Nakagami-m fading, and the postdetection of EGC (equal gain combining) is considered.

In this paper, the generic expression of BER performance for MC-CDMA system working in uncorrelated and correlated fading channels is evaluated. The general correlation of channels with Nakagami fading distribution is assumed. There is a closed-form obtained via the sum of Gamma variates to avoid the difficulty of explicitly obtaining the pdf for the SNR (signal-to-noise ratio) at the MRC output. The results analyzed how deep does the channel correlation affect the performance of MC-CDMA systems.

The rest of this paper is organized as follows: section

II gives a description of the MC-CDMA system model. The uncorrelated and correlated-Nakagami fading channel model are given in section III. Section IV describes the receiver model of MC-CDMA system. The performance of MC-CDMA operating in uncorrelated and correlated fading channel is carried out in section V. There are numerically results shown in section VI. Finally, section VII draws briefly conclusions.

II. System Model

The MC-CDMA system model is described in this section. It is assumed that there exist K simultaneously users with N subcarriers within a signal cell. Any effect of correlation among users is going to be ignored by assuming the number of users is uniformed of distribution. As shown in Fig.1, a signal data symbol is replicated into N parallel copies. The signature sequence chip with a spreading code of length L is used to BPSK (binary phase shift keying) modulated each of the N subscribers of the k -th user. Where the subcarrier has frequency F/T_b Hz, and where F is an integer number. [1][2]. The technical described above is same as to the performance of OFDM (Orthogonal Frequency Division Multiplexing) on a direct sequence spread-spectrum signal when set $F=1$. The larger values of F , the more transmit bandwidth increase. The transmitted signal the resulting transmitted baseband signal $S_k(t)$ corresponding to the M data bit size can be expressed as

$$S_k(t) = \sqrt{\frac{2P}{N}} \sum_{n=0}^{M-1} \sum_{m=0}^{N-1} a_k[n] b_k[m] P_{P_b}(t) \text{Re}\{e^{jn\omega_n t}\}, \quad \begin{matrix} a_k[n] \in \{-1,1\} \\ b_k[m] \in \{-1,1\} \end{matrix} \quad (1)$$

where P is the power of data bit, M denotes the number of data bit, N denotes the number of subcarriers, the sequencer $a_k[0], \dots, a_k[N-1]$ and $b_k[0], \dots, b_k[M-1]$ represent the signature sequence and the data bit of the k -th user, respectively. The $P_{P_b}(t)$ is defined as an unit amplitude pulse that is non-zero in the interval of $[0, T_b]$, and $\text{Re}\{\cdot\}$ denotes the real part of a complex number, $\omega_n = 2\pi(f_c + nF/T_b)$ is the angular frequency of the n -th subcarrier.

A frequency-selective channel with $1/T_b \ll BW_c \ll F/T_b$ is addressed in this paper, where BW_c is the coherence bandwidth. This channel model means that each modulated subcarrier does not experience significant dispersion and with transmission bandwidth of $1/T_b$, i.e. $T_b \gg T_d$, where $1/T_d$ is the Doppler shift typically in the range of 0.3~6.1 Hz [1] in the indoor environment, and the amplitude and phase remain constant even the symbol duration T_b . In addition to, the channel of interest has the transfer function of the continuous-time fading channel assumed for the k -th user can be represented as

$$H_k[f_c + i\frac{F}{T_b}] = \mathbf{b}_{k,i} e^{j\theta_{k,i}} \quad (2)$$

where $\mathbf{b}_{k,i}$ and $\mathbf{q}_{k,i}$ are the random amplitude and phase of the channel of the k -th user at frequency $f_c + i(F/T_b)$. In order to follow the real world case, the

random amplitude, $\mathbf{b}_{k,i}$ are assumed to be a set of N correlated not necessarily identically distributed in one of our scenarios.

III. Channel Model

In this section, there are two cases of propagating channels considered, that is, uncorrelated and correlated Nakagami channels. The equal fading severities are considered for all of the channels, namely $m_i = m, i=1, \dots, N$. The pdf of the fading amplitude for the k -th user with i -th channel, $\mathbf{b}_{k,i}$, are assumed as r.v. (random variable) with the Nakagami distribution, and given as [9]

$$P(\mathbf{b}) = \frac{2\mathbf{b}^{2m-1}}{\Gamma(m)} \cdot \left(\frac{m}{\Omega}\right)^m \cdot \exp\left(-\frac{m\mathbf{b}^2}{\Omega}\right), \quad \mathbf{b} \geq 0 \quad (3)$$

where $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad \Omega = E[\mathbf{b}^2] \text{ denoting expectation, the parameter } m \text{ of the amplitude distribution characterizes the severity of the fading, and it is defined as}$$

$$m = \frac{\Omega^2}{E[(\mathbf{b}^2 - \Omega)^2]} \geq 0.5 \quad (4)$$

It is well known that $m=0.5$ (one-sided Gaussian fading) corresponds to worst case fading condition, $m=1$ and $m = \infty$ correspond to Rayleigh fading (purely diffusive scattering) and the non-fading condition, respectively. As what follows, we consider these two cases.

A. Uncorrelated Channel

Firstly, if the propagation channels are assumed as i.i.d (identically independent distributed), then by use of the variable changing, the variable \mathbf{g} is assigned as the fading power of the channel, and let $\mathbf{g} = \mathbf{b}^2$, then the pdf of \mathbf{g} is given follows as a gamma distribution, can be obtained by the processing of random stochastic as

$$P_r(r) = \frac{r^{m-1} e^{-r/\Omega}}{\Gamma(m)\Omega^m} \quad (5)$$

B. Correlated Channel

Let $[\mathbf{g}_i], i=1, \dots, N$ be a set of N correlated identically distributed, and all the figure parameters and the average power are assumed equivalent, that is, $m_i = m_j = m$, and $\Omega_i = \Omega_j = \Omega$, where $i \neq j$, for $i, j=1, \dots, N$. The power at the output of the MRC is a function of the sum of the squares of signal strengths, and is given as $R = \sum_{i=1}^N \mathbf{g}_i$. Hence following the results extended from the [8] by M. S. Alouini, A. Abdi, and M. Kavehthe [7]. The pdf of

$R = \sum_{i=1}^N \mathbf{g}_i$, can be expressed as

$$P_R(r) = \prod_{i=1}^N \left(\frac{1}{I_i}\right) \sum_{q=0}^{\infty} \frac{\mathbf{m}_q \mathbf{g}^{mN+q-1} e^{-\mathbf{g}/I_1}}{I_1^{mN+q} \Gamma(mN+q)} \quad (6)$$

where the coefficients \mathbf{m}_q can be obtained recursively by the following formula

$$\begin{cases} \mathbf{m}_0 = 1 \\ \mathbf{m}_{q+1} = \frac{m}{q+1} \sum_{i=1}^{q+1} \left[\sum_{j=1}^N \left(1 - \frac{I_1}{I_j}\right)^i \right] \mathbf{m}_{q+1-i} \end{cases}, \quad q=0, 1, 2, \dots \quad (7)$$

where $I_1 = \min\{I_i\}$, and $I_i, i=1, \dots, N$ are the eigenvalues of the matrix $\mathbf{Z} = \mathbf{X}\mathbf{Y}$, where \mathbf{X} is the $N \times N$ diagonal matrix with the entries of average power $\Omega_i, i=1, \dots, N$,

when the subcarrier paths are correlated, the entries of Ω_i can be obtained by taking the minimum value of $\Omega_i = \mathbf{g}_i/m_i$. The matrix \mathbf{Y} is the $N \times N$ positive definite matrix defined by

$$\mathbf{Y} = \begin{bmatrix} 1 & \mathbf{r}_{12}^{1/2} & \dots & \mathbf{r}_{1N}^{1/2} \\ \mathbf{r}_{21}^{1/2} & 1 & \dots & \mathbf{r}_{2N}^{1/2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_{N1}^{1/2} & \dots & \dots & 1 \end{bmatrix}_{N \times N} \quad (8)$$

where \mathbf{r}_{lm} denotes the correlation coefficient between \mathbf{g}_l and $\mathbf{g}_m, l \neq m, l, m=1, 2, \dots, N$, and \mathbf{r}_{lm} can be expressed as

$$\mathbf{r}_{lm} = \frac{\text{Cov}(\mathbf{g}_l, \mathbf{g}_m)}{[\text{Var}(\mathbf{g}_l) \cdot \text{Var}(\mathbf{g}_m)]^{1/2}} = \mathbf{r}_{ml}, \quad 0 \leq \mathbf{r}_{ml} \leq 1 \quad (9)$$

where $\text{Var}(\cdot)$ and $\text{Cov}(\cdot)$ are the variance and the covariance operators, respectively.

IV. MC-CDMA Receiver Model

A slowly varying fading channel is considered in this paper, that is, the channel parameters are unchanged over one bit duration T_b . For K active transmitters, the received signal $r(t)$ can be written as

$$r(t) = \sqrt{\frac{2P}{N}} \sum_{k=1}^{K-1} \sum_{n=0}^{N-1} \mathbf{b}_{k,n} a_k[n] b_k[m] P_{t_k}(t - mT_b - t_k) \cos(\mathbf{w}_n t + \mathbf{q}_{m,n}) + n(t) \quad (10)$$

where $n(t)$ is the AWGN (additive white Gaussian noise) with a double-sided power spectral density of $N_0/2$. We can evaluate, the local-mean power, $P_{k,n}$, which is given as

$$P_{k,n} = E[\mathbf{b}_{k,n}^2] \frac{P}{N} \quad (11)$$

The total-mean power of the k -th user is defined to be $P_k = N \cdot P_{k,n}$, if the local-mean power of the subcarriers is assumed equal. Assuming that acquisition has been accomplished for the user of interesting ($k=0$). In addition, the system operates synchronously with each user having the same clock is assumed, and the MRC diversity reception technique is considered in this paper. For the reason of using MRC, it is assumed that perfect phase correction can be obtained, i.e., $\hat{\mathbf{q}}_{0,i} = \mathbf{q}_{0,i}$. Demodulating each subcarrier includes applying a phase correction, $\hat{\mathbf{q}}_{0,i}$, and a gain correction factor $d_{0,n} = \mathbf{b}_{0,n} \cdot a_0[n]$ is multiplied by the n -th subcarrier signal as shown in Fig. 2.

With all the assumptions for MRC combining, the decision variable D_0 of the m -th data bit reference user, and given by

$$D_0 = \frac{1}{T_b} \int_{mT_b}^{(m+1)T_b} r(t) \cdot \sum_{n=0}^{N-1} a_0[n] \cdot d_{0,i} \cdot \text{Re}\{e^{j(\mathbf{w}_n t + \mathbf{q}_{0,0})}\} dt = U_s + I_{MAI} + \mathbf{h}_0 \quad (12)$$

where $r(t)$ is the received signal shown in (10), $d_{0,i}$ is the gain factor for MRC diversity. The first term in second equivalent last equation represents the desired signal, can be expressed as

$$U_s = \sqrt{\frac{P}{2N}} \sum_{n=0}^{N-1} \mathbf{b}_{0,i}^2 a_0[m] \quad (13)$$

, and the second term, I_{MAI} , is the MAI (multiple access interference) contributed from all other users which can be written as

$$I_{MAI} = \sqrt{\frac{P}{2N}} \sum_{k=1}^{K-1} \sum_{n=0}^{N-1} a_k[m] \cdot b_k[m] \cdot a_0[m] \cdot \mathbf{b}_{k,m} \cdot \mathbf{b}_{0,n} \cdot \cos(\mathbf{q}_{k,n}^i) \quad (14)$$

where $\mathbf{q}_{k,n}^i = \mathbf{q}_{0,n} - \mathbf{q}_{k,n}$ and $\mathbf{q}_{k,n}$ are i.i.d uniformly distributed over $[0, 2\pi)$, \mathbf{h}_0 is the AWGN term.

V. Performance Analysis

A generalized average BER for the k -th user using coherent BPSK (binary phase shift keying) modulation scheme is derived in this section. For coherent demodulation in the presence of AWGN, the probability of error conditioned on the instantaneously SNR can be expressed as [10]

$$P_e(s) = \frac{\Gamma\left(\frac{1}{2}, \mathbf{s}_0 s\right)}{2\sqrt{p}} = 0.5Q\left(\sqrt{2\mathbf{s}_0 s}\right) \quad (15)$$

where the incomplete gamma function is defined as $\Gamma(z, t) = \int_t^\infty x^{z-1} e^{-x} dx$, and the received instantaneously SNR, which conditioned on $\mathbf{g}_{0,n} = \mathbf{b}_{0,n}^2$, at output of the receiver is calculated as

$$\frac{U_s^2}{\mathbf{s}_T^2} = \frac{P \sum_{n=0}^{N-1} \mathbf{b}_{0,n}^2}{\mathbf{s}_{MAI}^2 + \mathbf{s}_h^2} \quad (16)$$

where \mathbf{s}_{MAI}^2 is the variance of I_{MAI} , which is shown in (14). In the limiting case of large N and by the methods of central limit theory (CLT), the MAI can be approximated by a Gaussian r.v. with zero mean and the variance, \mathbf{s}_{MAI}^2 , can be determined as

$$\mathbf{s}_{MAI}^2 = E[I_{MAI}^2] = \frac{P}{2}(k-1) \cdot E[\mathbf{b}_{k,n}^2] \cdot E[\cos^2 \bar{\mathbf{q}}_{k,n}] = \frac{P}{4}(k-1)\Omega_{k,n} \quad (17)$$

where $\Omega_{k,n} = E[\mathbf{b}_{k,n}^2]$, $E[\cos^2 \bar{\mathbf{q}}_{k,n}] = 1/2$. On the other hand, the background noise term \mathbf{h}_0 is a random variable with zero mean and the variance can be calculated as

$$\mathbf{s}_{h_0}^2 = E[\mathbf{h}_0^2] = \frac{NN_0}{4T_b} \quad (18)$$

By substituting (17) and (18) into (16), which can be obtained as

$$\frac{U_s^2}{\mathbf{s}_T^2} = \frac{1}{2N} \cdot S \quad (19)$$

where $S = \sum_{n=0}^{N-1} \mathbf{b}_{0,n}^2 / \Omega_{k,n}$ (20)

, and $\mathbf{s}_0 = \frac{NN_0}{4PT_b\Omega_{k,n}} + \frac{k-1}{4} = \frac{N}{4g_0} + \frac{k-1}{4}$ (21)

where $\mathbf{g}_{k,n} = \frac{PT_b\Omega_{k,n}}{N_0} = \frac{E_b\Omega_{k,n}}{N_0}$ is the SNR of each

bit, and $E_b = PT_b$ denotes the bit energy.

It is known that the decision variable in (12) has a Gaussian distribution conditioned on the uncorrelated and correlated channel power $\mathbf{b}_{0,n}^2$, respectively, and the AWGN, \mathbf{h}_0 , and the MAI, \mathbf{h}_{MAI} are mutually independent. Therefore, the probability of error by means of BPSK modulation conditioned on the instantaneously SNR has been given in (15) can be evaluated as follows.

A. Uncorrelated Channel

We adopt the alternative expression for the Q-function, which is expressed as [11]

$$Q(t) = \frac{1}{p} \int_0^{\frac{t}{2}} e^{-\frac{t^2}{2\sin^2 j}} dj, \quad t \geq 0 \quad (22)$$

and the random variables $\{\mathbf{g}_i, i=1, \dots, N\}$ are assumed to be independent. Such that the equation (15) can be expressed as

$$P_e(s) = \frac{1}{p} \int_0^{\frac{t}{2}} \prod_{n=1}^N J_{0,n}(\Omega_{0,n}, j) dj \quad (23)$$

where [12]

$$J_{0,n}(\Omega_{0,n}, j) = \int_0^\infty e^{-\frac{r_0}{\sin^2 j}} P_r(r_{0,n}) dr_{0,n} \quad (24)$$

$$= \left(\frac{m}{\Omega_{0,n}}\right)^m \left(\frac{1}{\sin^2 j} + \frac{m}{\Omega_{0,n}}\right)^{-m}$$

where $\Omega_{0,n}$ denotes the average value of $\mathbf{g}_{0,n}$. In Nakagami fading channels, the $r_{0,n}$ follows the gamma distribution expressed as in (5). In this case, all N subcarriers are assumed i.i.d, then the average bit error probability can be calculated by using the simple form of a single integral with finite limits and obtained as

$$P_e = \frac{1}{p} \int_0^{\frac{t}{2}} [J_{0,n}(\Omega_{0,n}, j)]^N dj \quad (25)$$

B. Correlated Channel

If the condition of correlated channels are considered, then the average bit error probability for the case can be calculated by averaging (6) and (15), and yield as

$$P_e = \int_0^\infty P_{cs} \cdot P_e(S) dS$$

$$= \frac{1}{2\sqrt{p}} \prod_{n=1}^L \left(\frac{I_n}{I_n}\right)^m \sum_{q=0}^\infty I_1^{mN+q} \Gamma(mN+q) \cdot \frac{\mathbf{s}_0^{q/2} \cdot \Gamma(mN+q+\frac{1}{2})}{(mN+q)(\mathbf{s}_0+\frac{1}{I_1})^{mN+q+\frac{1}{2}}}$$

$$\times {}_2F_1(1, mN+q+\frac{1}{2}; mN+q+1; -\frac{1/I_1}{\mathbf{s}_0+1/I_1})$$

where \mathbf{m}_i is shown in (7), I_1 and \mathbf{s}_0 are shown in (6) and (21), respectively, and the symbol ${}_2F_1(\cdot, \cdot; \cdot)$ denotes the confluent hyper geometric function [12].

VI. Numerical Results

In this section we show some of the numerical results. Fig. 3 and Fig. 4 illustrate the results of SNR (dB) and user number versus BER for MC-CDMA system works over uncorrelated fading channels. On the other hands, Fig. 5 and Fig. 6 show the performance results for MC-CDMA system works over correlated channels. In Fig. 3 where is not only the different fading parameters, $m=2$ and 4, figure out the different BER performances curves but the different subcarrier numbers, $N=8, 16, 32$, present that the much more subcarrier the better BER performance. The user number, K , versus BER for MC-CDMA system operate in uncorrelated fading channels is illustrated in Fig. 4, in which the SNR=5dB is assumed. The user number is limited both of the SNR and the subcarrier number.

Next, the correlation coefficients between the channels are assumed to be $r=0.1, 0.4, 0.8$ in Fig. 5. The subcarriers number and the user number are assumed 16 and 8, respectively. To calculate the BER, the mean power of the desired signal is assumed equal to the mean power of each interfering mobile units. Based on Fig. 5, we should know that the performance of BER will be promoted gradually by the increasing of the correlation coefficient, r , between the correlated channels. It is reasonable that the system performance BER becomes much better when the fading parameter is increasing. The performance of BER versus user number is shown in Fig. 6, in which the correlation coefficients are assumed $r=0.4, 0.6$, and 0.8 , the average bit SNR is set as 5dB.

VII. Conclusion

The SNR and the user number versus probability error rate for MC-CDMA works in uncorrelated and correlated Nakagami fading channels has been evaluated in this paper. The results explicitly shown that the phenomena of channel correlation do degrade the performance of MC-CDMA communication systems. Therefore the consideration of correlation coefficient for channel fading should pay much attention while designing the MC-CDMA systems.

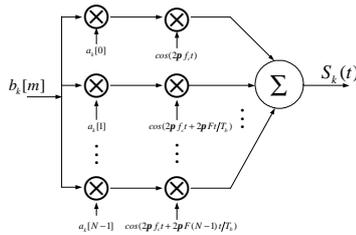


Fig. 1 The transmitter model of the MC-CDMA system

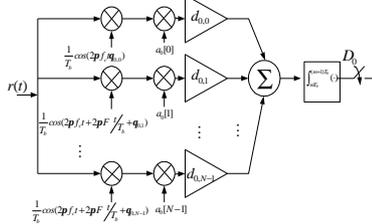


Fig. 2 The receiver model of the MC-CDMA system

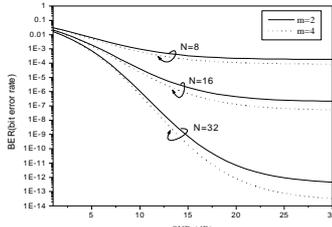


Fig. 3 The SNR vs BER for MC-CDMA system with uncorrelated channels

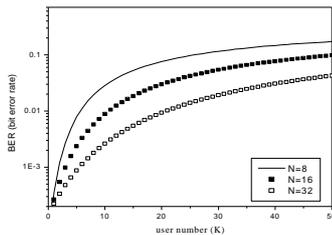


Fig. 4 The user number vs BER for MC-CDMA system with uncorrelated channels

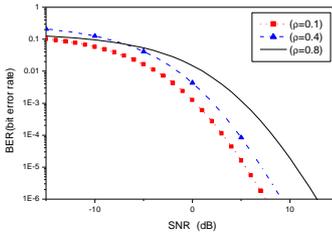


Fig. 5 The SNR vs BER for MC-CDMA system with the correlated fading channels

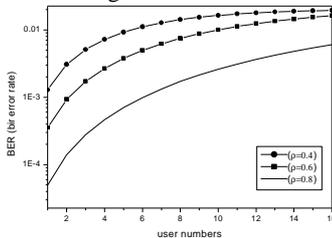


Fig. 6 The user number vs BER for MC-CDMA system with correlated fading channels

References

[1] N. Yee, J. -P. Linnartz, and G. Fettweis, "Multi-carrier CDMA in Indoor Wireless Radio Networks", IEICE trans. on Commun., Vol. E77-B, No.7, pp. 900-904, July 1994.

[2] E. A. Sourour and M. Nakagami, "Performance of Orthogonal Multicarrier CDMA in a Multipath Fading Channel," IEEE trans. on Commun., Vol. 44, pp. 356-367, Mar. 1996.

[3] T. Kim, Y. Kim, J. Park, K. Ko, S. Choi, C. Kang, and D. Hong, "Performance of an MC-CDMA System with Frequency Offsets in Correlated Fading", IEEE International Conference on ICC 2000, Vol. 2, pp. 1095 –1099, June 2000.

[4] J. Park, J. Kim, S. Choi, N. Cho, and D. Hong, "Performance of MC-CDMA Systems in Non-independent Rayleigh Fading", IEEE on ICC' 99, Vol. 1, pp. 506-510, 6-10 June 1999.

[5] Q. (Rock) Shi and M. Latva-aho, "Exact Error Floor for Downlink MC-CDMA with Maximal Ratio Combining in Correlated Nakagami Fading Channels", 2002 International Zurich Seminar on Broadband Communications, pp. 37-1-37-5, 20 Feb. 2002.

[6] Q. Shi and M. Latva-aho, "Performance Analysis of MC-CDMA in Rayleigh Fading Channels with Correlated Envelopes and Phases", IEE proc. Commun. Vol. 150, No. 3, pp.214-220, Jun. 2003.

[7] Z. Li, and M. Latva-aho, "Error Probability for MC-CDMA in Nakagami-m Fading Channels Using Equal Gain Combining", IEEE International Conference on ICC 2002, Vol. 1, pp. 227 –231, 28 April- 2 May 2002.

[8] M. -S. Alouini, A. Abdi, and M. Kaveth, "Sum of Gamma Variates and Performance of Wireless Communication Systems over Nakagami-Fading Channels", IEEE trans. on V. T., Vol. 50, No. 6, pp. 1471-1480, Nov. 2001.

[9] Nakagami, M., "The m-Distribution-A General Formula of Intensity Distribution of Rapid Fading", Statistic Methods of in Radio Wave Propagation, pp. 3-36, Pergamon Press, New York, 1960.

[10] M. Schwartz, W. R. Bennett, and S. Stein, "Communication Systems and Techniques", McGraw-Hill: New York, 1966.

[11] M. K. Simon and M. -S. Alouini, "A Unified Approach to the Performance Analysis of Digital Communication over Generalized Fading Channel," Proc. Of the IEEE, Vol. 86, pp. 1860-1877, Sept. 1998.

[12] I. S. Gradshteyn, and I. M. Ryzhik. Table of Integrals, series, and products, San Diego, CA: Academic Press, 5th Ed. 1994.

Performance of Adaptive Fuzzy-Based Partial Parallel Interference Canceller for Multi-Carrier CDMA Systems

Yung-Fa Huang,

Department of Computer Science and Information Engineering
 Chung Chou Institute of Technology,
 6, Lane 2, Sec. 3, Shan-Chou Rd., Yuanlin, Changhwa, Taiwan, R.O.C.
 Tel: +886-4-8311498 Ext. 4103
 E-mail: huangyf@dragon.ccut.edu.tw

Abstract

In this paper, the adaptive cancellation weights are inferred from a proposed multistage fuzzy inference system (FIS) to perform effective partial parallel interference cancellation (PPIC) multiuser detection under time-varying frequency selective fading channels for multi-carrier direct-sequence code-division multiple-access (MC-CDMA) communication systems. Simulation results show that the four-stage adaptive fuzzy-based PPIC (FB-PPIC) scheme outperforms the conventional PIC (CPIC) and constant weight PPIC (CW-PPIC), especially in near-far environments.

I. Introduction

The combined technique of OFDM and CDMA, multi-carrier code-division multiple-access (MC-CDMA) communication technique, which exploits the spreading feature of CDMA but without the adverse effect of increasing frequency selectivity in the channel, has received much attention among researcher and been adored as a favorite candidate for 4th generation cellular communication systems [1-3]. By using multi-carrier modulation, CDMA signal is spread over several carriers by which frequency-diversity is achieved similar to path diversity in RAKE receivers. The multiple narrowband channels in each subcarrier undergo nearly frequency flat fading. A suitable guard time can be inserted to eliminate the effect of delay spread. Then, with a sufficient cyclic extension, the PN codes spreading in each subcarrier can keep synchronously. Therefore, the MC-CDMA communication systems perform the multiple accesses and overcome the multipath frequency selective fading. However, the code orthogonality among users in MC-CDMA systems is highly distorted by the instantaneous frequency response of the channel. Therefore, the MAI due to the frequency selectivity in fading channels degrades the performance of MC-CDMA communication systems. Then, the performance of an MC-CDMA mobile system is primarily limited by MAIs. Thus multiuser detection (MUD) becomes very important for separating user's signals, even in the absence of the near/far effect [4].

Among the suboptimal multiuser detectors [5-7], the parallel interference cancellation (PIC) scheme simultaneously subtracts the interference from each

user's received signal. However, at heavy system load, the multistage conventional PIC (CPIC) approach suffers performance degradation due to a poor cancellation, which is brought about by the relatively high error rate of bit decisions in the preceding stage [6, 7]. Thus, the partial cancellation contrarily is a better policy than the complete cancellation. However, its performance obviously suffers degradations in the near-far channels. Moreover, the optimal weight of each interferer depends on their reliability, respectively. Consequently, to achieve the optimal performance of the PPIC, the PCWs of the interferers should be adjusted according to the reliability of their estimate statistics [6,7].

The fuzzy logic developed by Zadeh [8] has drawn a great deal of attention because of its universal approximation ability in the nonlinear problem [9]. Therefore, in this paper, we investigate the performance of the proposed FIS and the inference of the adequate PCWs for each interferer in the PPIC scheme over the frequency selective fading channels.

II. System Model

A quasi-synchronous uplink system is assumed for simplicity of the system model. A K -user MC-CDMA system with BPSK modulation is considered. After serial to parallel (S/P) conversion, the k th user transmits P parallel data symbols with the binary data signal $d_k(t) = \sum_{i=-\infty}^{\infty} \sum_{p=0}^{P-1} b_{k,p}(i) p_{T_b}(t - iT_b)$, where $b_{k,p}(i) \in \{-1, +1\}$ is the p th symbol of the k th user during the i th data duration and $p_{T_b}(t)$ is a rectangular pulse with amplitude 1 and duration T_b . On the $(Pn+p)$ th sub-carrier, the k th user's signal is spread by the spreading codes, $a_{k,n} = \{-1, +1\}$, the n th chip of the k th user. Then, the signal $c_k(t)$ transmitted by the k th user is given by

$$c_k(t) = \sum_{i=-\infty}^{\infty} \sum_{p=0}^{P-1} \sqrt{\frac{2P_{k,i}}{N}} b_{k,p}(i) \cdot \sum_{n=1}^N a_{k,n} \cos(\omega_n t + \theta_k) \cdot p_{T_b}(t - iT_b), \quad (1)$$

where $P_{k,i}$ and θ_k are the signal power and the phase of the k th user respectively, and $\omega_n = \omega_c + 2\pi(Pn+p)\Delta f$ is the n th subcarrier frequency of the p th symbol where ω_c is the common carrier frequency and $\Delta f = 1/T_b$ is the minimum carrier separation between subcarriers.

To increase the frequency diversity and to ensure that each sub-carrier undergoes frequency non-

selective fading [3], the number of sub-carriers is chosen to be a multiple, P , so that P symbols will be transmitted during one OFDM symbol of duration, T_b , comprising of $N_c = NP$ subcarriers. The transmission signal vector corresponding to the p th symbol, $p = 0, 1, \dots, P-1$. The N_c components of the k th user are scrambled by applying frequency interleaving in order to eliminate the correlation among fading between adjacent subcarriers. Then, the summation of N_c components is followed by inserting a guard interval, which is greater than the multipath delay spread, T_m of the radio channel.

As a frequency selective Rayleigh fading channel, we assume a wide-sense stationary uncorrelated scattering (WSSUS) channel [11]. Then, the received signal at the base station can be expressed as

$$r(t) = \sum_{i=-\infty}^{\infty} \sum_{k=1}^K \sum_{p=0}^{P-1} \hat{b}_{k,p}(i) \sum_{n=0}^{N-1} G_{k,Pn+p}(i) a_{k,n} \cos[\omega_m(t - \tau_{k,Pn+p}) + \phi_{k,Pn+p}(i)] \times p_{T_b}(t - iT_b) + n(t), \quad (2)$$

where $G_{k,Pn+p}(i)$ and $\phi_{k,Pn+p}(i)$ are the received signal amplitude and phase at the $(Pn+p)$ th subcarrier in the i th data bit for the p th symbol of k th user, $n(t)$ is an AWGN process with zero mean and two-sided power spectral density (PSD) $N_0/2$, $\tau_{k,Pn+p}$ is the propagation delay at the $(Pn+p)$ th subcarrier of the k th user, and $\phi_{k,Pn+p}(i)$ is a random variable with uniform distribution on $(0, 2\pi)$. After the discrete Fourier transform (DFT) operation, the received signal for the $(Pn+p)$ th subcarrier during the i th data interval at the base station can be expressed as

$$r_{Pn+p}(i) = \sum_{k=1}^K G_{k,Pn+p}(i) b_{k,p}(i) a_{k,n} + n(i), \quad (3)$$

where $G_{k,Pn+p}(i)$ is the complex gain of the received signal at the $(Pn+p)$ th subcarrier in the i th data bit for the p th symbol of k th user, and $n(i) = n_I(i) + j n_Q(i)$ is a complex-valued AWGN with independent real and imaginary components and the two-sided PSD equals to $N_0/2$. To acquire the advantage of frequency diversity, the maximal ratio combining (MRC) technique which maximizes the instantaneous signal to interference and noise ratio (SINR) [11] is adopted in this paper. Therefore, the decision statistic for the p th symbol of the i th data duration of the k th user for conventional MRC receivers is thus obtained as

$$Z_{k,p}(i) = \Re \left[\sum_{n=0}^{N-1} r_{Pn+p}(i) a_{k,n} G_{k,Pn+p}^*(i) \right], \quad (4)$$

where $[\]^*$ is the conjugate operation. Then, the p th data symbol can be estimated by hard decision as obtained by

$$\hat{b}_{k,p}(i) = Z_{k,p}(i) / |Z_{k,p}(i)|. \quad (5)$$

III. Multistage Partial Parallel Interference Cancellation

After the DFT operation, the signal obtained from the $(Pn+p)$ th subcarrier, for the p th symbol in the i th

bit at the first stage (i.e., before any interference cancellation) is obtained as (3). The decision statistic for the p th symbol of the k th user at the i th data duration is obtained by despreading and combining as (4). Then, the i th data bit can be estimated by hard decision as obtained by $\hat{b}_{k,p}^{(1)}(i) = Z_{k,p}^{(1)}(i) / |Z_{k,p}^{(1)}(i)|$. Therefore, we can construct an estimated signal of the $(Pn+p)$ th subcarrier in the i th bit of the k th user at stage s as

$$\hat{\mathbf{s}}_{k,Pn+p}^{(s)}(i) = \hat{b}_{k,p}^{(s-1)}(i) a_{k,n} \mathbf{G}_{k,Pn+p}. \quad (6)$$

The signal amplitude can be estimated by transmitting a training sequence in the pilot-tone channel.

The multistage PIC is performed by regenerating and simultaneously subtracting the estimated signals of the interfering users from the received subcarrier's signal $r_{Pn+p}(i)$ to form a new received subcarrier's signal $r_{k,Pn+p}^{(s)}(i)$ in the i th bit interval for the k th user after stage s , given by

$$r_{k,Pn+p}^{(s)}(i) = r_{Pn+p}(i) - \sum_{\kappa=1, \kappa \neq k}^K \hat{\mathbf{s}}_{\kappa,Pn+p}^{(s)}(i). \quad (7)$$

The decision statistic of the i th bit at stage s , $Z_{k,p}^{(s)}(i)$, is obtained by despreading and combining the new received (after interference cancellation) subcarrier's signal $r_{Pn+p}^{(s)}(i)$ with the k th user's signature signal

$$Z_{k,p}^{(s)}(i) = \Re \left\{ \sum_{n=0}^{N-1} r_{Pn+p}^{(s)}(i) \cdot a_{k,n} \mathbf{G}_{k,Pn+p}^* \right\}. \quad (8)$$

Using this procedure, an arbitrary number of stages of the PIC may be performed to obtain the data bits transmitted by each user.

Since the estimates of MAIs may not be completely correct, thus, adding a partial cancellation weight (PCW) on the path of the interference cancellation, which is called PPIC, would improve the performance of interference cancellation. With this modification, the new received signal $r_{k,Pn+p}^{(s)}(i)$ for the k th user after stage s is given by

$$r_{k,Pn+p}^{(s)}(i) = r_{Pn+p}(i) - \sum_{\kappa=1, \kappa \neq k}^K w_{\kappa}^{(s)}(i) \hat{\mathbf{s}}_{\kappa,Pn+p}^{(s)}(i), \quad (9)$$

where $w_{\kappa}^{(s)}(i)$ is the PCWs for the κ th user at stage s and $0 \leq w_{\kappa}^{(s)}(i) \leq 1$. It is easy to know that the optimal weight becomes larger as the number of user K decreases. Therefore, as the interference or noise increases the signal link becomes unreliable so that the PCW should be set to a smaller value. On the other hand, if the interference decreases or the signal power becomes higher, the weight should be set to a larger value to reflect the reliability of the signal link. Here we define the effective number of users, K_{eff} , which measures the amount of MAIs, obtained by

$$K_{eff,k,n}(i) = \frac{\sum_{\kappa=1}^K \sum_{p=0}^{P-1} G_{\kappa,Pn+p}(i) \cdot G_{\kappa,Pn+p}^*(i)}{\sum_{p=0}^{P-1} G_{k,Pn+p}(i) \cdot G_{k,Pn+p}^*(i)}. \quad (10)$$

We choose three and four triangular MBFs to cover the entire universe of discourse of two inputs, $K_{eff,k,n}$ and $SNR_{k,n}$, and one output $w_{k,n}$, respectively. The membership functions of the two inputs, $K_{eff,k,n}$ and $SNR_{k,n}$, and one output $w_{k,n}$, are illustrated in Fig. 1. In this paper, we consider the fuzzy control rules of a two-input-single-output fuzzy system as

$$R^j: \text{IF } SNR_{k,n} \text{ is } F_1^{l_1} \text{ AND } K_{eff,k,n} \text{ is } F_2^{l_2}, \\ \text{THEN } w_{k,n} = F_3^{l_3}, \quad (11)$$

where $F_1^{l_1}$, $F_2^{l_2}$ and $F_3^{l_3}$ are the linguistic terms of the input variables $SNR_{k,n}$, $K_{eff,k,n}$ and the output variable $w_{k,n}$, respectively, and $l_1, l_3 = 1, 2, 3, 4, l_2 = 1, 2, 3$, and the index of rule $j = 1, 2, \dots, 12$. The rule base, shown in Table 1, including 12 fuzzy IF-THEN rules for stage 2, can be established heuristically by the experimental results, which infer the relations between the adequate weights of interferers and their reliability. The defuzzification method with the centroid calculation [9], which returns the center of area under the aggregated MBFs curve, is being employed here. Thus, the FIS mechanism for the second stage of FB-PPIC scheme can be constructed. As for the latter stages, the adequate PCW is larger than the previous stage due to the increasing reliability of the MAI estimation. Thus, instead of modifying the rule bases, the membership functions of PCWs at stages 3 and 4 of FB-PPIC scheme are modified to adapt to the increasing reliability of MAI estimations. Therefore, a multistage PPIC multiuser detector based on FIS over frequency selective fading channels is shown in Fig. 2.

IV. Simulation Results

To examine the characteristics of MAI incurred by the out of orthogonality, the orthogonal Walsh Hadamard codes [10] are generated for each user with the processing gain 64. In order to focus much attention on the BER variations on the different partial cancellations, we assume perfect subcarrier synchronization with no frequency offset and no nonlinear distortion and perfect subcarrier amplitude/phase estimation for MC-CDMA systems. The terms CPIC (conventional PIC), CW-PPIC and FB-PPIC are used to represent the PIC schemes mentioned in Section III.

From Fig. 3, it is observed that at stage 2 the performance of FB-PPIC is the most robust among the PIC schemes. Moreover, the FB-PPIC outperforms CW-PPIC at light system load and CPIC at heavy system load respectively. Furthermore, at stages 3 and 4, the FB-PPIC outperforms slightly both CPIC and CW-PPIC. To investigate the near-far resistance of multiuser detectors, Fig. 4 depicts BER vs. near-far ratio (NFR) of the half high power users to the desired low power users with the active number of users $K=64$ for the multistage PIC schemes. From Fig. 4, we observe that the performance of FB-PPIC is more near-far resistant than both CPIC and CW-PPIC.

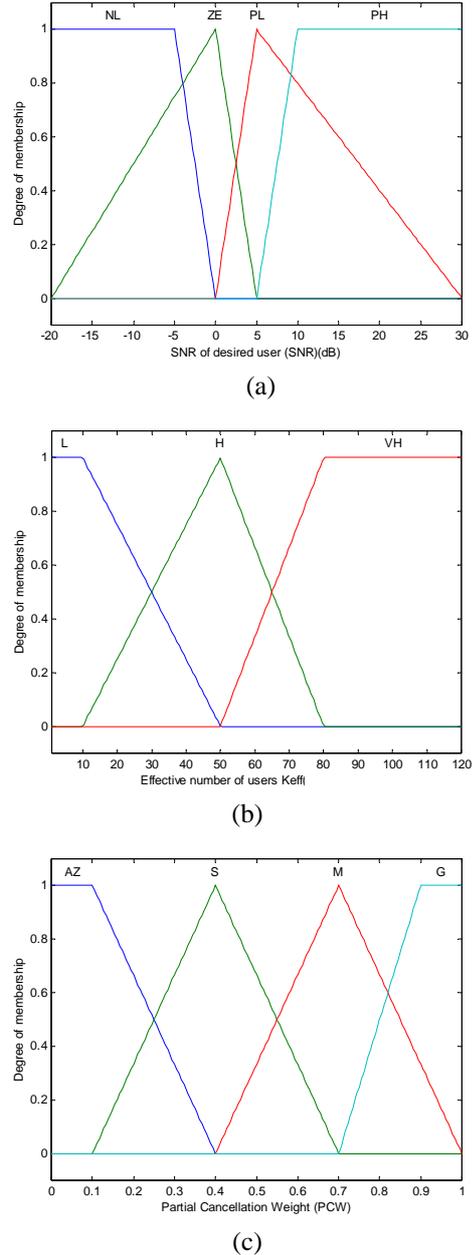


Fig. 1. The membership functions of FIS at stage 2 of FB-PPIC scheme: (a) SNR of interferer (b) effective number of users, K_{eff} (c) PCWs.

Table 1. The rule base of the FIS in stages 2, 3 and 4.

$K_{eff} \backslash SNR$	NL	ZE	PL	PH
L	M	M	G	G
H	S	M	M	G
VH	AZ	S	S	M

Then, the comparisons on the capacity of multistage FB-PPIC, CW-PPIC and CPIC are shown in Fig 5. From Fig. 5, with the performance requirements of $BER=0.01$, the capacity of FB-PPIC is 65, which is higher than both 54 of CW-PPIC and 52 of CPIC, respectively, at stage 4. Moreover, the CPIC exhibits a lower BER than CW-PPIC at low system load $K < 50$.

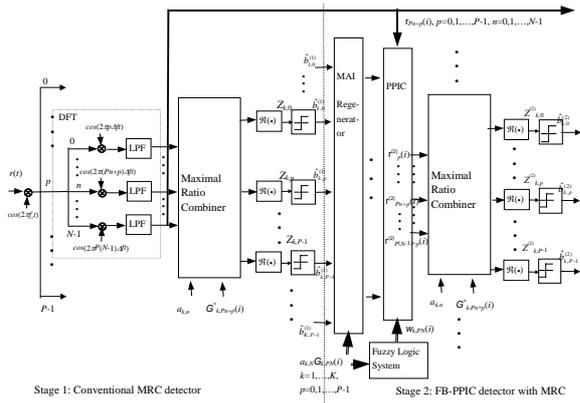


Fig. 2. Multistage fuzzy-based partial parallel interference cancellation detection scheme.

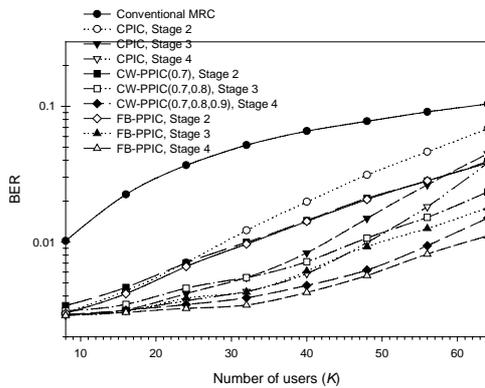


Fig. 3. Performance of multistage PIC multiuser receivers with $SNR=6dB$ and $N=64$.

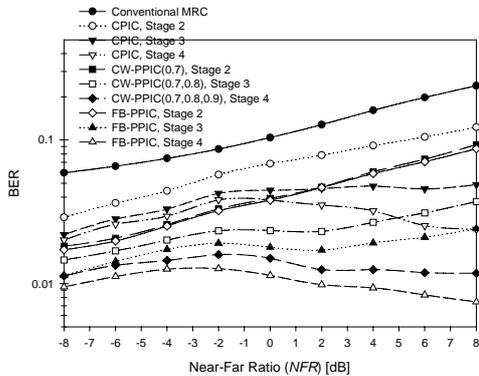


Fig. 4. The BER vs. near-far ratio for PIC detectors with the active number of users $K=64$ and $SNR=6dB$.

V. Conclusion

In this paper, we propose a multistage fuzzy-based PPIC multiuser detector to improve the performance of MC-CDMA systems over frequency selective fading channels. Simulation results show that with the inferred adequate PCWs by FIS, the FB-PPIC can then outperform the CPIC and CW-PPIC over frequency selective fading channels, especially at heavy load and in near-far environments.

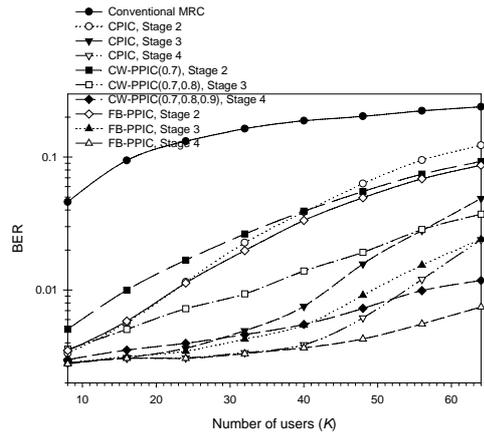


Fig. 5. The BER vs. the active number of users for PIC detectors with $NFR=8dB$, $SNR=6dB$ for the desired low power users.

References

- [1] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatley, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303-312, May 1991.
- [2] J. A. C. Bingham, "Multicarrier for data transmission: an idea whose time has come," *IEEE Communications Magazine*, pp. 5-14, May 1990.
- [3] S. Hara and R. Prasad, "Design and performance of multicarrier CDMA system in frequency-selective Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1584-1595, Sep. 1999.
- [4] S. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.
- [5] M. K. Varanasi and B. Aazgang, "Multistage detection in asynchronous code division multiple-access communications," *IEEE Trans. Commun.*, vol. COM-38, no. 4, pp. 509-519, April 1990.
- [6] J.-H. Wen and Y.-F. Huang, "Fuzzy-based adaptive partial parallel interference canceller for CDMA communication systems over fading channels," *IEE Proceedings-Communications*, vol. 149, no. 2, pp. 111-116, April 2002.
- [7] Y. F. Huang, "Performance of partial parallel interference cancellation for multi-carrier CDMA systems," to appear in *Proceedings of IEEE ISPACS 2003*, Japan, Dec. 7-10, 2003.
- [8] L. A. Zadeh, "Fuzzy set," *Inf. Control*, vol. 8, pp. 338-353, 1965.
- [9] C.-T. Lin and C. S. G. LEE, *Neural Fuzzy Systems: A Neural-Fuzzy Synergism to Intelligent Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [10] R. L. Peterson, R. E. Ziemer and D. E. Borth, *Introduction to Spread Spectrum Communications*, NJ: Prentice-Hall, 1995.
- [11] J. G. Proakis, *Digital Communications*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

Performance of Authentication and Encryption algorithms in IPSec Tunnelling

Siraj Shaikh

Broadlands 202, Park Campus,
Department of Multimedia and Computing,
University of Gloucestershire Business School,
Cheltenham Spa, GL52 2QF, UK

Email: sshaikh@glos.ac.uk

Samir Al-Khayatt

School of Computing and Management Sciences
Sheffield Hallam University
City - Campus, Howard St,
Sheffield, S1 2WB, UK

Email: s.alkhayatt@shu.ac.uk

Abstract. Internet Protocol Security (IPSec) is a protocol suite that provides a secure way for communicating by using the TCP/IP protocol. The IPSec protocol is a set of security extensions developed by the IETF, providing what may be termed “packet-level security”. It provides an extensive set of configurations and tunnelling techniques. The choice, however, of authentication and encryption algorithms and configurations affects issues such as data throughput and performance throughout the network. This paper investigates the use of various authentication and encryption algorithms used in IPSec tunnels and their effect on the overall performance. A test setup is established and IPSec is implemented. The setup used was inspired by similar attempts carried out in [1] and [2].

Keywords: *IPSec, Security protocol, Algorithm, Performance*

1 Introduction

Cryptographic algorithms are critically fundamental in providing security services, especially when used in security protocols. While their use provides a convenient way to encrypt and decrypt ciphertexts and plaintexts, it is not without compromising considerable processing resources. The purpose of this study is to deploy IPSec on a small scale and analyse the performance of the network for some of the widely used algorithms.

The paper is structured as follows. Section 2 describes the main components of IPSec, relevant to the performance tests, Section 3 describes the test setup, Section 4 presents the results, and Section 5 concludes the paper.

2 Overview of IPSec

The purpose of this section is to explain the underlying concepts of IPSec [3]. It is a suite of

protocols rather than a single protocol, providing a set of IP extensions for implementing efficient security at the IP-layer level. This is important as IP networks have provided the foundation for the modern worldwide networks, and the “network of the networks”, the Internet. It offers effective key negotiation and exchange, which makes it very suitable for scalable deployment.

It combines two sub-protocols, Authentication Header (AH) [4] and Encapsulating Security Payload (ESP) [5] to work together. Both of them play a different role and although they work together, at the same time they are independent protocols performing independent functions. The functionality of each and how they coordinate with each other to provide different aspects of a secured communication model is examined.

2.1 Authentication Header (AH)

The AH protocol [4] ensures that every IP packet is authenticated and maintains its integrity. It applies cryptographic hash to the packet data and identification information (source/destination address), and attaches it to the packet. The AH header comes after the basic IP header, as shown in Figure 1 below. The hash is applied to the shaded parts of the packet.

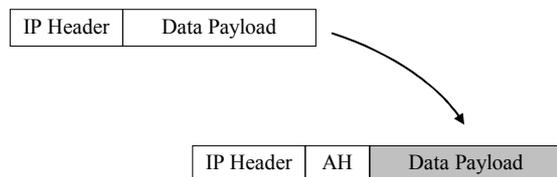


Figure 1. Packet processing in Authentication Header

SHA-160 [6] and MD5 [7] are used by AH protocol to provide message authentication and integrity. SHA-160 (Secure Hash Algorithm) is part of a Secure Hash Standard and is one of the effective

ways to provide data integrity with origin identification. It produces a “condensed representation of a message or a data file”; any message of any length (less than 264 bits) is fed to the algorithm and it produces a 160-bit output, known as the message digest.

The MD5 algorithm performs a similar operation of preserving the data’s integrity. It differs from SHA-1 in the way it processes the data for hashing and is claimed to be faster than SHA-1. It pads the message to ensure that its length in bits plus 64 is divisible by 512. A 64-bit binary representation of the original length of the message is then concatenated to the message. The message is then processed in 512-bit blocks in an iterative structure and each block is processed in four distinct rounds, each performing an operation over the 512-bit block, resulting in a message digest.

SHA-1 is 32-bits longer than MD5. Therefore the probability of producing two messages with the same message digest is 1 in 3.40×10^{38} in MD5 and 1 in 1.46×10^{48} in SHA-1; hence SHA-1 is considerably stronger than MD5. MD5 is generally considered to be slightly more vulnerable whereas SHA-1 is generally believed to be resistant to cryptanalysis. One reason is that the design criteria of SHA-1 are not public, so the security is difficult to comment on. In terms of execution, both algorithms are executed considerably faster on a 32-bit architecture. MD5 has 4 rounds of 16 steps and a bit-length of 128, while SHA-1 has 4 rounds of 20 steps and a bit-length of 160. Therefore, SHA-1 compared to MD5 in theory will execute slightly slower as it has more steps and a longer buffer.

2.2 Encapsulating Security Payload (ESP)

The ESP protocol [5] rewrites the IP payload in an encrypted form as shown in Figure 2. The ESP header does not consider the fields of the IP header in front of it and therefore guarantees nothing but the payload, leaving it for the AH layer to deal with. An ESP header may also provide authentication for the payload, but not the IP header as the AH provides it. The packet processing for ESP is exactly the same as AH. AH and ESP may not be applied together. If the AH and ESP are both applied concurrently, then the ESP header will follow the AH Header, allowing the AH to perform the hash function on the packet, including the ESP header.

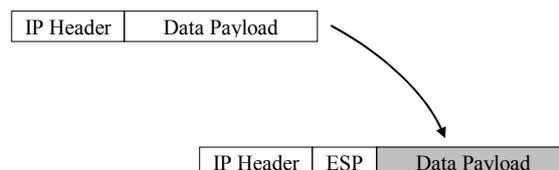


Figure 2. Packet processing in Encapsulating Security Payload

ESP uses DES (Data Encryption Standard) [8] and Triple-Des to encrypt the data at the packet level. DES is one of the most commonly used symmetric encryption algorithms. This means the sender and receiver of the data must know the same secret key being used to encrypt and decrypt the data. It is a block cipher that transforms 64-bit data blocks under a 56-bit key, by means of permutation and substitution. The key size for the single DES is actually 64-bit but 8 parity bits are removed from it. Triple-DES is the same algorithm applied three times, in order to further strengthen the security; the input data is, in effect, encrypted three times.

3 Performance Test Setup

The aim of these tests is to reveal the affects on performance by applying different encryption algorithms and authentication methods, used for tunnelling the data. This will be done by determining and comparing the performances of the VPN tunnelling applying different encryption algorithms and authentication methods. This section describes the test network that was used for our experiments.

Figure 3 shows the setup of the network. The set-up shows a very orthodox Firewall/VPN network set-up. A DMZ (De-MilitariZed) subnet can be seen positioned between two networks, the public network on the left and the private network on the right. It hosts a router-based-firewall in a Cisco 2621 router [9]. It also hosts a Windows 2000 Server [10] acting as a VPN Gateway/Server. The purpose of this is to provide a secure communication tunnel with the users who are placed on the public network (node A) accessing the resources on the private network (Node B). The two test machines used were using a Pentium III processor with a processor speed of 800 MHz, supported by 128 MB of RAM memory.

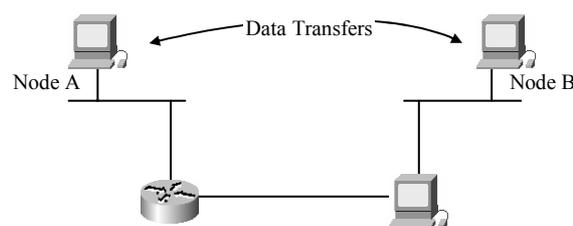


Figure 3. Test setup

The criteria used throughout the test were the channel data throughput. The throughput determined in Kilobytes per second (denoted by Kbps) reflects the efficiency of the communication channel. The throughputs calculated for different arrangements and configurations are then compared with each other in order to establish a comparative study. The higher the throughput, the more effective and quicker a set-up is. Differences in throughputs achieved are also calculated in percentages for the purpose of comparison.

To determine the performance of the network, two nodes were selected: node A and B (refer to figure 4). Data is being transmitted from node A to node B. To calculate the throughput, files were transferred using the File Transfer Protocol (FTP) mechanism with time recorded to transfer the complete file. Three different file sizes are transferred; 100, 200 and 400 MB. For each file size, the file was transferred three times with the average time of the three transfers being recorded only. The data throughput is, subsequently, calculated. The procedure is repeated using a different encryption algorithm and an authentication method every time. During these tests, the same rules applied as in the last tests. No configurations were altered at any time. Every time the algorithms were changed, a new SA was formed between the communicating machines; obvious as a new SA is formed every time any of the parameters change. The results obtained varied considerably and provide interesting observations as discussed in the next section.

4 Performance Results

The choice of algorithms in the Windows 2000 implementation of IPSec is between DES and Triple-DES for encryption and SHA-160 and MD5 for authentication methods. Table 1 below shows the throughputs for the four different combinations of algorithms.

Encryption	Authentication	Data Throughput (Kbps)
DES 64-bit	SHA-1 160-bit	3154.88
DES 64-bit	MD5 128-bit	2635.37
3DES 192-bit	SHA-1 160-bit	2348.47
3DES 192-bit	MD5 128-bit	2543.63

Table 1. Comparing the performance of algorithms

The Triple-DES is slower than DES as expected. When combined with Triple-DES, SHA-1 complies with the theoretical basis of its execution and is slower than MD5 but with DES the results suggest otherwise. One of the reasonable explanations can be the fact that due to it having a larger buffer, some

extra data may be processed in each payload unit, providing quicker processing of data.

4.1 Processing offload

It is interesting to observe that the test machines involved spent a large part of their processing resources in the encryption/decryption of the packets; incurring a huge load. When considering a large-scale deployment of IPSec, such features have to be carefully examined. One of the ways to deal with such a problem is to use some dedicated hardware to offload some of the cryptographic processing. The test machines used a unique network interface card (NIC), which offloads some of the IPSec processing. The packets pass through the NIC and vice versa. They work along with the computer processor to share some of the processing and hence, offload some of the IPSec processing. The performance of the same combination of algorithms with Windows 2000 IPSec-offloading is shown below in Table 2.

The IPSec-offloading does affect the performance for DES but not much difference is noted between the throughputs obtained with triple DES. The extra amount of processing does not allow the offloading NICs (dedicated hardware) to provide any better efficiency. When combined with DES, MD5 is slower than SHA-1; in contrast to that when MD5 is combined with triple DES, its apparently faster. The relation may be supported by the results obtained; here again, the pattern is repeated..

Processing with dedicated hardware		
Encryption	Authentication	Data Throughput (Kbps)
DES 64-bit	SHA-1 160-bit	3568.83
DES 64-bit	MD5 128-bit	2978.04
3DES 192-bit	SHA-1 160-bit	2548.60
3DES 192-bit	MD5 128-bit	2668.04

Table 2. Performance with dedicated hardware

The comparison between the hardware and software supported processing of different algorithm combination is shown in Table 3 below.

Encryption	Authentication	IPSec Processing		
		Software supported	Hardware supported	Gain (%)
DES 64-bit	SHA-1 160-bit	3154.88	3568.83	13.12
DES 64-bit	MD5 128-bit	2635.37	2978.04	13.00
3DES 192-bit	SHA-1 160-bit	2348.47	2548.60	8.52
3DES 192-bit	MD5 128-bit	2543.63	2668.04	4.89
Average improvement (due to dedicated hardware)				9.88

Table 3. Performance gain with dedicated hardware

An average improvement of nearly 10% only stresses the efficiency of the hardware offloading of IPSec processing. Any offloading of these functions improves the performance considerably. In the first two options (with 64-bit DES) the increase in performance is at least 13%.

5 Conclusion

This paper has focussed on a particular aspect of the performance of the IPSec protocol. It has looked into the various cryptographic algorithms used in the protocol and how they affect the throughput of the network in various combinations. The test results obtained are discussed and compared. The most interesting aspect of these tests have been the use of dedicated hardware to offload some of the cryptographic processing on the test nodes.

Performance testing secured communications and networks requires consideration and care in order for the results to be accurate and worthy. The test setups used in this study were inspired by similar attempts made in [1] and [2]. It is important to consider more controlled and accurate test setups and adopt more precise testing procedures.

For the purpose of reliable security, it is worthwhile to investigate and evaluate the cryptographic properties of the IPSec protocol; some important weaknesses have been highlighted elsewhere [11]. Whether these weaknesses relate to the security performance in the IPSec standard is of interest.

References

- [1] Al-Khayatt, S., Shaikh, S., Akhgar, B. and Siddiqi, J. *A Study of Encrypted, Tunnelling Models in Virtual Private Networks*. 3rd IEEE Conference on Information Technology ITCC-2002, Las Vegas, Nevada, USA, 2002
- [2] Al-Khayatt, S., Shaikh, S., Akhgar, B. and Siddiqi, J. *Performance of Multimedia Applications with IPSec Tunneling*. 3rd IEEE Conference on Information Technology ITCC-2002, Las Vegas, Nevada, USA, 2002
- [3] Kent, S. and Atkinson, R. *Security Architecture for the Internet Protocol*. Internet Request for Comments RFC 2401, Internet Society, Network Working Group, 1998
- [4] Kent, S. and Atkinson, R. *IP Authentication Header (AH)*. Internet Request for Comments RFC 2402, Internet Society, Network Working Group, 1998
- [5] Kent, S. and Atkinson, R. *IP Encapsulating Payload (ESP)*. Internet Request for Comments RFC 2406, Internet Society, Network Working Group, 1998
- [6] *The Use of HMAC-SHA-1-96 within ESP and AH*. RFC 2404. On the Web at <http://www.ietf.org/rfc/rfc2404.txt>
- [7] *The Use of HMAC-MD5-96 within ESP and AH*. RFC 2403. On the Web at <http://www.ietf.org/rfc/rfc2403.txt>
- [8] *Data Encryption Standard* (1999) FIPS PUB 46-3, U.S Department of Commerce/National Institute of Standards and Technology. On the Web at <http://csrc.nist.gov/fips/fips46-3.pdf>
- [9] *Cisco Systems Inc*. On the Web at <http://www.cisco.com>
- [10] Microsoft Windows 2000 product family. On the Web at <http://www.microsoft.com/windows2000/>
- [11] Ferguson, N and Schneier, B. *A Cryptographic Evaluation of IPSec*. Counterpane Internet Security, Inc. On the Web at <http://www.counterpane.com>, 2000

Performance of Quantized Turbo Decoding for Decoders With Maximum A Posteriori Algorithm

Omar Hasan
Princess Sumaya University
Amman, Jordan

ABSTRACT

This paper investigates Bit Error Performance for quantized turbo code over Gaussian channels. The turbo encoder is built using a parallel concatenation of two recursive systematic convolution encoders and a random interleaver. The turbo decoder employs two iterative symbol by symbol MAP based algorithm decoders for decoding the received data. For a given quantization level, the Bit Error Rates are examined by considering the influence of the code structure, the block size, and the number of decoding steps. Finally, to measure loss of power efficiency due to quantization, the results are compared to previously published results for the unquantized case under the same Bit to noise ratio and the code structure.

I. INTRODUCTION

In recent years, a new class of forward error correction code known as turbo code has been invented and showed to be so powerful [7] in terms of its power efficiency. The importance of turbo coding, arises from its ability for providing significant coding gain by employing two simple constituent convolution encoders and an interleaver. For instance, it was reported by [14], a required $E_b / N_0 = 0.8$ dB for bit-error rate $BER = 10^{-6}$ for a rate $1/2$ turbo code. This remarkable coding gain makes turbo coding suitable for low power applications such as; satellite communications and deep – space communications. In fact, turbo coding has been used by NASA since early 1960 on the Galileo space craft for deep space communications.

Two main decoding algorithms can be implemented for turbo decoding. The symbol-by-symbol maximum a posteriori (MAP) algorithm and the soft-output Viterbi (SOVA) algorithm. They differ by the way both algorithms estimates the states of the trellis. The SOVA estimates the states through a connected path, while the states estimated by the MAP algorithm need not to be connected. The MAP algorithm is more computationally involved than the SOVA algorithm, but performs better than the SOVA algorithm. For example, at $E_b / N_0 = 2.5$ dB , a bit-error rate

$BER \approx 10^{-5}$ and $BER \approx 10^{-4}$ can be attained by using the MAP decoding algorithm and the SOVA decoding algorithm respectively, for a rate $R=1/2$ turbo encoder with memory length $m=2$ (convolution encoder) and a 1024 bits random interleaver. In this paper, we will consider the MAP algorithm as the base decoding algorithm for evaluating the bit-error rates (BER).

Many papers have provided results for turbo code applications [3], [4], and [5]. However, in their analysis, they assumed that the received data levels entering the decoder take on any value. However, this assumption makes turbo coding unattractive for FPGA technology and digital processing. In practice, it is desired to limit the number of levels entering the decoder, which simply can be achieved by quantizing the received data according to a certain quantization error depending on the signal type and the desired system accuracy. This paper begins with an overview of turbo code structure followed by full understanding of the recursive systematic convolution codes (RSC), where RSC forms the essence of turbo code. Next, a description of the turbo coding-decoding process is presented, followed by discussion of the results obtained in this paper. The results obtained here will be compared against the unquantized case [14] using the same decoding algorithm (MAP) and the code structure. Finally, we conclude the paper.

II. System Model

A. Recursive Systematic convolution Codes (RSC)

Before the introduction of turbo codes, low rates convolution codes were implemented by many low power applications such as; deep space communications and satellite communications. It is well known, that at low signal to noise ratio Bit Error Rates (BER) of RSC outperforms the BER performance of non systematic convolution codes (NSC). It is also well known, that for any signal to noise ratio and for block size of about $N \geq 1000$ bits, RSC performs better than non-systematic (NSC) with the same code structure.

A systematic convolution code can be attained from a non systematic convolution code without affecting the minimum distance property, by feeding back one of the outputs to the input. Such a code is called recursive systematic convolution code, and it is the basic building block for turbo codes. Figure 1, shows a RSC encoder obtained from a NSC encoder with memory $M = 3$ and code generators expressed in the octal format $G = (g_0, g_1)_8 = (13, 17)$, with g_0 being the feedback code polynomial and g_1 is the forward code polynomial. It is interesting to note that the parity bits x_{1k} and x_{2k} generated by both encoder do not correspond to the same data input due to the existence of the interleaver at the input of the lower encoder.

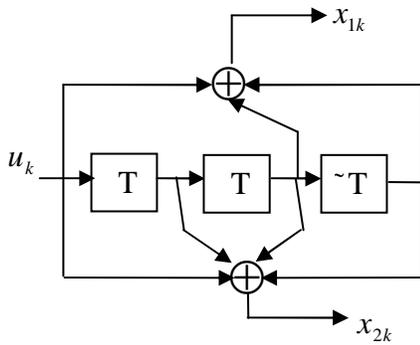


Fig. 1a. Non Systematic encoder

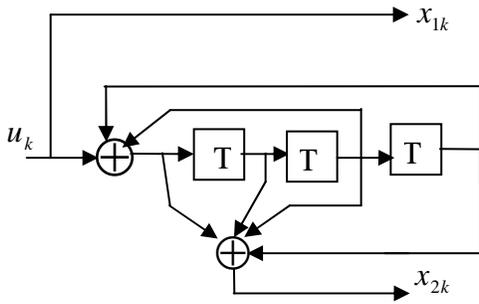


Fig. 1b. Recursive Systematic encoder

B. Turbo Encoder

A typical turbo encoder is shown in Fig. 2. As seen in the figure, the turbo encoder is constructed from cascading two binary rate $R = 1/2$ RSC encoders in parallel (parallel concatenation) separated by N-bit interleaver. For this encoder, it is apparent that a block of N data bits is mapped by the encoder to produce $3N$ coded bits with an overall code rate $R^* = 1/3$. Even though, the turbo encoder considered here utilizes two identical constituent encoders, but turbo code encoders can be constructed from a multiple of different constituent encoders as seen in [11]. However, due to the nature of the decoding process for turbo codes, multiple

constituent encoders increase the system complexity and the system delay tremendously for irrelevant coding.

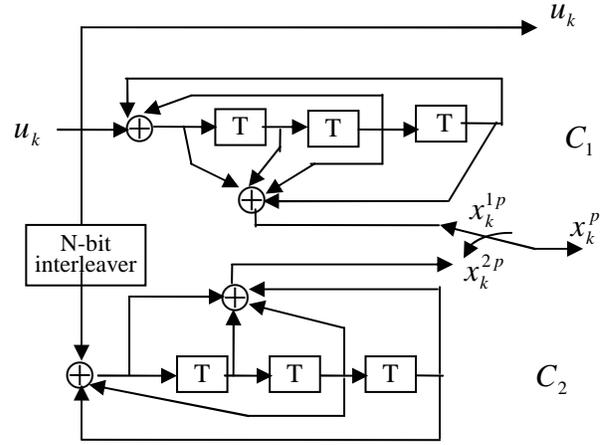


Fig. 2. Parallel Concatenated Recursive Systematic Encoder

Now, assume a turbo encoder with interleaver size of N-bits and an input data sequence u given by

$$u = (u_1, u_2, u_3, \dots, u_N) \quad (1)$$

then, for n -parity bit encoder, the equivalent output code word generated by the encoder is:

$$C = (C_1, C_2, C_3, \dots, C_N) \quad (2)$$

where C_k is the output code word at time k given by:

$$C_k = (u_k, x_k^{1p}, x_k^{2p}, \dots, x_k^{np}) \quad (3)$$

Usually, the systematic bit u_k is taken from the upper RSC encoder (C_1), while the generated parity bits x_k^{ip} are taken from the upper RSC encoder and lower encoder (C_2). However, for the case under study as predicted by figure 2, the output code word has the form

$$C_k = (u_k, x_k^{1p}, x_k^{2p}) \quad (4)$$

The three output streams are then multiplexed to form the code word $C = (C_1, C_2, C_3, \dots, C_N)$, with overall code rate $1/3$. If a puncturing mechanism is used to increase the data efficiency, some parity bits x_k^{1p} or x_k^{2p} are deleted according to some puncturing pattern defined by a matrix \mathbf{P} . For example, the following puncturing matrix \mathbf{P} with period six can be used to increase a rate $1/3$ code to rate $1/2$

$$\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

If such a puncturing scheme is used, then the code word C given by

$$C = (u_1, x_1^{1p}, x_1^{2p}, u_2, x_2^{1p}, x_2^{2p}, u_3, x_3^{1p}, \dots, u_N, x_N^{1p}, x_N^{2p})$$

becomes the output code word C_p

$$C_p = (u_1, x_1^{1p}, D, u_2, D, x_2^{2p}, u_3, x_3^{1p}, D, \dots)$$

where D is used to denote the deleted bits (punctured bits). For discrete memoryless Gaussian channel with noise variance σ^2 and binary modulation, the received signal at the quantizer input at time k is \hat{y}_k

$$\hat{y}_k = (\hat{y}_{1k}, \hat{y}_k^p) \quad (5)$$

where the systematic received component is given

$$\hat{y}_{1k} = u_k + n_1 \quad (5a)$$

and the received parity component is:

$$\hat{y}_k^p = x_k^p + n_2 \quad (5b)$$

Here, n_1 and n_2 are zero mean independent Gaussian noise components with identical variance σ^2 . The decoder input at time k is y_k :

$$y_k = (y_{1k}, y_k^p) \quad (6)$$

The received parity bit y_k^p is demultiplexed and sent to decoder C_1 when $x_k^p = x_k^{1p}$, and to decoder C_2 when $x_k^p = x_k^{2p}$. The deleted parity bits at the encoder is replaced by a zero at the decoder. Finally, the received data are decoded using the iterative MAP algorithm which will be explained next.

III. The MAP Decoding Algorithm

In this section, we first discuss the BCJR algorithm [12] for performing symbol-by-symbol MAP decoding. Then, we will show how to utilize this algorithm for iterative decoding employing two BCJR-MAP decoders. For binary channel with data bits $u_k = \pm 1$ and a received sequence \mathbf{y} , the MAP decoder decides $\hat{u}_k = +1$, according to

$$P_r(u_k = +1 | \mathbf{y}) > P_r(u_k = -1 | \mathbf{y}) \quad (7)$$

otherwise it decides $\hat{u}_k = -1$. The above probability can be written in the form of log likelihood ratio (LLR) $L(\hat{u}_k)$ as:

$$L(\hat{u}_k) = \log \frac{P_r(u_k = +1 | \mathbf{y})}{P_r(u_k = -1 | \mathbf{y})} \quad (8)$$

where $\mathbf{y} = (y_1, y_2, \dots, y_N)$, are the noisy received code words, and N is the length of the trellis. Thus, according to the above LLR in (8), the MAP decoder decides according to the following rule;

$$\hat{u}_k = \text{sign}[L(\hat{u}_k)] \quad (9)$$

For the problem under consideration, let the data binary bit u_k at time k is associated with state transition from state s' to s . Where the encoder states s' and s are the trellis levels at times $k-1$ and time k . Incorporating the code trellis, the LLR $L(\hat{u}_k)$ can be written as

$$L(\hat{u}_k) = \log \frac{\sum_{(s', s: u_k = +1)} p_r(s', s, \mathbf{y})}{\sum_{(s', s: u_k = -1)} p_r(s', s, \mathbf{y})} \quad (10)$$

where the summation in (10), is taken over all existing state transitions from state s' to state s caused by an input $u_k = +1$ or $u_k = -1$. The BCJR algorithm for writing $L(\hat{u}_k)$ has the form [1]

$$L(\hat{u}_k) = \log \frac{\sum_{(s', s: u_k = +1)} \alpha_{k-1}(s') \cdot \gamma_k(s', s) \cdot \beta_k(s)}{\sum_{(s', s: u_k = -1)} \alpha_{k-1}(s') \cdot \gamma_k(s', s) \cdot \beta_k(s)} \quad (11)$$

where $\alpha_k(s) = p_r(s, \mathbf{y}_1^k)$ with initial value $\alpha_0(s') = 1$, is the forward recursion of the BCJR-MAP algorithm computed from

$$\alpha_k(s) = \sum_{s'} \gamma_k(s', s) \cdot \alpha_{k-1}(s') \quad (12)$$

and $\beta_k(s) = p_r(s, \mathbf{y}_{k+1}^N)$ initialized with $\beta_N(s) = 1$, is the backward recursion of the BCJR-MAP algorithm evaluated from

$$\beta_{k-1}(s') = \sum_s \gamma_k(s', s) \cdot \beta_k(s) \quad (13)$$

The probability $\gamma_k(s', s) = p_r(y_k | s', s) p_r(s | s')$ is the state transition probability corresponds to the transition event ($s' \rightarrow s$) at time k . By realizing the event u_k corresponds to the transition event ($s' \rightarrow s$), the transition probability $\gamma_k(s', s)$ can be written as:

$$\gamma_k(s', s) = p_r(y_k | u_k) p_r(u_k) \quad (14)$$

Defining the log-likelihood ratio of the bit u_k , $L(u_k)$ as

$$L(u_k) = \log \frac{p_r(u_k = +1)}{p_r(u_k = -1)} \quad (15)$$

The a priori probability $p_r(u_k)$ in (14) can be expressed as

$$\begin{aligned} p_r(u_k \pm 1) &= \frac{e^{\pm L(u_k)}}{1 + e^{\pm L(u_k)}} = \left(\frac{e^{-L(u_k)/2}}{1 + e^{-L(u_k)}} \right) \cdot e^{u_k L(u_k)/2} \\ &= A_k \cdot e^{u_k L(u_k)/2} \end{aligned} \quad (16)$$

and the conditional probability $p_r(y_k | u_k)$ for systematic convolution codes can be expressed as

$$\begin{aligned} pr(y_k | u_k) &= (1/\sqrt{2\pi}) \exp\left[-\frac{2E_c}{N_0} \{(y_{1k} - u_k) \right. \\ &\quad \left. - (\sum_{i=1}^n y_k^p - x_k^p)^2\}\right] \\ &= B_k \exp\left[\frac{2E_c}{N_0} \{u_k y_{1k} + \sum_{i=1}^n y_k^{ip} x_k^{ip}\}\right] \end{aligned} \quad (17)$$

where $E_c = N_0 / 2\sigma^2 = R_c \cdot E_b$, is the coded bit energy and R_c is the coding rate. The quantity B_k is computed from

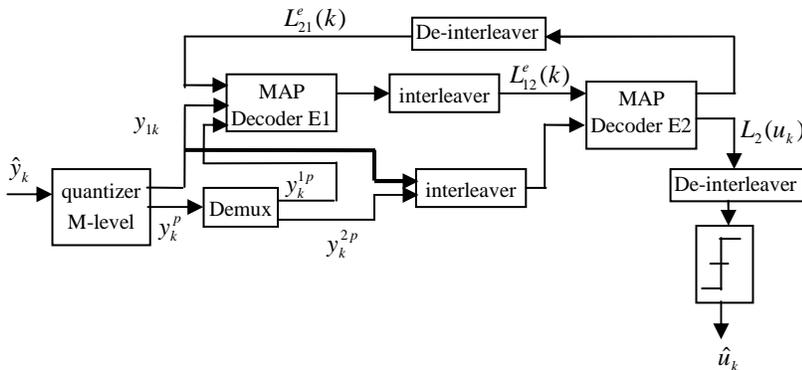
$$\begin{aligned} B_k &= (1/\sqrt{2\pi}) \exp\left[-\frac{2E_c}{N_0} \{y_{1k}^2 + u_k^2 \right. \\ &\quad \left. + \sum_{i=1}^n (y_k^{ip})^2 - (x_k^{ip})^2\}\right] \end{aligned} \quad (18)$$

Combining (16) with (18), the state transition $\gamma_k(s', s)$ can now be expressed as

$$\begin{aligned} \gamma_k(s', s) &= A_k \cdot B_k \exp\left[0.5L_c \{u_k y_{1k} + \sum_{i=1}^n y_k^{ip} x_k^{ip}\}\right] \\ &\quad \cdot e^{u_k L(u_k)/2} \end{aligned} \quad (19)$$

The terms A_k and B_k in (16) and (17) are equal for all state transitions $(s', s : u_k = \pm 1)$, and hence will cancel out on the ratio of (11). Therefore, the LLR $L(\hat{u}_k)$ in (11) can be expressed as

$$L(\hat{u}_k) = L_c y_{1k} + L(u_k) + \log \frac{\sum_{(s', s: u_k = +1)} \alpha_{k-1}(s') \cdot \beta_k(s) \cdot \tilde{\gamma}_k(s', s)}{\sum_{(s', s: u_k = -1)} \alpha_{k-1}(s') \cdot \beta_k(s) \cdot \tilde{\gamma}_k(s', s)} \quad (20)$$



where $L_c = (4E_c / N_0)$ is the channel reliability value, and

$$\tilde{\gamma}_k(s', s) = \exp\left[\frac{1}{2} \sum_{i=1}^n L_c y_k^{ip} x_k^{ip}\right] \quad (21)$$

However, for the case under consideration (figure 2), the quantity $\tilde{\gamma}_k(s', s)$ can be simplified to have the form

$$\tilde{\gamma}_k(s', s) = \exp\left[\frac{1}{2} L_c y_k^p x_k^p\right] \quad (22)$$

Furthermore, if we let

$$L^e(u_k) = \log \frac{\sum_{(s', s: u_k = +1)} \alpha_{k-1}(s') \cdot \beta_k(s) \cdot \tilde{\gamma}_k(s', s)}{\sum_{(s', s: u_k = -1)} \alpha_{k-1}(s') \cdot \beta_k(s) \cdot \tilde{\gamma}_k(s', s)} \quad (23)$$

then, the LLR can be expressed as

$$L(\hat{u}_k) = L_c y_{1k} + L(u_k) + L^e(u_k) \quad (24)$$

The second term represents a priori information about the current bit u_k provided by a previous decoder, and the third term represents an extrinsic information to be passed on to a subsequent decoder during the next half iteration as an additional piece of information (which was not available during the previous iteration) about the present bit u_k .

IV. Turbo Decoder

A typical turbo decoder is shown in Figure 4.

This decoder implements two BCJR-MAP decoders and operates as follows; at any given iteration, decoder one calculates an extrinsic value $L_{12}^e(u_k)$ for each u_k and deliver it to decoder two as an additional piece of information which was not available to decoder two during the information which was not available to decoder two during the previous iteration. Then, this extrinsic value is used by decoder two as a priori information about u_k during the next half iteration step. Similarly, decoder two delivers to decoder one an extrinsic information $L_{21}^e(u_k)$ for each u_k , which will be used by decoder one as a priori information about u_k during the next half iteration. This process continues until the last iteration step is reached, and finally decoder two delivers a hard decision about u_k according to the rule developed by (9). Thus, for any iteration step at time k , decoder two delivers a hard decision for each u_k utilizing three independent estimates about the current bit u_k : the channel value $L_c y_{1k}$, the priori information $L(u_k) = L_{12}^e(u_k)$ provided by decoder one, and an extrinsic value $L_{21}^e(u_k)$ generated by decoder two:

$$L(\hat{u}_k) = L_c y_{1k} + L_{12}^e(u_k) + L_{21}^e(u_k) \quad (25)$$

It is interesting to observe how the interleaver and the de-interleaver are set up to deliver the systematic and the a priori information in the proper sequence for each decoder. The dependency between the three different components in (25) is valid for iteration steps values $i < 8$ (considering the length of the trellis to be within the size used here).

V. Results

The turbo-code encoder used in this section is shown in figure 2. This encoder with an overall rate 1/2 is built using a parallel concatenation of two identical recursive systematic convolution encoders with generating polynomial $g = (g_0, g_1)_8$. The generating polynomials g_0 and g_1 are the feedback polynomial and the forward generating polynomial respectively in the octal format. The turbo decoder shown in figure 4 implements two BCJR-MAP algorithms and operates as explained by equation 25.

The interleaver used in this section is a random interleaver, and the interleaver is randomly generated for each frame. The code is punctured by alternating the coded bits as shown in figure 2. Finally, the **quantizer** used here is a uniform quantizer with peak-to-peak value $v_p = 2$.

For a given quantization level M , the Bit Error Rates are computed after each decoding step as a function of E_b/N_0 , where E_b is the bit energy and N_0 (w/Hz) is the noise power spectral density. For each decoding step, the Bit Error Rates (BER) is considered by allowing a minimum of 50 frames in error to occur. For example, for block size of $N = 1024$ bits and a BER of $p = 10^{-5}$, we attained 50 frames in errors after transmitting 8000 frames. This amounts to 8.2×10^6 transmitted bits. In all of our results, the simulation was terminated by allowing the number of transmitted bits NB , to be $NB \geq 80/p$ and the number of frames in error NF , $NF \geq 50$.

In figures 5-8, the BER are shown for a code generating polynomial $g = (15,17)_8$ as a function of the iteration step i and E_b/N_0 , for quantization levels $M=4,8,16,32$. For any E_b/N_0 value, the BER improves as the value of iteration step i increases. For example, to attain $p = 10^{-4}$ using $M=32$, going from $i=4$ to $i=5$, the coding gain is about 0.125dB. Figure 9, predicts BER for $M=4,8,16,32$, with $g = (15,17)_8$ and an interleaver size $N=2048$ as compared against the unquantized case ($M = \infty$) for iteration value $i=5$. To attain a $BER = 8 \cdot 10^{-5}$, the loss on using $M=32$ is about 0.18dB and is about 0.22dB for $M=16$. Figure 9, also predicts that as the quantization level increases, the gain tends to increase as compared to the ideal case ($M = \infty$). However, the gain is only about 0.04db on going from quantization level $M=32$ to $M=16$, and the gain will be less than 0.04db for quantization level $M > 32$.

The influence of the quantization level M on the BER can be easily examined from figures 5-9. More gain can be attained by increasing M . However, the gain is about 0.04dB going from $M=16$ to $M=32$. This suggests that using a quantization level greater than $M=32$, the gain attained becomes irrelevant. Table I shows the power loss due to the quantization process.

VI. Conclusion

In this paper, we have presented Bit Error Rates for the quantized turbo code which employees the MAP algorithm in Gaussian channel. The results were shown using the recursive Systematic Code for frame size $N=2048$. For each case, the results are shown as a function of the decoding step and the bit energy to noise ratio. As the results depicted, for any turbo code structure and suitable BER value, a sufficient quantized level that can be employed is $M=32$. The gain becomes irrelevant for quantization level greater than $M=32$.

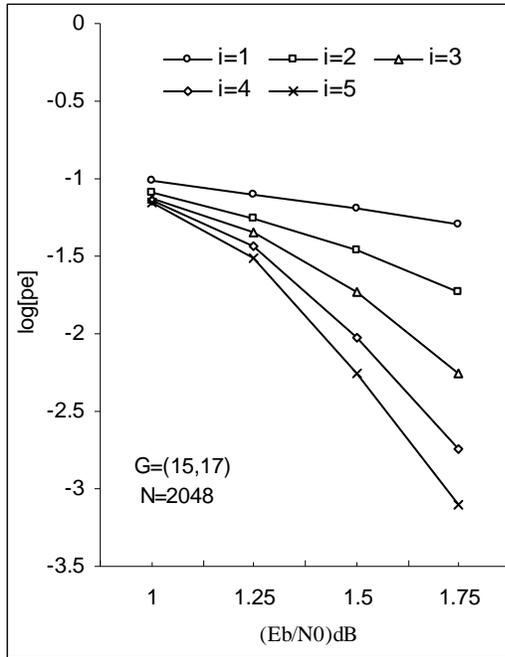


Fig. 5. Turbo code with quantized level $M=4$ for different iteration values

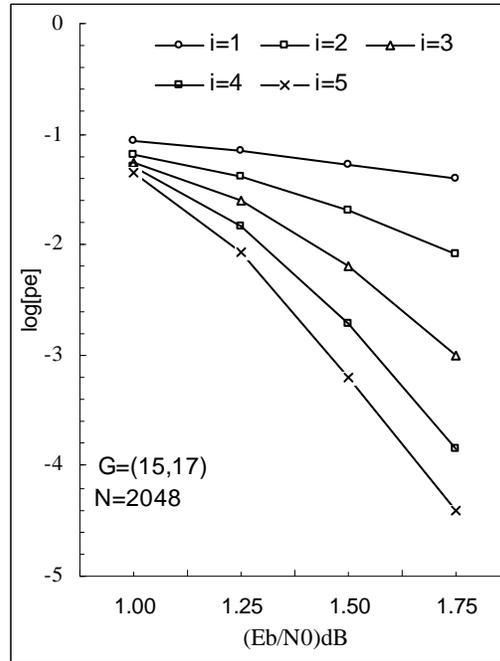


Fig. 6. Turbo code with quantized level $M=16$ for different iteration values

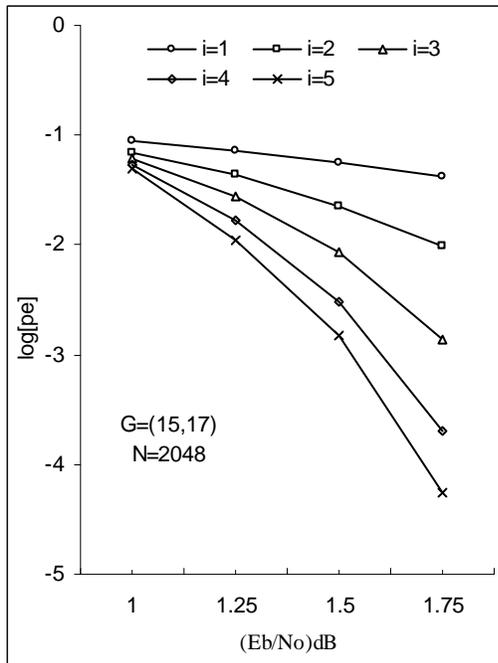


Fig. 6. Turbo code with quantized level $M=8$ for different iteration values

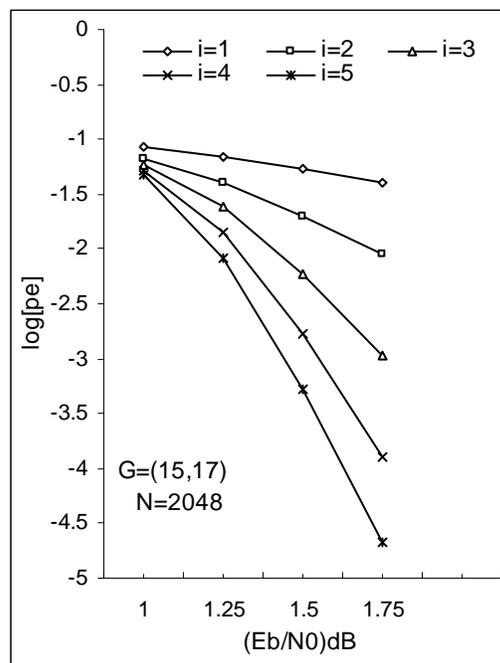


Fig. 6. Turbo code with quantized level $M=16$ for different iteration values

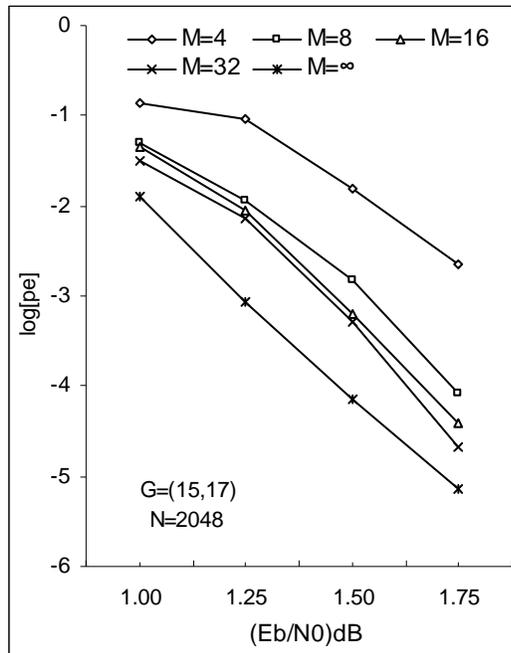


Fig. 9. Quantized VS. unquantized Bit Error Rate performance for N=2048 and iteration step I=5

References

- [1] A. C. Reid, T. A. Gulliver, and D. P. Taylor, "Convergence and errors Turbo- Decoding ,"IEEE Trans. On Comm, Vol. 49, No. 12, pp. 2045-2051, Dec. 2001.
- [2] P. Moqvist, and T. M. Aulin, "Serially concatenated continuous phase modulation with iterative decoding , "IEEE Trans. On Comm, Vol. 49, No. 11, pp. 1901-1915, Nov. 2001.
- [3] J. Hagenauer, E. Offer, and L. Pappe, " Iterative Decoding Of Binary Block and Convolutional Code," IEEE Transactions on Information Theory, Vol. 42, No. 2, pp. 429-445, March 1996.
- [4] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error – Correcting Coding and Decoding: Turbo-codes (1)," In Proc., IEEE Inter. Conf. on Communications, Geneva, Switzerland, pp. 1064-1070, May 1993.
- [5] S. Benedetto, and G. Montostri, " Design of Parallel Concatenated Codes," IEEE Trans. Comm., pp. 591-600, May 1996.
- [6] J. K. Wolf, "Efficient maximum likelihood decoding of linear block code using trellis," Proc. IEEE Trans. Inform. Theory, Vol. 24, pp. 76-80, Jan 1978.
- [7] S. Benedetto, D. Montostri, and F. pollara, " Soft-Output decoding algorithms in iterative decoding of Turbo codes," TDA Progress Report 42-124, pp. 63-87, Feb 1996.
- [8] L. Perez, J. Seghers, and D. Castello, " A distance spectrum interpretation of turbo codes," IEEE Trans. Inf. Theory, pp. 1698-1709, Nov. 1996.
- [9] P. Robertson, E. Villebrum, and P. Hoeher, " A comparison of optimal and suboptimal MAP decoding algorithm in the log domain," Proc. 1995. Int. Conf. on Comm., pp. 1009-1013.
- [10] M. C. Valenti, and B. D. Woerner, " Variable latency turbo codes for wireless multimedia applications," Proc. Int. Symp. On Turbo Codes and Related Topics, (Brest, France), pp. 216-219, Sept. 1997.
- [11] B. Sklar, " A primer on turbo code concepts," IEEE Comm. Magazine, Vol. 35, pp. 94-102, Dec. 1997.
- [12] L. R. bahl, J. Cocke, F. Jeinek, and J. Raviv, " Optimal decoding of linear codes for minimizing symbol error rates," IEEE Trans. Inform. Theory, Vol. 20, pp. 284-287, Mar. 1974
- [13] D. Divsalar, and F. Pollara, " Turbo codes for deep-space communications,"The Telecommunication and Data Acquisitions Progress Report 42-120, Jet Propulsion Lab., Pasadena, CA, pp. 29-39, Feb. 1995.
- [14] D. Divsalar, and F. Pollara, " On the design of Turbo codes ,"The Telecommunication and Data Acquisitions Progress Report 42-123, Jet Propulsion Lab., Pasadena, CA, pp. 99-121, Nov.1995.

Pre-Post-Rake Schemes for DS/SS Mobile Communications

Gaetano Giunta^{*}, Luc Vandendorpe^{**}, Francesco Benedetto^{*}

^{*} *Signal Processing and Multimedia Communications Laboratory, Department of Applied Electronics, University of ROMA TRE, Via della Vasca Navale 84, 00146 Rome, Italy.*

^{**} *Communications and Remote Sensing Laboratory, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.*

e-mails: giunta@ieee.org, vandendorpe@tele.ucl.ac.be, fbenedet@uniroma3.it

Abstract. - This contribution focuses on the Pre-Post-Rake procedure by proposing a new Pre-Post-Rake scheme that uses phase-only filters. According to our results, the new scheme improves the performance of the conventional approach in both single user and multi-user environments, as well as in terms of bit error rate, by a little extra computational complexity.

1. INTRODUCTION

Direct-Sequence Spread-Spectrum Code-Division Multiple-Access (DS/SS CDMA) has raised a significant importance in mobile communications of the last few years due to its resilience to multi-path propagation. This feature can be used to obtain signal diversity by employing a filter, matched to the channel response (see Fig. 1a), and implemented by a rake receiver [1]. Another technique, which an emerging trend has focused on, is Time Division Duplexing (TDD) in which one of the main advantages is the reciprocity between the link channel characteristics, implying the knowledge of the channel impulse response [2]. This statement allows achieving multi-path diversity gain using signal processing at the transmitter, instead of employing a classical (Post-)Rake receiver at the mobile station (see Fig. 1b). This technique, known as Pre-Rake, has been well discussed in [3-5]. In this contribution, we will focus on the Pre-Post-Rake procedure as an alternative solution to the standard Rake processing and to the Pre-Rake approach. The Pre-Post-Rake is a well-known technique in TDD systems in downlink [6]. The Post-Rake in the Pre-Post-Rake scheme acts essentially like a conventional Rake Receiver, but it is conversely matched to the combination of Pre-Rake and multi-path channel (see Fig. 1c). In this work, we propose a new Pre-Post-Rake scheme which, using phase-only filters (Fig. 1d), improves the performance of the conventional approach in terms of bit error rate

(BER) in both single user and multi-user environments.

2. FRAMEWORK OF RAKE SCHEMES

Let us consider a communication link consisting of the concatenation of transmit filter, radio channel and receive filter, with frequency response $H(f)$ and additive white Gaussian noise; then a conventional Pre-Post-Rake scheme will obtain better performance in relation to a simple Pre-Rake, and consequently in relation to a classical Rake Receiver. This solution is not yet the optimal one because of the new scheme that we propose will achieve better results, as underlined in the next sections. The new scheme is composed by a Pre-Rake with a phase-only filter and by a Post-Rake matched to the combination of the Pre-Rake itself and the response of the multi-path channel. Referring again to Fig. 1, showing a block diagram of the four different approaches, focusing on the conventional one (Fig. 1c) and the new scheme (Fig. 1d), we define the new phase-only Pre-Rake filter with the following frequency response:

$$<H^*(f) = \left[\frac{H^*(f)}{H(f)} \right]^{\frac{1}{2}} \quad (1)$$

where $[\cdot]^*$ denotes complex conjugate. The filter's gain A that multiplies $<H^*(f)$ in Fig. 1d is properly set in order to have the same power level at the receiver's input in both conventional and new schemes.

Like in [7], we employ a base-band tapped delay line with tap spacing equal to the chip period T_c as a testing channel model. For sake of simplicity, we also consider data transmission with pulse shaping in band-limited channels, where the shaping filter is designed accordingly to the Nyquist criterion [6]. In addition, we assume that the channel impulse response changes slowly in relation to a transmission frame, considering a TDD system under slow fading conditions (which are typical for portable communication systems), in order to assume the channel time-invariance within a frame

This work has been partially supported by the Italian *Ministry for Telecommunications* under the *ISCTI-CIRIAF* research project.

[8]. Moreover, we assume that both transmission and reception filters have unit energy and, in order to avoid inter-symbol interference due to the filter, ideal sampling is also assumed [6]. Hence, the complex low-pass impulse response of the channel is given by:

$$h(t) = \sum_{l=0}^{L-1} K_l e^{j\gamma_l} \delta(t - lT_c) \quad (2)$$

where L is the number of channel resolvable paths; the path gains K_l are independent identically distributed (i.i.d.) Rayleigh random variables for all l ; the angles γ_l are i.i.d. (uniformly distributed) in $[0, 2\pi)$. Let now $s(k)$ be the discrete transmitted signal, hence, the sampled signal after the filtering will be:

$$r(k) = \sum_{l=0}^{L-1} g_l s(k-l) + n(k) \quad (3)$$

where $g_l = K_l e^{j\gamma_l}$ are the complex gain (with amplitude K_l and phase γ_l) of the l path and $n(k)$ represents the filtered complex Gaussian noise with zero mean and variance N_0 . We assume the modulation is a coherent QPSK with chip energy E_c , processing gain G complex chips/bit and the data bits are processed by an error correction encoder of rate R , whose output is repeated $K=GR$ times. The discrete transmitted signal is:

$$s(k) = \sqrt{\frac{E_c}{2}} d(k) c(k) \quad (4)$$

where $c(k)$ is the complex scrambling code and $d(k)$ the signal at the chip rate. Hence, with a conventional Rake Receiver approach the received signal can be written as:

$$r(k) = \sqrt{\frac{E_c}{2}} \sum_{l=0}^{L-1} g_l d(k-l) c(k-l) + n(k). \quad (5)$$

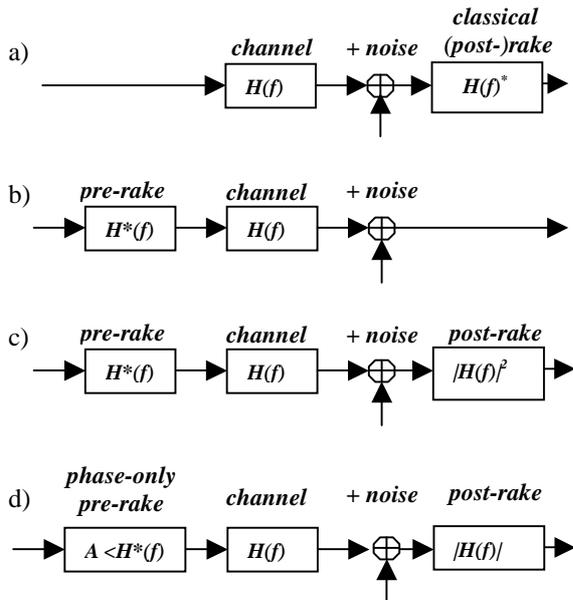


Fig. 1: Principle block models of classical (Post-) Rake, Pre-Rake, conventional, and new Pre-Post-Rake schemes for DS/CDMA communications.

3. SINGLE-USER ENVIRONMENT

Using a Pre-Rake, the discrete transmitted signal expressed by (5) is pre-processed like in a Rake Receiver according to the channel characteristics. Following the approach of [6], the transmitted signal normalized to the chip energy E_c is (under the hypothesis of perfect channel knowledge):

$$s_{pR}(k) = \sqrt{\frac{E_c}{2 \sum_{m=0}^{M-1} |g_m|^2}} \sum_{m=0}^{M-1} g_m^* d(k+m) c(k+m) \quad (6)$$

The received signal will be, after the channel:

$$r(k) = \sum_{l=0}^{L-1} g_l s_{pR}(k-l) + n(k) \quad (7)$$

To simplify the analysis let us consider the modified received signal given by:

$$r'(k) = r(k) \sqrt{\sum_{m=0}^{M-1} |g_m|^2} \quad (8)$$

and considering the components with the same delay in the same summation term, it can be expanded to:

$$r'(k) = \sqrt{\frac{E_c}{2}} \sum_{\lambda=-M+1}^{L-1} d(k-\lambda) c(k-\lambda) \sum_{\substack{l,m \\ l-m=\lambda}} g_m^* g_l + n(k) \sqrt{\sum_{m=0}^{M-1} |g_m|^2} \quad (9)$$

After the Post-Rake processing we have:

$$y(k) = \sum_{\mu=-M+1}^{L-1} \left(\sum_{\substack{l,m \\ l-m=\mu}} g_m^* g_l \right) r'(k+\mu) \quad (10)$$

The signal obtained from (10) is then correlated with the scrambling code in order to obtain the decision variable:

$$z_{pRp} = K \sqrt{2E_c} \sum_{\mu=-M+1}^{L-1} \left| \sum_{\substack{l,m \\ l-m=\mu}} g_m^* g_l \right|^2 d + n_{pRp} + n_{self} \quad (11)$$

where n_{pRp} is the Gaussian noise component whose variance is given by:

$$\text{var}(n_{pRp}) = KN_0 \left(\sum_{m=0}^{M-1} |g_m|^2 \right) \sum_{\mu=-M+1}^{L-1} \left| \sum_{\substack{l,m \\ l-m=\mu}} g_m^* g_l \right|^2 \quad (12)$$

and n_{self} is the component referring to the self-interference that will be ignored for high processing gain values. The bit error rate is a function of the signal-to-noise ratio (SNR). For a system employing both a Pre-Rake and a Post-Rake, the SNR according to the (11) and (12) is given by:

$$SNR_{pre-post-Rake} = \sqrt{\frac{2KE_c \sum_{\mu=-M+1}^{L-1} \left| \sum_{\substack{l,m \\ l-m=\mu}} g_{l,m}' g_l \right|^2}{N_0 \sum_{m=0}^{M-1} |g_m'|^2}} \quad (13)$$

For our analysis, we have considered the same 6-path Rayleigh-faded channel model as in [6], with exponential path decay. The obtained results are presented in Fig. 2, with the BER on the vertical axis and the ratio E_b/N_0 on the horizontal one, where E_b is the bit energy. The results, obtained through the theoretical analysis, are very similar to the ones obtained from simulation, which indicate their correctness. We can see, in this figure, that the performance of the new Pre-Post-Rake approach using phase-only filters is better than the one obtained with the conventional Pre-Post-Rake scheme. This behaviour is an obvious consequence of using less selective reception filters.

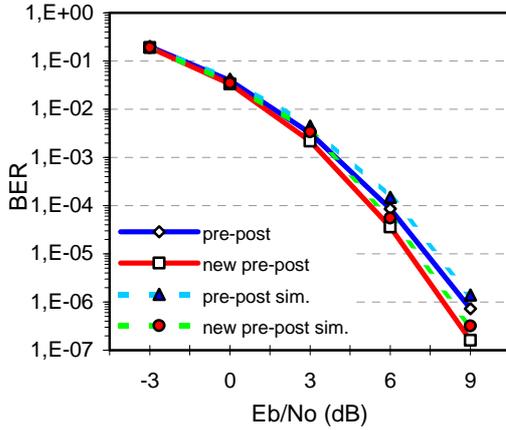


Fig. 2: Analytic and simulated performance of the Pre-Post-Rake schemes in a single-user environment

4. MULTI-USER ENVIRONMENT

In a multi-user environment, each user i has a different spreading code $c_i(k)$, transmits different coded bits $d_i(k)$, and has a different channel impulse response determined by the coefficients $g_{i,l}$ with $0 \leq l < L$. Following again the approach of [6], the transmitted signal is now the sum of the signals of all the users and it is given by:

$$s_{pR}(k) = \sum_{i=0}^{N-1} s_i(k) = \sum_{i=0}^{N-1} \sqrt{\frac{E_c}{2}} \beta_i \sum_{m=0}^{M-1} g_{i,m}' d_i(k+m) c_i(k+m) \quad (14)$$

where β_i is a normalisation factor to ensure that all the users transmit with the same chip energy E_c :

$$\beta_i = \sqrt{\frac{1}{\sum_{m=0}^{M-1} |g_{i,m}'|^2}} \quad (15)$$

and N is the number of the users. Without loss of generality, we can proceed as in the case of the single user 0 ; considering the modified received signal. As we made before, in order to simplify the analysis, the components with the same delay are considered in the same summation term, hence the signal can be expanded, using equation (14), to:

$$r_0'(k) = \sqrt{\frac{E_c}{2}} \sum_{i=0}^{N-1} \sum_{\lambda=-(L+M)-1}^{L+M-1} d_i(k-\lambda) c_i(k-\lambda) \cdot \sum_{\substack{l,m \\ l-m=\lambda}} g_{l,m}' g_{0,l} + n(k) \sqrt{\sum_{l=0}^{L-1} |g_l|^2} \quad (16)$$

The signal obtained from (16) is then processed by a Post-Rake, matched to the combination of the Pre-Rake and the channel, giving:

$$y_0(k) = \sum_{\mu=-M+1}^{L-1} \left(\sum_{\substack{l,m \\ l-m=\mu}} g_{0,m}'^* g_{0,l}' \right) r_0'(k+\mu) \quad (17)$$

which can be expanded as:

$$y_0(k) = \sqrt{\frac{E_c}{2}} d_0 \sum_{\lambda=-M+1}^{L-1} \left| \sum_{\substack{l,m \\ l-m=\mu}} g_{0,m}' g_{0,l}' \right|^2 + \sqrt{\frac{E_c}{2}} \sum_{\lambda=-M+1}^{L-1} \sum_{\substack{\mu=-M+1 \\ \lambda \neq \mu}}^{L-1} a_{0,\lambda-\mu}(k) \gamma_0'(\lambda) \gamma_0^*(\mu) + \sqrt{\frac{E_c}{2}} \sum_{i=1}^{N-1} \frac{\beta_0^2}{\beta_i^2} \sum_{\lambda=-M+1}^{L-1} \sum_{\mu=-M+1}^{L-1} a_{i,\lambda-\mu}(k) \gamma_i'(\lambda) \gamma_i^*(\mu) + \sqrt{\sum_{m=0}^{M-1} |g_{0,m}'|^2} \sum_{\lambda=-M+1}^{L-1} n(k-\lambda) c(k) \gamma_0'(\mu) \quad (18)$$

where:

$$\gamma_i'(\mu) = \sum_{\substack{l,m \\ l-m=\mu}} g_{i,m}' g_{0,l}' \quad (19)$$

The first term in (18) is the desired signal, the second term represents the self-interference, the third corresponds to the multi-user interference and the fourth to the white Gaussian noise. The decision variable plus the multi-user interference is now:

$$z_{pRp} = K \sqrt{2E_c} \sum_{\mu=-M+1}^{L-1} \left| \sum_{\substack{l,m \\ l-m=\mu}} g_{0,m}' g_{0,l}' \right|^2 d + n_{self} + n_{pRp} + n_{pRp} \quad (20)$$

The variance of the Gaussian noise is already given in (12), while the variances of the interference components are:

$$\text{var}(n_{self, pRp}) = KE_c \sum_{\substack{p=-2(M-1) \\ p \neq 0}}^{2(L-1)} \left| \sum_{\substack{\lambda, \mu \\ \lambda - \mu = p}} \gamma'_o(\lambda) \gamma_o^*(\mu) \right|^2 \quad (21)$$

$$\begin{aligned} \text{var}(n_{mu, pRp}) = & KE_c \sum_{i=1}^{N-1} \left(\beta_i \left| \sum_{\substack{\lambda, \mu \\ \lambda - \mu = p}} \gamma'_i(\lambda) \gamma_o^*(\mu) \right|^2 \alpha_i + \right. \\ & \left. + KE_c \sum_{i=1}^{N-1} \beta_i \sum_{\substack{p=-2(M-1) \\ p \neq 0}}^{2(L-1)} \left| \sum_{\substack{\lambda, \mu \\ \lambda - \mu = p}} \gamma'_i(\lambda) \gamma_o^*(\mu) \right|^2 \right) \end{aligned} \quad (22)$$

We can also define the interference power spectral density given by:

$$I_{0, pRp} = \frac{\text{var}(n_{self, pRp}) + \text{var}(n_{mu, pRp})}{K \sum_{m=0}^{M-1} |g_{0,m}|^2 \sum_{\substack{\mu=-M+1 \\ l,m \\ l-m=\mu}}^{L-1} \left| \sum_{l,m} g_{0,m} g_{0,l} \right|^2} \quad (23)$$

As done before for the single user environment, we can use the above equations to calculate the BER. Once again, the analytical results match the simulation results. The obtained outcomes evidence better performance of the new scheme versus the conventional Pre-Post-Rake. This behaviour, illustrated in figure 3, reflects in a consistent reduction of the multi-user interference due to the phase-only filter used in the processing. Hence, it is exactly in such situations, i.e. multi-cellular systems, that the Pre-Post-Rake approach outperforms the other techniques. In these configurations, a significant part of the interference is due to the surrounding cells and, since the signals from different cells are not synchronised on chip level, the other-cell interference can be modelled as white Gaussian noise.

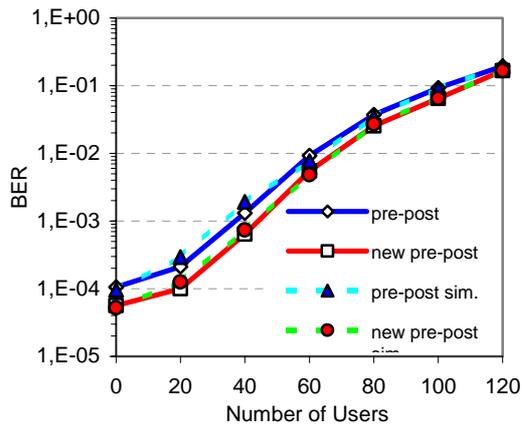


Fig. 3: Analytic and simulated BER of the two Pre-Post-Rake schemes in a multi-user environment ($E_b/N_0 = 6\text{dB}$)

5. CONCLUSIONS

In this contribution, we have proposed a new scheme for DS/SS mobile communications based on the Pre-Post-Rake approach involving phase-only filters. We have verified that the proposed technique offers better performance, in relation to the conventional solutions, in terms of BER, reducing the interference in both single user and multi-user environments. Hence, such a promising technique is able to turn out a sensible capacity increase in the downlink with little additional complexity. Future investigations are to be carried out to parametrically quantify the gain and analyse the reasons of such a trend from a theoretical viewpoint.

REFERENCES

- [1] Esmailzadeh R., Nakagawa M., "Pre-Rake Diversity Combination for Direct Sequence Spread Spectrum Mobile Communications Systems", *IEICE Trans. on Comm.*, Aug. 1993, pp. 1008-1015.
- [2] Choi R.L., Letaief K.B., Murch R.D., "Transmit Diversity Combining for Wideband DS/CDMA in Wireless Communication Systems", *Proc. ISCC 2000 Fifth IEEE Symposium on Computers and Communications*, 3-6 July 2000, pp. 717-722.
- [3] Choi R.L., Letaief K.B., Murch R.D., "CDMA Pre-Rake Diversity System with Base Station Transmit Diversity", *Proc. VTC2000-Spring*, Tokyo, Japan, May 2000, pp. 951-955.
- [4] Choi R.L., Letaief K.B., Murch R.D., "MISO CDMA Transmission with Simplified Receiver for Wireless Communication Handsets", *IEEE Transactions on Communications*, vol. 49, no. 5, May 2001, pp. 888-898.
- [5] Choi R.L., Murch R.D., "Evaluation of a Pre-Rake Smart Antenna System for TDD CDMA Systems", *Vehicular Technology Conference VTC 2001 Fall. IEEE VTS 54th*, Vol. 3, 7-11 Oct. 2001, pp.1543-1547.
- [6] Barreto A.N., Fettweis G., "Performance improvement in DS-Spread Spectrum CDMA System using a PRE-and a POST-RAKE", *Proc. International Zurich Seminar on Broadband Communications*, 15-17 Feb. 2000, pp. 39-46.
- [7] Kadous T.A., Sourour E.E., El-Khamy S.E., "Comparison Between Various Diversity Techniques of the Pre-Rake Combining System in TDD/CDMA", *Vehicular Technology Conference*, 4-7 May 1997 IEEE 47th, Vol: 3, pp.2210-2214.
- [8] Esmailzadeh R., Sourour E., Nakagawa M., "Pre-Rake Diversity Combination in Time-Division Duplex CDMA Mobile Communications", *IEEE Trans. on Vehicular Technology*, vol. 48, no. 3, May 1999, pp. 795-801.

QoS-based Call Admission Control using Multi-code for Voice/Data Traffic in CDMA Networks

Fariborz Mahmoudi, Mahmoud Ferdowsizadeh-Naeni, Saleh Yousefi, Hosein Pedram
TMN Dept., Iran Telecommunication Research Center, North Kargar Ave., Tehran, Iran
{mahmoudi, mferdosi, syousefi, pedram}@itrc.ac.ir

Abstract

In traditional mixed voice/data traffic call admission control algorithms, a single code assignment policy is considered for data users' packets transmission. When voice traffic of the network is low this strategy wastes the capacity of the network. For removing this drawback, in this paper we propose a multi-code call admission control for voice/data traffic in CDMA networks. The main motivation of this CAC algorithm is to use the capacity of network for data transmission efficiently. For better evaluation with considering some QoS parameters, the superiority of proposed CAC algorithm is shown with compare of a single-code CAC.

1. Introduction

Next generations mobile networks will be required to accommodate various services such as voice, data, and multimedia calls. Each of these services has different specific QoS requirement parameters and also different resource requirements. One of the most critical resources of the network is radio link resource. In the third generation mobile networks CDMA is basic technique used for radio link and multiple accesses. Because of the soft capacity characteristic of the CDMA systems, radio resource allocation is more efficient and also more difficult than for the FDMA and TDMA based systems. However the main problem in a multi service network is how many users of different services must be accepted in each period, in order to network throughput be maximized and users of different services experience a suitable QoS. Many researchers have been answering this question in form of the different call admission control (CAC) algorithms.

In some services such as non-real-time data the variable bit rate (VBR) allocation techniques can be used in the CAC algorithms in order to increase the network throughput. There are two approaches in the VBR-based CAC algorithms. The first one is based on using multiple chip rates in the multiple RF channels [5-10]. Because of the calculation complexity, there are a few papers published with

this approach. The second approach is based on using one or more DS-CDMA codes in single RF channel. In this way, there are many works with single code assignment techniques for CAC algorithms [2-4]. As Ayyagari and Ephremides showed multi code allocation techniques has more network capacity to manage data traffic, in comparison of single code techniques [1]. Then it is expected that a multi-code CAC algorithm will provide better QoS specifications.

Main contribution of this paper is presentation of a CAC algorithm which named Maximum Packet Transmission (MPT) for a multi-code CDMA system. In this algorithm the number of codes per data user is chosen so that the numbers of transmitted packets in each time slot become maximized. Our experimental results show superiority of this algorithm against traditional single code (SC) CAC algorithm.

This paper is organized as follow: In section 2 the system model for call generation of voice and data users and the mobility modeling of voice users is presented. Considered CDMA capacity model is introduced in section 3. The proposed single and multi code CAC algorithms are introduced in section 4. Section 5 presents the experimental results for evaluating the CACs, and finally the paper is concluded in section 6.

2. Voice and Data Traffic Model

We will use the model introduced in [11, 12] for voice and data call arrival traffic. In this section specification of the models are explained.

a) *Voice traffic model:* In this model the call arrival process to cell i is assumed to be a Poisson process with rate λ_{vi} and the call dwell time has the exponential distribution with mean $1/\mu$. At the end of a dwell time a call may stays in the same cell with probability of q_{ii} , or moves to the adjacent cell j with probability of q_{ij} , or leaves the network with probability q_i . If a call in progress in cell i remains in cell i after completing its dwell time, a new dwell time independent of the previous dwell time begins immediately. The network traffic is assumed to be

Table1. The mobility model parameters

$\ A_i\ $	q_{ij}	q_{ii}	q_i
3	0.1	0	0.7
4	0.075	0	0.7
5	0.060	0	0.7
6	0.050	0	0.7

* $\|A_i\|$ is the number of cells adjacent to cell i .
 * q_{ij} is handoff probability from cell i to j .
 * q_{ii} is the probability of continuing a call in cell i .
 * q_i is the probability of a call in cell i leaves the network.

homogenous (λ_{vi} s are equal for all of the cells). The parameters of the mobility model used in this paper are shown in table 1.

b) Data traffic model: Data traffic is assumed as burst data traffic. The generation of the messages for this data traffic is assumed to be a Poisson process with mean rate λ_d message per second. The probability density function of the message size y is defined by *k-Erlang* distribution [13],

$$P(y) = \frac{(k/L_b)^k}{(k-1)!} x^{k-1} \exp\left(-\frac{kx}{L_b}\right) \quad (1)$$

Where L_b is the average message size and k is assumed to be 6.

3. CDMA capacity model

The capacity of CDMA system is limited by interference. When a new call is accepted the interference level will be increased, but the effect of acceptance of a voice call and a data call on the interference level is not the same. On the other hand, it is dependent on the activity factor and the transmission power.

Ayyagari and Ephremides introduced a capacity model based on the number of CDMA code assignment to data users [1]. In this paper we consider their model for network capacity. Fig 1 shows the capacity diagram of networks which calculated based on their model for the single and double code data users. The diagram shows the relation between maximum numbers of acceptable data users N_d , versus the number of accepted voice users N_v .

If the number of accepted voice and data users in the cell i denoted as N_{vi} and N_{di} respectively, then the point (N_{vi}, N_{di}) which called *state point* must be below the capacity diagram, otherwise the interference level exceeds the desired level.

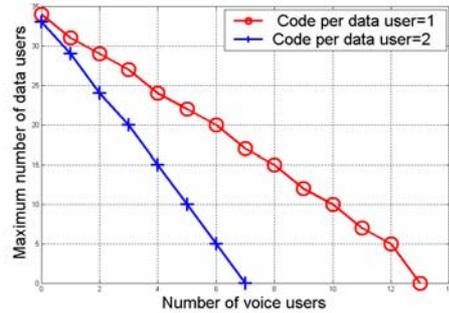


Fig 1. Capacity diagram of single/double code data users

4. Proposed CAC algorithms

This section introduces the multi code CAC algorithm (MPT) for improving the network performance. For better evaluation of this algorithm we also developed a traditional single code (SC) CAC algorithm. For better presentation of MPT superiority the parameters of two algorithms are similar, except the number of code assignment for data users. The following parts present details of these algorithms

4.1. Single-code CAC

The proposed single code (SC) call admission control algorithm is developed based on the capacity model presented with single code curve in fig 1. The main rules of this algorithm are as follow:

1- As we see in fig 1 the maximum number of acceptable voice users is 13, thus a new voice or handoff call will be accepted if the number of existing voice users is less than 13.

2- When a new data call request is received it will be queued if the buffer is not full, otherwise it will be dropped.

3-After each following events the CAC Algorithm investigates the possibility of the admission of blocked data users:

- a) The transmission of packets of a data user is finished.
- b) A handoff is occurred.
- c) A voice call session is completed and the user leaves the network.
- d) A voice user is accepted.
- e) A data user is entered to the buffer.

When one of these events is occurred the number of active data users is chosen so that the state point remains under the capacity diagram of the single code data users in fig 1.

4.2. MPT CAC

The main idea of the maximum packet transmission (MPT) algorithm is double code assignment for the data users to transmit their packets with higher speed, when the voice users' traffic is low. By this modification against SC CAC, the buffered request will be serviced earlier and thus the data user dropping probability will be decreased. In this case we have two capacity diagram; one for the maximum number of double code data users versus the number of voice users, and other for the single code data users.

The rules of this algorithm are as for the SC except that when one of the events mentioned in part 3 of section 4.1 is occurred, the possibility of changing the speed of data users also will be investigated.

At first we define two Packet Transmission Capacity function PTC for single and double code case as follow:

$$PTC_1(N_v) = \min\{N_d^1, N_d^{active} + N_d^{inactive}\} \quad (2)$$

$$PTC_2(N_v) = 2 \times \min\{N_d^2, N_d^{active} + N_d^{inactive}\} \quad (3)$$

where N_d^1 and N_d^2 are the maximum number of single and double code data users while the number of existing voice users is N_v . N_d^{active} and $N_d^{inactive}$ are numbers of active and inactive data users existing in buffer. In fact these functions shows number of transmittable packets in each time slot while all of data users are single or double code.

This CAC algorithm assumes the number of voice users is determined in the beginning of each time slot. Then, the packet transmission capacity functions are calculated and the suitable data transmission speed and the number of active and inactive data users will be computed by algorithm shown in fig 2.

if ($PTC_1(N_v) \geq PTC_2(N_v)$)
Codes per data user = 1
 $N_d^{active} = \min\{N_d^1, N_d^{active} + N_d^{inactive}\}$

else
Codes per data user = 2
 $N_d^{active} = \min\{N_d^2, N_d^{active} + N_d^{inactive}\}$

Fig 2. algorithm for determining the number and speed of active data user

5. Experimental Results

For evaluating the performance of two mentioned CAC algorithm we compared the CACs in a simulation test-bed under unique conditions. In the rest of this section, we describe specifications of

Table 2. Simulation parameters

CDMA channel bandwidth	1.228 MHz
Basic transmission rate of a single code	9.6 kbit/sec
processing gain	128
Max. path loss in cell	-130dB
Max. voice handset transmission power	200 mw
Frame Error Rate (FER) for voice calls	1%
Service Availability Probability (SAP) for data users	90%
Low bound of SIR for voice/data users	7 dB
Voice activity factor	3/8
The mean dwell time of the voice calls	1.5 minutes
Phase index of k -Erlange distribution	6
Average message size	250 packets
Packet transmission time duration	20ms
Buffer length	35

simulation test-bed, traffic conditions which are considered and results of simulation.

a) *Simulation test-bed*: In this test-bed we considered a conventional cellular system like IS-95 with 27 cells. To ensure uniform coverage for all users in this network, the required received power of the voice users is set at -106.98dB. Table 2 shows other assumptions of the simulation test-bed. Under these assumptions the capacity diagrams of fig 1 have been obtained for the single and double code data users.

b) *Traffic conditions*: To determine the traffic conditions where MPT algorithm has better performance than the SC one, we defined a Packet Transmission Gain as follow:

$$PTG(N_v) = \frac{\max\{PTC_2\}}{\max\{PTC_1\}} = \frac{2N_d^2}{N_d^1} \quad (4)$$

In fact PTG is the maximum ratio of the packet transmitted per time slot, for the MPT and SC call admission control algorithms.

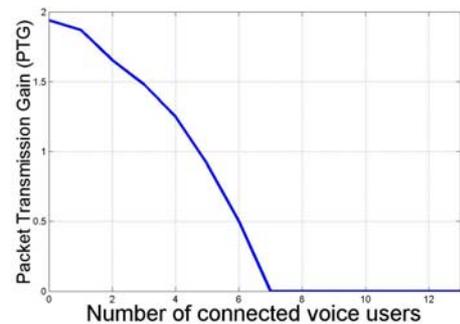


Fig 3. Packet Transmission Gain versus the number of connected voice users

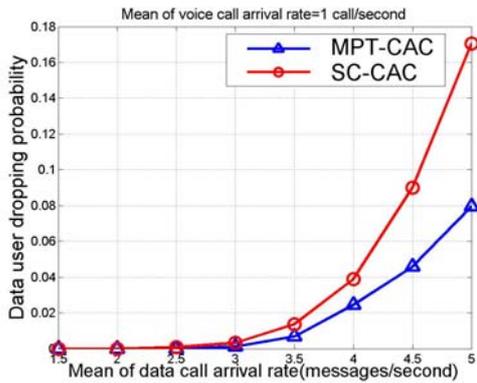


Fig 4. dropping probability versus data traffic

In fig 3 the *PTG* versus the number of connected voice users has been plotted. This diagram shows, when the number of connected voice users is less than 4.2 the *PTG* is greater than unit. Thus the MPT algorithm achieve higher performance than SC one for high data traffic. Our experimental results show for existing 4.2 voice users per cell, the mean call arrival rate is 1.6 calls/minute. We conclude when the mean rate of voice call arrival is less than 1.6 calls/minute then the MPT has a better performance than the SC with respect to the data request dropping probability.

c) *Simulation results:* Fig 4 shows data request dropping probability versus the data request mean arrival rate when the mean arrival rate of voice call requests is 1 call/minute. This diagram presents superiority of MPT CAC against SC CAC in reduction of data users' dropping probability.

We also define a network throughput function which is the average of the sum of connected voice users and number of transmitted data packets in each slot. Fig 5 shows for high data traffic the network throughput using the MPT CAC is higher than when we are using the simple SC CAC. This figure is also in voice traffic of 1call/minute.

6. Conclusions

In this paper we proposed a multi code call admission control called MPT for mixed voice and data traffic in a CDMA network. In this algorithm the speed of data users is determined so that the number of transmitted packets in each slot becomes maximized. Simulations results show that the data user dropping probability and network throughput of this algorithm is better than the simple Single Code (SC) algorithm in high data traffic.

References

[1] Deepak Ayyagari and Anthony Ephremides, "Cellular multi code CDMA capacity for integrated (voice and Data) services", IEEE journal on selected areas in communication, VOL.17, NO.5, May 1999.

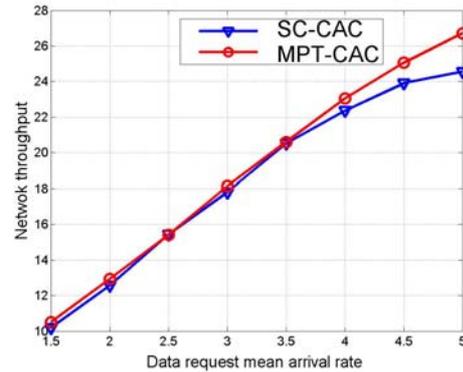


Fig 5. Network throughput versus data traffic

[2]- S. Oh and K.M. Wasserman, "Integration of Voice and Data Traffic in CDMA Networks Using Dynamic Spreading Gain Control," Proceedings of the IEEE ICC'98, Atlanta, June 7-11, 1998.

[3]C-L.I and K.K.Sabnani, "Variable spreading gain CDMA with adaptive control for true packet switching wireless network", Proc. IEEE ICC. pp. 725-730, 1995.

[4] C-L.I and R.D.Gitlin. "Multi-Code CDMA wireless personal communication network", Proc. IEEE ICC, pp.1060-1064, June 1995.

[5]P.G.Andermo and L.M.Ewerbring, "A CDMA-based radio access design for UMTS", IEEE personal Commun., vol.2, no.1, pp.48-53, Feb.1995.

[6] A.Baier, U.C.Fiebig, W.Granzow, W.Koch, P.Teder, and J.Thielecke, "Design study for a CDMA-based third generation mobile radio system," IEEE J.Sel.Areas Com., vol.12, no.4, pp.733-743, May 1994.

[7] S.C.Bang, "ETRI IMT 2000 multiband CDMA system for terrestrial environments," Mobile Comm. Research Div., ETRI, Taejon, Korea, 1997.

[8] T.H.Wu, "Issues in multimedia packet CDMA personal communication networks," Ph.D Thesis, Dept. of Elec. Eng., Univ. of Maryland, College Park, 1994.

[9] D. Lyu, I.Song, Y. Han, and H.M. Kim, "Analysis of the performance of an asynchronous multiple-chip-rate DS/CDMA system", Int. J.Electron. Commun., vol.51, issue 4, pp.213-218, July 1997.

[10] Y.W.Kim, S.J.Lee, M.Y.Chung, J.H.Kim, D.K.Sung, "Radio Resource Assignment in Multi-Chip Rate DS/CDMA Systems Supporting Multimedia Services," IEICE Trans. on Communications, Vol. J82-B, No. 1, pp.145-155, 1999.

[11] R.G. Akl, M.V. Hegde, M. Naraghi-Pour, P.S. Min, "Call Admission Control Scheme for Arbitrary Traffic Distribution in CDMA Cellular Systems," IEEE WCNC, vol. 1, pp. 465-470, September 2000.

[12]. Robert G. Akl., Manju V. Hegde, , Mort Naraghi-Pour, and Paul S. Min. " Multicell CDMA Network Design "IEEE Transactions on Vehicular Technology, VOL. 50, NO. 3, MAY 2001.

[13]L. Kleinrock, Queuing System Theory: Volume I. New York: John Wiley & Sons, 1976.

ROBUST BLIND CHANNEL ESTIMATION FOR SPACE-TIME CODED SYSTEMS

Stylianos Manioudakis, Wai Lok Woo and Bayan Sharif
 School of Electrical, Electronic and Computer Engineering
 University of Newcastle upon Tyne
 Newcastle upon Tyne
 United Kingdom

Email: Stylianos.Manioudakis@ncl.ac.uk, W.L.Woo@ncl.ac.uk, Bayan.Sharif@ncl.ac.uk

ABSTRACT

Space-time coding is an effective approach to increase the data rate and capacity of a wireless communication system that employs multiple transmit and multiple receive antennas. It involves coding techniques that are designed for multiple transmit antennas. This paper explores the class of blind channel estimation algorithms (no use of training sequences) that exploit the constant modulus (CM) of the transmitted space-time codes (STCs). A robust scheme that gives good performance in Gaussian and non-Gaussian noise environments is proposed. Computer simulation results in both single-user and multiuser cases show the improvement in error performance.

1. INTRODUCTION

Towards the search for bandwidth efficient and reliable wireless networks, transmit antenna diversity has been exploited recently. By the use of STCs, a dramatic increase in quality and data rates is evident [1]-[2]. The use of simple maximum-likelihood decoders, based on channel estimates via training symbols, results in a reasonably low overall system complexity. However, in GSM systems for example, 20% of the bandwidth is devoted to training sequences, a fact that reduces GSM bandwidth efficiency. On the other hand, even by employing training, the assumption of perfect channel estimates at the receiver is usually violated in the vast majority of real life applications.

In this paper, blind channel estimation algorithms are employed at the receiver assuming no channel state information. By the term "blind" here, it is implied that very little assumptions are made on the transmitted signals [3]. *Notation:* Matrices (vectors) are denoted by boldface upper (lower) case letters, the superscripts \wedge , $*$, denote estimation and complex conjugate transposition respectively and $\mathbf{0}$ denotes a matrix of zeros.

The rest of this paper is organized as follows. In section 2, the system model is explored and a mathematical analysis is given. Computer simulation results and discussions are given in section 3. Concluding remarks are given in section 4.

2. SYSTEM MODEL

Consider a K user model where each user employs T_x transmit antennas. At the receiver, R_x antennas are used and the receive array matrix \mathbf{X} is formed as

$$\mathbf{X} = \mathbf{H}\mathbf{S} + \mathbf{N} \quad (1)$$

Throughout this paper, \mathbf{X} is the $R_x \times N$ receive array matrix where N is the number of snapshots. \mathbf{H} is the $R_x \times KT_x$ channel matrix of full rank where $R_x \geq KT_x$. \mathbf{S} is the $KT_x \times N$ source matrix, and \mathbf{N} is the noise matrix consisting of independent and identically distributed (i.i.d.) complex random variables. A frequency-flat and quasi-static Rayleigh fading channel is assumed. User data as well as channel information are unknown to the receiver. The only available information is \mathbf{X} . The receiver's task is to perform the required signal processing on \mathbf{X} resulting in a $KT_x \times N$ matrix \mathbf{Y} , as shown in fig. 1, from which the user data can be obtained.

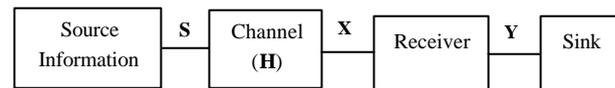


Fig. 1. The system model

The algorithm that is proposed in [4] uses an analytical method for solving the blind CM problem called the analytical CM algorithm (ACMA). It is a deterministic method that encapsulates all the conditions on the weight vectors into a single matrix and finds all independent vectors in the kernel of this matrix that have a Kronecker product structure. It can achieve good performance when recovering signals in noiseless channels. In a noisy environment, however, phenomena like coefficient biases and the existence of noise in the recovered signals reduce the effectiveness of the algorithm substantially.

In this paper, the retrieval of independent signals in noisy channels is viewed as an ill-posed problem [5]. Once \mathbf{H} is estimated by ACMA, the following rules may be employed. Let $\mathbf{y}(t)$ and $\mathbf{x}(t)$ represent the t^{th} column of \mathbf{Y} and \mathbf{X} respectively for $t = 1, \dots, N$. The cost function proposed is

$$J = \min_{\hat{\mathbf{y}}(t+1)} \left\{ \left\| \mathbf{x}(t) - \mathbf{H} \mathbf{y}(t+1) \right\|_2^2 + \right. \\ \left. + (\hat{\mathbf{y}}(t+1) - \hat{\mathbf{y}}(t))^* \mathbf{O} (\hat{\mathbf{y}}(t+1) - \hat{\mathbf{y}}(t)) \right\} \quad (2)$$

$$\mathbf{O}(t) = \mathbf{s}_n^2 \mathbf{S}^{-1}(t) \quad (8)$$

where \mathbf{s}_n^2 is the noise variance and \mathbf{S} is the error covariance matrix of $\hat{\mathbf{y}}(t)$ estimated adaptively as

$$\mathbf{S}(t+1) = \mathbf{S}(t) - \mathbf{h} \left\{ \mathbf{O}(t) + \mathbf{H}^* \mathbf{H} \right\}^{-1} \mathbf{H}^* \mathbf{H} \mathbf{S}(t) \quad (9)$$

Matrix \mathbf{O} is a $KT_x \times KT_x$ symmetric non-negative definite weighting matrix and $\|\cdot\|_2$ denotes the Euclidean norm.

The term on the top in (2) can be considered as a standard least squares (LS) optimization while the bottom term is the regularizing term. The latter is an auxiliary function that imposes a priori constraints on the blind recovery problem. This converts the ill-posed problem into a well-posed problem, a conversion that is a key step for robust performance [5]. Substituting for $\hat{\mathbf{y}}(t+1) = \hat{\mathbf{y}}(t)$ in (2) leads to the standard LS optimization. Thus, the regularizing term can be viewed as a smoothing operator acting on the estimated noisy time series $\hat{\mathbf{y}}(t)$ and the solution to the resulting LS optimization is proved in appendix A to be given by

$$\bar{\hat{\mathbf{y}}}(t) = \left\{ \mathbf{O} + \mathbf{H}^* \mathbf{H} \right\}^{-1} \mathbf{H}^* \bar{\mathbf{x}}(t) \quad (3)$$

Substituting (3) into its original coordinates and introducing a step size for tracking and stability we get

$$\hat{\mathbf{y}}(t+1) = \hat{\mathbf{y}}(t) + \mathbf{h} \left\{ \mathbf{O} + \mathbf{H}^* \mathbf{H} \right\}^{-1} \mathbf{H}^* (\mathbf{x}(t) - \mathbf{H} \hat{\mathbf{y}}(t)) \quad (4)$$

A further generalization can be deduced by

$$\hat{\mathbf{y}}(t+1) = \hat{\mathbf{y}}(t) + \mathbf{h} \left\{ \mathbf{O} + (\mathbf{H}(t))^* \mathbf{H}(t) \right\}^{-1} (\mathbf{H}(t))^* \mathbf{y}(\mathbf{x}(t) - \mathbf{H}(t) \hat{\mathbf{y}}(t)) \quad (5)$$

where $\mathbf{y}(\mathbf{g}(t))$ is the activation function of the noise distribution of $\mathbf{g}(t) = \mathbf{x}(t) - \mathbf{H}(t) \hat{\mathbf{y}}(t)$ and $\mathbf{H}(t)$ denotes the $R_x \times KT_x$ channel matrix at time instant t . The proposed activation function is defined as

$$\mathbf{y}(\mathbf{g}(t)) = \begin{bmatrix} \mathbf{y}(g_1(t)) \\ \vdots \\ \mathbf{y}(g_{R_x}(t)) \end{bmatrix} = \begin{bmatrix} -\frac{d}{d\{g_1(t)\}} \log p_n(g_1(t)) \\ \vdots \\ -\frac{d}{d\{g_{R_x}(t)\}} \log p_n(g_{R_x}(t)) \end{bmatrix} \quad (6)$$

The term $p_n(g_i(t))$ is the noise distribution perturbing the receive antenna i for $i = 1, \dots, R_x$. The step size or learning rate parameter \mathbf{h} is defined as

$$\mathbf{h} = \frac{1}{\|\hat{\mathbf{y}}(t)\|_2^2} \quad (7)$$

The weighting matrix \mathbf{O} can be estimated adaptively as

The robustness of the proposed framework can be realized by the fact that it is able to effectively ameliorate a class of frequently encountered noise called non-Gaussian noise. Non-Gaussian noise can be either natural or man-made and it can be encountered in applications like mobile radio communications, radar systems, acoustic channels etc. If not properly modeled and accounted for, it can dramatically degrade the performance of wireless systems [6]. This phenomenon is manifested by the existence of irregular noise spikes of short duration with randomly varying (and relatively high) magnitudes. The probability density function (pdf) of the non-Gaussian noise considered here is given by

$$f = (1 - \mathbf{e}) N(0, \mathbf{s}_1^2) + \mathbf{e} N(0, \mathbf{s}_2^2) \quad (10)$$

where $\mathbf{s}_1, \mathbf{s}_2 > 0$, $0 \leq \mathbf{e} \leq 1$, $\mathbf{s}_2^2 = k \mathbf{s}_1^2$ and $k \geq 1$. The $N(0, \mathbf{s}_1^2)$ term represents the nominal background noise, and the $N(0, \mathbf{s}_2^2)$ term represents the impulsive component with \mathbf{e} representing the probability that impulses occur. This is a two-term Gaussian mixture model for the additive noise samples. The noise variance is defined in [6] to be given by

$$\frac{\mathbf{s}_n^2}{2} = (1 - \mathbf{e}) \mathbf{s}_1^2 + \mathbf{e} \mathbf{s}_2^2 \quad (11)$$

In the non-Gaussian noise case, the activation function for receive antenna i is proved in appendix B to be given by

$$\mathbf{y}(g_i(t)) = \frac{g_i(t)}{\alpha(g_i(t)) + b(g_i(t))} \left\{ \frac{1}{\mathbf{s}_1^2} \alpha(g_i(t)) + \frac{1}{\mathbf{s}_2^2} b(g_i(t)) \right\} \quad (12)$$

To conclude with the regularized framework for noise cancellation in space-time coded systems, the framework can be summarized by the adaptation rules in (5) through (12). By using (5) in ACMA the algorithm is actually made more robust and efficient. Robust, since it can work satisfactorily in both Gaussian and non-Gaussian noise environments. Efficient since it achieves substantially better performance with respect to the original (non-regularized) realization in [4] when applied in space-time coded systems.

3. SIMULATION RESULTS

For all simulations, $T_x = 2$ antennas are assumed per user employing the full rate STC [2]. Blind identifiability of this STC may be achieved via appropriate code modifications [7]. QPSK modulation is used although the generalization to other modulation schemes is straightforward. In the single-user case $R_x = 2$ antennas exist at the receiver. In the multiuser case two synchronous equal power users are assumed, $R_x = 4$ and the user of interest's curve is shown. For a fair comparison the ST-GOFDM scheme in [8] is also simulated for $T_x = R_x = 2$. Furthermore, no multiuser generalizations of the ST-GOFDM are given in [8] and hence it is simulated here only for the single-user case.

3.1. Gaussian noise environment

In the case of Gaussian noise shown in fig. 2 the proposed scheme has approximately the same performance as the ST-GOFDM scheme at low signal to noise ratio (SNR). However, at SNR higher than 8 dB the proposed scheme outperforms the ST-GOFDM scheme. This is due to the improved convergence of ACMA in the LS sense, which results from the robustness to noise of the proposed regularizations. Although ST-GOFDM can cope with multipath propagation impairments, it is generally not robust in frequency-flat environments. At a bit error rate (BER) of 10^{-3} the proposed scheme outperforms the ST-GOFDM scheme by 1.2 dB and the non-regularized scheme by 4 dB. In the non-regularized case, noise prevents ACMA to obtain good convergence. In the multiuser case, convergence is an even more challenging task due to the interfering user that reduces the number of subchannels that actively participate in conveying the user of interest's information. At a BER of 10^{-3} the proposed scheme outperforms the non-regularized scheme by more than 2 dB.

3.2. Non-Gaussian noise environment

In the case of non-Gaussian noise shown in fig. 3 the proposed scheme's performance superiority becomes even more pronounced in both single user and multiuser cases. The impulsiveness parameter is set to $\epsilon = 0.01$ and $k = 100$. The noise distribution is given by (11). At a BER of 10^{-3} , in the single-user case, the proposed scheme outperforms the ST-GOFDM scheme by 4 dB and the non-regularized scheme by more than 6 dB. In the multiuser case the non-regularized scheme gives very poor performance mainly due to the fact that the system is designed based on the Gaussian noise assumption. At a BER of 10^{-2} the proposed multiuser scheme outperforms the non-regularized scheme by 6 dB and the gap increases as we further increase the SNR. Due to the impulsive nature of noise all schemes based on the Gaussian assumption perform worse than the proposed scheme. This phenomenon occurs since the proposed regularizations allow the use of the activation

function in (12) and hence impulsiveness can be substantially ameliorated.

4. CONCLUSIONS

In this paper, the blind channel estimation of space-time coded systems is exploited. Inspired by the philosophy of solving ill-posed problems a regularized scheme is proposed for boosting the performance of existing algebraic blind channel estimation algorithms. By using appropriate activation functions the scheme gives good performance in both Gaussian and non-Gaussian noise environments and it has been explored and compared with other schemes. Robustness is evident in both single-user and multiuser cases via computer simulation results in frequency-flat and quasi-static Rayleigh fading channels.

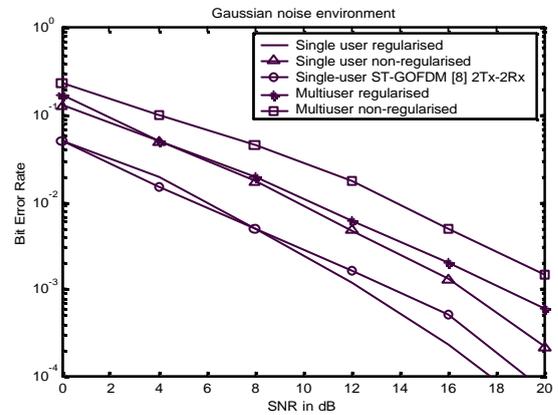


Fig. 2. Performance of the proposed scheme in Gaussian noise

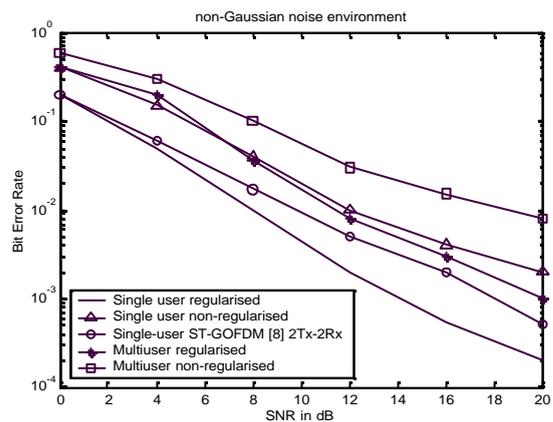


Fig. 3. Performance of the proposed scheme in non-Gaussian noise for $k = 100$ and $\epsilon = 0.01$

5. APPENDIX A

Substituting $\bar{\mathbf{y}}(t) = \hat{\mathbf{y}}(t+1) - \hat{\mathbf{y}}(t)$ and $\bar{\mathbf{x}}(t) = \mathbf{x}(t) - \mathbf{H}\hat{\mathbf{y}}(t)$ into (2), it follows that

$$J = \min_{\hat{\mathbf{y}}(t)} \left\{ \left\| \bar{\mathbf{x}}(t) - \mathbf{H}\bar{\mathbf{y}}(t) \right\|_2^2 + \left(\bar{\mathbf{y}}(t) \right)^* \mathbf{O} \bar{\mathbf{y}}(t) \right\} \quad (\text{A1})$$

Since \mathbf{O} is non-negative symmetric, then $\left(\bar{\mathbf{y}}(t) \right)^* \mathbf{O} \bar{\mathbf{y}}(t) = \left\| \mathbf{O}^{1/2} \bar{\mathbf{y}}(t) \right\|_2^2$. Substituting this into (A1), we arrive at

$$J = \min_{\hat{\mathbf{y}}(t)} \left\{ \left\| \bar{\mathbf{x}}(t) - \mathbf{H}\bar{\mathbf{y}}(t) \right\|_2^2 + \left\| \mathbf{O}^{1/2} \bar{\mathbf{y}}(t) \right\|_2^2 \right\}$$

$$= \min_{\hat{\mathbf{y}}(t)} \left\{ \left\| \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{x}}(t) \end{bmatrix} - \begin{bmatrix} \mathbf{O}^{1/2} \\ \mathbf{H} \end{bmatrix} \bar{\mathbf{y}}(t) \right\|_2^2 \right\} = \min_{\hat{\mathbf{y}}(t)} \left\{ \left\| \bar{\mathbf{g}}(t) \right\|_2^2 \right\} \quad (\text{A2})$$

The next step is to solve (A2) in order to derive (3).

$$\text{Let } \bar{\mathbf{?}}(t) = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{x}}(t) \end{bmatrix}, \mathbf{?} = \begin{bmatrix} \mathbf{O}^{1/2} \\ \mathbf{H} \end{bmatrix} \quad (\text{A3})$$

By substituting (A3) into (A2) it follows that

$$J = \min_{\hat{\mathbf{y}}(t)} \left\{ \left\| \bar{\mathbf{g}}(t) \right\|_2^2 \right\} = \min_{\hat{\mathbf{y}}(t)} \left\{ \left\| \bar{\mathbf{?}}(t) - \mathbf{?} \bar{\mathbf{y}}(t) \right\|_2^2 \right\} \\ = \min_{\hat{\mathbf{y}}(t)} \left\{ \left(\bar{\mathbf{?}}(t) - \mathbf{?} \bar{\mathbf{y}}(t) \right)^* \left(\bar{\mathbf{?}}(t) - \mathbf{?} \bar{\mathbf{y}}(t) \right) \right\} \quad (\text{A4})$$

Differentiating the argument of (A4) with respect to $\bar{\mathbf{y}}(t)$ and setting the result equal to zero we arrive at

$$\bar{\mathbf{y}}(t) = (\mathbf{?}^* \mathbf{?})^{-1} \mathbf{?}^* \bar{\mathbf{?}}(t) \quad (\text{A5})$$

By substituting (A3) into (A5) we get

$$\bar{\mathbf{y}}(t) = \left\{ \begin{bmatrix} \mathbf{O}^{1/2} & \mathbf{H}^* \end{bmatrix} \begin{bmatrix} \mathbf{O}^{1/2} \\ \mathbf{H} \end{bmatrix} \right\}^{-1} \begin{bmatrix} \mathbf{O}^{1/2} & \mathbf{H}^* \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{x}}(t) \end{bmatrix} \quad (\text{A6})$$

Given (A6) and via matrix multiplication, we arrive at

$$\bar{\mathbf{y}}(t) = \left\{ \mathbf{O} + \mathbf{H}^* \mathbf{H} \right\}^{-1} \mathbf{H}^* \bar{\mathbf{x}}(t) \quad (3)$$

5.2. APPENDIX B

The activation function for the receive antenna i in (12) can be derived by letting

$$a(g_i(t)) = -\frac{(1-\mathbf{e})}{\mathbf{s}_1 \sqrt{2\mathbf{p}}} e^{-\frac{(g_i(t))^2}{2\mathbf{s}_1^2}} \quad (\text{B1})$$

$$b(g_i(t)) = -\frac{\mathbf{e}}{\mathbf{s}_2 \sqrt{2\mathbf{p}}} e^{-\frac{(g_i(t))^2}{2\mathbf{s}_2^2}} \quad (\text{B2})$$

Differentiating the rule in (10) and substituting (B1) and (B2) we arrive at

$$-\frac{df}{d\{g_i(t)\}} = -\left\{ \frac{g_i(t)}{\mathbf{s}_1^2} a(g_i(t)) + \frac{g_i(t)}{\mathbf{s}_2^2} b(g_i(t)) \right\} \quad (\text{B3})$$

Using (6), (B3) and with further manipulations we get

$$\mathcal{Y}(g_i(t)) = -\frac{g_i(t)}{a(g_i(t)) + b(g_i(t))} \left\{ \frac{1}{\mathbf{s}_1^2} a(g_i(t)) + \frac{1}{\mathbf{s}_2^2} b(g_i(t)) \right\} \quad (12)$$

6. REFERENCES

- [1] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281-302, April 2003.
- [2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451-1458, Oct. 1998.
- [3] S. Manioudakis, W. L. Woo and B. S. Sharif, "Code concatenations for blind space-time coded systems," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Orlando USA, Oct. 2003.
- [4] A.-J. Van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, no. 5, pp. 1136-1155, May 1996.
- [5] A. V. Morozof, *Regularization Methods for Ill-Posed Problems*. USA: CRC Press, 1993.
- [6] X. Wang and H. V. Poor, "Robust adaptive array for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1352-1366, Oct. 1998.
- [7] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2489-2498, Oct. 2002.
- [8] Z. Liu, G. Giannakis, S. Barbarossa, A. Scaglione, "Transmit-antennae space-time block coding for generalized OFDM in the presence of unknown multipath," *IEEE J. Select. Areas Commun.*, vol. 19, no. 7, pp. 1352-1364, July 2001.

Simulating Heterogeneously Timed Networks in Network Simulator NS

Guang-Yeu Luo, Fei Xia, Ian G. Clark, Albert M. Koelmans, Alex V. Yakovlev
School of EECE, University of Newcastle upon Tyne, UK

Abstract

This paper describes the implementation of 3-slot and 4-slot Pool/Signal asynchronous communication mechanism (ACM) agents in NS (Network Simulator) for simulating hets (heterogeneously timed networks). This is the first step in the process of making it possible to employ NS to reveal the behaviours of networks containing ACMs. The rationale of including ACMs in networking is that by including these devices a “shared memory fabric” can be superimposed onto the conventional packet switching network framework to improve the direct data communications between close neighbouring processes which are not in the same timing domain. ACM models are compared with standard TCP solutions in NS simulations which demonstrate the potential usefulness of the former when blocking needs to be avoided in data communications.

1 Introduction

Motivation: Asynchronous communication mechanisms (ACMs) have been developing for several decades. They can be modelled and simulated by Petri nets and presented by algorithmic codes. The properties of 3-slot and 4-slot ACM algorithms have been thoroughly examined and proved and been applied in industry [CB97] [Sim90]. However, there is no tool to precisely simulate and monitor transportation behaviours for every single data packet in hets that make use of ACMs. This is not conducive to the wider use of ACMs in network systems.

Network Simulator: NS (version 2.26) is an object-oriented, discrete event driven network simulator that simulates a variety of IP networks. NS was developed at UC Berkley and is written in C++ and OTcl [Pro99]. It is a powerful tool for network simulation. NS is primarily designed for simulating LAN and WAN. During the development of NS, many researchers have used NS to validate algorithms and protocols for network research. NS provides support for simulation of TCP, routing, and multicast protocols over wired and wireless networks. It also provides a topology generator to create a topology from a number of nodes to thousands of nodes. However, the network components for simulating hets have not been developed.

To setup and run a simulation network, a user should write an OTcl script that initiates an event scheduler, sets up the network topology using the network objects and the plumbing functions in the library, and

tells traffic sources when to start and stop transmitting packets through the event scheduler.

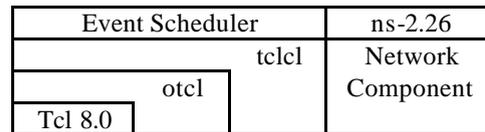


Figure 1: Architecture View of NS

In Figure 1, a general user (not an NS developer) can be regarded as standing at the bottom left corner, designing and running simulations in Tcl using the simulator objects in the OTcl library. The event schedulers and most of the network components are implemented in C++ and available to OTcl through an OTcl linkage that is implemented using tclcl. The whole thing together makes NS, which is an object extended Tcl interpreter with network simulator libraries [CC03]. This property makes NS easy to extend by developers and user friendly.

2 3-slot and 4-slot ACM Algorithms

An ACM is a data connector between two communicating processes which are not necessarily synchronized for the purpose of communicating with each other. The process generating the data is called the writer of the ACM and that making use of the data is called the reader of the ACM. By allowing the writer to write to the ACM and not directly to the reader and the reader to read from the ACM and not directly from the writer, an ACM provides more scope for asynchrony in the data communications between the two processes. This also implies shared memory in the ACM. An ACM may allow different degrees of asynchrony to one or both of the communicating processes and ACMs are classified into four types depending on the qualitative asynchrony between an ACM and its communicating processes [XY02]. The highest degree of asynchrony that can exist between a process and an ACM is known as full asynchrony. This means that the process has full temporal independence with regard to the ACM (i.e. can write or read without regard to the ACM’s state). The ACM is also said to be “non-blocking” to the process in this case.

In the terminology of multi-slot ACMs, a “data slot” is a unique portion of the shared memory that may contain one item of data. Multi-slot ACMs contain more than one slot for transmitting data between a writer and a reader process. Each of these processes is assumed to be a cyclic thread and during each cycle, a

single item of data is either written to or read from the ACM. A Pool ACM potentially allows both the writer and the reader processes full independent timing. A Signal ACM allows the writer temporal independence (i.e. the writer does not have to wait to begin writing, whatever the data state of the ACM) but requires the reader to wait when no new data item is available. A fully asynchronous Pool cannot be implemented with only one data slot and still provide data coherence, due to concurrent reading and writing accesses of multi-bit data items or packets from the same data slot which will probably cause modification of data contents after leaving the writer [XY02]. It has been shown in [Sim90] and [Cla00] that a fully asynchronous Pool can only be implemented using three or more data slots. The 3-slot and 4-slot ACM algorithms can be found in [XY02] and [MX99].

3 Implementation

Topology: The first step is to define the topology. Equivalent to a blueprint for building a house, to define a topology and to set up simulation scenarios allow the developer to build a new protocol and network component for NS. The scenario for simulating ACM networks is shown in Figure 2. The idea is to create a node for the writer, a node for the reader, and another node for the ACM agent that deals with shared memory and heterogeneous timing between the writer and the reader. To simulate the writer, Node 0 could be considered a UDP agent, with one of the traffic generators for simulating data transmission. NS supports three kinds of traffic generators which are CBR, Exponential, and Pareto. NS also allows users to develop traffic generator according to their needs. Node 1 performs the behaviours of shared memory and ACM algorithms. We have built ACM agents corresponding to 3-Slot and 4-slot Pool and Signal ACMs so far. Node 2 is used to simulate the reader. It can apply different traffic generators for sending request. It also needs to receive packets and record important ACM information for analysis.

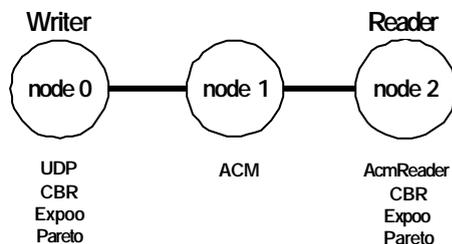


Figure 2: Topology of ACM simulation

Implementing 3-slot and 4-slot ACM agents: The implementations of 3-slot and 4-slot Pool/Signal ACM agents are shown in Figure 3 and Figure 4. An ACM agent will receive two different kinds of packets in heterogeneous timing. When an ACM agent receives a packet from the writer, it will distribute this packet in accordance with control variable next_. As soon as the

ACM receives request packets from the reader, it will access and send packets to the reader according to control variable read_.

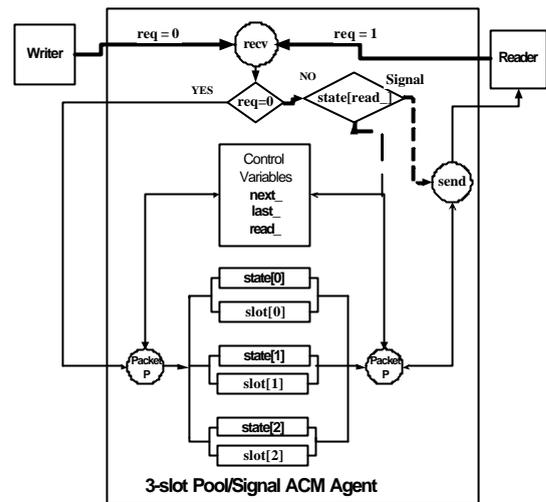


Figure 3: Algorithm of implementing a 3-slot Pool/Signal ACM agent

A Signal ACM can be read from when the data state of the slot chosen for the reader is '1'. This data state means that the packet (item) in this slot is fresh (has not been read before). If the data state is '0', this packet has been read during a previous reader cycle and the reader must wait until the data state changes. It will then obtain the new item replacing the previously read one in the same slot.

The ACM component is named "AcmAgent" and "Acm4slotAgent" for 3-slot and 4-slot ACM respectively. They are implemented as a derived class of "Agent" in Figure 5. The matching OTcl hierarchy names are "Agent/Acm" and "Agent/Acm4slot" for the 3-slot and 4-slot ACM agents. For the reader, we created an agent called "AcmReaderAgent", which is implemented as a derived class of "AcmAgent". The matching OTcl hierarchy name is "Agent/AcmReader". The root of the NS hierarchy is the TclObject class that is the base class of all OTcl library objects (scheduler, network components, timers and the other objects including NAM related ones). As a parent class of TclObject, NsObject class is the base class of all basic network component objects that handle packets, which may compose compound network objects such as nodes and links. The basic network components are further divided into two derived classes, Connector and Classifier, based on the number of the possible output data paths. The basic network objects that have only one output data path are under the Connector class, and switching objects that have possible multiple output data paths are under the Classifier class. An ACM network component should inherit from the Connector class and have general characteristics inherited from Agent class.

recv() method: This method function deals with the actions of "write to data slots" and "listen for requests" from the reader. Three writer steps are performed in this method. As mentioned before, each data slot has a

Packets are sent at a fixed rate during burst periods, and no packets are sent during idle periods. The lengths of burst and idle periods are assumed to conform to exponential distributions. Packet size is set to be constant.

Scenario: Writer's burst rate is 800k, mean burst time 25ms, mean idle time 50ms. Reader's access rate is 500k and the other conditions are the same as the writer.

Result: Figure 6 shows the simulation for a 3-slot Pool ACM agent. The intervals of both the writer and reader are irregular according to exponential distribution. Both reader and writer are independent (wait-free), as should be for a Pool. Data coherence and freshness are maintained in this test. Both data overwriting and rereading happened when the speed of the writer differs from that of the reader.

5 Conclusion and Future Work

The 3-slot and 4-slot Pool/Signal ACM agents have been successfully built into NS. Initial simulation results confirm that

- The usage of data slots in the mechanism are fairly symmetrical. Reader and writer do not clash on the same data slot (thus not violating data coherence requirements)
- No assumptions are made about the speed and the duration of either the read operation or the write operation (thus conforming to the requirement of temporal independence between the processes).
- The latest completed version of the shared data produced by the writer always is available for accessing by the reader (thus not violating data freshness requirements)
- Overwriting and rereading may occur during a simulation. It depends on the relative speed of the reader and the writer.

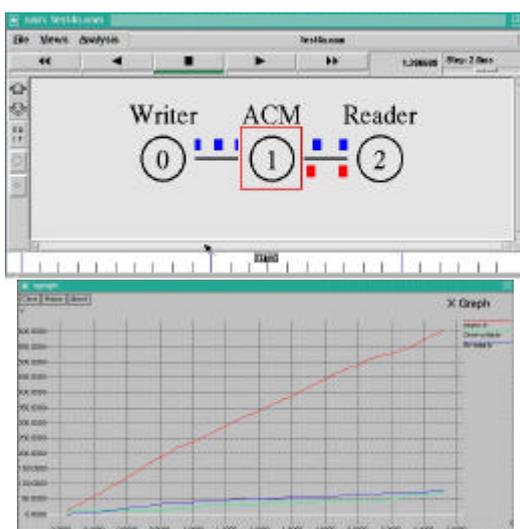


Figure 7: NAM & XGRAPH for TEST

NS is shown to be capable of being employed to simulate hets. However, it still can be enhanced in

many ways. For instance, we may apply multi-cascade ACM agents for simulating networks with, for instance, single thread, mesh and star network structures. ACM agents can also be further developed in order to simulate a LAN or WAN containing ACMs. This will help such work as the definition of an ACM network transport protocol to compete with the synchronous transport protocols (those based on request/acknowledgement handshakes, for instance) which have been widely used in the current Internet environment. Today's Internet, the majority (80-90%) of traffic is generated and controlled by the TCP protocol. The non-blocking provided by ACMs at the lowest levels of detail (gate-level hardware) may turn out to be very important for low-power, high speed data communications between near neighbouring processes in such distributed systems as sensor networks and such highly integrated solutions as networks on chip.

The Signal ACM may turn out to be useful in tackling the blocking inherent in channels by introducing share memory and retransmission. For instance, the sender may behave as the writer of a Signal until receiving a nack (negative acknowledgement), and then resends the specific packet before going back to normal actions. This way if a receiver dies or a connection dies, the sender just goes on its own merry way without affecting other tasks it may be doing etc. In other words, the temporal independence provided by ACMs to their accessing processes can directly prevent local deadlocks from spreading through a network at the lowest level.

We will carry out further investigations to reveal whether the Signal ACM is applicable in networking in general (as opposed to specifically in real-time systems), where the main concern is the overall performance. This research will also show whether Signal is unique in its way or people working in the networking area have already used mechanisms functionally similar to Signal. Whatever the answer is, even if something similar is known to be used in networking, it would be possible to compare and contrast those mechanisms with the Signal.

6 Acknowledgement

This work is supported by the EPSRC through the Coherent project (GR/R32666 and GR/R32895). The authors wish to thank H. Simpson and E. Campbell for helpful discussions.

References

- [XY02] F. Xia, A. Yakovlev, I.G. Clark, D. Shang. 'Data communication in systems with heterogeneous timing', IEEE Micro, Vol.22, No.6, pp.58-69, November/ December 2002.
- [CB97] J. Chen, A. Burns. 'A Three-Slot Asynchronous Reader/Writer Mechanism

- for Multiprocessor Real-Time Systems*', Department of Computer Science, University of York, January 1997.
- [MX99]** A. Madalinski, F. Xia and A. Yakovlev '*Data Studying the data loss and data re-reading behaviour of a four-slot asynchronous communication mechanism using stochastic Petri net techniques*', Department of Computing Science, University of Newcastle, November 1999.
- [Pro99]** The VINT Project. The '*ns manual*' (formerly NS notes and documentation). Technical report, UC Berkeley, LBL, USC/ISI, and Xerox PARC, October 1999. The current version of the NS manual is available at <http://www.isi.edu/nsnam/ns/doc/index.html>, [l,ns_doc.ps.gz,ns_doc.pdf](http://www.isi.edu/nsnam/ns/doc/index.html).
- [CC03]** Jae Chung, Mark Claypool. '*ns by example*'. Computer Science, Worcester Polytechnic Institute (WPI), 2003. Available from <http://nile.wpi.edu/NS/>.
- [Cri03]** Marc Cries. '*ns tutorial*', 2003. Available from <http://www.isi.edu/nsnam/ns/tutorial/index.html>
- [Mrk01]** Aleksandar Mrkaic. '*Porting a WFQ Scheduler into ns2's Diffserv Environment*', Student Thesis, Swiss Federal Institute of Technology Zurich, 2001
- [Cla00]** Clark I.G. '*A Unified Approach to the Study of Asynchronous Communication Mechanisms in Real Time Systems*', PhD thesis, King's College, London University, May 2000.
- [XL98]** Fei Xia, Bin Liu. '*Quantitative investigation of certain properties of slot mechanisms*', Proc. 5th UK Asynchronous Forum, Computer Laboratory, University of Cambridge, December 1998.
- [Sim90]** H.R. Simpson. '*Four-slot fully asynchronous communication mechanism*', IEE Proceedings Vol.137, Pt. E, No.1, pp.17-30, January 1990
- [Cap97]** D.M. Capper, '*C++ for Scientists, Engineers and Mathematicians*', Springer-Verlag, 4th ED, London, 1997
- [Sch98]** H. Schildt, '*Teach Yourself C++*', McGraw-Hill, 2nd ED, CA, 1997
- [VHB01]** A. Veres, B. Hungary, M. Boda, '*The Chaotic Nature of TCP Congestion Control*', Ericsson Research, Stockholm, Sweden, 2001

Simulation of All-Optical Recirculating Fibre Loop Buffers employing Soliton Pulses and Optical Regeneration

A. Als⁺ and Z. Ghassemlooy*

+ University of West Indies, Cave Hill Campus, Barbados
Phone: +1 246 4174793, Fax: +1 246 4174597, Email: aals@uwichill.edu.bb

* Optical Communications Research Labs., University of Northumbria, Newcastle upon Tyne, NE1 8ST, U.K.
Phone: +44 191 2274902, Fax: +44 191 2273684, Email: fary@ieee.org

Abstract

All optical buffering within high-speed optical time division multiplexing (OTDM) based networks is becoming a necessity. In this paper, an active all-optical recirculating fibre loop buffer (RFLB) employing an all-optical regenerator to compensate for pulse distortion is introduced. The OTDM pulse shape is soliton and the data rate used is 20 Gb/s. Simulation results for the bit error rate (BER) performance for different regeneration periods are presented.

1. Introduction

The continuing growth in the demand for communication bandwidth has prompted the progression from the 2nd generation transparent optical networks (TONs) to the 3rd generation all-optical networks (AONs). Consequently, the bottlenecks imposed by the tardy electrically controlled switching (/routing) mechanisms are expressly avoided¹⁻³. In such networks, all-optical buffering (AOB) option then becomes crucial in resolving contention issues where data signals from multiple sources compete for a single destination port.

Additionally, AOB may be employed at both the transmitter and receiver sites as illustrated in Fig. 1. In each of these scenarios the buffer is employed to effect flow control whilst ensuring that both the

data and control signals remain in the optical domain. In the absence of dynamic all-optical memory devices, fixed lengths of optical fibre are typically used to facilitate buffering. However, RFLB architectures may be employed to offer greater flexibility by providing a mechanism for variable buffering times. The inherent difficulty with such architectures is the pulse shape distortion that results from non-linear pulse propagation, inter-pulse interactions and the accumulation of amplified spontaneous emission noise (ASE) produced by the in-loop optical amplifier. In this work an optical regeneration feature is introduced into the RFLB architecture to compensate for these performance degrading factors. In section 2, the simulation model and fundamental operation of the RFLB is described. Next, the simulation results are presented in section 3, and finally, the conclusions are presented in section 4.

2. Simulation model

Figure 2 shows the simulation model system block diagram, which consists of an OTDM optical packet generator, an active RFLB and a receiver unit. The OTDM generator produces fundamental soliton ($N = 1$) data pulses as defined by⁴:

$$S(t) = N \operatorname{sech}(\tau/T_0), \quad (1)$$

where $\tau = t - z/V_g$, $N = \sqrt{\gamma P_0 T_0^2 / |\beta_2|}$, T_0 = full width half maximum pulse width/1.763, V_g the group velocity, and P_0 the peak incident power required for fundamental soliton propagation.

The logical representation of the optical buffer architecture is illustrated in Fig. 3. Note that in VPI simulation package a ‘ringinitialiser’ module is used to implement the buffer, therefore there is no need to use optical splitter and combiner. The operation of the buffer is as follow. Pulses entering the buffer via port A circulate the loop of dispersion shifted fibre for a finite period of time until they are switched out via port B. Conversely, unbuffered pulses traverse along the path AC. In the case where regeneration is employed, the buffered pulses exiting the loop, at port C, are redirected to port A via the all-optical regeneration unit. This unit functions as an optical logical AND-gate where the input data signal is gated with a synchronised clock pulse. Moreover, the unit also consists of an optical amplifier with an adjustable gain to ensure that the peak power of the regenerated pulses is sufficient enough to support the soliton propagation. For a given fibre of length L , the delay per loop traversal T_{Loop} is defined as:

$$T_{loop} = nL/c, \quad (2)$$

where c , n and L are the speed of light in a vacuum, the refractive index and the length of the fibre, respectively.

The receiver unit consists of an attenuator, an ideal clock recovery unit, a PIN photodetector, an electrical third order low-pass Bessel filter, a sampler, a threshold detector, and a BER module. The attenuator is used as a control to obtain the BER over a range of received signal powers. The clock extractor determines the timing delay between the transmitted and received signals for synchronisation purposes. Finally, the photodetector is assumed to be ideal with unity responsivity.

3. Simulation Results

The model in Fig. 2 was simulated using the VPItransmission Maker Photonic simulation software with the key parameters listed in Table 1. The soliton pulse propagation is modelled using the beam propagation technique.

Parameter	Value
Pulse width	5 ps
Dispersion Slope	1.01 sm ⁻³
Switch insertion loss	1 dB
Switch crosstalk	30 dB
In-loop amplifier gain	~1.1 dB
Peak power	15.4 mW
Attenuation	30-40 dB
Amplifier noise figure	4.0
Number of loop traversals	400

Table 1. Simulation parameters

Figures 4a and 4b show the packet waveform at the input and output of the RFLB with 200 loop traversals without any regeneration, respectively. The noisy waveform in Fig. 4b is due to several factors. First, the accumulated ASE contributions on each amplification pass around the dispersion shifted fibre (DSF) loop. Second, the loss induced pulse broadening and the associated peak power fluctuations, which are bit pattern dependent. Last, the switch loss disturbs the delicate balance of self-phase modulation and group velocity dispersion required to effect the fundamental soliton propagation⁴. These factors result in the soliton ‘breathing’ in an attempt to stabilise their peak power to pulse width ratios.

In all cases it is considered that the packet has a maximum total circulation time of 1 ms (i.e. 400 loop traversal) and use a regeneration period R_p to define the number of times packets are regenerated within 1 ms. Figure 5 shows the BER against the received signal power for different values of R_p . Observe that for $R_p = 50$ loops (i.e. 8 packet regeneration instances) the BER performance is much worse than for large values of R_p (i.e. lower number of packet regeneration instances). This is explained as

follows. After 50 iterations, the regenerated pulses are sufficiently clean and need no further regeneration. However, on passing through the optical amplifier, a significant amount of ASE noise is added to the pulses before being feedback into the fibre loop. This process then has implication on successive regeneration periods, and as a result is important to find the optimum R_p when regeneration becomes beneficial. Note that the BER performance is improved when the R_p is increased to 100 and above. For BER of 1×10^{-9} the minimum received optical power requirement of ~ -28 dBm is achieved for R_p is equal to 200 and 400.

Finally, Figs. 6a and 6b show the eye diagrams after 200 iterations with and without optical regeneration, respectively. The benefits of regeneration can be observed, as the eye is more open and less noisy. This therefore translates into an improved BER performance.

4. Conclusion

In this paper, the performance of an active all-optical recirculating fibre loop buffer (RFLB), equipped with an all-optical regenerative capability for OTDM transmission system was investigated. Results shows that the optimal regeneration period for a BER of $\sim 1 \times 10^{-9}$ to be achieved is 200 or 400 iterations. However, one may opt of the latter to reduce any overhead associated with the regeneration process. The results also suggest that $R_p = 200$ is better suited for lower received powers. At this stage a simple regenerator based on the

logical AND gate principle was employed. This resulted in some of the peak power fluctuations within the input signals being transferred to the regenerated output signal. Future work will seek to optimise the regeneration process to minimise this effect and thereby improve the BER performance.

5. Acknowledgement

A. Als wishes to thank Mr. R. Ngah for his contribution to the research efforts and VPI systems for licensing the simulation package used in generating the results presented in this paper.

6. Reference

1. I. Glesk, K. I. Kang, and P.R. Prucnal, "Demonstration of ultrafast all-optical packet routing", *Electronics Letters* 33 (9),1997, pp. 794-795.
2. G. P. Agrawal, *Fiber-Optic Communication Systems* (Wiley, 1997).
3. G. P. Agrawal, *Applications of Nonlinear Fiber Optics* (Academic Press, 2001).
4. Taylor JR: *Optical Solitons - Theory and Experiment.*, Cambridge University Press, 1992
5. T. Yabu, M. Geshiro, T. Kitamura, K. Nishida, and S. Sawa, "All-Optical Logic Gates Containing a Two-Mode Nonlinear Waveguide", *IEEE Journal of Quantum electronics* Vol. 38, No. 1, 2002, pp.37-46

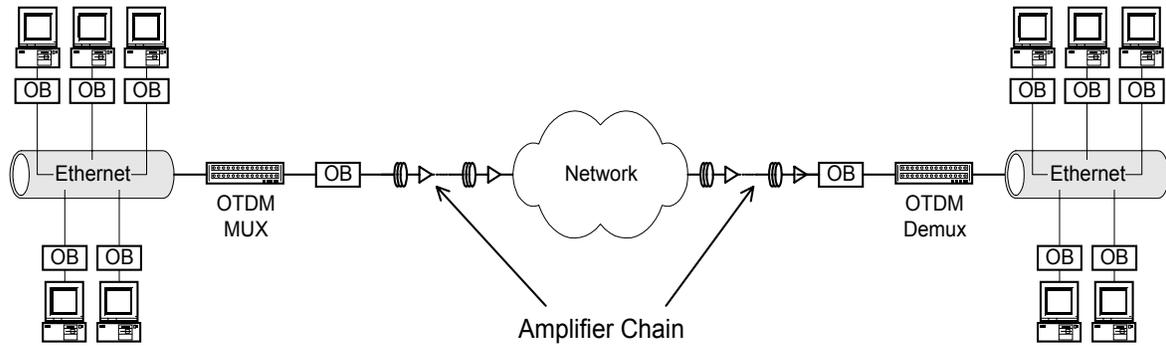


Fig. 1 Arbitrary network architecture with optical buffers (OBs)

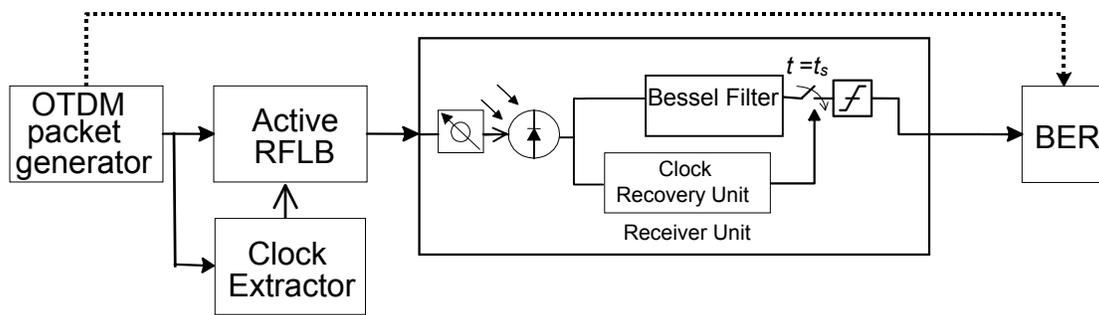


Fig. 2 System block diagram highlighting the receiver unit

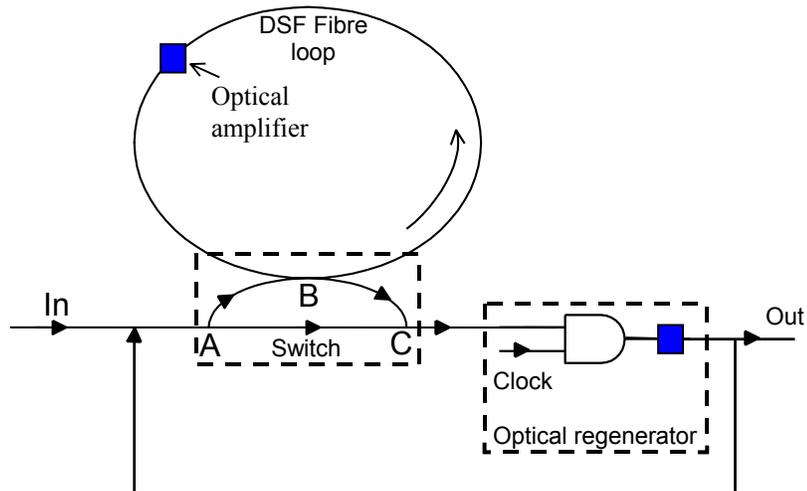


Fig. 3 Logical representation of active RFLB

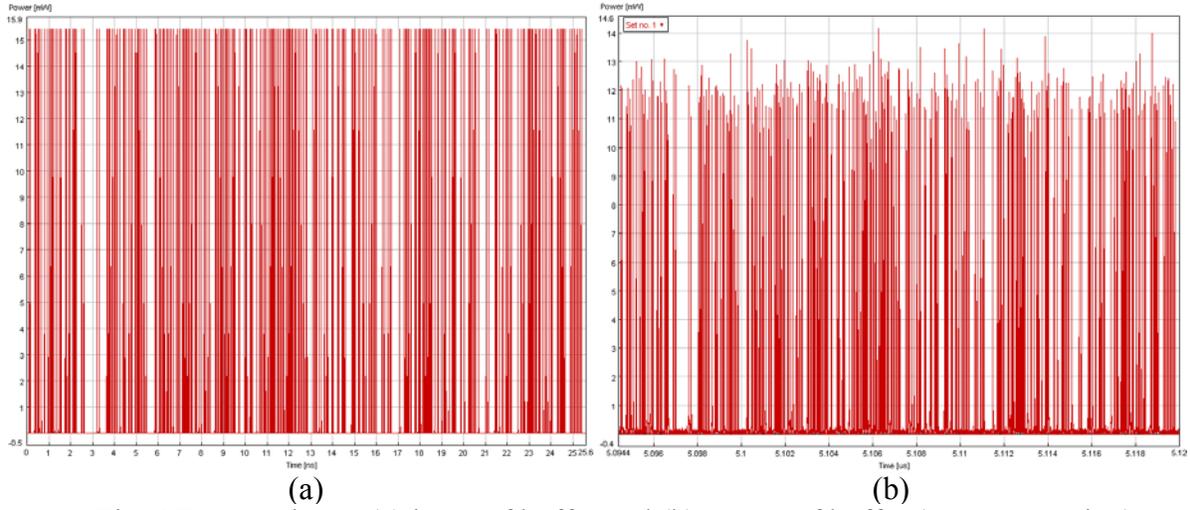


Fig. 4 Data packet at (a) input of buffer and (b) output of buffer (no regeneration)

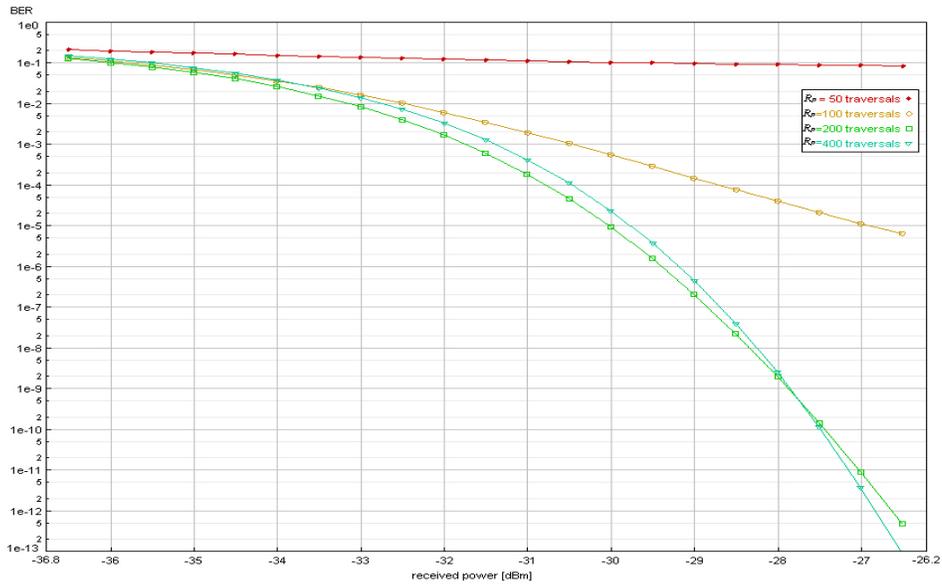


Figure 5 BER versus received power for different values of R_p

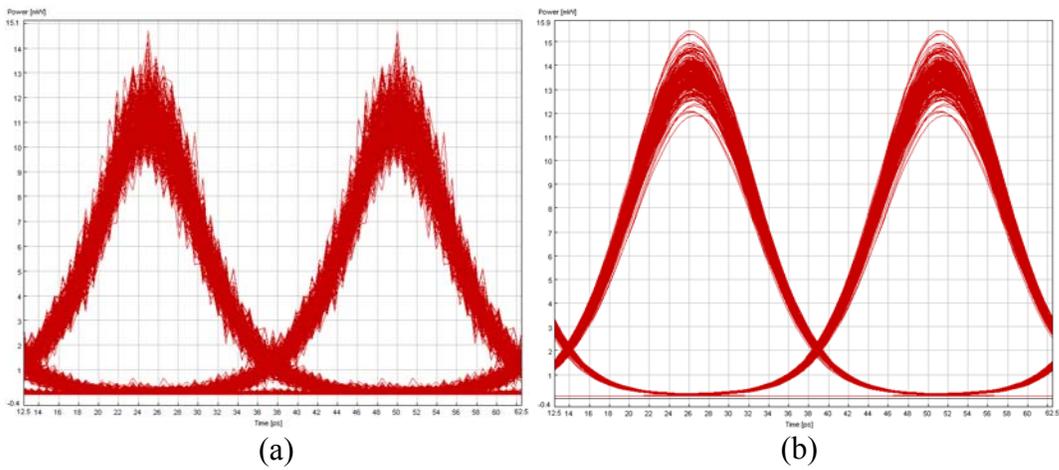


Fig. 6 Eye diagrams after 200 iterations (a) without regeneration and (b) with regeneration

Simulation Study of WFQ Packet Scheduling Mechanism for different coding schemes in GPRS network

Colin Chai and Jean Armstrong

Department of Electronic Engineering
La Trobe University
Victoria 3086 AUSTRALIA

Email: c.chai@latrobe.edu.au, j.armstrong@latrobe.edu.au

Abstract

General Packet Radio Service (GPRS) is a packet data technology based on the ETSI standard for exchanging packet-mode data using the Global System for Mobile communications (GSM) network. In this paper we look at ways of introducing different service classes to different GPRS users by using the Weighted Fair Queuing (WFQ) mechanism at the IP layer. The simulation results show that with the WFQ algorithm applied in the GPRS core network, better quality of service in web-page download throughput can be achieved by higher class users provided the bandwidth asymmetry is kept at a moderate level.

I. Introduction

GPRS radio resources are used only when users send or receive data. The available radio channels can be concurrently shared among several users to ensure effective spectrum utilization [1], [2], [3]. Therefore GPRS is designed to support transfers ranging from intermittent and bursty data (e.g. web browsing) to occasional transmission of large volumes of data (e.g. file transfer). GPRS is intended to provide better services for Internet applications compared to the existing circuit-switched data services of GSM. In order to support packet-switched services, additional network elements have been devised. The Gateway GPRS Support Nodes (GGSNs) act as the logical interface to external packet data networks. The Serving GSN (SGSN) is responsible for the delivery of packets to the mobile stations (MSs) within its service area, via base station subsystems (BSSs) serving individual cell sites.

Different coding schemes (CS-1 to CS-4) have been adopted in the GPRS network to offer varying levels of error protection to account for variations in the radio channel conditions. The more robust the coding scheme the lower the data rate and hence user throughput.

II. Wireless WFQ Algorithm

Multimedia applications over wireless networks have been a major topic of research in recent years. These applications have specific requirements in terms of delay and bandwidth that challenge the quality of service the wireless network can provide to the individual users.

Almost all the traditional packet-scheduling algorithms have proven to be effective in the wireline environment. In the wireline environment, the data streams rate is kept constant for a long period of time and has very stringent delay requirements. However, the effect is unknown when applied to GPRS wireless network with bursty data sources.

Excess queuing can lead to gross unfairness in the wireless network between the competing IP data flows associated with each connection. As a result, the weighted fair queuing packet scheduling algorithm (WFQ) is being studied to address the queuing problems that lead to bandwidth unfairness [4]. It aims to provide multiple data service classes for different users by assigning each traffic-source/client a weight determined by its class. The weight controls the amount of traffic a source may deliver to the client relative to other active sources during some period of time.

III. OPNET Implementation

OPNET is a network simulation tool that is highly suitable for modeling and simulating network environments [5].

A. Simulation Setup

Fig. 1 illustrates the GPRS network architecture that is implemented using OPNET simulation software [5].

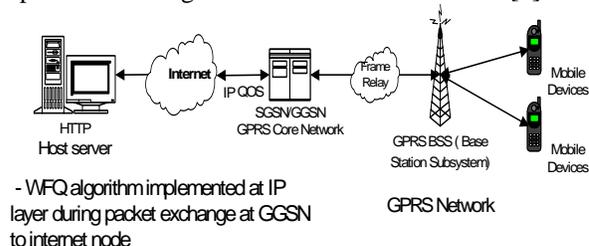


Fig. 1 OPNET Simulations: Smaller GPRS Network

Mobile stations (MSs), base station subsystems (BSSs) and the core network (CN) are the three main elements modeled in the OPNET GPRS network.

The MSs handle the mobility management between the mobile station and BSS and its GPRS capability by

assigning time slots in both uplink and downlink channels. In the simulations, single or multiple timeslots can be allocated at each MS for throughput comparisons when the WFQ algorithm is applied.

In addition, there is an underlying frame relay network that handles the communication between the BSS and CN. The BSS process module handles the interworking functions between the GPRS network of the MS and the frame relay network of the core network.

The CN is a single node entity consisting of both Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN) networks. The extra signaling between the SGSN and GGSN is not modeled here. The GGSN node is the gateway node between an external packet data network (IP) and the GPRS core network.

B. Traffic Source Simulation

To date, there has been no measurement of source traffic models for wireless data users. As a result, the traffic measurements are taken from the wire-line network environment and rely on several publications [6], [7]. The source traffic can be modeled with two layers. The first layer is the session layer, which contains the arrival process for web sessions of the user activities. The second layer is the layer where the actual activity process takes place – i.e. the number of pages in a web session etc. According to [7], the page arrival process has a Gamma distribution and the number of pages belonging to one session follows a Lognormal distribution.

In the OPNET simulation environment, each mobile handset/client is assigned with a particular Type of Service (TOS) for all packets originated from the individual client. It represents a session attribute, which allows packets to be processed faster in IP queues in the core network. Each TOS at the client is assigned as an integer value between 0 - 252, with 252 being the highest priority. As an example, the lowest TOS, TOS 0 has a value of 0. TOS 1 at the client has a value of 32 and so on and so forth.

Also at the router/intermediate node end, we use the concept of WFQ algorithm, which is largely based on the type-of-service (TOS) within the IPv4 header. This technique is being applied at the interface between GGSN node (the GPRS core network) and the internet node as indicated in Fig.1. It is used to indicate a certain way by which a given IP datagram should be queued at the routers/GPRS core network devices in terms its priority level over an end-to-end connection. Within the WFQ algorithm, a higher weight indicates larger allocated bandwidth and shorter delays. In a

similar fashion, lower weight in the IP queue indicates less bandwidth and throughput that will come through that queue. In this simulation, TOS 0 or best effort service defines the weight as 1.0 whereas TOS 1 or background traffic defines the weight as 10. Hence ratio of TOS 1 to TOS 0 is 10.

IV. Simulation Scenarios and Results

Multiple simulation scenarios were performed based on the traffic sources from the previous section. In these simulations we assume there is no packet error or channel error, as, the main purpose of this study is to determine the effectiveness of packet scheduling techniques used in the GPRS networking environment. The bandwidth assignments for both uplink and downlink channels are asymmetrical because World Wide Web (WWW) applications generate a much higher traffic loads in the downlink direction than in uplink.

A. Symmetrical Network Configurations Scenario

In this section, we first compare the case where symmetrical time slots are assigned both in the uplink and downlink directions.

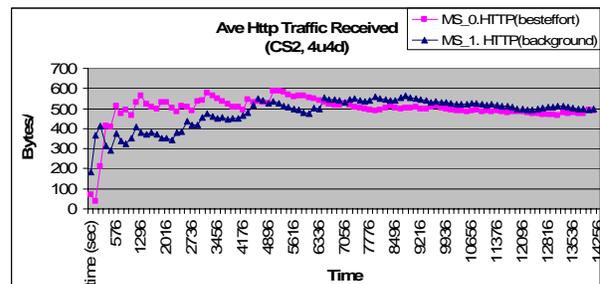


Fig. 2 Average Http Traffic Received for CS-2 with four uplink and four downlink channels assigned

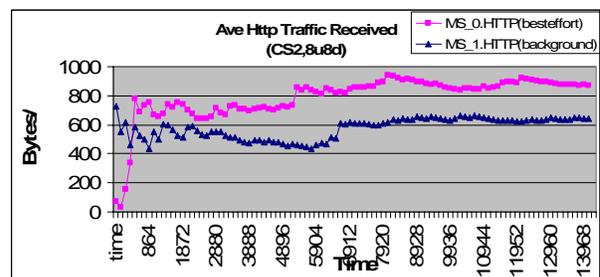


Fig. 3 Average Http Traffic Received for CS-2 with eight uplink and eight downlink channels assigned

The simulation results have indicated for both cases in Fig.2 and Fig.3 there is no throughput improvement for MS1 with higher TOS (TOS1) compared to MS0 (TOS 0). In addition, it is also impractical to have equal time

slots assigned in both up and downlink directions due to wastage of bandwidth in the uplink directions. The uplink channels are used for sending dataless acknowledgements (ACKs).

B. Asymmetrical Network Configurations Scenario

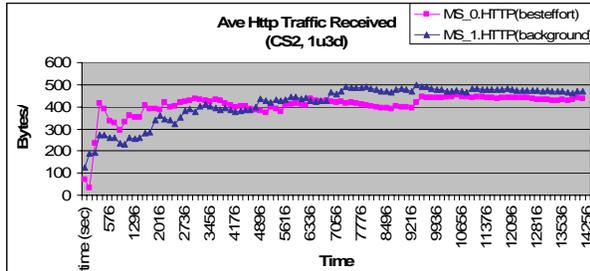


Fig. 4 Average Http Traffic Received for CS-2 with one uplink and three downlink channels assigned

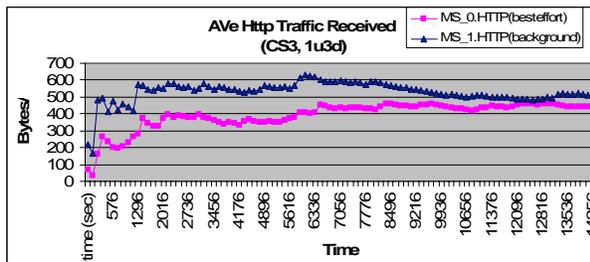


Fig. 5 Average Http Traffic Received for CS-3 with one uplink and three downlink channels assigned

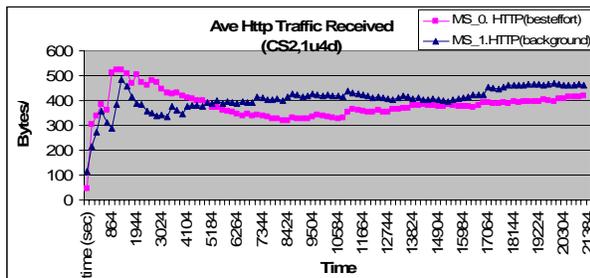


Fig. 6 Average Http Traffic Received for CS-2 with one uplink and four downlink channels assigned

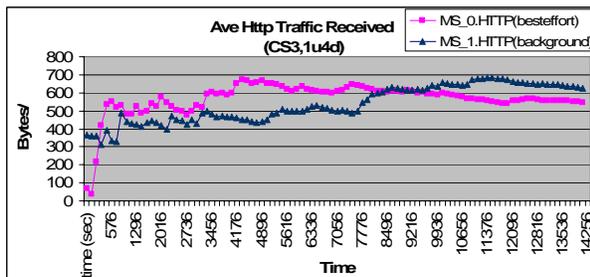


Fig. 7 Average Http Traffic Received for CS-3 with one uplink and four downlink channels assigned

In this section, we adopt coding schemes 2 and 3 in our simulations. It has been proven [3] that CS2 shows very good behavior for low and medium channel conditions. As for CS3, it offers higher throughput but with lower error correction compared to CS2.

Here we perform the simulations in pairs. Fig.4 corresponds to CS-2 with one uplink and three downlink channels whereas Fig.5 corresponds to CS3 with the same configurations. We can clearly see that steady state is achieved after 7000 secs of simulation in Fig.4. In both cases, throughput improvements are achieved for MS1 with higher TOS assigned via WFQ algorithm in the IP queues.

In Fig.6 and Fig.7, one uplink and four downlink channels are assigned in each coding schemes CS2 and CS3. The results show that with slightly higher bandwidth asymmetry in the channel assignment compared to previous case, it has not affected the throughput gains for MS1 in both cases.

In the simulations, consistency of the results is further tested with longer simulations which are not included here. For longer simulations, the results also show that MS1 achieved better throughput in comparison to MS0.

On the other hand, as bandwidth asymmetry increases, in the case of simulations in Fig.8 & 9, the throughput for MS1 using CS-2 dipped below the throughput for MS0. Similarly, there is no clear and distinct throughput improvement for CS-3 case.

In the case of Fig.10 & 11, one timeslot in uplink and eight timeslots in the downlink channels are used in the configurations, throughput degradation is observed for both CS-2 and CS-3 cases.

Hence, it can be concluded that the WFQ algorithm does not work well if there is a large bandwidth asymmetry in the assignment of GPRS channels. The results have indicated that a large bandwidth asymmetry between downstream and upstream forces the sender to clock out more slowly than optimal, thus decreasing the effectiveness of WFQ algorithm and hence the overall WWW downloading throughput.

In addition, the TCP sender also depends on the receipt of ACKs to grow its congestion window, which governs its transmission rate. If the ACKs get bunched up in the uplink direction due to large bandwidth assigned, the sender may burst out data in the downlink direction from time to time, which could lead to the disruption of TCP window growth and degrade performance to a fraction of the available bandwidth. As a result, the bandwidth is underutilized and hence the throughput

improvement cannot be achieved for MS/client with higher TOS value assigned.

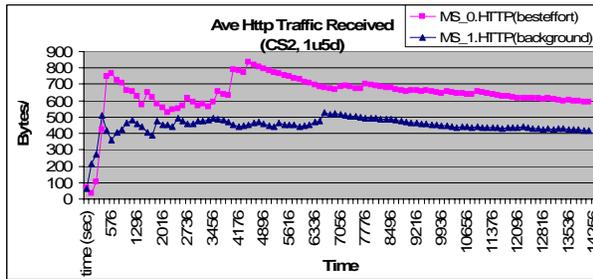


Fig. 8 Average Http Traffic Received for CS-2 with one uplink and five downlink channels assigned

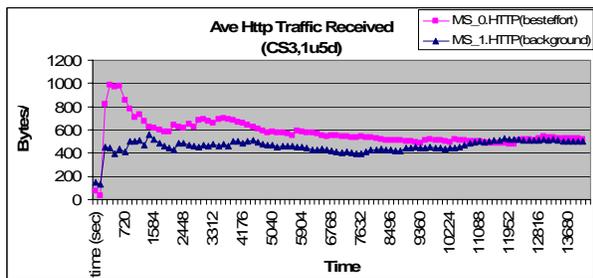


Fig. 9 Average Http Traffic Received for CS-3 with one uplink and five downlink channels assigned

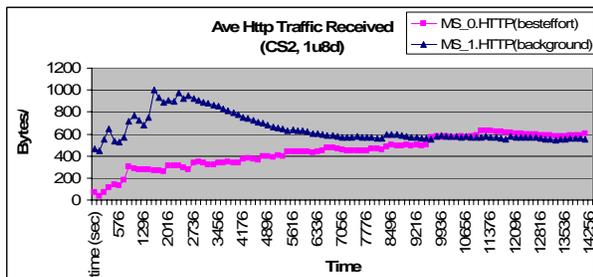


Fig. 10 Average Http Traffic Received for CS-2 with one uplink and eight downlink channels assigned

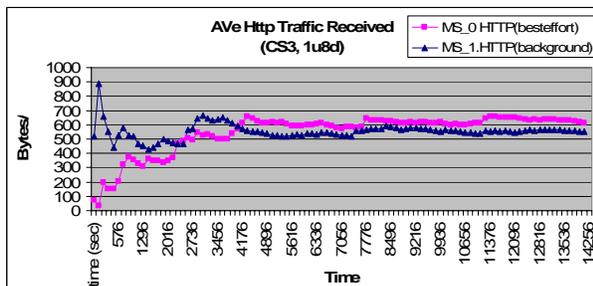


Fig. 11 Average Http Traffic Received for CS-3 with one uplink and eight downlink channels assigned

V. Conclusions

This paper presents a throughput analysis of GPRS network when the WFQ algorithm is used at the IP network layer. Our simulation results demonstrate that

a GPRS user achieves better throughput for Mobile Station with higher classes if lower bandwidth asymmetry in uplink/downlink channels allocations are used. Large bandwidth asymmetry in channels assignment results in the disruption in TCP data and ACKs flow in both directions in the TCP layer. Hence it has hindered the effectiveness in WFQ algorithm used at the TCP/IP layer.

Data traffic burstiness is also another contributing factor in throughput degradation for higher-class users in the simulated cases.

Due to the burstiness of the data traffic, WFQ can provide differentiated services only to a limited extent. However, by assigning the bandwidth appropriately both in downlink and uplink channels, performance improvement can be achieved reasonably well with IP WFQ algorithm implemented in the GPRS network.

References

- [1] R. J. Bates, "GPRS: General Packet Radio Service". New York, NY: McGraw-Hill 2001.
- [2] G. Brasche., B. Walke: "Concepts, Services and Protocols of the new GSM Phase 2+ General Packet Radio Service". IEEE Communications Magazine, August 1997, pp 94-104.
- [3] R. Kalden, I. Meirick, M. Meyer, "Wireless Internet Access Based on GPRS", IEEE Personal Communications, April 2000.
- [4] V. Bharghavan, S. Lu and T. Nandagopal, "Fair Queueing in Wireless Networks: Issues and Approaches," IEEE Personal Communications Magazine, February 1999.
- [5] <http://www.opnet.com>.
- [6] S. Hoff, M. Meyer, and J. Sachs, "Analysis of the General Packet Radio Service (GPRS) of GSM as Access to the Internet," ICUPC '98, Florence, Italy, October 1998.
- [7] A. Reyes-Lecuona, E. Gonzalez-Parada, E. Casilari and A. Diaz-Estrella, "A page-oriented WWW traffic model for wireless system simulations", in: Proceedings of ITC'16, Edinburgh, UK, June 1999, pp.1271-1280.
- [8] [DiffServ Working Group Web site: http://www.ietf.org/html.charters/diffserv-charter.html](http://www.ietf.org/html.charters/diffserv-charter.html)
- [9] M. Meyer, "TCP performance over GPRS", in: IEEE Wireless Communications and Networks Conference 1999 (WCNC'99), Vol. 3 (1999), pp. 1248-1252.
- [10] Z. Wang, and P. Cao "Persistent Connection Behavior of Popular Browsers", <http://www.cs.wisc.edu/cao/papers/persistent-connection.html>.

Super-Orthogonal Space-Time Trellis Codes together with Blind Channel Identification Algorithm based on Symmetric Noise Subspace

G. Ferré, J.P. Cances, V. Meghdadi, G. R. Mohammad-Khani

University of Limoges, ENSIL-GESTE, Parc ESTER Technopole, 87068 Limoges, France

E-mail : {ferre, cances, meghdadi, reza} @ ensil.unilim.fr

Abstract

This paper introduces a new class of space-time codes called super-orthogonal space-time trellis codes. These codes combine the coding advantage of STTC's together with the advantage diversity of STBC. Using the determinant criterion introduced by Tarokh & al the set partitioning rules of Ungerboeck still applies. The obtained subsets enable to affect different STBC's to each subset parallel transition branches. Hence codes operating at different rates, up to the highest theoretically possible rate, for different number of states can be designed thanks to these partitioning rules. The idea was first developed by Fitz & al and extended by Jafarkhani & al. We present here new code constructions in a common general framework and propose a blind channel identification which exhibits losses less than 1.5 dB when compared to the perfect CSI case.

I. Introduction

ST-TCM has recently emerged as a commonly accepted effective coding strategy for wireless communication systems with multiple transmit antennas. Since the first studies presented by Tarokh & al [1] there has been a great deal of research aiming at improving the designs of STTC [2] among many others. The main drawback of the previous mentioned STTC schemes consists in the necessary tradeoff between transmission rate, constellation size, diversity advantage and trellis complexity which prevents the constructed codes to reach the maximum diversity advantage for a given trellis complexity. This can be overcome by using the revolutionary idea of Ungerboeck that is constellation expansion and set partitioning and by combining STTC with STBC technique which offers the maximum diversity advantage. Set partitioning can be realized using the determinant criterion [1] examining the different pairs of possible transmitted symbols if we use, for example, a STBC with two transmit antennas. Having partitioned the constellation into different subsets we affect a particular STBC from a set of possible candidates to transitions originating from a state. Therefore, in general, for a $N \times T$ STBC, picking a trellis branch emanating from a state is equivalent to transmitting NT symbols from N transmit antennas in T time intervals. By doing so, it is guaranteed that we get the diversity of the corresponding STBC while we show how to design the trellis code for the highest possible rate to get the maximum coding gain

as well. These ideas originated by Fitz & al [3] and extended by Jafarkhani & al [4] lead to the so-called SOSTTC : Super Orthogonal STTC. In addition to a general framework for the construction rules of this new class of STTC codes we propose a blind channel identification algorithm based on symmetric noise subspace which exhibits outstanding performances in a MIMO context. The loss exhibited when comparing its performance to the case of ideal CSI is less than 1.5 dB. The paper is organized as follows. In section II the set partitioning rules are explained for a two transmit antennas scheme example. Section III presents examples of SOSTTC designed by the previously mentioned rules. Section IV describes the blind channel identification algorithm while simulation results are exhibited in Section V. Section VI is the conclusion which summarizes the main results.

II. Set partitioning rules

In this paper we concentrate on a two-transmit antennas scheme. We use the following class of orthogonal designs as transmission STBC matrices :

$$C(x_1, x_2, \theta) = \begin{pmatrix} x_1 e^{j\theta} & x_2 \\ -x_2^* e^{j\theta} & x_1^* \end{pmatrix} \quad (1)$$

The case $\theta = 0$ provides the well known scheme of Alamouti [5]. We pick θ such that for any choice of x_1 and x_2 from the original constellation points, the resulting transmitted signals are also from the same constellation. For example, if we use M -PSK constellation signals, the constellation signals and thus x_1 and x_2 can be represented by $\exp(j2\pi l/L)$, $l = 0, \dots, L-1$. One can pick $\theta = 2\pi l' / L$, $l' = 0, \dots, L-1$. In this case, the resulting transmitted signals are also members of the M -PSK constellation and, therefore, do not expand the constellation signals. For BPSK we'll use $C(x_1, x_2, 0)$ and $C(x_1, x_2, \pi)$ and $C(x_1, x_2, 0)$, $C(x_1, x_2, \pi/2)$, $C(x_1, x_2, \pi)$ and $C(x_1, x_2, 3\pi/2)$ for QPSK.

We provide now the partitioning rule for transmitting symbols with the class of orthogonal designs given in (1). We denote the difference of the transmission matrices for codewords c_1 and c_2 by $B(c_1, c_2)$ and its Hermitian, complex conjugate and transpose, by $B^H(c_1, c_2)$. Following the definitions in [1], the diversity of such a code is defined by the minimum rank of the matrix

$B(c_1, c_2)$. For a full-diversity code, the minimum of the determinant of the matrix $A(c_1, c_2) = B(c_1, c_2) \cdot B^H(c_1, c_2)$ over all possible pairs of distinct codewords c_1 and c_2 corresponds to the coding gain. We define, similar as in [4], the coding gain distance (CGD) between codewords c_1 and c_2 as $d^2(c_1, c_2) = \det(A(c_1, c_2))$, where $\det(A)$ is the determinant of matrix A . Then, we use CGD instead of Euclidean distance to define a set partitioning similar to Ungerboeck's set partitioning [4].

III. Examples of SOSTTC construction

We begin our partitioning with the simplest case of BPSK. There is a four way partitioning of the orthogonal code as shown in Fig.1. At the root of the tree, the minimum determinant is 16. At the first level of partitioning, the highest determinant that can be obtained is 64. This is obtained by a set partitioning in which subsets S_0 and S_1 use different transmitted signal elements for different transmit antennas. At the next level of partitioning, we have four subsets S_{00} , S_{11} , S_{01} and S_{10} with only one element per set.

We show now how to maximize the coding gain without sacrificing the transmission rate. We assign a constituent space-time block code to all transitions from a state. Adjacent states are typically assigned to one of the constituent space-time block codes from the super-orthogonal code. Similarly, we can assign the same space-time block code to branches that are merging into a state. This rule guarantees that any path that diverges from (or merges to) the correct path differs by rank 2. In other words, every pair of codewords diverging from (or merging to) a state achieves full diversity because the pair is from the same orthogonal code. Fig. 2-a shows a four-state example of a new Super Orthogonal Space Time Trellis Code (SOSTTC). If we use a QPSK constellation and the set partitioning in Fig.1 a), the result is a four state SOSTTC at rate 2-bits/s/Hz. The minimum CGD for this code is equal to 16, which is greater than the CGD of the corresponding STTC in [1].

Fig. 2-b exhibits a 3 bits/s/Hz SOSTTC using 8-PSK and the corresponding set partitioning of Fig. 1-c. The minimum CGD for this code is equal to 2.69 and there is no four-state space-time trellis code for 8-PSK in [1].

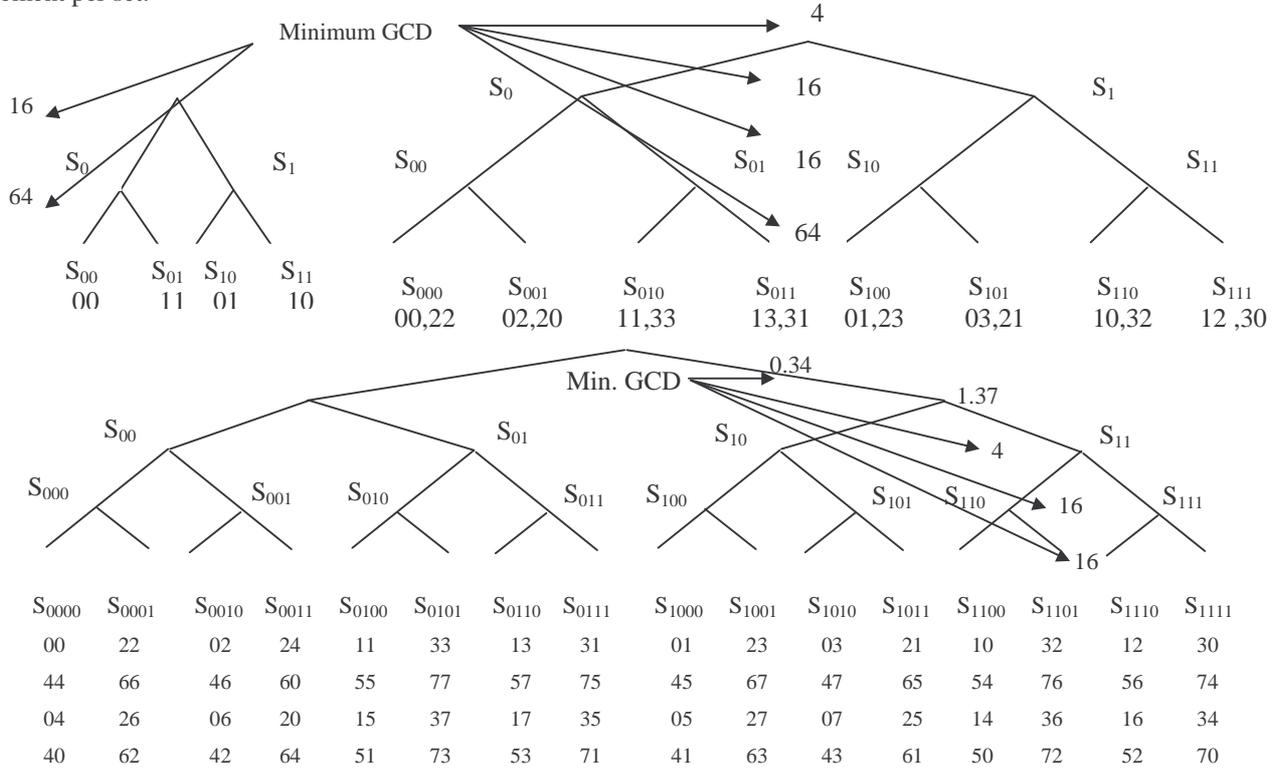


Fig. 1: Set partitioning for BPSK (a), QPSK (b) and 8-PSK (c)

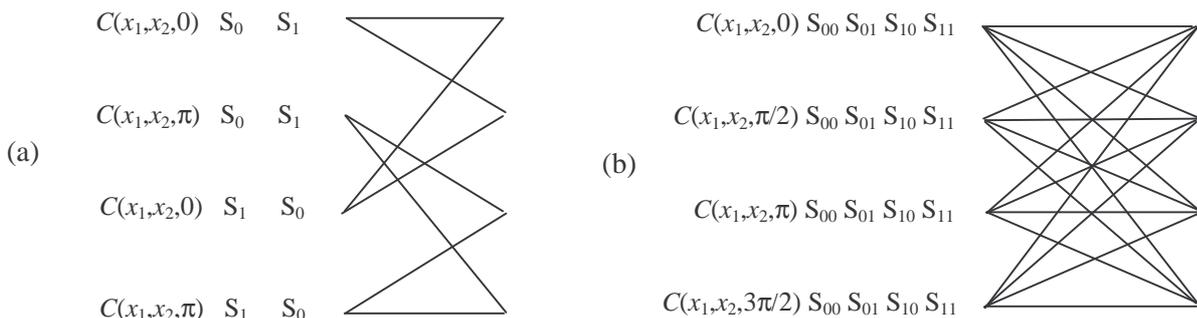


Fig. 2: a) four-state code, $r = 2$ bits/s/Hz (QPSK) b) four-state code, $r = 3$ bits/s/Hz (8-PSK)

IV. Blind Channel Identification Algorithm

We consider here the use of the subspace method SS since it appears to be one of the most robust to noise. A major drawback of the SS method is its high computational cost due to the eigendecomposition of a large dimensional matrix. Fortunately, a computationally attractive subspace method called Minimum Noise Subspace (MNS) has been proposed for MIMO system identification [6]. This method computes the noise subspace via a set of noise vectors which are computed in parallel from a set of combinations of system outputs that form a basis of the noise subspace. The algorithm works as follows. Let $\mathbf{y}(n)$ be a q -variate discrete-time stationary time-series given by

$$\mathbf{y}(n) = \sum_{k=0}^M \mathbf{H}(k) \cdot s(n-k) + \mathbf{w}(n) \quad (2)$$

where $\mathbf{H}(z) = \sum_{k=0}^M \mathbf{H}(k) \cdot z^{-k}$ is an unknown causal FIR $q \times p$ transfer function satisfying $\text{rank}(\mathbf{H}(z)) = p$ for each z and $\text{rank}(\mathbf{H}(M)) = p$ ($\mathbf{H}(z)$ is column reduced), $s(n)$ is a p -dimensional unknown process and $\mathbf{w}(n)$ is an additive q -dimensional white noise i.e.

$$E(\mathbf{w}(l) \cdot \mathbf{w}^*(l)) = \delta(n-l) \cdot \sigma^2 \mathbf{I}_q$$

We introduce the spatio-temporal variables

$$\begin{aligned} \mathbf{y}_N^{(i)} &= \Gamma_N(\mathbf{H}_{i,:}) \mathbf{s}_N + \mathbf{w}_N^{(i)} \\ \mathbf{y}_N &= \Gamma_N(\mathbf{H}) \mathbf{s}_N + \mathbf{w}_N \end{aligned} \quad (3)$$

where

$$\begin{aligned} \mathbf{y}_N^{(i)} &= [y_i(n), \dots, y_i(n+N-1)]^T, \quad \mathbf{y}_N = [\mathbf{y}_N^{(1)T}, \dots, \mathbf{y}_N^{(q)T}]^T, \\ \mathbf{s}_N &= [s_N^{(1)}(n), \dots, s_N^{(p)}(n)]^T, \quad s_N^{(i)}(n) = [s_i(n), \dots, s_i(n+N+M_i-1)]^T \\ \Gamma_N(\mathbf{H}_{i,:}) &= [\Gamma_N(\mathbf{H}_{i,1}), \dots, \Gamma_N(\mathbf{H}_{i,p})] \quad \text{and} \\ \Gamma_N(\mathbf{H}) &= [\Gamma_N(\mathbf{H}_{1,:})^T, \dots, \Gamma_N(\mathbf{H}_{q,:})^T]^T. \end{aligned}$$

$\Gamma_N(\mathbf{H}_{i,j})$ being the $N \times (N+M)$ Sylvester matrix associated with the scalar polynomial $\mathbf{H}_{i,j}(z)$ [6].

Let \mathbf{R}_N be the covariance matrix of $\mathbf{y}_N(n)$:

$$\mathbf{R}_N = E(\mathbf{y}_N(n) \mathbf{y}_N^*(n)) = \Gamma_N(\mathbf{H}) \mathbf{S} \Gamma_N^*(\mathbf{H}) + \sigma^2 \mathbf{I}_{qN} \quad (4)$$

where $\mathbf{S} = E(\mathbf{s}_N(n) \mathbf{s}_N^*(n))$ is assumed to be positive definite. For $N > pM$, the noise-free covariance matrix is rank deficient and its eigen-decomposition is given by:

$$\mathbf{R}_N = \mathbf{U}_s \mathbf{A}_s \mathbf{U}_s^* + \sigma^2 \mathbf{U}_n \mathbf{U}_n^* \quad (5)$$

where $\text{range}(\mathbf{U}_s) = \text{range}(\Gamma_N(\mathbf{H}))$ is the signal subspace (\mathbf{A}_s is the diagonal matrix of the $p(N+M)$ largest eigenvalues of \mathbf{R}_N) and $\text{range}(\mathbf{U}_n) = \text{range}(\Gamma_N(\mathbf{H}))^\perp$ is the noise subspace. The following orthogonality relation

$$\mathbf{U}_n^* \Gamma_N(\mathbf{H}) = 0 \quad (6)$$

is the keystone of the SS method to identify the transfer function $\mathbf{H}(z)$ up to a $p \times p$ constant matrix [6]. The major contribution of [6] is to show that first, only $q-p$ properly chosen noise vectors are as sufficient as the whole noise subspace $\text{range}(\mathbf{U}_n)$ for (6) to yield a consistent estimate of $\mathbf{H}(z)$ and, second, under a mild additional assumption, each of the $q-p$ noise vectors can be found by computing the least eigenvector of a covariance matrix corresponding to a distinct $(p+1)$ -tuples of channel outputs given by a *properly connected sequence* (PCS). The choice of this PCS is crucial to obtain satisfactory performances: the PCS must guarantee that all the system outputs are used in the same way with the same weight importance. This is the case when choosing the following indexes (7).

Note that the first $q-p$ tuples correspond to a PCS and the last p tuples correspond to the additional redundancy we introduce to guarantee that all system outputs are used $p+1$ times. More precisely, the MNS estimation method proceeds as follows (see (8)).

$$\begin{cases} (1, \dots, p+1)(2, \dots, p+2) \dots (q-p, \dots, q) \\ (q-p+1, \dots, q, 1) \dots (q, 1, \dots, p) \end{cases} \quad (7)$$

MNS procedure:

- For each tuple of channel outputs $(m_{i_1}, \dots, m_{i_{p+1}})$, we compute the covariance matrix:

$$\mathbf{R}^{(i)} = \frac{1}{T-N+1} \sum_{n=1}^{T-N+1} \mathbf{y}_i(n) \mathbf{y}_i^*(n) \quad (8)$$

T being the sample size and

$$\mathbf{y}_i(n) = \left[\mathbf{y}_N^{(m_{i_1})}(n)^T, \dots, \mathbf{y}_N^{(m_{i_{p+1}})}(n)^T \right]^T$$

Then we compute its least dominant eigenvector \mathbf{v}_i .

- Let $\mathbf{v}_i = [\mathbf{v}_{i_1}^T, \dots, \mathbf{v}_{i_{p+1}}^T]^T$ where each subvector has the dimension $N \times 1$ (i.e. $\mathbf{v}_{i_k} = [\mathbf{v}_{i_k}(0), \dots, \mathbf{v}_{i_k}(N-1)]^T$ for $k=1, \dots, p+1$). Then define “zero-padded” $q \cdot N \times 1$ vectors

$$\mathbf{v}_i = \begin{bmatrix} \mathbf{v}_{i,1} \\ \vdots \\ \mathbf{v}_{i,q} \end{bmatrix} \quad \text{where } \mathbf{v}_i(k) = \begin{cases} \mathbf{v}_{i_j} & \text{if } k = i_j \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

and form a $qN \times q$ matrix \mathbf{V}_n of the q vectors $\{\mathbf{v}_i\}$, i.e., $\mathbf{V}_n = [\mathbf{v}_1, \dots, \mathbf{v}_q]$.

- Estimate the channel parameter vector:

$$\mathbf{h} = \text{vec} \left(\left[\mathbf{H}_1^T, \dots, \mathbf{H}_q^T \right]^T \right)$$

where $\mathbf{H}_i = [\mathbf{H}_i^T(0), \dots, \mathbf{H}_i^T(M)]^T$ by minimizing the least squares criterion:

$$\hat{\mathbf{h}} = \arg \min \left\| \Gamma_N^*(\mathbf{F}) \mathbf{V}_n \right\|^2 \quad (10)$$

under a suitable constraint.

V. Simulation results

We provide here simulation results for SOSTTC provided in section III in the context of a two-transmit, two-receive antennas system. We compare our results with those of the existing space-time trellis codes in the literature when a comparable code exists. In all simulations, similar to the results in [1], a frame consists of 130 transmissions out of each transmit antenna. We first show the simulation results in the case of ideal CSI (Channel State Information). Fig. 3 shows the simulation results for transmitting 2 bits/s/Hz using a QPSK constellation with ideal CSI. We also include the results for four-state codes form [7] ("YB"), [8] ("BBH"), and [9] ("CYV"). For each couple of transmit-receive antennas the channel is assumed to be a flat quasi-static Rayleigh fading as in [1]. It is clear that the proposed SOSTTC outperforms the other space-time codes by more than 1 dB (CYV). Fig. 4 depicts the results obtained for SOSTTC in the case of MNS tracking use. The loss encountered when compared to the case of perfect CSI is less than 1.5 dB. The lost symbols due to the convergence length of the algorithm which is less than 10 iterations are not taken into account in the computation of the SNR.

VI. Conclusion

We have proposed the construction of new space-time codes named SOSTTC which outperform the classical structures originally proposed by Tarokh & al. This is due to the fact that the SOSTTC combine the coding advantage of STTC with the maximum diversity advantage of STBC. Furthermore, we propose a blind channel identification algorithm which performs very well in a MIMO environment exhibiting losses less than 1.5 dB when compared to the ideal channel state information.

References

[1] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance analysis and code construction", *IEEE Trans. Inform. Theory*, vol. 44, pp. 744-765, Mar. 1998.
 [2] H. El Gamal and A. R. Hammons, Jr, "Algebraic space-time codes for block fading channels", *ISIT* June 2001.
 [3] J. Grimm, M. P. Fitz and J. V. Krogmeier, "Further results on space-time codes for Rayleigh fading", in *Proc. 36th Annu. Allerton Conf.*, Sept. 1998, pp. 391-400.

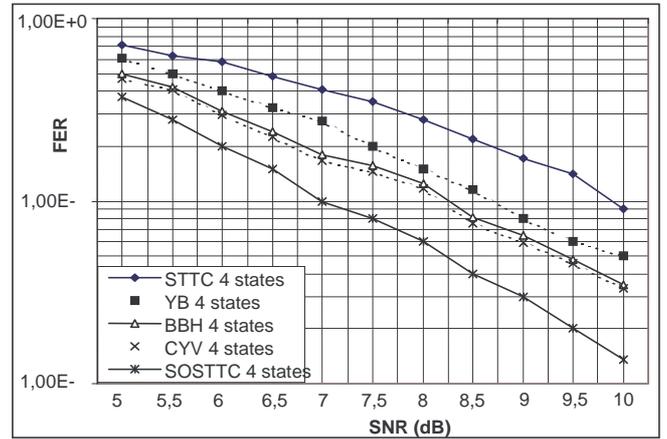


Fig.3: Simulation results for $r = 2$ bits/s/Hz (QPSK) (perfect CSI)

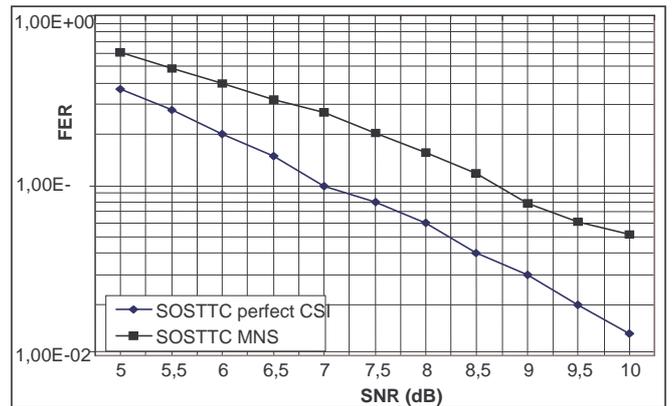


Fig.4: Simulation results for $r = 2$ bits/s/Hz (QPSK) (MNS tracking)

[4] H. Jafarkhani, and N. Seshadri, "Super-Orthogonal Space-Time Trellis Codes", *IEEE Trans. Inform. Theory*, vol. 49, n° 4, pp. 937-950.
 [5] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications", *IEEE Journal Select. Areas. Commun.*, vol. 16, pp 1451-1458, Oct. 1998.
 [6] K. Abed-Meraim and Y. Hua, "Blind identification of MIMO systems using minimum noise subspace", *IEEE Trans. On Signal. Process.*, vol. 45, pp. 254-258, n°1, Jan. 1997.
 [7] Q. Yan and R. S. Blum, "Optimum space-time convolutional codes for quasi-static slow fading channels", in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, vol. 3, Sept. 2000, pp. 1351-1355.
 [8] S. Baro, G. Bauch and A. Hansmann, "Improved codes for space-time trellis coded modulation", *IEEE Commun. Letters*, vol. 4, pp ; 20-22, Jan. 2000.
 [9] Z. Chen, J. Yuan and B. Vucetic, "Improved space-time trellis coded modulation scheme on slow fading channels", *Electron. Lett.*, vol. 37, 440-441, Mar. 2001.

The combination of hyperelliptic curves and block codes for data encryption

I. Jirón

I. Soto

R. Carrasco

Abstract:

An efficient cryptographic system is built, which is based on the combination of a hyperelliptic curve with linear block codes. Its strength is based on the simplicity to perform the data compression through a Jacobian construction. Then, construction data compression criterion is established for select a block code.

Introduction:

Perhaps the only place in the literature where non-linear structures are combined is in chaos and communication systems, with the aim of providing levels of encryption and privacy in the transmission. However, the analysis of the Bit Error Rate (BER) for most of the traditional chaotic schemes (DCSK) fails at approximately 8.5 dB for $BER = 10^{-5}$ in comparison with uncoded BPSK [2].

In 1989, Koblitz presented the idea of using Hyperelliptic Curves (HC), which are non-linear structures as a base to construct cryptosystems [3]. In 1998, Smart compared the performance of cryptosystems based on elliptic curve (EC) and HC, but he concluded that HC do not offer a better performance than the elliptic ones [4]. However, recent work has shown that the HC offer smaller size operands compared with the EC, since the underlying field for a HC is smaller. Therefore, if an 80-bit security level is desired, the underlying field should have an order of approximately 2^{160} with genus 1, 2^{80} with genus 2 and $2^{54} \cdot 3$. The HC also offer faster arithmetic than EC or at least achieves the same computational time] for the same security level.

The main contributions of this paper against Smart's results is to show a simple construction criterion for the combination of a non-linear hyperelliptic curve with a Linear Block Codes (LBC), and provide encrypted transmission and error correction capabilities over an AWGN channel for reasonable level of complexity and data compression to select a block code for a fixed HC.

System description: Figure 1 represents the diagram of the proposed system,

where $\mathbf{d} = \{d_{l-1} \dots d_1 d_0\}$ is a binary sequence of arbitrary length l . The Non-linear Hyperelliptic encrypt/encoder mapping has been built using a HC and a block code. A hyperelliptic curve C of genus $g \geq 1$ over a

finite field F_p is formed by the points $P = (u, v) \in \overline{F_p} \times \overline{F_p}$, which satisfy the Equation (1), with a point ∞ at infinity [9]:

$$C: v^2 + h(u)v = f(u) \quad (1)$$

where h and f are polynomials with $\deg(h(u)) \leq g$, $\deg(f(u)) = 2g + 1$ and $f(u)$ is monic. A semireduced divisor is defined as:

$$D = \sum m_i P_i - \left(\sum m_i \right) \infty \quad (2)$$

where P_i is a point on C and $m_i \in \mathbb{Z}$, $m_i \geq 0$. Additionally, a reduced divisor D satisfies

$\sum m_i \leq g$ on C . A reduced divisor only belongs to a unique equivalence class in the Jacobian of C . The Jacobian is the quotient group $J(C) = \mathcal{D}^0 / \mathcal{P}$, where \mathcal{D}^0 is the group of divisors of order zero on C and \mathcal{P} is the group of the principal divisors on C . \mathcal{P} is a subgroup of \mathcal{D}^0 [3].

Assuming that data message \mathbf{d} can be split in z different data blocks as in Equation (3):

$$\begin{aligned} \mathbf{d} = & \{d^z_{m-1} \dots d^z_1 d^z_0\} \\ & \cup \dots \\ & \cup \{d^2_{m-1} \dots d^2_1 d^2_0\} \\ & \cup \{d^1_{m-1} \dots d^1_1 d^1_0\} \end{aligned} \quad (3)$$

Equation (5) represents the way in which the proposed system produces data compression by the application of a mapping over b_j and a representation in terms of D^j is obtained.

$$\begin{aligned} M_r: B_m & \rightarrow J(C) \\ b_j & \rightarrow M_r(b_j) = D^j \end{aligned} \quad (4)$$

Where $B_m = \{b_j / b_j = \{d^j_{m-1} \dots d^j_1 d^j_0\}\}$ is a generic data block of m bits in agreement with Equation (3). In this paper, case studies 1 and 2 will show how different mappings $M_r(\cdot)$ produce different data compression and performance levels.

Equation (4) represents the Hasse-Weil theorem, which is used to estimate the order of the Jacobian denoted by $\#J(C)$ [3].

$$(\sqrt{q} - 1)^{2g} \leq \#J(C) \leq (\sqrt{q} + 1)^{2g} \quad (5)$$

Where $q = \#F_p$ and represent the cardinality of the Field F_p . Therefore, each equivalence class can be enumerated with an identifier denoted by Id^j , whose numbers are chosen from the range $\{0, 1, 2, 3, \dots, K-1\}$ with $K = (\sqrt{q} + 1)^{2g}$. Therefore each reduced divisor can be addressed with an identifier number Id^j from a look up table in agreement with Equations (5) and (6), which transforms into a binary sequence of k bits denoted by \mathbf{c}^j . Using the Hasse-Weil boundary, Equation (6) represents how to obtain the best k for the block code (n, k) :

$$k = \begin{cases} 2 & K = 1, 2, 3, 4 \\ \lceil \log_2 K \rceil & \forall K, K \geq 5 \end{cases} \quad (6)$$

where n is constructed according to the protection level of the code through the parity bits and the data compression level necessary to maintain a good performance in the system.

Suppose that a LBC (n, k) will be used, which is defined with a generator matrix G with order $k \times n$, such that a message \mathbf{c}^j is encoded using $\mathbf{v}^j = \mathbf{c}^j \cdot G$, where \mathbf{v}^j is a codeword.

The algorithm will produce a binary sequence of k bits denoted by $\hat{\mathbf{c}}^j$, which is transformed to a identifier number $\hat{\text{Id}}^j \in \{0, 1, 2, \dots, K-1\}$, by the using $\hat{\text{Id}}^j$ as a cursor in a look up table to search the divisor \hat{D}^j . In order to get the data block $\hat{\mathbf{b}}_j$ an inverse mapping M_r^{-1} of Equation (4) is applied to \hat{D}^j . Finally by the concatenation of $\hat{\mathbf{b}}_j$ with the previous iteration, denoted by $\hat{\mathbf{d}} \leftarrow \hat{\mathbf{d}} \cup \hat{\mathbf{b}}_j$, an approximation of \mathbf{d} represented by $\hat{\mathbf{d}}$ is obtained. Algorithm 1 summarises the above steps.

Case study:

Let $C_3 : v^2 + uv = u^5 + 5u^4 + 6u^2 + u + 3$ a HC with genus $g=2$ over the finite field. By using Equations (5) and (6) a value $7.34 \leq \#J(C_3) \leq 176.66$ and $k=8$ can be obtained. The reduced divisors can be addressed with a look up table using the second and third columns of table 1. Equation

(7) represents the generator matrix of $(10, 8)$ LBC, which is used to implement the encoder.

$$G = [P^T \ I_{8 \times 8}] \quad (7)$$

Where $P = \begin{bmatrix} 01101100 \\ 10010011 \end{bmatrix}_{2 \times 8}$ and $I_{8 \times 8}$ is the identity matrix.

The first mapping M_1 is defined using the first and second columns in table 1. This function transforms a data block with three bits b_j into an element that belongs to F_8 .

$$M_1 : \quad B_3 \rightarrow J(C_3) \quad (8)$$

$$b_j \rightarrow M_1 \begin{pmatrix} b_j \end{pmatrix} = \begin{pmatrix} u, g(b_j) \end{pmatrix} + 6^\infty$$

Where $g(b_j) \in F_8$.

The mapping M_1 can be applied to Equation (9).

$$\begin{aligned} \mathbf{d} &= 0010011010101001010110 \\ &= 001 \ 001 \ 101 \ 011 \ 010 \ 010 \ 101 \ 110 \quad (9) \\ &= b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8 \end{aligned}$$

When $z=8$, the first data block

$$M_1(001) = (1, g(001)) + 6^\infty = (1, 1) + 6^\infty = D^1,$$

which is addressed with $\text{Id}^1 = 33$, according to the look up table 1. Then, the first identifier Id^1 is transformed into the binary sequence denoted by $\mathbf{c}^1 = 00100001$, which is encoded producing $\mathbf{v}^1 = \mathbf{c}^1 * G = 1100100001$. This codeword passes through the modulator, the AWGN channel and the demodulator producing the first received codeword denoted by $\hat{\mathbf{v}}^1 = 1100100101$, which will be decoded the Viterbi algorithm. Then an estimation of first identifier $\hat{\mathbf{c}}^1 = 00100001$ will be obtained, which will be converted into $\hat{\text{Id}}^1 = 33 \in \{0, 1, 2, \dots, 175\}$, addressing the divisor $\hat{D}^1 = (1, 1) + 6^\infty$. From this the data block $\hat{\mathbf{b}}_1 = 001$ is obtained using an inverse mapping M_1^{-1} . Finally, the data block $\hat{\mathbf{b}}_1 = 001$ is concatenated with $\hat{\mathbf{d}}$, using $\hat{\mathbf{d}} \leftarrow \hat{\mathbf{d}} \cup \hat{\mathbf{b}}_1 = \hat{\mathbf{d}} \cup \{001\}$, and this process will be repeated for each data block b_j with $j = 2, 3, \dots, 8$.

Since the HC is defined on a small field, the reduced divisors restriction can be relaxed in order to embed the data in the semi-reduced divisors and in this way the capacity of the HC

is increased. In this case the data can be split into data blocks of 9 bits. So, a second mapping M_2 , more complex than M_1 , is introduced in Equation (10):

$$\begin{aligned} M_2 : B_3 \times B_3 \times B_3 &\rightarrow \mathcal{J}(C_3) \\ b_i b_j b_k &\rightarrow M_2(b_i b_j b_k) \\ &= (u_i, g(b_i)) + (u_j, g(b_j)) \\ &\quad + (u_k, g(b_k)) - 3^\infty \end{aligned} \quad (10)$$

Equation (11) represents how the three elements from B_3 are embedded into a semireduced divisor defined in Equation (2):

$$\begin{aligned} b_i b_j b_k = 010\ 001\ 111 &\rightarrow M_2(b_i b_j b_k) \\ &= (2,2) + (1,1) + (5,3) - 3^\infty \end{aligned} \quad (11)$$

In this case another look up table must be defined. The nine bit divisors have been addressed using an identifier number represented by a sequence of $k = 8$ bits, introducing a benefit of one bit of compression in regards to M_1 .

Experiment and analysis:

Two experiments have been conducted: the first experiment uses a LBC and the second uses a Reed-Solomon code. In order to compare the new system Figure 2 shows a comparison of BER v/s SNR curves for different LBC against BPSK and DCSK [2].

Assuming that the curve C_3 is fixed, Figure 3 shows the BER for the combination of a HC with two different LBC, $(14,11)$, $(10,8)$. The two codes have been selected according to the number of bits required to map the different reduced divisors of the Jacobian. Of course a different curve will introduce a different level of security.

The $(14,11)$ LBC is represented as a matrix of order 11×14 , where the 11-bits sequences address 2^{11} elements of the Jacobian. As was discussed before, this curve C_3 has at most 176 reduced divisors, but there are 2^{11} identifier numbers available and therefore the combination is more complex for the intruder. These sequences are transformed to 14-bits sequences, but this combination has a worse spectral efficiency since it introduces more bits to address the reduced divisors. In this case, the mapping does not produce data compression. However, this combination produces a gain of 5.90 dB compared with BPSK and a gain of 14.4 dB compared with DCSK for a BER = 10^{-5} .

The $(10,8)$ LBC is represented as a matrix of order 8×10 , where the 8-bits sequences address 2^8 elements of the Jacobian which introduces less security than, the previous case,

since there are less identifier numbers to address the reduced divisors and therefore less complex to the intruder. These sequences are transformed into 10-bit sequences. The mappings M_1 and M_2 are used in this combination, but M_2 only produces the data compression. So this combination has better spectral efficiency compared to the previous one, since it introduces less bits to address the reduced divisors producing a gain of 6.3 dB for a BER = 10^{-5} compared to BPSK and a gain of 15.0 dB compared to DCSK for a BER = 10^{-5} .

Conclusions:

A new cryptographic system has been constructed using two different combinations from a HC and a block code. The strength of this system is formed by the combination of a HC and a block code, which introduces information into the HC for safe transmission, the data compression mapping and how the elements in the Jacobian are embedded in the block code. When this system is compared with a BPSK system an improvement of 5.9 dB to 8.0 dB for a BER = 10^{-5} can be achieved. It is also compared with a DCSK system and an improvement of 15 dB to 16.5 dB for a BER = 10^{-5} can be reached. Additionally, this system approaches Shannon's limit, from 3.1 dB to 1.0 dB for a BER = 10^{-5} . Finally, the obtained criterion to combine a HC with a block code establishes that given a HC a block code must be chosen that maintains the data compression.

Acknowledgment: The authors would like to thank FONDECYT (Project 1030149/2003) Chile for their financial support.

References

- [1] Schimming, T.; Hasler, M. "Optimal detection of Differential Chaos shift keying". IEEE Transactions on circuits and systems-I: Fundamental theory and applications, Vol. 47, N° 12, dec. 2000.
- [2] Hasler, M.; Shimming, T. "Optimal and Suboptimal Chaos Receivers". Proceedings of the IEEE, Vol. 90, N° 5, may 2002.
- [3] Koblitz, N. "Hyperelliptic Cryptosystems". Journal of Cryptology N° 1, pp 139 - 150, 1989.
- [4] Smart, N.P. "On the performance of hyperelliptic cryptosystems". HPL-98-162. September, 1998. Available at www.hpl.com/techreports/98/HPL-98-126.html

[5] Jacobson M., Jr., Menezes A., Stein, A. "Hyperelliptic Curves and Cryptography". Available at: <http://pages.cpsc.ucalgary.ca/~jacobs/publications.html>

[6] Lange, T., "Efficient arithmetic on genus 2 hyperelliptic curves over finite fields via explicit formulae". December 15, 2003. Available at <http://www.itsc.ruhr-uni-bochum.de/tanja>

[7] McLeman, C. "Hyperelliptic curves", August 13, 2001. Available at <http://crypto.math.hmc.edu/papers/Hyper.pdf>

[8] Hess, F.; Seroussi, G.; Smart, N.P. "Two topics in hyperelliptic cryptography". University of Bristol, Department of Computer Science, June 2000. Available at: <http://citeseer.nj.nec.com/357558.html>.

[9] Koblitz, N. "Algebraic Aspect of Cryptography. Algorithms and Computation in Mathematics". ISBN 3-540-63446-0 Springer-Verlag, 1998.

[10] Sklar, Bernard. "Digital Communications Fundamentals and Applications". ISBN 0130847887. Second edition, 2001. Prentice Hall, Inc.

Table 1. Look up table of the mapping M_1 for

$$C_3 : v^2 + uv = u^5 + 5u^4 + 6u^2 + u + 3.$$

m -blocks in bits	Reduced Divisor	Identifier numbers
000	$(0,0) + 6\infty$	100
001	$(1,1) + 6\infty$	33
010	$(2,2) + 6\infty$	90
011	$(2,3) + 6\infty$	210
100	$(6,4) + 6\infty$	112
101	$(1,5) + 6\infty$	120
110	$(5,6) + 6\infty$	57
111	$(5,3) + 6\infty$	40

Figure 2. Comparison of BER v/s SNR curves for different LBC against BPSK and DCSK

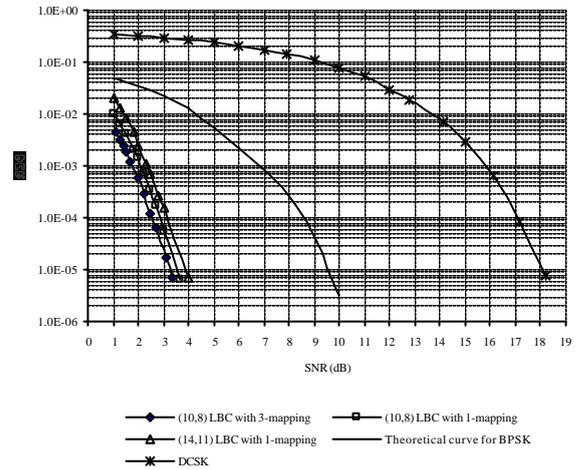
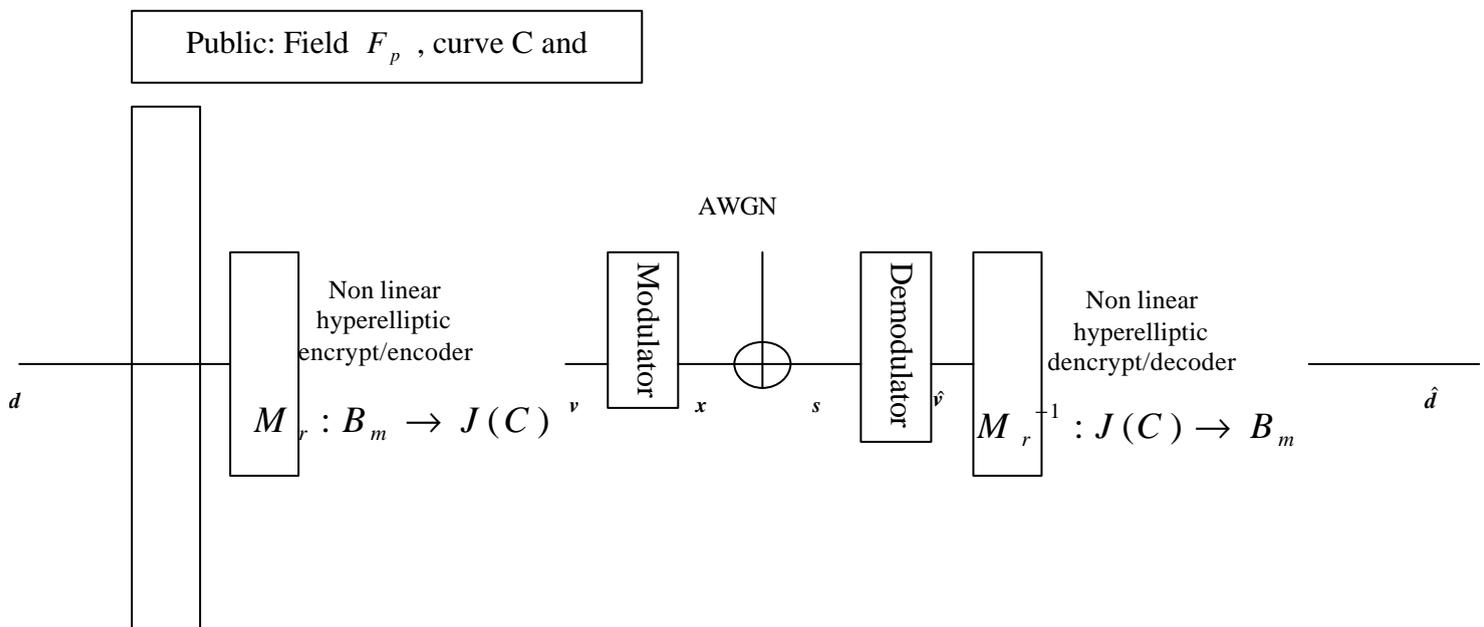


Figure 1. System diagram.



The Effect of Timing and Frequency Offsets on OFDM Systems Using One and Two-Dimensional Modulation Schemes

A. J. Al-Dweik and S. R. Al-Araji
 Department of Communications Engineering,
 Etisalat University, P.O.Box 980, Sharjah, UAE
 (e-mail: dweik@fulbrightweb.org)

ABSTRACT

In this work, the sensitivity of OFDM systems to timing and frequency offsets is investigated using one and two dimensional modulation schemes. Theoretical results confirmed by simulation have shown that one dimensional modulation schemes are more sensitive to synchronization errors compared to two dimensional schemes. The sensitivity of both systems to timing offsets is much higher than their sensitivity to frequency offsets.

Index Terms— OFDM, sensitivity, timing offset, frequency offset

I. Introduction

Orthogonal frequency division multiplexing (OFDM) is a modulation technique that employs N orthogonal subcarriers modulated by N parallel data streams [1]. Compared to single carrier systems, OFDM shows significant immunity in multipath and impulsive noise channels. Due to these advantages, OFDM was adapted as the modulation scheme for audio digital broadcasting (ADB) system [2] and is also proposed as the terrestrial HDTV transport in Europe [3]. Due to its high bandwidth efficiency OFDM was also proposed for several systems over AWGN channels [4].

In an effort to optimize the performance of OFDM systems in various environments, several modulation techniques were proposed in the literature. For example [5] has proposed to combine noncoherent MASK with OFDM for frequency hopping networks, coherent MASK was proposed in [6] using tone spacing of $1/2T_s$ to reduce the system complexity by replacing the conventional two-dimensional modulation techniques such as MPSK and QAM with one-dimensional modulation such as MASK given that no significant bandwidth expansion or bit error rate (BER) degradation will result.

In this work, the performance of OFDM systems was analyzed and compared using one and two-dimensional modulation schemes with imperfect synchronization conditions. The two-dimensional schemes to be considered are QAM and MPSK, the one-dimensional scheme is the MASK with $1/2T_s$ tone spacing proposed by [6].

II. OFDM-MASK Systems

Consider an OFDM system with N subcarriers $C_0 \dots C_{N-1}$, modulated by N parallel data streams with

symbol rate (R_s). The equivalent baseband representation of the transmitted signal during the k th OFDM symbol can be expressed as [6]

$$S(t) = \sum_{i=0}^{N-1} a_i \cos\left(2\pi i \frac{R_s}{2} t\right) \quad (1)$$

where $T_s = 1/R_s$, a_i is the i th data symbol modulating the i th subcarrier during the k th OFDM symbol duration. The k th symbol of subchannel- x ($a_{x,k}$), $0 \leq x \leq N-1$, can be extracted by correlating the received signal ($r(t)$) with C_x ,

$$a_{x,k} = \frac{1}{T_s} \int_0^{T_s} r(t) C_x(t) dt \quad (2)$$

assume initially that the channel is perfect and all signal parameters are known at the receiver except for the time delay (ε) and a frequency offset (Δf). A crucial condition for (2) to be true is the knowledge of the beginning and the end of the symbol k (symbol timing) and the exact carrier frequency. In the absence of such information, the correlator window will be working partially on two consecutive symbols, at the same time the baseband symbols will have residual frequency offset, thus

$$\hat{a}_{x,k} = \frac{1}{T_s} \int_{\varepsilon}^{T_s} S_1(t, \Delta f) C_x(t) dt + \frac{1}{T_s} \int_{T_s}^{T_s+\varepsilon} S_2(t, \Delta f) C_x(t) dt \quad (3)$$

where $\hat{a}_{x,k}$ is an estimate of $a_{x,k}$, however, if ε and Δf are equal to zero, then $\hat{a}_{x,k}$ and $a_{x,k}$ will be identical. The signals $S_1(t)$ and $S_2(t)$ represent the OFDM signal during the time periods of the symbols $a_{x,k}$ and $a_{x,k+1}$ respectively.

Since subchannel- x is the subchannel of interest, then $a_{x,k}$ can be recovered by substituting $C_x = 2 \cos(x \frac{R_s}{2} t)$ in (3). Thus

$$\hat{a}_{x,k} = \sum_{i=0}^{N-1} a_{i,k} \left(\frac{\gamma_{i,x}}{Z_{i,x}} + \beta_{i,x} \right) - a_{i,k+1} \beta_{i,x} \quad (4)$$

where $\gamma_{i,x} = (-1)^{i+x} \sin(2\pi u)(-4\mu - 2i)$, $\beta_{i,x} = (\sin(\pi \Delta V) \bar{V} + \sin(\pi \Delta \bar{V}) V) / Z_{i,x}$, Δ , the relative timing offset, is equal to ε/T_s , $V = 2\mu + i + x$, $\bar{V} = 2\mu + i - x$, $Z = \pi(-i^2 - 4i\mu - 4\mu^2 + x^2)$, and μ , the

relative frequency offset, is equal to $\frac{\Delta f}{R_s}$. Taking the term that corresponds to $x = i$ out of the summation, thus

$$\hat{a}_{x,k} = a_{x,k}\alpha_{x,x} - a_{x,k+1}\beta_{x,x} + \sum_{\substack{i=0 \\ i \neq x}}^{N-1} a_{i,k}\alpha_{i,x} - a_{i,k+1}\beta_{i,x} \quad (5)$$

where $\alpha_{i,x} = \frac{\gamma_{i,x}}{Z_{i,x}} + \beta_{i,x}$, the second term in (5) represents the ISI and the third term represents the inter channel interference (ICI). Therefore, the received signal in AWGN channels can be expressed as

$$\hat{a}_{x,k} = a_{x,k}\alpha_{x,x} + \dot{n} + n, \quad \Delta \ll 0.5 \quad (6)$$

where $\dot{n} = ISI + ICI$ which represents the distortion resulted from the loss of orthogonality between subcarriers, and n is an AWGN with two sided power spectral density of $N_0/2$. In order to evaluate the statistical properties of \dot{n} , further assumptions are necessary. Particularly, it will be assumed that $E\{a_i\} = 0$ and $E\{a_i a_j\} = 0$ [9], that is the modulation symbols will have zero mean and are uncorrelated. Therefore, for small values of Δ and μ , and large values of N the central limit theorem can be applied where the interference distortion (\dot{n}) can be approximated by a normally distributed random variable with zero mean and a variance ($\sigma_{\dot{n}}^2$) that can be expressed as

$$\sigma_{\dot{n}}^2 = \beta_{x,x}^2 P_{av,x} + \sum_{\substack{i=0 \\ i \neq x}}^{N-1} P_{av,i} (\alpha_{i,x}^2 + \beta_{i,x}^2) \quad (7)$$

where $P_{av,i} = E\{a_{i,k}^2\} = E\{a_{i,k+1}^2\}$ which is the individual OFDM symbol average power. For an MASK system that has a minimum distance of $2d$ between adjacent symbols, $P_{av,MASK} = \frac{d^2}{M} \sum_{m=1}^M (2m-1-M)^2$ [7]. The multiplicative term ($\alpha_{x,x}$) represents the attenuation. Notice that the most sensitive channel is the one located in the center of the spectrum since it will be affected by the interference from both sides. Since N is an even number, subchannels $\frac{N}{2}$ and $\frac{N}{2} - 1$ will be the most sensitive, any other subchannel will be affected less than these channels.

III. OFDM-MPSK, QAM

OFDM with QAM or MPSK will have N ($2L$) complex subcarriers $C_{-L} \dots C_{L-1}$, modulated by N complex parallel data streams with symbol rate (R_s). The equivalent complex baseband representation of the transmitted signal during the k th symbol can be expressed as [5][7]

$$S(t) = \sum_{i=-L}^{L-1} a_i e^{jR_s i t} \quad (8)$$

where $T_s = 1/R_s$, a_i is the complex data symbol modulating the i th subcarrier during the k th OFDM symbol duration, $a_i = a_{i,I} + j a_{i,Q}$. The real and imaginary parts of $S(t)$ represent the inphase and the quadrature components respectively. The k th symbol of subchannel- i ($a_{i,k}$)

can be extracted from (8) by correlating the received signal ($r(t)$) with C_i ,

$$a_{i,k} = \frac{1}{T_s} \int_0^{T_s} r(t) C_i(t) dt \quad (9)$$

In AWGN channels, the received signal with frequency offset can be expressed as

$$r(t) = \sum_{i=-L}^{L-1} a_i e^{jR_s(i+\mu)t} + n(t) \quad (10)$$

It should be emphasized that (10) represents the complete OFDM symbol when CP is not used, or the useful part of the OFDM symbol when the CP is used [10]. Without loss of generality it is assumed that subchannel-0 is the subchannel of interest, thus, following the same procedure used for MASK and noticing that $C_0 = 1$ for subchannel-0, the inphase and quadrature components can be expressed as

$$\hat{a}_{0,k} = \sum_{i=-L}^{L-1} a_{i,k} \chi_{1,i} + a_{i,k}^* \chi_{2,i} + a_{i,k+1} \chi_{3,i} + a_{i,k+1}^* \chi_{4,i} + n \quad (11)$$

where a^* is the complex conjugate of the data symbol a , $\chi_{1,i} = \frac{\sin(2\pi\mu) - \sin(2\pi\Delta(i+\mu))}{2\pi(i+\mu)}$, $\chi_{2,i} = \frac{\cos(2\pi\Delta(i+\mu)) - \cos(2\pi\mu)}{2\pi(i+\mu)}$, $\chi_{3,i} = \frac{\sin(2\pi\Delta(i+\mu))}{2\pi(i+\mu)}$ and $\chi_{4,i} = \frac{1 - \cos(2\pi\Delta(i+\mu))}{2\pi(i+\mu)}$. separating subchannel-0

$$\hat{a}_{0,k} = a_{0,k} \chi_{1,0} + a_{0,k+1}^* \chi_{2,0} + a_{0,k} \chi_{3,0} + a_{0,k+1}^* \chi_{4,0} + n_{ICI} + n \quad (12)$$

where n_{ICI} is similar to (11) except that n and the terms that correspond to $i = 0$ are omitted. For simplicity $\hat{a}_{0,k}$ can be expressed as

$$\hat{a}_{0,k} = a_{0,k} \chi_{1,0} + \dot{n} + n \quad (13)$$

where $\dot{n} = ISI + ICI$ and its variance ($\sigma_{\dot{n}}^2$) can be expressed as [8],

$$\sigma_{\dot{n}}^2 = \frac{P_{av,0}}{2} \sum_{j=2}^4 \chi_{j,0}^2 + \sum_{\substack{i=-L \\ i \neq 0}}^{L-1} \frac{P_{av,i}}{2} \sum_{j=1}^4 \chi_{j,i}^2 \quad (14)$$

where P_{av} is the average symbol power, for MPSK $P_{av,MPSK} = d^2 / \sin^2(\pi/M)$ and for rectangular QAM $P_{av,QAM} = P_{av,MASK}(\sqrt{M})$ given that \sqrt{M} is an integer, i.e., P_{av} for an M -QAM is equal to P_{av} for \sqrt{M} -ASK. Thus, for a fixed d , M -QAM and \sqrt{M} -ASK will have the same BER performance since they have the same average power [6].

IV. Performance Degradation

Performance degradation is usually measured by calculating the overall signal-to-noise ration (SNR). The sensitivity of OFDM using two-dimensional modulations to frequency offsets was discussed in [9], a lower bound

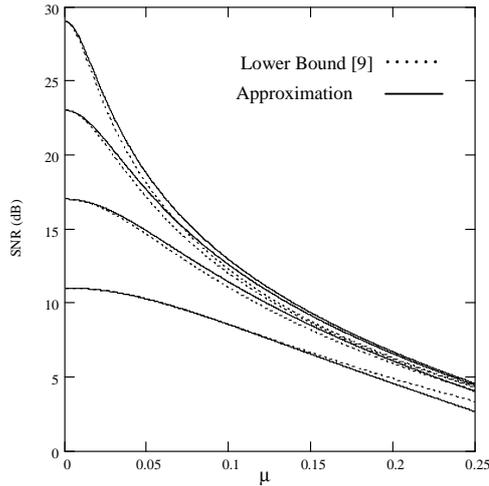


Fig. 1. Lower bound and approximation of the SNR versus μ

was derived for the SNR at the output of the receiver, symbol timing is assumed to be perfect. In this work we have derived an approximation for the SNR and included the effect of timing offsets given that no CP is used which is usually the case for flat channels. If CP is used and the timing offset is smaller than the CP duration, the timing offset effect can be suppressed by substituting $\Delta = 0$ in the approximation.

Equations (7) and (14) can be used to give an approximation for the SNR. Thus the SNR for OFDM using two-dimensional modulations [9]

$$SNR \approx \frac{\chi_{1,0}^2 (E_{s,0}/N_0)}{1 + \frac{E_{s,0}}{2N_0} \sum_{j=2}^4 \chi_{j,0}^2 + \sum_{\substack{i=-L \\ i \neq 0}}^{L-1} \frac{E_{s,i}}{N_0} \sum_{j=1}^4 \frac{\chi_{j,i}^2}{2}} \quad (15)$$

if all subcarriers have equal energy [9], then

$$SNR \approx \frac{\chi_{1,0}^2 (E_s/N_0)}{1 + \frac{E_s}{2N_0} \left(\sum_{j=2}^4 \chi_{j,0}^2 + \sum_{\substack{i=-L \\ i \neq 0}}^{L-1} \sum_{j=1}^4 \chi_{j,i}^2 \right)} \quad (16)$$

Equation (16) was compared with the lower bound derived in [9], Fig. 1 shows the results obtained using (16) and the results obtained in [9] for values of E_s/N_0 that equal to 11, 17, 23 and 29 dB using 8-PSK modulation. The approximation derived in this work gives roughly 0.7 dB improvement in the SNR estimation for values of μ less or equal to 0.25.

Following the same procedure for one-dimensional modulations,

$$SNR \approx \frac{\alpha_{x,x}^2 (E_s/N_0)}{1 + \frac{E_s}{N_0} \left(\beta_{x,x}^2 + \sum_{\substack{i=0 \\ i \neq x}}^{N-1} (\alpha_{i,x}^2 + \beta_{i,x}^2) \right)} \quad (17)$$

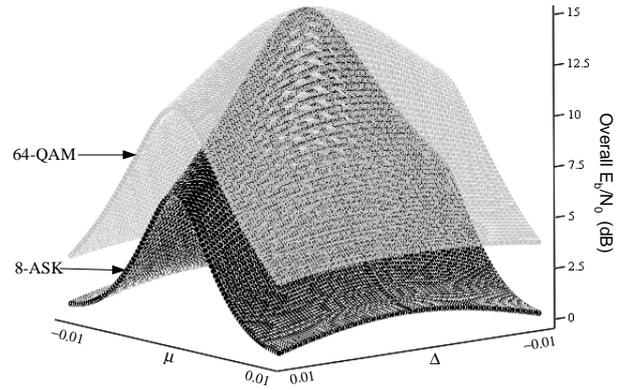


Fig. 2. The over all E_b/N_0 for an 8-ASK and 64-QAM systems, $E_b/N_0 = 15$ dB

Using (16), (17) and assuming that $\Delta = 0$, the maximum relative frequency offset allowed to maintain a signal-to-interference ratio greater than 20 dB equals 2% and 4% for one and two-dimensional modulations respectively. However, if the system does not use the CP or the timing offset exceeds the CP duration, extra degradation will result from the timing offset. Hence, both μ and Δ will contribute to the interference and they have to share the available tolerance. For example, if $\Delta = 0.34\%$ the SNR value is equal to 20 dB and no additional interference is permitted as a result of frequency offset.

In this work, the losses in E_b/N_0 (E_b is the individual OFDM bit energy) is considered as the metric for the system sensitivity to synchronization offsets, the overall E_b/N_0 (\tilde{E}_b/\tilde{N}_0) for one and two-dimensional modulations can be expressed as

$$\frac{\tilde{E}_b}{\tilde{N}_0} = \frac{\lambda^2 E_b}{2(\sigma_n^2) + N_0} \quad (18)$$

where λ is the attenuation and σ_n^2 is the variance of the interference, this equation applies for both one and two-dimensional schemes. Fig. 2 shows \tilde{E}_b/\tilde{N}_0 as a function of μ and Δ for an 8-ASK and 64-QAM systems. It is clear that the sensitivity of OFDM systems to timing offsets is much higher than the sensitivity to frequency offsets regardless of the modulation scheme used.

To compare the performance of the three modulation schemes, an E_b/N_0 of 12 dB was used as an initial operating point for all techniques, the QAM and the \sqrt{M} -ASK will have the same BER (1.38×10^{-4}) value in ideal conditions while MPSK will have higher BER (7×10^{-3}). As the value of Δ or μ is increased the BER will increase accordingly as shown in Fig. 3 and Fig. 4. However, the rate at which the BER of \sqrt{M} -ASK increases is much higher than those of the QAM and the MPSK. For this reason, at a certain values of Δ and μ a crossover point will occur where the MPSK will have lower BER values compared to the BER values of the \sqrt{M} -ASK modula-

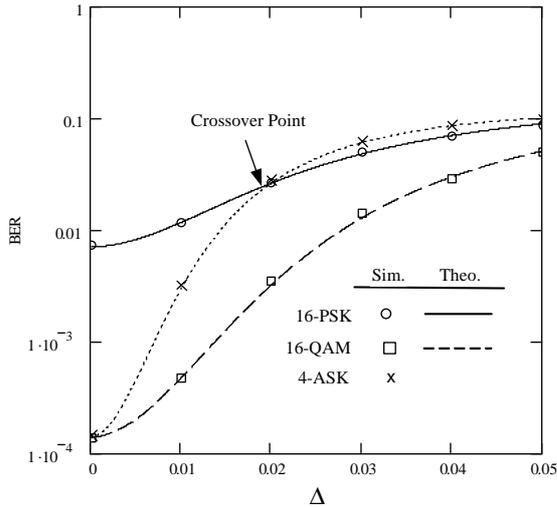


Fig. 3. BER as a function of Δ , $\mu = 0$, $N = 16$, $E_b/N_0 = 12$ dB

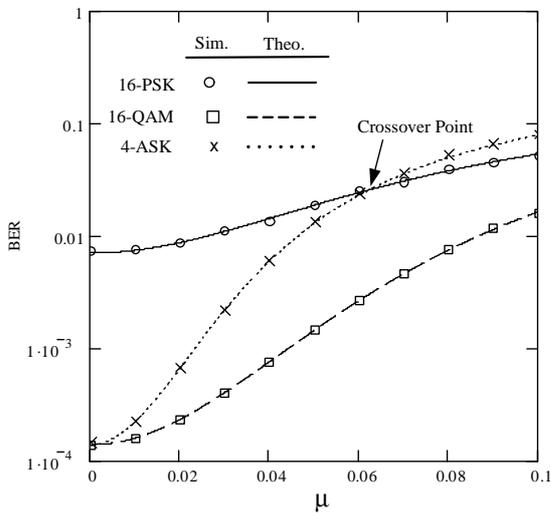


Fig. 4. BER as a function of μ , $\Delta = 0$, $N = 16$, $E_b/N_0 = 12$ dB

tion. This point occurs for several combinations of Δ and μ . The crossover point is shown in Fig. 5 for 256-PSK/16-ASK and 64-PSK/8-ASK at E_b/N_0 of 10 and 15 dB. This figure shows that increasing the modulation order or the system E_b/N_0 will decrease the values of μ and Δ at which the crossover point occurs.

V. Conclusion

The sensitivity of OFDM systems using one and two-dimensional modulation schemes was analyzed in imperfect synchronization conditions. Theoretical results, confirmed by simulations have shown that OFDM systems sensitivity to timing errors is much higher than their sensitivity to frequency offsets. The performance degradation of OFDM systems using one-dimensional modulation schemes is greater than the degradation resulted for OFDM systems with two-dimensional schemes. Although \sqrt{M} -ASK will have better BER performance with perfect synchronization conditions, MPSK will have

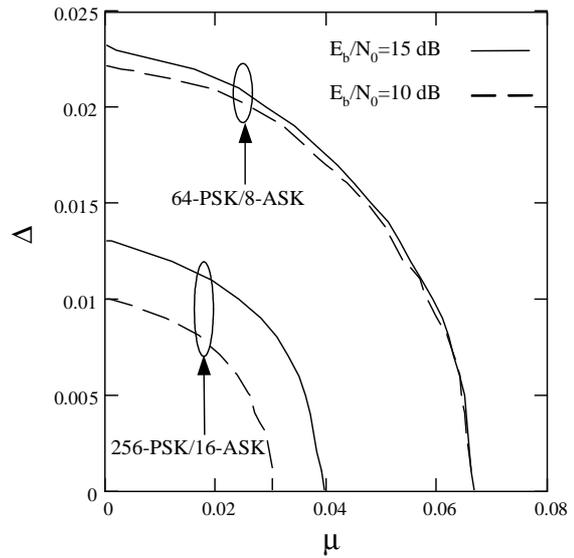


Fig. 5. Crossover point for 16-ASK and 256-PSK at 10 and 15 dB E_b/N_0

a slightly better performance in practical applications since the crossover point occurs within the range of normal operation for practical systems.

References

- [1] R. Nee and R. Prasad, *OFDM Wireless Multimedia Communications*. Boston: Artech House, 2000.
- [2] B. Floch, R. Lassale, D. Castelain, "Digital sound broadcasting to mobile receivers," *IEEE Trans. Consumer Electron.*, vol. 35, pp. 493-503, Aug. 1989.
- [3] P. Tourtier, R. Monnier and P. Lopez, "Multicarrier modem for digital HDTV terrestrial broadcasting," *Signal Processing, Image Commun.*, vol.5, pp. 379-403, Dec. 1993.
- [4] L. Wald, "Feasibility Activities Completed for the Direct Data Distribution Experiment," <http://www.grc.nasa.gov/WWW/RT1999/6000/6150wald2.html>.
- [5] A. Al-Dweik and Fuqin Xiong, "Frequency-Hopped Multiple Access Communications with Noncoherent M-ary OFDM-ASK," *IEEE Trans. Commun.*, Vol. 51, No. 1, January 2003.
- [6] F. Xiong, "M-ary amplitude shift keying OFDM system Communications", *IEEE Trans. Commun.*, Volume: 51, Issue: 10, Oct. 2003.
- [7] J. Proakis, *Digital Communications*, New York: McGraw-Hill, 1989.
- [8] A. J. Al-Dweik, "Robust Non Data-Aided Symbol Synchronization Technique for OFDM Systems". The IEEE Military Communications Conference, Boston, MA, USA, October 2003, pp. 1219-1224.
- [9] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Tans. Commun.*, vol. 42, no. 10, Oct. 1994.
- [10] T. Schmidl and D. Cox, "Robust Frequency and Timing Synchronization for OFDM," *IEEE Trans. Commun.*, Volume: 45, no 12, Dec 1997, pp. 1613-1621.

The Evolution of a Wireless RFID and a Cell Phone

Katariina Penttilä^{*}, Reino Aarinen^{**}, Lauri Sydänheimo, Markku Kivikoski
 Tampere University of Technology, Institute of Electronics, Rauma Research Unit, Kalliokatu 2, FIN-26100
 Rauma, Finland. (^{*}katariina.penttila@tut.fi
^{**}Hollming Ltd. Electronics, P.O. Box 14, FIN-26101 Rauma, Finland.

Abstract— The evolutionary development of RFID (Radio Frequency Identification) technology has moved closer to handheld readers. The technology has begun to spread through global markets and therefore needs to be made more convenient and suitable for everyday use.

To avoid the growth of separate terminals this paper presents the concept of the multi-terminal, consisting of a common RF (Radio Frequency) block of an RFID reader and a cell phone. The use of one multi-terminal facilitates everyday life, making movement easier and the technology more convenient.

In order to realise the integration, the basic structure of both RF blocks are discussed and analysed. The fusion turns out to be feasible despite the existence of several differences between the functionalities of these two systems.

Index Terms—Cell phone, GSM, Multi-terminal, RF block, RFID, RFID reader

1. Introduction

With the use of radio frequency identification (RFID) standards becoming more established, they will be applied more commonly in the future, and not only to tailored special applications as is the case at present. Moreover, the increase in the reading distance as well as the simultaneous reading of several tags will make it possible to use the tags in an ever growing number of applications. As the tags and readers as well as the technology used in them become less expensive, they will become more readily available to consumers.

An integrated RFID reader will further enhance the multifunction terminal capability of a mobile phone. Current technology makes it possible already to combine the similar functions of the RFID reader and the radio technology used in the mobile phones. This article provides an insight into this issue.

The research work done at the RFID reader area is based on MIT (Massachusetts Institute of Technology) Auto-ID Center's RFID specifications of class I RFID tags at 860-930 MHz band, published on 14 November 2002 [1]. Information about cell phones is based on both academic articles and commonly known transceiver structures explained by the relevant literature [2]-[6].

This paper is organised as follows: section 2 introduces the communication principles of an

RFID reader and a cell phone. Section 3 presents simplified diagrams of the RF (Radio Frequency) blocks of both technologies. Requirements of the evolution is analysed in section 4. Finally, section 5 provides brief conclusions on the points discussed in this paper.

2. Communication principles

A. Radio frequency identification

RFID systems consist of RF transponders, namely tags, RF transceivers, namely readers, and a data processing system. A general tag implementation consists of a microchip storing data and implementing the tag's functionality, an antenna and a substrate for the tag. The cheapest RFID tags with the largest commercial potential are passive, harvesting energy from the reader's communication signal to power their operation and communication with the reader.

A general reader implementation consists of a RF module, a control unit and an antenna. The basic structure of data communication between an RFID reader and a tag is shown in Fig.1.[7]. The reader sends a data command to the tags, which reply by modulating the backscattering field. The tag carries out modulation by switching its load resistor on and off. Reader uses amplitude shift keying (ASK) and FM0 coding for data communication. Table I presents modulation parameters of the specification [1].

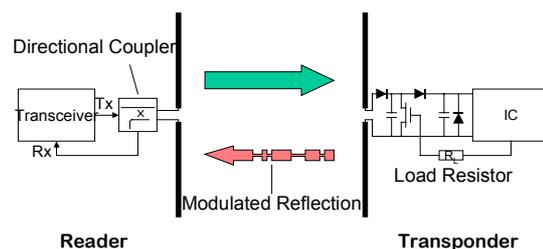


Fig.1 Communication between an RFID reader and a tag

Table I The modulation parameters of RFID technology [1]

Modulation parameter	Frequency region I (Europe)	Frequency region II (North America)
Data rate from reader to tag	15 kbps	70,18 kbps
Data rate from tag to reader	30 kbps	140, 35 kbps
Clock rate	66,67 ms ± 1% max.	14,25 ms ± 1% max.
Modulation depth	50 %	90 %
Rise/fall time	300 ns	300 ns

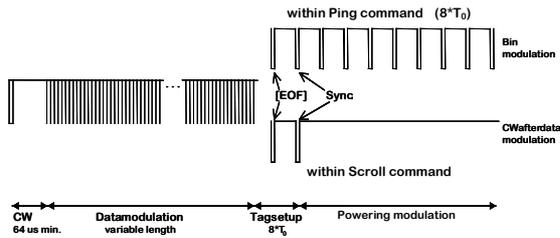


Fig.2 The structure of the reading commands [1]

Tag identification is based on an EPC™ code, defined by MIT Auto-ID Center. It consists of 16 bits cyclic redundancy check code, an identification code with a variable number of bits, within the cheapest tag solutions 64 or 96 bits.

The specification defines both single and multiple tag reading methods, presented in Fig.2. Scroll command requests the tag's identification code in whole and if several tags are in the reading range, the reading operation is likely to be unsuccessful. Ping command solves this problem by going through all possible code combinations with the minimum number of reading cycles as a binary tree.[1]

B. Cell phone

A cell phone network consists of several mobile phones and only few base stations. Both ends of the system are active, having a battery or wired power connection.

A general cell phone implementation consists of an RF-front-end, a base band with DSP and memory unit, a SIM (Subscriber Identity Module) card unit, a display, a keyboard and a power management unit. More in-depth structure is manufacturer dependant and therefore the research is based on an example of a cell phone implementation.

It is important to notice that cell phones modulates the carrier commonly either with quadrature shift keying or with Gaussian minimum shift keying, which both differ from the simple ASK used with RFID. [5]

3. RF block diagram examples

RF block integration of two different systems requires in-depth analysis of their system level operation. Several functions can be carried out in the same block, but some; such as modulation and coding must take place in separate blocks. In addition, the antenna connection, i.e. the use of a duplexer, directional coupler, circulator or simple switch methods differs from each other.

In the remainder of this section, the basic RF block diagrams of both an RFID reader and a cell phone are presented and analysed.

A. RFID reader

In order to minimize RFID transceiver's components, a direct conversion topology is considered. Fig.3 presents a basic RF block diagram of an RFID reader [8]. Due to ASK modulation, phase information and therefore quadrature channelisation is not needed.

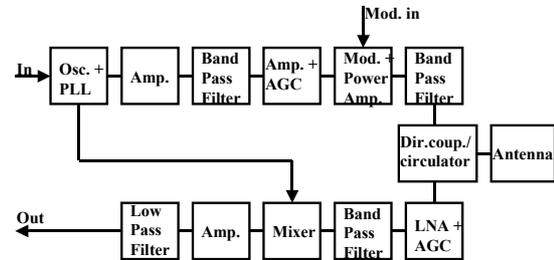


Fig.3. An example of a RFID reader RF block diagram

The transmitter arm consists of a local oscillator (LO) block and few amplification and filtering stages.

After first stage amplification and filtration, the signal is amplified with an AGC amplifier (Automatic Gain Control) to control the signal intensity level. All amplifiers, except the power amplifier of the transceiver should have both high intercept points, in the order of 30 dB, and low NF (Noise Figure), in the order of 1,5 dB, to have sufficient linearity.

Next, the signal is modulated and amplified with a power amplifier to guarantee that the transmission signal level is high enough. The NF of the power amplifier is allowed to have higher NF, in the order of 3 dB, since high gain is now a priority need. Finally the signal is filtered, again with a band pass filter and fed via a directional coupler to the antenna.

Due to the relatively long wavelength of an RFID system at UHF (Ultra High Frequency) range compared to the communication distance, the communication to both transmission and receiving arms must take place at the same time. The use of a switch is therefore not possible. Both a directional coupler and a circulator offer the possibility of full duplex communication, having low losses, an order of 0,15 dB. Directional couplers are preferred due to their generally higher isolation. The total gain of the transmitter arm should to be an order of 30 dB.

At the receiver arm the signal is received via an antenna and the directional coupler by a LNA (Low Noise Amplifier) with AGC. In order to separate the weak incoming signal from noise, NF of the LNA must be low enough (an order of 1 dB).

After the LNA a narrow band pass filter filters the signal so that the third order harmonics will be blocked out. Filter losses should be less than 1 dB.

Finally the carrier is down converted to base band. The conversion loss of the mixer should be around 5dB to maintain its sufficient quality factor. Due to the relatively high transmission gain needed, the intercept point should be as high as possible. Today several available mixers have the intercept point around at 23 dBm at UHF range. In order to amplify the signal level up to 0 dBm, the total gain of the receiver arm should be at least 70 dB.

Patch antennas are commonly used as reader antennas due to their easy implementation and mounting. However, a patch antenna has a wide radiation pattern and is not therefore optimal choice for handheld readers. In order to use a reader as a pointer, the antenna should have relatively high directivity. [8]

B. Cell phone (GSM)

Two commonly known RF block topologies of a cell phone are a superheterodyne transceiver and a direct conversion transceiver. This paper focuses on the direct conversion transceiver, due to its growing popularity and simpler implementation and suitability for application at hand. Fig. 4 introduces PCB and a dual-band antenna structure of a cell phone (Nokia 3310). The RF components show that it is a direct conversion transceiver.

In order to realize modulation methods used in cell phones, the phase information must be stored throughout the frequency conversion. Therefore I/Q channels for both transmission and receiving arms are needed. Typically, the phase difference is 90° .

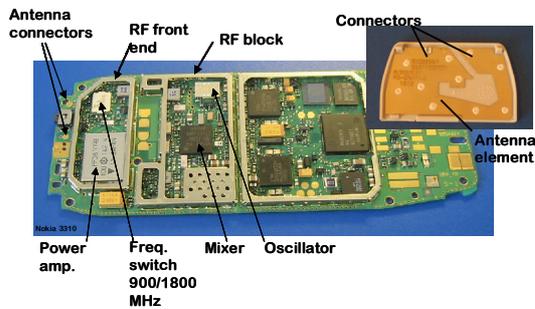


Fig.4 An example of a GSM PCB and a dual-band antenna element

A basic block diagram of GSM direct conversion receiver is presented in Fig.5 [5]. Both transmitter and receiver arm consist of two amplification and filtration stages. Unlike with RFID reader, there is now no need for full duplex data transmission due to the relatively long distance compared to the wavelength used, and therefore a switch between the transmitter and the receiver is enough. [8]

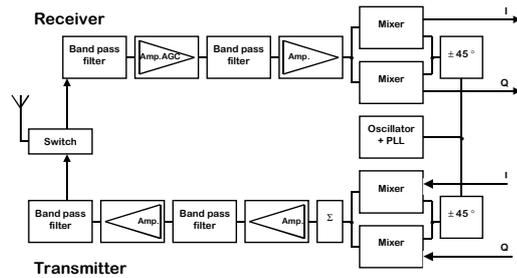


Fig.5. An example of a direct conversion GSM RF block diagram

4. Requirements of the evolution

In order to develop a profitable mobile multi-terminal, several requirements must be realised. Due to guaranty of market sales, the terminal implementation must be low cost having small size and weight. The number of common components must be maximized to minimize cost, size and heat production.

Fig. 6 presents an RF block example of the multi-terminal. Switches connect the common and the separate parts of the RF block for reasons of cost and simplicity. The different frequencies of the systems are generated at the oscillator/PLL block, which consists of two separate feedback loops, one for the RFID reader and another for the cell phone, controlled by switches. Due to different modulation methods of the RFID reader and the cell phone, isolated blocks, separated by switches, realize the required modulations. The switch between the antenna and the RF front end of a cell phone must be replaced with a directional coupler (or a circulator) to guarantee the functionality of the RFID reader of the multi-terminal. This solution is clearly simpler than a switch on the both sides of the cell phone's switch and the RFID reader's directional coupler. All the switches of the multi-terminal are controlled by a background system, which is automatically controlled when operating with the user interface.

Transmission power of both an RFID reader and a cell phone is typically around 30 dBm. The theoretical receiving power of the RFID reader typically varies between -70 dBm and -30 dBm, while the sensitivity of a cell phone is more than -100 dBm. The sensitivity of the cell phone is therefore adequate considering the RFID reader.

The amplifiers used for the multi-terminal must be specified carefully. The power amplifier at the end of the transmission arm must amplify the signal intensity as much as the technical specification allow to guarantee the maximum efficiency and the maximum identification range for RFID reader. The LNA, at the receiving arm, is needed to amplify the very low level received signals without adding noise to the signal, which can reduce the receiver sensitivity. By choosing the LNA with sufficient

features the SNR (Signal-to-Noise Ratio) of the RF front end will be kept relatively high. The use of AGCs on both arms of the transceiver is needed to maintain the wide range of signal intensity levels at a constant level.

The last filter stage at the receiver arm in Fig.6 separates the signal from all noise signals and therefore has an extremely narrow pass band. This operation cannot be carried out at an earlier stage since all these filters are common for both the RFID reader and the cell phone and therefore needed to have relatively large pass band to guarantee communication for both systems.

The most important requirements of the multi-terminal's antenna are sufficient directivity and gain, feasibility and relatively small size. Due to the size requirements, only one antenna element with multiple frequency bands is recommended. The requirement of the antenna size is at present hard to implement, however, with new antenna solutions and new materials it ought to become possible.

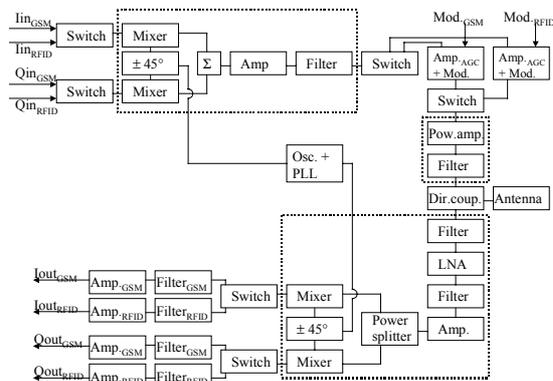


Fig.6 An example of a RF block of the multi-terminal

5. Conclusion

This paper has addressed the evolutionary integration of the RFID reader and the cell phone. An overview of both an RFID reader and a cell phone RF block are presented and analysed. The concept of the multi-terminal has arisen from the idea of a common RF block to meet both the requirements of user and manufacturer, in terms of convenience and practicality.

Integration of several technologies into same cell leads to more convenient, user-friendly terminal. Item identification will grow dramatically in the future due to conscious consumers and more efficient processes. The combination of an RFID reader and a mobile phone as described above could be used, for example, for supply chain management. Identification could assist a cell phone's web service to obtain the item related information from www to bring the item individuality close to each participants of the chain. The revolution of automated object identification is on the way but still requires more research and

development work before handheld RFID readers will identify the objects in their environment in their entirety.

In the future scenario fast digital signal processors will replace an ever greater part of the radio technology used in the mobile phone, whereby every last complexity connected with the integration of analogue functions, as described in this article, will be diminished or eliminated.

Acknowledgement

The authors would like to thank the Finnish National Technology Agency and Hollming Ltd. Electronics for financing the research work done for this paper.

References

- [1] Auto-ID Center, Massachusetts Institute of Technology. Technical Report: 860 MHz - 930 MHz Class I Radio Frequency Identification Tag Radio Frequency & Logical Communication Interface Specification. Version 1.01s. 14th November 2001
- [2] D. Leenarts, J. van der Tang, C. Vaucher. Circuit Design for RF Transceivers. (Kluwer Academic Publisher 2001)
- [3] B. Razavi. RF Microelectronics (Prentice Hall PTR 1998)
- [4] J. Wiart, C. Dale, A. Bosisio, A. Corne. Analysis of the influence of the power control and discontinuous transmission on RF exposure with GSM mobile phones. *IEEE Transactions on Electromagnetic Compatibility, CNET*. Issy-les-Moulineaux. Vol.42, Issue: 4, Nov 2000
- [5] B. Henshaw. Design of an RF transceiver. *IEE Colloquium on Analog Signal Processing* (Ref. No. 1998/472), TTP Commun. 10/28/1998 Oxford, UK
- [6] T. Stetzler, I. Post, J. Havens, M. Koyama. A 2.7 V to 4.5 V single-chip GSM transceiver RF integrated circuit. *IEEE International Solid-State Circuits Conference, 1995*. Digest of Technical Papers. 42nd ISSCC, 1995, San Francisco, CA, USA
- [7] K. Finkenzeller, RFID Handbook (John Wiley & Sons Inc, 1999)
- [8] S. Reynolds, B. Floyd, T. Beukema, T. Zwick, U. Pfeiffer, H. Ainspan. A Direct Conversion Receiver Integrated Circuit for WCDMA Mobile Systems. *IBM Journal on research & development Vol. 47 no. 2/3 March/May 2003*.

The use of clustering methods for combining equalization and decoding in the GSM system

Adina Burian and Markku Renfors
Tampere University of Technology, Institute of Communications Engineering
Korkeakoulunkatu 1, FIN-33720, Tampere, Finland
E-mail: adina @cs.tut.fi

Abstract: Traditionally, channel equalization and decoding are carried out separately, in cascade. This paper investigates the possibility of performing jointly equalization and decoding, using clustering methods. The received data is clustered around specific points, determined by the coding type and the spread of the channel, and a Viterbi-type procedure is applied in order to find out these clusters and the connections between them.

This method is investigated in a realistic environment, the Global System for Mobile Communications (GSM). Separate equalization and decoding is used as benchmark, through two conventionally equalizers: a Maximum Likelihood Sequence Equalizer which employs Viterbi algorithm (MLSE/VA) and a decision feedback equalizer (DFE).

1. INTRODUCTION

Data transmission over realistic systems is often corrupted by various distortions, such as intersymbol interference (ISI) due to finite bandwidth of the channel and noise. In order to reduce these distortions, channel equalization is used to compensate for ISI and channel coding to correct the remaining errors. Conventional channel equalizers employ sequence estimation or symbol decision equalization. The optimal solution is given by the MLSE, which has high computational complexity also. In addition, the Viterbi algorithm requires the usage of the nonlinear model of the channel. In practical situations, this information is not available, hence a channel estimator must be included.

A channel equalizer can be viewed also as a classifier, which assigns each received symbol to the correct decision region. A maximum-likelihood method exploiting the classification nature of an equalizer, which avoids explicit channel estimation is presented in [1]. The proposed technique, named Clustering Based Sequence Equalizer (CBSE), operates on a sequence mode and implements the Viterbi searching algorithm. In [3] the interdependence of successive observation was exploited, and a maximum-likelihood procedure based on a nearest neighbour distance metric and using a Viterbi type procedure was proposed.

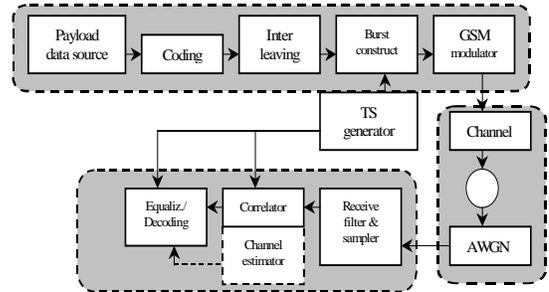


Figure 1- Block diagram of simulated GSM system

In a conventional receiver scheme, the equalizer does not use the redundancy introduced by the channel coding, which is suboptimal. Several approaches to performed jointly equalization and coding have been proposed. A multilayer neural network which performs equalization and decoding simultaneously is presented in [8]. A receiver which performs jointly decoding and equalization of trellis-encoded signals is described in [6]. Also a family of Viterbi decoders which combine equalization and decoding and offer a tradeoff between complexity and performance, is introduced in [7]. A major disadvantage of sequence equalizers performing jointly equalization and decoding is their increased complexity. Several suboptimal techniques which can reduce complexity are proposed in [2][6].

The aim of this paper is to illustrate the possibility of replacing channel equalization and decoding solely by classification in the realistic GSM environment. This paper is organized as follows. Section 2 introduces the GSM simulation environment. The clustering approach for combined equalization and decoding in the GSM system is presented in Section 3. The obtained results are given and discussed in Section 4. The final conclusions are given in Section 5.

2. SYSTEM MODEL

Experiments were carried out in a simulated GSM environment [5], which was modeled using Cossap software [4]. The block diagram of simulation system is presented in Figure 1.

We concentrate only on the full rate speech traffic channel (TCH/FS), which is used to carry speech at 13 kbit/s input rate, by blocks of 260 bits.

An interleaver is used in order to avoid the risk of losing consecutive data bits. At the transmitter, the data bits are coded and arranged into bursts according to GSM recommendations. For the presented simulations, we have used just the GSM normal burst. The midamble is one of the eight known training sequences used in GSM. The modulation technique is Gaussian minimum shift keying (GMSK). The Rayleigh fading propagation profile is implemented using a 6-tap channel model. Next the signal is corrupted by a block adding Additive White Gaussian Noise (AWGN) of variance:

$$\sigma_{\eta}^2 = 10^{(-SNR/10)} N_s P \quad (1)$$

where $N_s=4$ is the number of samples per symbol and $P=1$ is the average power of the signal. The receive filter consists of two fourth order Butterworth filters. The correlator is used to prepare the input data for the equalizer. The channel model is considered to be a five tap FIR filter, with a spacing of one symbol period T :

$$c(t) = \sum_{i=0}^4 a(t-i)h(i) + \eta(t), \quad (2)$$

where $a(t) = (1 - 2i(t))j^n$ and j denotes the imaginary unit, $c(t)$ is the input to the correlator, $\eta(t)$ is the AWGN and $h(i)$ is the impulse response of the spread due to GMSK modulation, ISI and the receiver filter. The coefficients are assumed to be sufficiently constant for the period of half a burstlength, and therefore no adaptation during a burst is needed. The input signal is multiplied by a derotation factor j^k in order to remove part of the modulation and to simplify the reception. The signal is next inputed to the block which combines equalization and decoding.

Channel estimation, required just by the MLSE/VA, is performed by using the 16 centermost bits of the training sequence. The timing estimation is also based on channel estimation, 5 successive channel coefficients with maximum energy define which 148 out of 156 samples are passed to the equalizer.

Three different channel models, defined in [5], which implement the propagation Rayleigh distributed amplitude of each tap varying according to a Doppler spectrum were studied. These channel models are the typical urban channel model with mobile velocity 50 km/h (TU50), the hilly terrain channel model with mobile velocity 100 km/h (HT100) and the rural area channel model with velocity 250 km/h (RA250).

3. CLUSTERING METHODS

Clustering is referred as the process of identifying clusters, or natural groupings in the feature space. Usually the number of classes is unknown beforehand. It was observed that symbol-by-symbol equalizers disregard the dependence that

exists between successive observations, and therefore are suboptimal. On the other hand, a maximum-likelihood procedure takes into account that successive feature vectors are not independent and different classes can be closely related.

A method which takes this observation into consideration is Clustering Based Sequence Equalizer (CBSE), proposed first for channel equalization purposes in [1][2] and then for joining decoding and equalization of trellis-encoded signals in [6]. Observations are clustered around specific points, determined by the finite spread of ISI and the coding type. This procedure tries to recover the uncoded transmitted information bits based on the corrupted coded bits.

We denote by $\mathbf{y} = [y_k, \dots, y_{k-D+1}]^T$ the vector of D successively received samples. Usually in CBSE implementation $D=2$ [6]. The number of states of convolutional decoder is $2^{r(L_c-1)}$, where r is the number of input bits in the convolutional encoder and L_c is the constraint length. The number of states for equalizer trellis structure L^{ISI} is determined by the span of channel over transmitted symbols. In the resulting combined trellis structure there are $2^{r(L_c-1)+L_{ISI}+D-1}$ states corresponding to one of possible values of sequences $\{I_k, \dots, I_{k-r(L_c-1)-L_{ISI}-D+1}\}$. Each of these sequences (or states) corresponds to a cluster representative, i.e. a cluster center, which is the noiseless channel response vector in D -dimensional space. Consequently, the number of clusters is given by:

$$C = 2^{r(L_c-1)+L_{ISI}+D-1}, \quad (3)$$

Due to interdependence on successive received data, imposed by ISI and coding, only certain transitions among different clusters are possible. Therefore a Viterbi path between the cluster centers, based on an appropriate distance metric, can be constructed. The transition from a state s_{k-1} to a state s_k corresponds to emission of a specific cluster representative, which is indicated by the sequence of bits formed by the current state and the new information bits transmitted. The decisions are then taken according to the path with the minimum distance.

Training of the clusters' centers is based on uncoded data, which gives the number of each cluster and is done during a training period. The position of each cluster is obtained by averaging of all data vectors belonging to the respective cluster. We have used the Mahalanobis distance, a popular distance metric in classification problems [9], which is defined as:

$$d_i = (\mathbf{y} - \mathbf{c}_i)^H \sum_i^{-1} (\mathbf{y} - \mathbf{c}_i), \quad (4)$$

where \sum_i is the covariance matrix of each cluster, defined as:

$$\sum_i = E[(y - c_i)(y - c_i)^H], \quad (5)$$

where H denotes the Hermitian transpose. The use of nondiagonal covariance matrix permits exploitation of the clusters shapes in case that distribution is not spherical, i.e. in the presence of nonwhite noise or in case of clusters grouping.

This method has the advantages that the trellis assumed for equalization/decoding operations is based on uncoded data, and in the training period the formed data clusters are directly mapped to uncoded data labels. Therefore the learning procedure is based on uncoded data, and the code knowledge is not needed at the receiver.

On the other hand, this method has high computational complexity, which grows exponentially with channel memory and constraint length of encoder. Different suboptimal techniques, as clusters grouping or the M algorithm, are proposed to reduce its complexity [1][2].

3.1 Clustering approach for combined equalization and decoding in GSM system

We encountered several major problems when trying to apply combined equalization and decoding in the GSM realistic environment. The eight training sequences from GSM have been chosen for their low correlation between one another, as well as for the special shape of their autocorrelation function [5]. Because of low correlation of training sequences, there is no sequence on 13 bits that when using the convolutional encoders of GSM to obtain the full 26 bits of all the used eight training sequences. Therefore the information regarding coding, needed for the CBSE method, can not be extracted from the training data. Another impediment is that in GSM each data block has both uncoded and coded data. So, different treatment should be ensured for these blocks of data.

To extract the coding information, we have chosen to use the training sequence code five corresponding to the sequence $17ac13c_{16}$, because it has a succession of 16 bits which can be obtained at the output of the convolutional encoder for the input $C_i = 104$ (01101000_2). Therefore eight bits of uncoded data are used to establish the number of centers.

It is desirable to use known very short training sequences. In order to obtain this for good performances, the map of clusters should be known 'a priori'. For example, this can be obtained in an 'initialization' phase as follows: the transmitter sends a known sequences that contains all 256 combinations on 8 bits, if 256 clusters are being used. The receiver uses this information to store the cluster positions map. Only the center 104 is used for training.

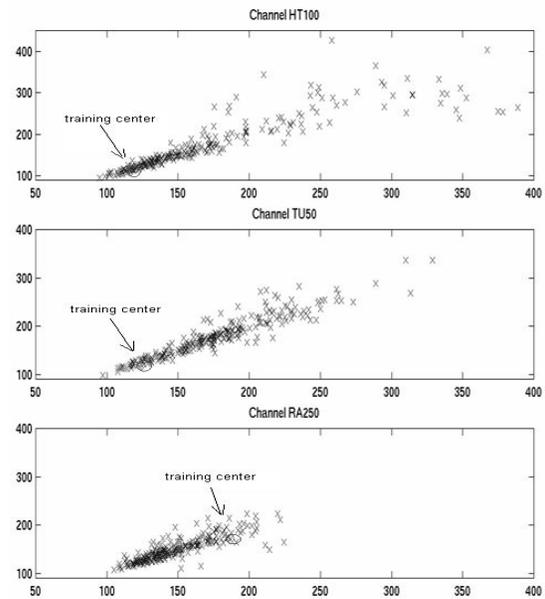


Figure 2- Cluster maps for HT100, TU50 and RA250 channels

The training becomes simple: the computed cluster position is compared with the position of cluster 104. The difference is used to translate the whole cluster map to the new location. This translated cluster map is used to equalize and decode the received data.

The least significant bit of the sequence corresponding to respective cluster gives the decoded output. The obtained cluster maps for the three used channels types are presented in Figure 2. The received samples were normalized with the energy on each burst. To widen the cluster positions, when computing them, the odd samples were multiplied with 2^p while the even samples by 2^{8-p} , with $p=0\div7$.

4. SIMULATION RESULTS

For comparison purposes we implemented a 16-state MLSE/VA. The bursts are processed in two parts, starting in both cases from the middle and proceeding then towards one or the other tail of the burst. Because the training sequence is known, all but one path metrics can be initialized to infinity. Starting with the state that has the lower path metric, the Viterbi algorithm traverses back the best path.

As an additional benchmarking method we implemented a Decision Feedback Equalizer with the feedforward order 5 and feedback order 3. No stopping criterion is used and the adaptation is completed after processing the training sequence once.

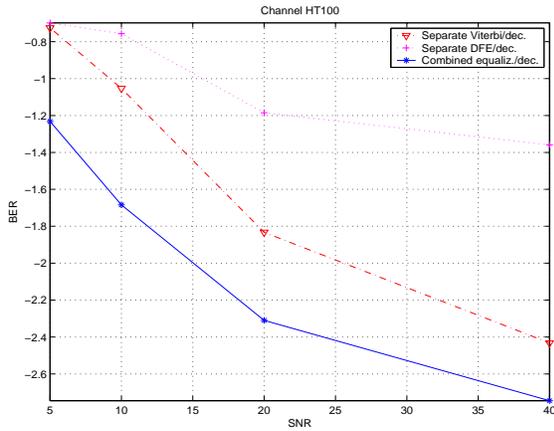


Figure 3- Performance for channel HT100

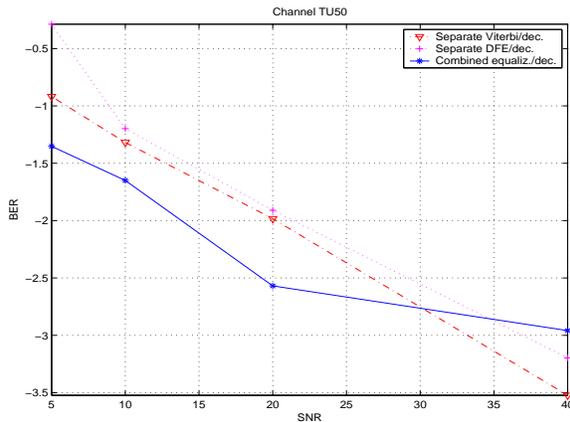


Figure 4 – Performance for channel TU 50

As performance criteria we have used the bit error rate (BER), which is computed as the percentage number of incorrect decisions made on the information sequence of each burst after equalization. The obtained results for HT 100 channel, TU50 channel, and RA250 channel, respectively, are presented in Figures 3, 4 and 5.

It can be observed that for all three channel types the combined method performs, in general, better than separate MLSE/VA equalizer/decoder. Another observation is that for HT100 and RA250 separate VA equalizer/decoder and combined method obviously have better performances than DFE technique, but for TU50 channel the performance of DFE comes very close to the separate MLSE/VA method. It is interesting to notice that only for TU50 channel, at high SNR (e.g. greater than 30 dB), both DFE and MLSE/VA perform slightly better than the combined method.

5. CONCLUSION

We studied the usage of clustering methods as channel equalizers in a realistic time-varying communication scheme. We proposed a new procedure, which performs the equalization and decoding task simultaneously.

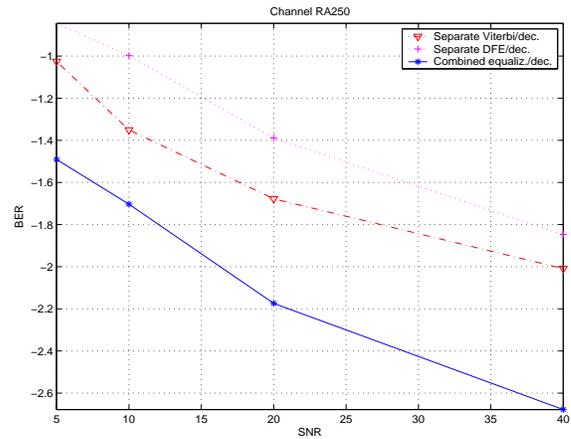


Figure 5 – Performance for channel RA250

Experiments were performed in an GSM simulated environment, for several channel types. The introduced method gives better results in BER terms. The disadvantage of the proposed method is the need of an initialization step. If the initialization step is available, the training set can be reduced to the information needed just for one single cluster center.

6. REFERENCES

- [1] Theodoridis, S., Georgoulakis, K., "Efficient clustering techniques for supervised and blind channel equalization in hostile environments", *Proc. EUSIPCO*, Trieste, Italy, 1996, pp. 611-614.
- [2] Kopsinis, Y., Theodoridis, S., "Reduced-complexity clustering techniques for nonlinear channel equalization", *Proc. ICSP*, 2000, pp. 1492-1495.
- [3] Theodoridis, S., See, C.M.S., Cowan, C.F.N., "Nonlinear channel equalization using clustering techniques", *Proc. ICC*, 1992, pp. 1277-1279.
- [4] COSSAP v.1998.08, Model Library gsmeg, Synopsys Inc., USA, 1998.
- [5] GSM Technical Specifications, ETSI, 1997.
- [6] Georgoulakis, K., Theodoridis, S., "Channel equalization for coded signals in hostile environments", *IEEE Trans. Sig. Proc.*, vol. 47, no. 6, June 1999, pp. 1783-1787.
- [7] Chevillat, P.R., Eleftheriou, E., "Decoding of trellis-encoded signals in the presence of intersymbol interference and noise", *IEEE Trans. Comm.*, vol. 37, July 1989, pp. 669-676.
- [8] Al-Mashouq, K. A., "The use of neural nets to combine equalization with decoding for severe intersymbol interference channels", *IEEE Trans. Neural Net.*, vol. 5, no. 6, Nov. 1994, pp. 982-988.
- [9] Callender, C.P., Cowan, C.F.N. "Equalizers for digital communication using generalized distance measurement", *Proc. 26th IEEE Asilomar Conf. Signals, Systems and Computers*, 1992, pp. 521-525.

Transmission of M -ary Digital Signals Using Controlled Projective Synchronisation of Chaos

Chin Yi Chee* and Daolin Xu

*School of Mechanical & Production Engineering, Nanyang Technological University,
Singapore 639798*

ABSTRACT

We report a chaos-based communication technique for transmitting M -ary digital signals using controlled projective synchronisation of coupled chaotic system. This development shifted chaos-based digital communication to a truly M -ary transmission where the communication system can transmit digital symbols that consist of more than one bit and thus allowing faster transmission rate without increasing the system complexity. The proposed technique is proven by numerical simulation results to be superior over conventional and other chaos-based digital communication techniques in term of bit error rate at low E_b/N_0 .

1. INTRODUCTION

The idea of using chaotic signal as the carrier to transmit digital information was first proposed in a technique known as chaotic shift keying [1]. This technique was developed to explore the potential utilisation of chaotic signal for encryption and spread spectrum communication applications. It is known that spread spectrum technique was first introduced as an anti-jamming technique for military purpose and the motivation to achieve highly jam-resistant communications systems subsequently resulted in assortment of the technique in areas such as multipath access.

In recent years, a number of chaos-based schemes [2, 3] have been proposed with the intention of providing a simple solution for robust wideband communications. For spread spectrum communication, chaotic signal with its inherent wideband characteristic offers the cheaper alternative to conventional methods. Furthermore due to a lower cross correlations between segments of chaotic waveforms than between segments of periodic waveforms, chaotic signal will also generates better performance under multipath propagation conditions.

However chaos-based techniques still face major challenges from conventional techniques in term of bit error rate (BER) and transmission rate. So far in

term of BER performances, the two more promising chaos-based techniques, namely differential chaos shift keying (DCSK) [4] and ergodic chaos shift keying (ECSK) [5] are still inferior to conventional techniques such as phase shift keying (PSK) and frequency shift keying (FSK) when comparison are made over additive white gaussian noise (AWGN) channel. Some of these BER evaluations and comparisons can be found in the literatures [5,7,8]. Furthermore most chaos-based techniques were developed for binary-only transmission where each transmitted symbol consists only 1 bit of information. This reduces the efficiency of chaos-based techniques in term of transmission rate. Therefore transmission of symbols with more than one bit, refers to as M -ary transmission, has been proposed in conjunction with chaos-based techniques [6-8]. However implementing these techniques result in further degradation of BER performance relative to the conventional techniques and require an increase of the system complexity.

In this paper, we describe a technique for M -ary transmission of digital information using projective synchronisation [9] of coupled chaotic system with control [10]. This technique allows the flexibility in choosing higher bits per symbol without increasing the system complexity. This is the first chaos-based technique that outperforms the conventional techniques. Furthermore the proposed technique is executable in simple chaotic systems, without the increase of system complexity that is usually associated with an increase of bits per symbol.

2. CONTROLLED PROJECTIVE SYNCHRONISATION

In projective synchronisation, the state variables of two systems in a coupled form synchronise up to a common scaling factor that defines a proportional relation between the synchronised states. The proportionality is the main feature of projective synchronisation. Two coupled partially linear systems can be expressed as

$$\left. \begin{aligned} \dot{\mathbf{u}}_m &= \mathbf{M}(z_m) \cdot \mathbf{u}_m + \xi \\ \dot{z}_m &= g(\mathbf{u}_m, z_m) \end{aligned} \right\} \text{drive subsystem} \quad (1)$$
$$\left. \begin{aligned} \dot{\mathbf{u}}_s &= \mathbf{M}(z_s) \cdot \mathbf{u}_s \\ \dot{z}_s &= g(\mathbf{u}_s, z_s) \end{aligned} \right\} \text{response subsystem}$$

*Corresponding Author

Email: chinyi_chee@pmail.ntu.edu.sg

In the coupled system (1), the drive system, denoted by the subscript m , drives the response system, denoted by the subscript s , through the variable z_m . A partially linear system refers to an autonomous system in which the state vector \mathbf{u} associates linearly with its time derivatives $\dot{\mathbf{u}}$ through Jacobian matrix $\mathbf{M}(z)$, where $\mathbf{M}(z) \in \mathfrak{R}^{n \times n}$ contains a variable z that is nonlinearly related to the state vector \mathbf{u} . Here the state vectors are $\mathbf{u}_m = (x_1, x_2, \dots, x_n)^T$ and $\mathbf{u}_s = (y_1, y_2, \dots, y_n)^T$.

From the third equation of system (1), it shows that the variable z_m governed by the drive subsystem is used to excite (drive) the response subsystem by uni-directional coupling. Consequently the two subsystems may synchronise in a proportional relationship between the drive and response states. This relation which is refers to as the scaling factor α , is expressed as

$$\alpha = \lim_{t \rightarrow \infty} \alpha(t) = \lim_{t \rightarrow \infty} \|\mathbf{u}_s\| / \|\mathbf{u}_m\| \quad (2)$$

Here $\|\bullet\|$ denotes a norm of a vector. The relationship between the variable z is given as

$$z_s / z_m = \alpha_z = \alpha^b \quad (3)$$

where b is a real number upon the nonlinear system used.

The scaling factor (2) depends on the initial conditions and system parameters of the underlying system. However without control, the ultimate state of projective synchronisation is hardly predictable. Thus a control vector $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_n)^T$ is added to the drive subsystem in (1) to direct the scaling factor to any desired value [10]. The control function is formulated as

$$\xi_i = -[\mathbf{m}_i \mathbf{e} + k e_i] / \alpha^*, \text{ for } i = 1, 2, \dots, n \quad (4)$$

Here \mathbf{m}_i is the i th row of the Jacobian matrix \mathbf{M} . α^* is the desired scaling factor, which is manipulated by the control such that $\lim_{t \rightarrow \infty} [\alpha(t)] = \alpha^*$.

The column vector, $\mathbf{e} = (e_1, e_2, \dots, e_n)^T = \alpha^* \mathbf{u}_m - \mathbf{u}_s$, is an error vector. The constant k is the slack constant that can take any positive value. The slack constant is used to adjust the convergence rate of synchronisation with a larger value used to bring about faster convergence rate [10].

3. DIGITAL COMMUNICATION TECHNIQUE USING PS

For M -ary digital communication, we can assign the distinct value of scaling factor to different digital symbols to be transmitted. The control function (4) can be used to assist in fulfilling the assignment of the scaling factor according to the digital symbols. Here the controlled Lorenz system, which is a

partially linear system ($\dot{x}_1 = \sigma(x_2 - x_1) + \xi_1$, $\dot{x}_2 = (\mu - z_m)x_1 - x_2 + \xi_2$, $\dot{z}_m = x_1 x_2 - \rho z_m$), is used to facilitate further description of the technique.

Fig.1 shows a configuration of this communication scheme with the system parameters set at $[\sigma, \mu, \rho] = [10, 60, 2]$ and $k = 100$. The transmitter sends out the information signal z_m from the drive subsystem to drive the response subsystem at the receiver end. In order to transmit the intended information, the digital symbols to be transmitted must first be encoded onto pre-assigned desired scaling factors and then input into the control function (4). Here the assignment of desired scaling factors to digital symbols is carried out in accordance to Gray coding. An example is shown in Table 1 for $M = 8$, where $M = 2^\beta$ and $\beta = 3$ is the number of bits per symbol.

Digital Symbol	Decimal Value	Desired Scaling Factor, α^*
000	0	1
001	1	2
011	3	3
010	2	4
110	6	5
111	7	6
101	5	7
100	4	8

Table 1. Gray coding for M -ary transmission, $M = 8$.

At the receiver end, the receiver sends feedback signals, y_1 and y_2 to the transmitter end that is required for eliminating the error vector in the control function (4), so that projective synchronisation can occur according to the desired scaling factors. Besides being a necessary input to eliminate the error vector, the feedback signals are also desirable when the technique is used for anti-jamming applications. Considering that the feedback signals can only be generated by the intended receiver due to the coupling effect and the access to the chaotic system and system parameters, the signals could therefore be used as a form of authentication and check to ensure the information is received at the receiver end. Meanwhile, the scaling factors are computed by using Eq. (3) with input from the received signal $z_m + N$ and the produced signal z_s . For Lorenz system, Eq. (3) is defined by $z_s / z_m = \alpha_z = \alpha^2$ as shown in Fig. 1 and N denotes channel noise. Decoding is then carried out to retrieve the transmitted symbols from the computed scaling factors.

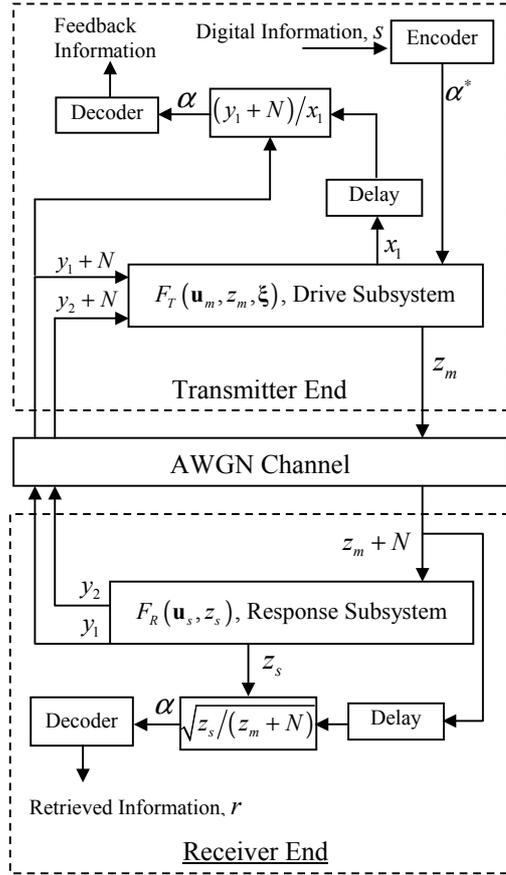


Fig. 1. The coupled Lorenz system with feedback control for digital communication using controlled projective synchronisation.

4. NUMERICAL EXPERIMENT

In this example with $M = 8$, Fig. 2(a) shows the intended digital symbols transmitted to the receiver. The symbol duration used here is equal to 5 time-units. A sampled transmitted signal, $z_m(t)$ of 5 time-units duration from the Lorenz system over AWGN channel with $E_b/N_0 = 10\text{dB}$ is shown in Fig. 2(b). Fig. 2(c) shows the computed scaling factor at the receiver end with AWGN of $E_b/N_0 = 10\text{dB}$ added to the channel. Under the influence of noise the scaling factor fluctuates around a constant within each symbol duration. To retrieve the transmitted symbols from the varying scaling factor and reduce the effect of noise, the average value of the scaling factor is computed for each symbol duration before rounding off to the nearest integer that is bounded between 0 and $M - 1$. As shown in Fig. 2(d), the average value before rounding off is indicated by the vertical line at the end of each symbol duration and the cross indicates the rounded off value, which is the retrieved scaling factor. There is an error bit occurred at $t = 35$. The transmitted signal shown in Fig. 2(a) and the corre-

sponding retrieved information shown in Fig. 2(d) can be used to compare with the applied Gray coding shown in Table 1.

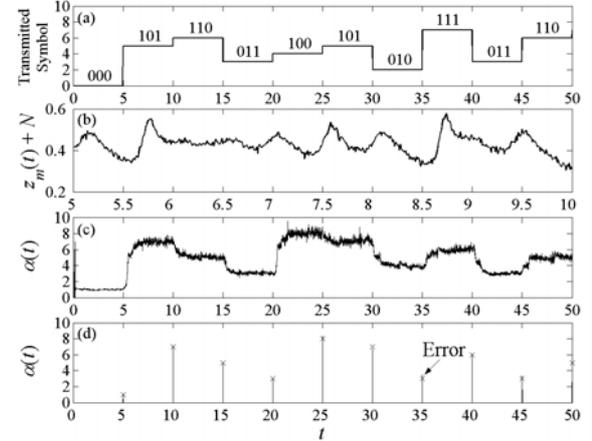


Fig. 2. Decoding process at the receiver end with symbol duration of 5 time-units. (a) The intended digital symbols transmitted to the receiver. (b) Sample transmitted signal, $z_m(t)$, of 5 time-units transmitted over AWGN channel with $E_b/N_0 = 10\text{dB}$. (c) The computed scaling factors at the receiver end with AWGN of $E_b/N_0 = 10\text{dB}$ added to the channel. (d) The average of the scaling factors at the end of each symbol duration indicated by the vertical lines with the crosses indicating the rounded off values.

We wish to discuss the BER performance of the proposed digital communication technique by comparing it with the BER performances of the conventional techniques and other chaos-based techniques [2,5,8]. Fig. 3 shows the BER performances of different techniques for transmitting binary symbols, $M = 2$. In this figure, the BER performance of digital communication using controlled projective synchronisation (PS) is compared with coherent binary phase shift keying (BPSK), which is a conventional technique and ergodic chaotic parameter modulation (ECPM) [5] and differential chaos shift keying (DCSK) [4], which are chaos-based techniques. These techniques are selected for comparison due to their outstanding BER performances under their respective classification. From Fig. 3, the results showed that PS clearly outperformed the other techniques in term of BER performance for transmitting binary symbols.

In the comparison of BER performances for transmitting M -ary symbols, $M > 2$, shown in Fig. 4, the performance of PS as compared to the other techniques is not as straight forward as the binary case. Here the conventional technique is represented by coherent M -ary phase shift keying (MPSK) and the chaos-based technique is represented by quadrature chaos shift keying (QCSK) [8] are used to compare with PS. In Fig. 4, the BER

performance curves for PS are plotted for M equals to 4, 8 and 16. For MPSK, the curves are plotted for M equals to 4 and 8. QCSK as the name suggested has $M = 4$. In the figure, when PS with $M = 4$ is compared to chaos-based QCSK, PS outperformed QCSK across the whole range of E_b/N_0 shown on the graph. However when compared to MPSK the BER performance of PS with M equal to 4 and 8 is only superior at lower E_b/N_0 range up to 4.6dB and 7.6dB respectively.

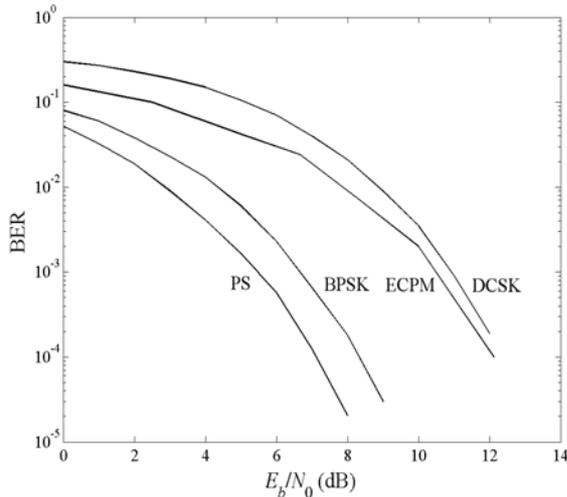


Fig. 3. The BER performance results for binary symbols transmission. The BER performance of controlled projective synchronisation technique (PS) is compared with the conventional technique represented by coherent binary phase shift keying (BPSK) and two chaos-based techniques represented by ergodic chaotic parameter modulation (ECPM) and differential chaos shift keying (DCSK).

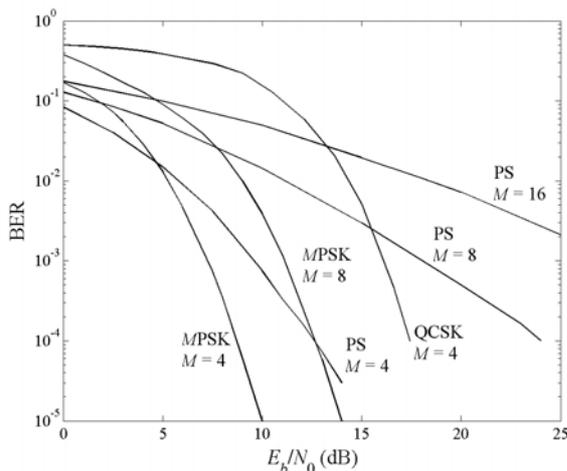


Fig. 4. The BER performance results for M -ary symbols transmission. The BER performance of controlled projective synchronisation technique (PS) is compared with the conventional technique represented by M -ary phase shift keying (MPSK) and chaos-based technique represented by quadrature chaos shift keying (QCSK).

5. CONCLUSIONS

In this paper we have proposed a chaos-based digital communication technique using controlled projective synchronization. This technique is able to transmit binary symbols as well as M -ary symbols by utilizing only a simple coupled chaotic system. The development of this technique contributed to chaos-based digital communication in two areas. Firstly, so far this is the first and only chaos-based technique that has superior BER performance over the conventional techniques. Secondly, the proposed technique shifted chaos-based digital communication to a truly M -ary transmission process. This allows faster transmission rate by increasing bits per symbol without increasing the system complexity.

6. REFERENCES

- [1] U. Parlitz, L.O. Chua, Lj. Kocarev, K.S. Halle & A. Shang, "Transmission of digital signals by chaotic synchronization," *Int. J. Bifurcation and Chaos* 2, 973-977, 1992.
- [2] M. Sushchik, L.S. Tsimring & A.R. Volkovskii, "Performance analysis of correlation-based communication schemes utilising chaos," *IEEE Trans. Circuits Systems. I* 47(12), 1684-1691, 2000.
- [3] G. Kolumbán, "Theoretical noise performance of correlator-based chaotic communications schemes," *IEEE Trans. Circuits Systems. I* 47(12), 1692-1701, 2000.
- [4] G. Kolumbán, M.P. Kennedy & L.O. Chua, "The role of synchronization in digital communications using chaos Part II: Chaotic modulation and chaotic synchronization," *IEEE Trans. Circuits Systems. I* 45(11), 1129-1140, 1998.
- [5] H. Leung, H. Yu, K. Murali "Ergodic chaos-based communication schemes," *Phys. Rev. E* 66, 036203-036210, 2002.
- [6] G. Kolumbán, M.P. Kennedy & G. Kis, "Multilevel differential chaos shift keying," *Proc. IEEE NDES*, 191-196, 1997.
- [7] M.P. Kennedy, G. Kolumbán, G. Kis & Z. Jákó, "Recent advances in communicating with chaos," *Proc. IEEE ISCAS IV*, 461-464, 1998.
- [8] Z. Galias & G.M. Maggio, "Quadrature chaos-shift keying: theory and performance analysis," *IEEE Trans. Circuits Systems. I* 48(12), 1510-1519, 2001.
- [9] R. Mainieri & J. Rehacek, "Projective synchronization in three-dimensional chaotic systems," *Phys. Rev. Lett.* 82, 3042-3045, 1999.
- [10] D. Xu & C.Y. Chee, "Controlling the ultimate state of projective synchronization in chaotic systems of arbitrary dimension," *Phys. Rev. E* 66, 046218-046222, 2002.

TURBO TRELLIS CODED MODULATION WITH MODIFIED TRANSITION METRIC FOR RAYLEIGH FADING CHANNELS

Christos Dimakis¹ Phd, Kostas Koutsouvelis¹ M.Sc

¹Aristotle University of Thessaloniki, Dept. Electrical and Computer Engineering, Telecom. Division. 54 124 Thessaloniki GREECE dimakis@auth.gr , kkouts@tee.gr

Abstract - This paper investigates the performance of a new Turbo Trellis Coded Modulation scheme over correlated flat fading channels with channel interleaving. The novelty of the scheme is based on the application of a new modified transition metric incorporated in the symbol-by-symbol MAP algorithm. We consider frequency non-selective slow, Rayleigh fading channels. Extensive simulation results show that the proposed scheme achieves better performance compared to the conventional transition metric when channel state information is no available at the receiver.

Keywords – Turbo Trellis Coded Modulation (TTCM), Maximum a Posteriori Probability (MAP), Viterbi, Rayleigh fading, Channel State Information (CSI), Trellis Coded Modulation (TCM)

I. INTRODUCTION

Turbo Trellis Coded Modulation (TTCM) scheme is the most powerful bandwidth efficient technique known today. The scheme proposed by Robertson and Wörz, incorporates two Ungerboeck codes as component codes with an 8-PSK redundant signal set and employs a puncturing technique in order to achieve the required bandwidth efficiencies [1]. An odd-even symbol interleaver separates the component codes. This structure allows the use of the symbol-by-symbol MAP algorithm in the decoder stage.

In case of a coded scheme designed for the Additive White Gaussian Noise channel, the best performance over a frequency non-selective slow Rayleigh fading channel is achieved when channel state information (CSI) is available at the receiver. However, when no CSI is provided, a new metric has been proposed in [2] that allows significant coding gains for the case of conventional Trellis Coded Modulation (TCM). The use of an 8-PSK constellation at the decoder, adapted at the signal energy of the received symbol at each decoding step, is incorporated at the estimation of the new transition metric.

In this paper we investigate a TTCM scheme with 8-PSK modulation scheme and bandwidth efficiency of 2 bits/sec/Hz [1]. A channel block interleaver is used in order to break the correlation between consecutive symbols. The decoder is based on the symbol by symbol MAP algorithm enhanced with the transition metric proposed in [2].

II. SYSTEM MODEL

A block diagram of the proposed system is shown in Fig. 1.

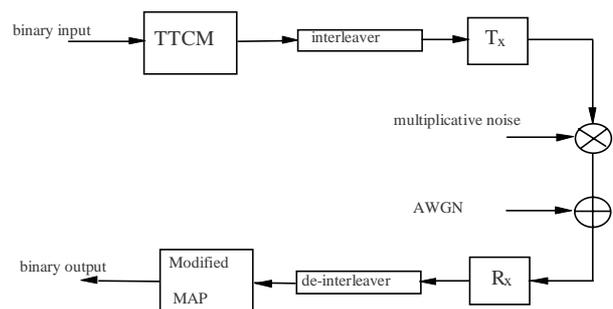


Fig . 1. Block diagram of the proposed TTCM system.

The binary input data is fed into the TTCM encoder. The TTCM consists of two 8-state Ungerboeck trellis codes with 8-PSK redundant modulation set [1]. The polynomials of the codes are:

Table 1

Punctured TTCM for 8-PSK [1]

H^0	H^1	H^2	d_{free}^2 / Δ_0^2
11	02	04	3

An odd-even, random, 1020 symbol interleaver separates the two component codes. A deinterleaver is used at the output of the second coder to permit alignment of the output symbols of the two codes so that after puncturing, information bits will be carried by one output symbol (either of the first or the second coder). The bandwidth efficiency of this scheme is 2 bits/sec/Hz.

The communication channel is a frequency non-selective slow Rayleigh fading channel. The fading is modelled as a complex narrowband stationary Gaussian process where the envelope and the phase exhibit a Rayleigh and uniform distribution correspondingly. In all cases the effect of phase distortion is neglected, so coherent detection of the signal at the decoder is assumed.

The channel block interleaver is used to eliminate the correlation between consecutive symbols.

If the transmitted 8-PSK signal at the k_{th} symbol interval is x_k , then the received signal y_k will be given by:

$$y_k = a_k x_k + n_k \quad (1)$$

where a_k is the Rayleigh fading envelope and n_k is the AWGN noise with variance σ .

III. THE MODIFIED TRANSITION METRIC

We will briefly present the modified Viterbi metric for the case of TCM as described in [2].

It is well known that the conventional Viterbi metric involves the estimation of the Euclidean distances between the received signal y_k and the symbols of the transmitted signal constellation. If x_k is the transmitted 8-PSK symbol associated with branch transition in the trellis at the k_{th} symbol interval, the Euclidean distance is :

$$\gamma_k \sim |y_k - x_k|^2 \quad (2)$$

The idea in [2] was to simply modify the statistics of the received sequence in order to match better the characteristics of the Rayleigh fading channel. They slightly modified relation (2) and proposed a new transition metric according to the formula below:

$$\gamma'_k \sim |y_k - |y_k| \cdot x_k|^2 \quad (3)$$

where $|y_k|$ is the amplitude of the received symbol which has been distorted by the multiplicative (fading) and additive noise. This value of course is readily available at the decoder and should not be confused with the case of CSI where the fading envelope is recovered with special circuitry.

By careful observation of (3) and since the mean energy of the transmitted 8-PSK signal constellation is 1, it is obvious that the new distances are estimated from an equivalent 8-PSK signal constellation with a mean energy always matched to the signal energy of the received symbol, see Fig. 2. In other words the decoder employs for each k_{th} symbol interval, an 8-PSK signal constellation of variable energy.

According to a detailed statistical analysis in [2], the implementation of the new metric increases the relative distances of the received signal from the correct regarding the rest of the 8-PSK symbols. Simulation results showed that in the case of TCM in Rayleigh fading channels, the relative increase of the distances leads to impressive gains of the order of 5 dBs comparable with the conventional Viterbi metric [2].

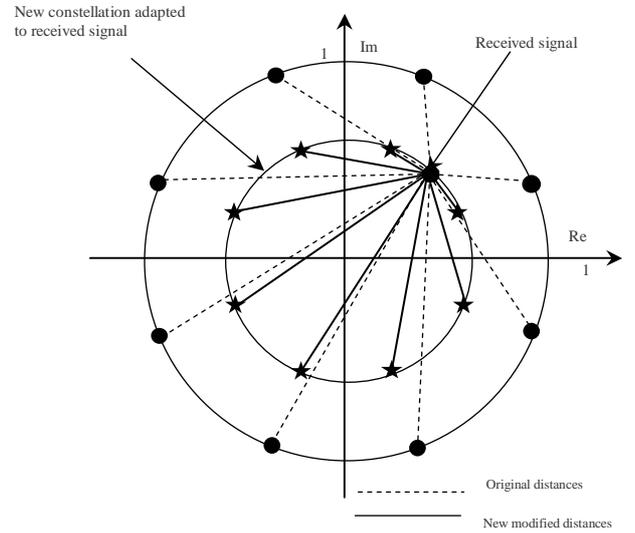


Fig. 2. Distance representation of the two metrics for 8-PSK constellation

For the case of TTCM, instead of the Viterbi algorithm, the symbol by symbol Maximum A posteriori Probability Algorithm (MAP) is used according to the turbo coding principle [1],[3]. This MAP provides the logarithm of the ratio of the a posteriori probabilities for each pair of information bits consisting a quaternary symbol u:

$$L(d_k = u) = \log \frac{P(d_k = u/y)}{P(d_k = 0/y)} = \log \frac{\sum_{d_k=u, m, m'} a_{k-1}(m) \gamma_k(y_k/d_k, m, m') \beta_k(m')}{\sum_{d_k=0, m, m'} a_{k-1}(m) \gamma_k(y_k/d_k, m, m') \beta_k(m')} \quad (4)$$

where the forward and backward metrics are:

$$a_k = \sum_{m, m'} a_{k-1}(m) \gamma_k(y_k, m, m') \quad (5)$$

$$\beta_k = \sum_{m,m'} \beta_{k+1}(m) \gamma_k(y_k, m, m') \quad (6)$$

Apparently the transition metric is incorporated in all (4), (5) and (6) relations.

If we replace this metric with the new modified transition metric γ' from (3), the log-likelihood ratio in the MAP algorithm will become:

$$L'(d_k = u) \sim \log \frac{\sum_{d_k=u, m, m'} a'_{k-1}(m) |y_k - |y_k| \cdot x_k|^2 \beta'_k(m')}{\sum_{d_k=0, m, m'} a'_{k-1}(m) |y_k - |y_k| \cdot 0|^2 \beta'_k(m')} \quad (7)$$

where 0 is the zero quaternary information symbol. The alphas and betas will also be given by:

$$a'_k \sim \sum_{m,m'} a'_{k-1}(m) |y_k - |y_k| \cdot x_k|^2 \quad (8)$$

$$\beta'_k \sim \sum_{m,m'} \beta'_{k+1}(m) |y_k - |y_k| \cdot x_k|^2 \quad (9)$$

The new modified MAP algorithm is based on relations (7), (8) and (9).

IV. SIMULATION RESULTS

Extensive simulation results have been carried in order to estimate the performance of the proposed system. The BER curves show the performance of the system for two different Rayleigh channels. The first one corresponds to a channel with factor $BT_s=0.01$ while the second one to a factor $BT_s=0.1$. The Turbo interleaver has 1020 symbols length and the channel interleaver is 120x80 block interleaver. The Turbo decoder outputs the estimated symbols after 8 full iterations.

From both graphs we can conclude that the Turbo decoder with the modified metric is a compromise between the case of having CSI available in the decoder and the case where CSI is no available. From Fig. 3 and Fig. 4 we notice that for $BER=10^{-4}$ a gain of 0.75 to 1 dB is achieved when no CSI is provided to the decoder. From the same figures we can observe that the proposed coded scheme is approximately 1 dB inferior comparable to the code which takes advantage of CSI information.

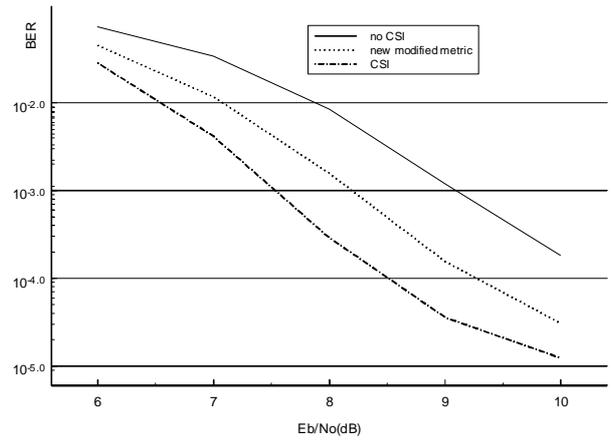


Fig.3. BER performance of the proposed TTCM scheme, $BT_s=0.01$

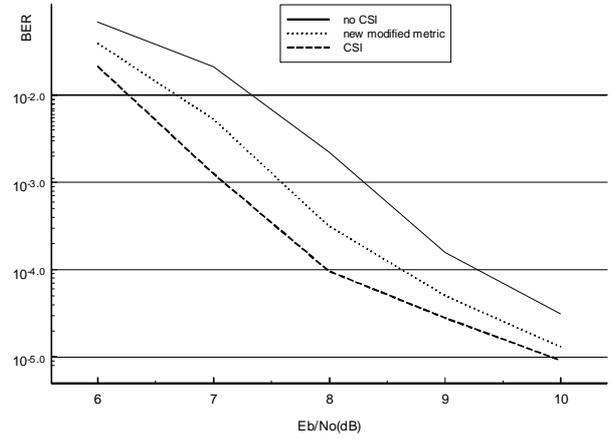


Fig.4 BER performance of the proposed TTCM scheme, $BT_s=0.1$

REFERENCES

- [1] P.Robertson and T.Wörz, "Bandwidth-Efficient Turbo Trellis-Coded Modulation Using Punctured Component Codes", *IEEE Journal on Selected Areas in Communications*, vol.16, No.2, pp. 206-218, Feb. 1998.
- [2] I.Nissopoulos, C.Dimakis and S.Kouris, "The Effect of Metric Modification on the Performance of TCM and Concatenated Systems over Rayleigh Fading Channels", *International Journal of Wireless Information Networks*, vo.7, No.3, pp. 167-174, 2000.
- [3] C.Berrou, A.Glavieux and P.Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes", *Proc.1993 IEEE International Conference on Communications*, Geneva, Switzerland, pp. 1064-1070, May 1993.

Title Using Invisible Region Wave Vectors for Determining the Properties of Microwave Antennas and Imaging Antenna Fields

Authors D. Smith*, M. Leach*, M. Elsdon*, S. J. Foti†

*School of Engineering and Technology Northumbria University Newcastle upon Tyne

†Pele Cottage, Wingates Longhorsley, Morpeth Northumberland NE65 8RW

Introduction

The determination of antenna far field radiation patterns from measurements taken in the near field is well established. Complex near field measurements are normally taken over a 2D scanned aperture which intercepts most of the radiation emitted by the antenna under test, AUT. Traditionally results have been taken at sample spacing, $\Delta x = \lambda / 2$, to satisfy the Nyquist criterion, sampling at spacing less than the Nyquist rate does not provide any additional information and increases the measurement time.

Recently an alternative indirect holographic technique for the determination of antenna radiation patterns and the reconstruction of antenna aperture fields from scalar intensity patterns has been described. [1] This method offers significant simplification over the direct measurement of complex fields as it replaces the vector measurement equipment by simple scalar power measurements.

In the indirect technique a holographic intensity pattern, $I(x,y)$, formed by combining the sampled antenna near field, $E(x,y)$, with a phase coherent offset reference wave, $R(x,y)$, is measured.

$$I(x,y) = |E(x,y) + R(x,y)|^2 = |E(x,y)|^2 + |R(x,y)|^2 + E(x,y)R^*(x,y) + E^*(x,y)R(x,y) \quad (1)$$

The far field radiation pattern of the AUT can be recovered from the Fourier Transform of this intensity pattern.

$$F\{I(x,y)\} = F\{|E(x,y)|^2\} + F\{|R(x,y)|^2\} + F\{E(x,y)\} \otimes F\{R^*(x,y)\} + F\{E^*(x,y)\} \otimes F\{R(x,y)\} \quad (2)$$

Provided that the outer terms of this transform can be isolated the far field radiation pattern of the AUT can be determined. The two outer terms enable a normal image and an inverted image to be reconstructed.

Additionally the original complex antenna aperture fields can also be reconstructed from the outer terms by back propagating the plane wave spectrum to the plane of the antenna aperture. [2]

If a radiated reference wave is used, as is normal at optical frequencies, the maximum offset wave vector is restricted to values within the visible region, $|k_r| < k_0$, where $k_0 = 2\pi / \lambda$. This restricts the ability to separate the terms of the Fourier Transform and has limited this technique to the determination of radiation patterns of large highly directive antennas.[3]

An outline of the recently developed technique is shown in Figure 1.

The signal from the source is divided into two parts, one part is fed directly to the antenna under test, the second part is fed forward to provide a reference signal.

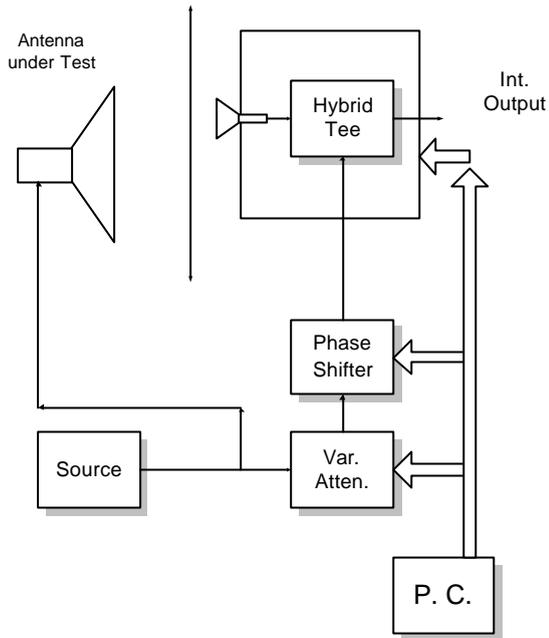


Figure 1. System for Synthesising Offset Reference Wave

A significant feature of this method is the electronically synthesised reference wave. As the reference signal is fed forward to be combined with the sampled antenna near field signal a controllable phase shift can be applied to this signal. By applying a linear phase shift, $\Delta\theta$ rads, as the probing antenna is moved a distance, Δx m, between samples a planar reference wave with offset wave vector, $k_r = \Delta\theta / \Delta x$, rads / m can be synthesised. As an example by reducing the sample spacing to, $\Delta x = \lambda / 10$, with a phase shift between samples of, $\Delta\theta = \pi / 2$, results in an offset wave vector of, $k_r = \Delta\theta / \Delta x = \pi/2 / \lambda/10 = 5.k_0$. This has a practical application in that the use of an offset wave vector in the invisible region increases the separation between the terms of the Fourier Transform and enables the radiation patterns of smaller antennas to be determined.

Simulation Results

The improvements achievable by using the invisible region can be demonstrated by

consideration of a 1D uniform distribution of 4 lambda length. Sampling this at a traditional spacing of half wavelength an offset wave vector in the visible region, $k_r = 2.k_0 / 3$, results in a pattern with overlap between the terms of the Fourier Transform, as shown in Figure 2, and the far field radiation pattern cannot be determined.

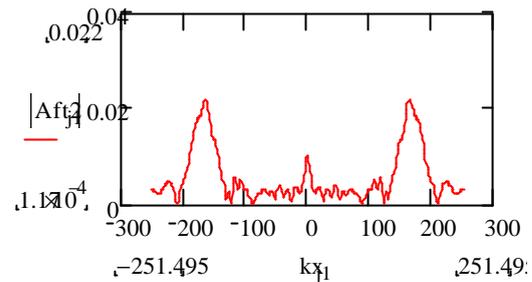


Figure 2. Fourier Transform of Intensity Pattern – Half Wavelength Sampling

By decreasing the sample spacing to, $\Delta x = \lambda / 2$, with a phase shift, $\Delta\theta = 2.\pi / 3$, moves the wave vector into the invisible region, $k_0 = 10 / 3 . k_0$. This enables the required term of the Fourier Transform to be adequately separated, as shown in Figure 3, and the far field radiation pattern can be determined.

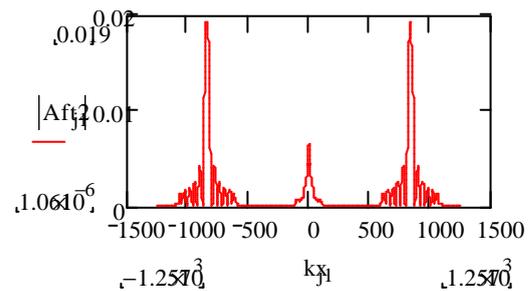


Figure 3. Fourier Transform of Intensity Pattern – Tenth Wavelength Sampling

The remaining unwanted terms of the Fourier Transform can be filtered off to leave

$$F\{I(x)\} = F\{E(x)\} \otimes F\{R^*(x)\} \quad (3)$$

The effect of using a constant amplitude reference signal with linear phase shift,

$$R(x) = R_0 e^{jk_x x} \quad (4)$$

is to shift this term in the spatial frequency domain by, $-k_r$.

In the optical case the original complex wavefront, $E(x)$, can be reconstructed by reintroducing the original reference signal.[4] The same effect can be achieved in this technique by using the computer to shift the filtered transform term by, $+k_r$. The original complex field, $E(x)$, at the measurement plane can now be reconstructed by forming the Inverse Fourier Transform of the filtered and shifted intensity pattern. Additionally when the distance to the antenna is known the plane wave spectrum can be back-propagated to reconstruct the original complex aperture field of the antenna.[5] The reconstructed image of the 1D uniform aperture distribution together with the original distribution is shown in Figure 4.

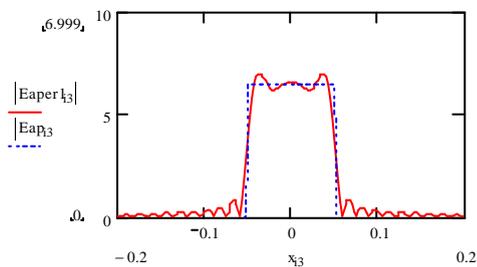


Figure 4 Reconstructed Uniform Field Distribution

Experimental results

This technique has recently been applied to the measurement of a small pyramidal horn antenna operating at 10 GHz. Using the technique described above a scalar holographic intensity pattern was recorded at a distance of 30 cm in front of the antenna aperture. Results were taken at sample spacing, $\Delta x = \lambda / 10$, with phase shift between samples, $\Delta\theta = 2\pi / 3$, to produce an offset wave vector in the invisible region, $k_r = 10 / 3 \cdot k_0$. A slice

through this holographic interference pattern is shown in Figure 5.

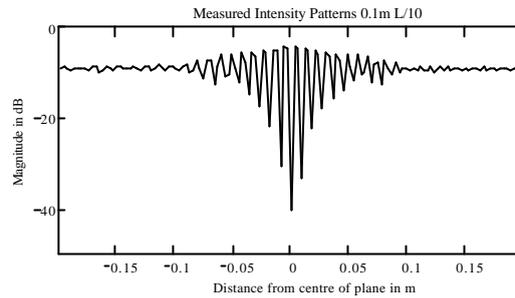


Figure 5 Measured Intensity Pattern – Pyramidal Horn Antenna

The holographic intensity pattern can be processed, as described above, to produce the plane wave spectrum and far field radiation pattern of the horn antenna. Results obtained from the near field scalar intensity pattern are shown for comparison with outdoor far field measurements in Figure 6.

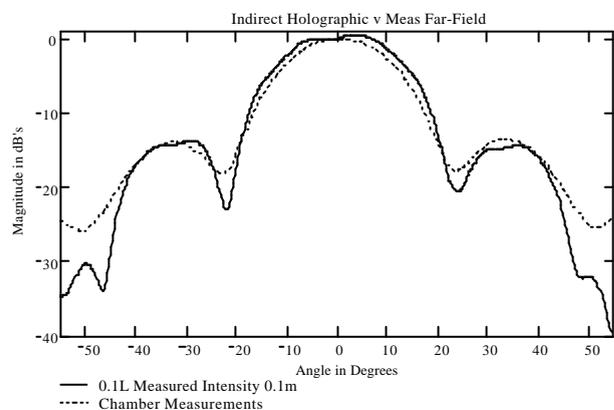


Figure 6. Far Field Radiation Pattern of Pyramidal Horn Antenna

Additionally, also following the method outlined above, the horn antenna aperture fields can be reconstructed from the holographic interference pattern. Figure 7 shows reconstructed antenna aperture fields together with theoretical values for

comparison. A good level of agreement is apparent from these results.

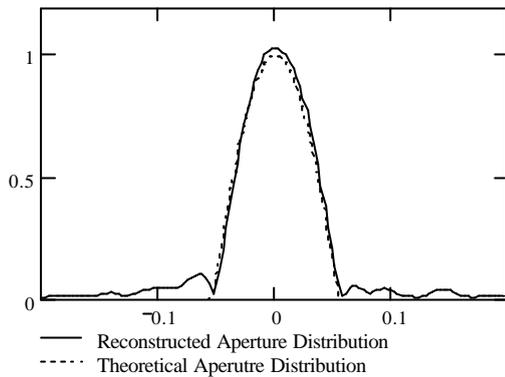


Figure 7a H Plane Aperture Distribution

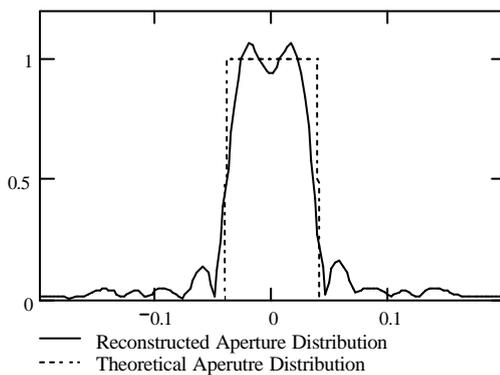


Figure 7a E Plane Aperture Distribution

Filtering the term in the spatial frequency domain has the effect of band limiting the image term and restricts the maximum spatial frequencies for reconstruction. This is equivalent to band pass filtering a voltage pulse and similarly results in a smoothing of the sharp edges.

Conclusion

An indirect holographic technique which utilises wave vectors in excess of the free space wave vector has been described. The use of these vectors, in what can be called the invisible region, to determine the far field radiation patterns of a microwave horn antenna has been described. This

work has also demonstrated how images of the original complex aperture fields can be reconstructed from simple scalar holographic interference patterns.

References

- [1] D. Smith, M. Leach, A. J. Sambell, "An Indirect Holographic Method for Determining Antenna Radiation Patterns and Imaging Antenna Fields", IEEE Symposium 2002, June 16 – 21, San Antonio, Texas, USA
- [2] A. P. Anderson, "Microwave Holography", Proc. IEE, vol. 124, No. 11R, pp946 – 962, 1977
- [3] J. C. Bennett, A. P. Anderson, P. A. McInnes, A. J. T. Whittaker, "Microwave Metrology of Large Reflector Antennas", IEEE Trans. Antennas Propagat., vol. AP-24, No. 3, pp 295 – 303, May 1976
- [4] Hariharan, "Optical Holography" Cambridge University Press, Second Edition, 1996, ISBN 0-521-43965-5
- [5] Y. Rahmat-Samii, L. I. Williams, R. G. Yaccarino, "The UCLA Bi-polar Planar Near-Field Antenna Measurement and Diagnostic Range", IEEE Antennas and Propagation Mag., vol. 37, No. 6, pp 16 – 35, Dec. 1995

Wireless Communication System Performance Enhancement Using Adaptive Array Antennas

Amin H. Al-Ka'bi and John Homer

Abstract—The performance of the steered beam adaptive array can be affected by many factors like the signal to noise ratio of the desired and undesired signals, the number and spacing between the sensor elements, feedback loop gain of the array and the accuracy of our knowledge of the direction of the desired signal. This paper discusses the effect of these factors on the performance of the array.

Index Terms—Steered beam adaptive array, Pointing errors, smart antennas, Wireless Communication.

I. INTRODUCTION

An adaptive antenna array is an array that controls its own radiation pattern by means of feedback control [1]. In steered beam adaptive arrays, it is desirable that the signal to interference plus noise ratio at the output of the array (SINR) exceeds a certain threshold, hence the adaptive algorithm of SINR maximization is used to set the sensors' weights of the array. But this goal cannot be achieved if there is an error in steering the main beam of the array towards the direction of the desired signal [2],[3]. The use of adaptive antenna array in communication systems is an emerging research area. Hence, the main objective of this research is to enhance the performance of the steered beam adaptive array in communication system, in presence of pointing errors, interference signals, and white noise.

II. PRINCIPAL SYSTEM ELEMENTS

The functional diagram of the adaptive array is depicted in Figure 1, which shows the principal system elements that an adaptive array must possess if it is to achieve successfully the twin objectives of enhancing the desired signal reception and rejecting the undesired interference signals [2]. The principal system elements of the adaptive array consist of array of sensors, pattern forming network, and adaptive pattern control unit or adaptive processor that adjusts the variable weights in the pattern forming network. The output of each of the N elements is directed to the pattern forming network, where the output of each sensor element is first multiplied by a complex weight and then summed with all other weighted outputs of the sensor elements to form the overall output of adaptive array. The weight values within the pattern forming network in conjunction with the radiation patterns of the individual sensor elements and their physical arrangement can determine the overall array beam sensitivity pattern. A problem with steering vector arrays is that its performance can be very sensitive to errors in the steering

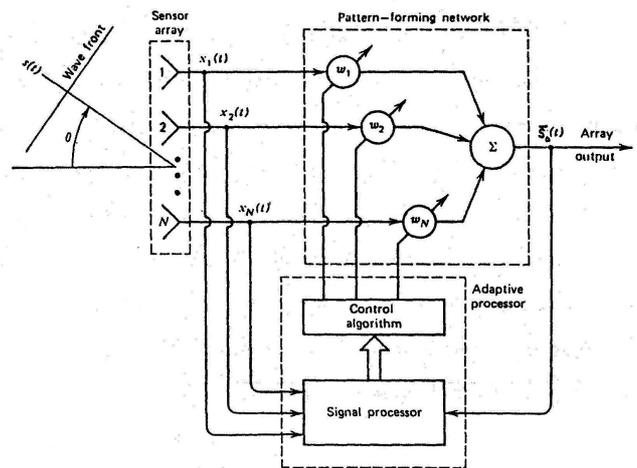


Fig. 1. Adaptive array functional diagram.

vector. The effect of these errors can be significant in array applications where the desired signal is present continuously, as in communication system applications [2],[5].

III. STEERING VECTOR ARRAYS

Consider an N-element adaptive array receiving complex analytic signals¹ where the i^{th} complex analytic signal is denoted by $\bar{x}_i(t)$ and the i^{th} complex weight is denoted by w_i as shown in Fig.1. We define the signal vector,

$$X = [\bar{x}_1(t), \bar{x}_2(t), \dots, \bar{x}_N(t)]^T \quad (1)$$

This signal vector can be split into a desired signal term, an interference term, and a thermal noise term,

$$X = X_d + X_I + X_n \quad (2)$$

The output signal of the array may be split in the same way,

$$\bar{s}_o(t) = \omega^T X = \bar{s}_d(t) + \bar{s}_I(t) + \bar{s}_n(t) \quad (3)$$

The desired signal power, interference power, and noise power at the output of the array will then be

$$P_d = E|\bar{s}_d(t)|^2 \quad (4)$$

$$P_I = E|\bar{s}_I(t)|^2 \quad (5)$$

The authors are with the School of Information Technology and Electrical Engineering, The University of Queensland, Qld, Australia, e-mail:{alkabi, homerj}@itee.uq.edu.au

¹Complex analytic signals correspond to real passband signals which have a baseband processed.

$$P_n = E|\bar{S}_n(t)|^2 \quad (6)$$

Where $E\{\cdot\}$ denotes expectation. The steering vector adaptive array is based on the concept of maximizing the ratio of the desired signal power P_d to the undesired signals powers P_u . The optimum criterion is to adjust the weights in order to maximize the quantity

$$SINR = P_d/P_u = P_d/(P_I + P_n) \quad (7)$$

IV. PROBLEM FORMULATION

Consider the N-element steered beam adaptive array shown in Figure 1. These elements are assumed isotropic and separated by non-uniform distances l_1, l_2, \dots, l_{N-1} . The output signal from each element $\bar{x}_i(t)$ which is assumed to be a complex random process is multiplied by a complex weight ω_i and summed to produce the array output $\bar{S}_o(t)$.

In such an array the steady-state weight vector $\omega = [\omega_1, \omega_2, \dots, \omega_N]^T$ is given by [2], [4]:-

$$\omega = [I + k\Phi]^{-1}\omega_o \quad (8)$$

where Φ is the covariance matrix,

$$\Phi = E\{X^* X^T\} \quad (9)$$

and ω_o is the steering vector

$$\omega_o = [\omega_{10}, \omega_{20}, \omega_{30}, \dots, \omega_{N0}]^T \quad (10)$$

In these equations, I is the identity matrix, k is the feedback loop gain, T denotes transpose, $*$ is complex conjugate, and X is the signal vector, where

$$X = [\bar{x}_1(t), \bar{x}_2(t), \dots, \bar{x}_N(t)]^T \quad (11)$$

and $E\{\cdot\}$ denotes the expectation. The output signal of the j^{th} element $\bar{x}_j(t)$ is given by:-

$$\bar{x}_j(t) = \bar{S}_{d_j}(t) + \sum_{i=1}^M \bar{S}_{I_{ij}}(t) + \bar{S}_{n_j}(t), j = 1, 2, \dots, N \quad (12)$$

where $\bar{S}_{d_j}(t)$ is the output of the j^{th} element due to the desired signal, $\bar{S}_{I_{ij}}(t)$ is output of the j^{th} element due to the i^{th} interference signal, M is the number of interference signals, and $\bar{S}_{n_j}(t)$ is the thermal noise signal which has zero mean, and uncorrelated with other thermal noise signals such that

$$E[\bar{S}_{n_i}(t)\bar{S}_{n_j}(t)] = \sigma^2\delta_{ij} \quad (13)$$

Where σ^2 is the variance of the thermal white noise, and δ_{ij} is the Kronecker delta. In this array, uniform plane waves are assumed to be incident on the sensors of the array, the medium is assumed to be linear and its only effect on the signals is time delay, hence the desired signals $\bar{S}_{d_1}(t), \bar{S}_{d_2}(t), \dots, \bar{S}_{d_N}(t)$ differ by an inter-element propagation time

$$T_{d_i} = (l_i/c)\sin\theta_d \quad (14)$$

where c denotes the velocity of light, θ_d is the arrival angle of the desired signal, l_i is the inter-element spacing

i^{th} element and the $(i+1)^{th}$ element, T_{d_i} and denotes the propagation time between the i^{th} element and the $(i+1)^{th}$ element. Hence

$$\bar{S}_{d_i}(t) = \bar{S}_d(t - \sum_{i=0}^{j-1} T_{d_i}), j = 1, 2, \dots, N \quad (15)$$

Similarly, $\bar{S}_{I_{ij}}(t)$ is given by:-

$$\bar{S}_{I_{ij}}(t) = \bar{S}_{I_{ij}}(t - \sum_{p=0}^{j-1} T_{I_{ip}}), j = 1, 2, \dots, N \quad (16)$$

where, i denotes the i^{th} interference signal, j denotes the j^{th} sensor element of the array, and

$$T_{I_{ip}} = (l_p/c)\sin\theta_{I_i} \quad (17)$$

where θ_{I_i} is the arrival angle of the i^{th} interference signal. The desired and the interference signals are assumed to have zero mean, stationary, and statistically independent from each other and from the thermal noise.

To determine ω_o , it can be noted that a CW signal from an angle θ_{max} will produce a signal vector:-

$$X = [1, e^{-j\beta l_1 \sin\theta_{max}}, e^{-j\beta(l_1+l_2)\sin\theta_{max}}, \dots, e^{-j\beta \sum_{i=1}^{N-1} l_i \sin\theta_{max}}] e^{j\omega_c t} \quad (18)$$

where ω_c is the center frequency of the power spectral density of the desired signal and $\beta (= 2\pi/\lambda)$ is the wave-number of the desired signal. The output of the array for such a signal would be

$$\bar{S}_o(t) = X^T \omega \quad (19)$$

so,

$$\bar{S}_o(t) = [\omega_1 + \omega_2 e^{-j\mu_1} + \omega_3 e^{-j\mu_2} + \dots, \dots + \omega_N e^{-j\mu_{N-1}}] e^{j\omega_c t} \quad (19a)$$

where

$$\mu_i = \left(\sum_{j=1}^i l_j \right) \beta \sin\theta_{max} \quad (19b)$$

The quiescent pattern of the array will have a maximum on this signal if

$$\omega_1 = \omega_2 e^{-j\mu_1} = \omega_3 e^{-j\mu_2} = \omega_N e^{-j\mu_{N-1}} \quad (20)$$

Therefore, for a given θ_{max} , and according to the phased antenna theory, ω_o could be chosen as

$$\omega_o = [e^{-j\mu_{N-1}}, \dots, e^{-j\mu_2}, e^{-j\mu_1}, 1]^T \quad (21)$$

From this ω_o , the steady-state weight vector ω may then be calculated from equation (8), that is $\omega = [I + k\Phi]^{-1}\omega_o$

The signals $\bar{S}_d(t)$ and $\bar{S}_{I_1}(t), \bar{S}_{I_2}(t), \dots, \bar{S}_{I_{N-1}}(t)$ are assumed to be wide sense stationary with flat band-limited power

spectral density centered at ω_c . The desired signal will have bandwidth $\Delta\omega_d$, so its autocorrelation function is given by:-

$$R_d(\tau) = \frac{S_d \sin(\Delta\omega_d \tau / 2)}{\Delta\omega_d \tau / 2} e^{j\omega_c \tau} \quad (22)$$

By substituting $\tau = T_d$ and noting that

$$\frac{\Delta\omega_d T_d}{2} = \frac{\Delta\omega_d}{2\omega_c} \cdot \omega_c T_d = \frac{B_d \phi_d}{2} \quad (23)$$

where,

$$B_d = \frac{\Delta\omega_d}{\omega_c} \quad (23a)$$

and ϕ_{d_i} is the inter-element phase shift of the desired signal between the i^{th} element and the $(i+1)^{th}$ element,

$$\phi_{d_i} = \omega_c T_{d_i} \quad (23b)$$

where T_{d_i} is the time delay of the desired signal between the i^{th} element and the $(i+1)^{th}$ element. So, normalized autocorrelation function of the desired signal is

$$\rho_d(p) = \left[\frac{\sin(\sum_{i=1}^p \frac{B_d \phi_{d_i}}{2})}{\sum_{i=1}^p \frac{B_d \phi_{d_i}}{2}} \right] e^{j \sum_{i=1}^p \phi_{d_i}} \quad p = 1, 2, 3, \dots \quad (24)$$

Similarly, if the i^{th} interference signal has bandwidth $\Delta\omega_{I_i}$, the fractional bandwidth could be defined as:

$$B_{I_i} = \frac{\Delta\omega_{I_i}}{\omega_c} \quad (25)$$

and phase shift due to the i^{th} interference signal between the n^{th} element and the $(n+1)^{th}$ element is

$$\phi_{I_{in}} = \omega_c T_{I_{in}} \quad (25a)$$

Where $T_{I_{in}}$ is the time delay of the i^{th} interference signal between the n^{th} element and the $(n+1)^{th}$ element. So,

$$\rho_{I_i}(p) = \left[\frac{\sin(\sum_{n=1}^p \frac{B_{I_i} \phi_{I_{in}}}{2})}{\sum_{n=1}^p \frac{B_{I_i} \phi_{I_{in}}}{2}} \right] e^{j \sum_{n=1}^p \phi_{I_{in}}}, \quad p = 1, 2, \dots \quad (25b)$$

Using the steady state weight vector ω in equation 8, the output power of the array due to the desired signal, interference signals and the thermal noise can be calculated, so the desired signal the output of the array is given by:

$$\bar{S}_{d_o}(t) = [\bar{S}_d(t), \bar{S}_d(t - T_{d_1}), \bar{S}_d(t - T_{d_1} - T_{d_2}), \dots, \bar{S}_d(t - \sum_{i=1}^{N-1} T_{d_i})] \omega \quad (26)$$

and output power of the desired signal is then

$$P_d = E[\bar{S}_{d_o}^*(t) \bar{S}_{d_o}(t)] \quad (27)$$

which can be written as

$$P_d = S_d \left[\sum_{n=1}^N |\omega_n|^2 + \sum_{n=1}^{N-1} \sum_{m=2}^N 2\Re\{\omega_n \omega_m^* \rho_{d_{ij}}\} \right] \quad (28)$$

for $m > n$

where $\rho_{d_{ij}}$ is the normalized autocorrelation between the desired signals coming from the i^{th} and j^{th} sensors. Similarly, the output of the i^{th} interference signal is

$$\bar{S}_{I_{i_o}}(t) = [\bar{S}_{I_i}(t), \bar{S}_{I_i}(t - T_{I_{i1}}), \bar{S}_{I_i}(t - T_{I_{i1}} - T_{I_{i2}}), \dots, \bar{S}_{I_i}(t - \sum_{n=1}^{N-1} T_{I_{in}})] \omega \quad (29)$$

and the output power of the i^{th} interference signal is

$$P_{I_i} = S_{I_i} \left[\sum_{n=1}^N |\omega_n|^2 + \sum_{n=1}^{N-1} \sum_{m=2}^N 2\Re\{\omega_n \omega_m^* \rho_{I_{ipj}}\} \right] \quad (30)$$

for $m > n$

where $\rho_{I_{ipj}}$ is the normalized autocorrelation of the i^{th} interference signal coming from the p^{th} and j^{th} sensor elements of the array. While the output power of the thermal noise can be given by:

$$P_n = \sigma^2 \sum_{i=1}^N |\omega_i|^2 \quad (31)$$

From P_d , $\sum_{i=1}^M P_{I_i}$, and P_n the output SINR can be computed as

$$SINR = 10 * \log\left(\frac{P_d}{\sum_{i=1}^M P_{I_i} + P_n}\right) \text{ dB} \quad (32)$$

which can be used as a measure of the performance of the adaptive array in the sequel.

V. PERFORMANCE OF THE ARRAY

Figure 2 shows the radiation pattern of a 10 element adaptive array located vertically along the y-axis, with a non-uniform spacing. The separations between the elements are assumed to be 0.4λ , 0.3λ , 0.5λ , 0.6λ , 0.7λ , 0.5λ , 0.4λ , 0.6λ , 0.3λ respectively. The desired signal with SNR of 20dB is inci-

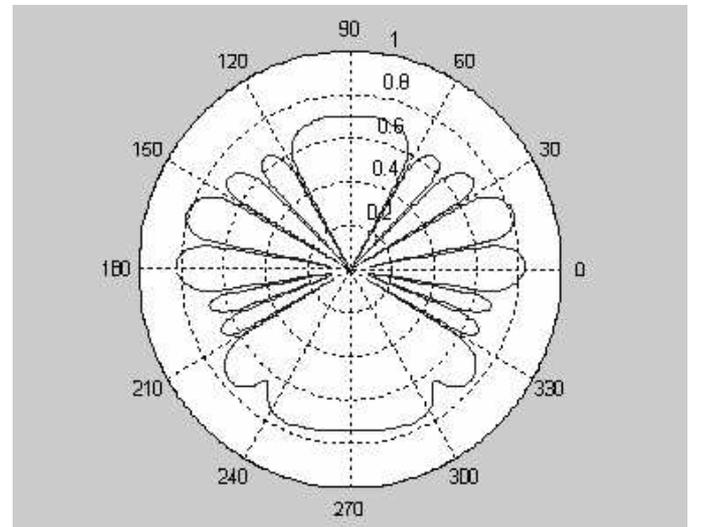


Fig. 2. Radiation pattern of a 10 element adaptive array.

dent from 0° . Additionally, there are seven interference signals incident on the array with INR's=40dB, and arrival angles of $-30^\circ, -20^\circ, -10^\circ, 10^\circ, 30^\circ, 45^\circ$ and 60° . The desired signal and the interference signals are assumed to have zero bandwidth and the feedback loop gain k is assumed to be 0.1. In this figure we have assumed that there is no pointing error (i.e. the steering vector of the array points exactly towards the direction of the desired signal (i.e. 0°)). It can be seen from this figure that array constructed a maximum towards the desired signal, and nulls towards the seven interference signals.

Figure 3 illustrates the effect of the pointing error on the radiation pattern of the adaptive array. Here we assume the same conditions as in Figure 2, except the existence of a 5° pointing error. It can be seen that a maximum is formed towards the assumed arrival angle of the desired signal (i.e. 0°), while the desired signal is arriving actually from 0° . This will degrade the performance of the array by decreasing its output SINR. We can notice also that the array still forms nulls towards the seven interference signals. To quantitatively illustrate the degradation

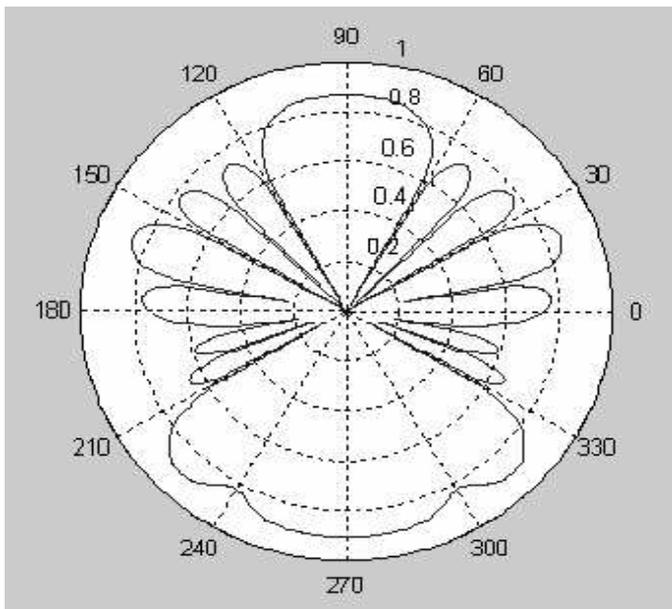


Fig. 3. Radiation pattern of a 10 element adaptive array with 5° pointing error

of the output SINR, we can draw the output SINR as a function of pointing error as depicted in figure 4. In this figure we have assumed an array of 10 elements with non-uniform spacing of $0.4\lambda, 0.3\lambda, 0.5\lambda, 0.6\lambda, 0.7\lambda, 0.5\lambda, 0.4\lambda, 0.6\lambda, 0.3\lambda$ respectively. The input SNR is 0dB, 10dB and 20dB while the interference signals do not exist. This figure shows that the output SINR has a maximum at 0° (i.e. when the pointing error is zero), while it decreases as pointing error increases. Also, we can notice that the higher the input SNR, the closer must be to to achieve maximum gain from the adaptive array. For example, if the input SNR= 0 dB, the array provides output SINR > 5 dB as long as the pointing error ranges between $-4.1^\circ < \theta_{max} < 4.1^\circ$, but for SNR=20dB, the array gain exceeds 5 dB only if the pointing error falls in the range $-1.2^\circ < \theta_{max} < 1.2^\circ$.

To find out how the existence of the interference signals affects the performance of the array, we have red

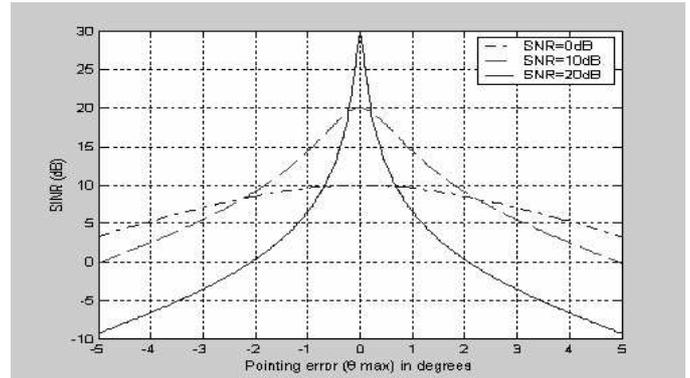


Fig. 4. Output SINR vs. Pointing error (no interference)

into figure 5 but with the existence of two interference signals incident from -8° and 4° . We can see that the output SINR deteriorates rapidly as the steering vector points towards an interference signal instead of the desired signal.

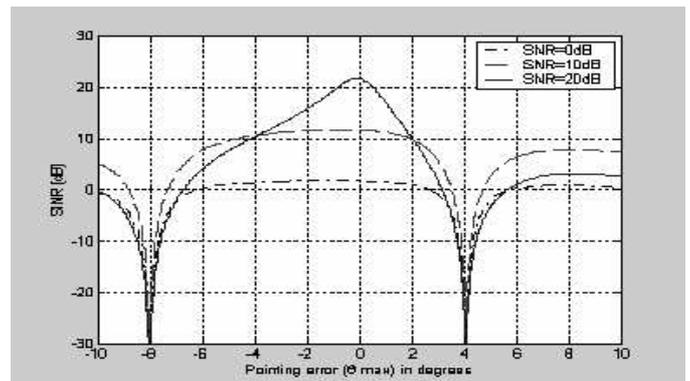


Fig. 5. Output SINR vs. Pointing error (with interference)

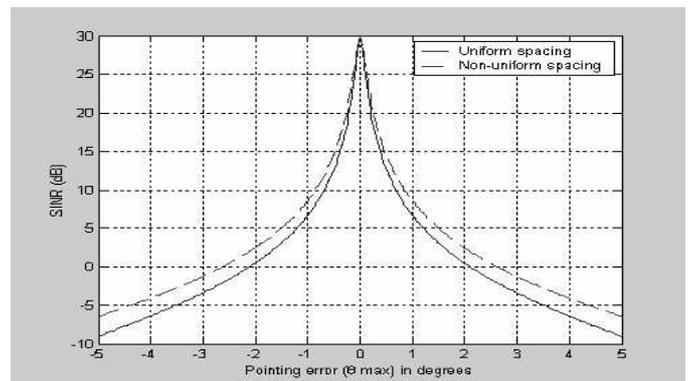


Fig. 6. Output SINR vs. Pointing error (with different spacings)

The arrangement of the elements in the array also affects its performance. This can be depicted in figure 6, where the performance of two arrays with different element arrangements is illustrated. Here, the performance of a uniformly spaced array of (0.6λ) is plotted against a non-uniform array with separations of $0.4\lambda, 0.3\lambda, 0.5\lambda, 0.6\lambda, 0.7\lambda, 0.5\lambda, 0.4\lambda, 0.6\lambda, 0.3\lambda$ respectively. In both cases interference signals do not exist and the input SNR = 20dB.

Furthermore, the performance of the array is affected by the number of the sensors in the array and the feed back loop gain as depicted in figures 7 and 8 respectively. In Fig 7, we can notice that by increasing the number of sensors the sensitivity of the array to pointing errors decreases. While in Fig 8, we can

- [4] S. P. Applebaum, "Adaptive Arrays", IEEE Trans. Antennas propagation, Vol. Ap-24, p.585, Sept. 1976.
 [5] J. W. Kim and C.K. Un, "An adaptive array robust to beam pointing error", Signal Processing, IEEE Transactions on , Volume: 40 Issue: 6, Jun 1992 , pp: 1582 -1584

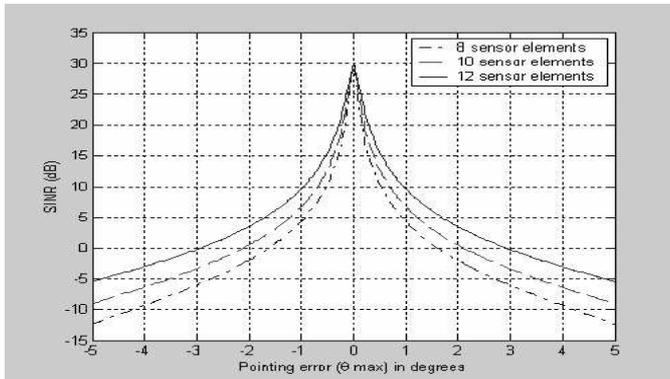


Fig. 7. Output SINR vs. Pointing error (with different number of sensors)

see that by increasing the feedback loop gain the sensitivity of the array to pointing errors increases. In both cases interference signals do not exist, the sensors are uniformly spaced by $\lambda/2$ and the input SNR=20dB.

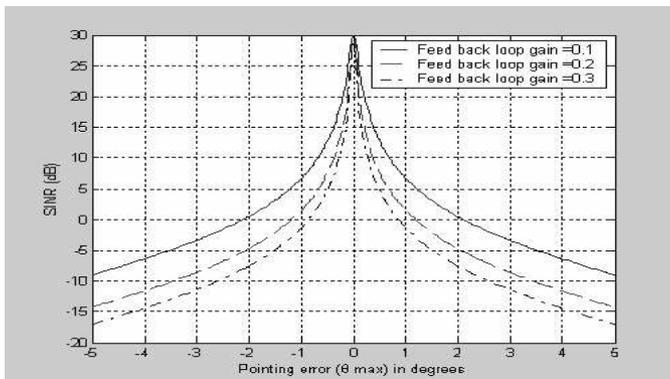


Fig. 8. Output SINR vs. Pointing error (with different feedback loop gains)

VI. CONCLUSIONS

We found that the performance of the steered beam adaptive array can be affected by many factors like the signal to noise ratio of the desired and undesired signals, the number and spacing between the sensor elements, feedback loop gain of the array and the accuracy of our knowledge of the direction of the desired signal. It was shown that some of these factors can be controlled by the array designer to enhance the performance of the array

REFERENCES

- [1] Raymond Steele and Lajos Hanzo, "Mobile Radio Communications", second edition, published by John Wiley and Sons, Ltd, 1999.
 [2] T. Compton, Jr., "Pointing Accuracy and Dynamic Range in a steered Beam Adaptive Array", IEEE Trans. On AES, vol. AES-16, May 1980, pp. 280-287.
 [3] R. T. Compton, Jr., "Adaptive Antennas, Concepts and Performance", Prentice-Hall, New Jersey, 1988, pp. 45-110 and 355-389.

A comparative study of combinations of concealment techniques with error correcting codes for voice packet recovery

K.M.S. Soyjaudah and S. Jharee

Faculty of Engineering, University of Mauritius, Reduit, Mauritius

s.soyjaudah@uom.ac.mu

s.jharee@uom.ac.mu

Abstract

IP telephony is gaining momentum and popularity, however the quality is inferior to the conventional telephone, PSTN, mainly because of high transmission delay and packet loss. When packets are lost, a gap is produced which, has to be filled. The easiest method to replace a missing packet is to use Insertion-based techniques. They require little or no computational cost but generally result in poor performance. Including modifications both on the sender and the receiver sides can give more powerful recovery schemes. In this paper, different error correcting codes are combined with simple packet loss concealment techniques and Comparative objective and subjective voice quality measures have been investigated. It is found that the RS code combined with packet repetition technique gives the best results.

1. Introduction

The growth of IP-based services in the last few years has been explosive. Internet telephony allows integration of many applications and services. However the quality is inferior to the conventional telephone, PSTN, mainly because of high transmission delay and packet loss. To compensate for Internet packet loss, network error control schemes, such as FEC [1,2] and automatic repeat request (ARQ) [3] exist. With ARQ, all packets originated from the source arrive at the destination without errors. However, due to the high latency, ARQ is not practical for the delay-stringent application, like VoIP. On the other hand, Forward Error Correction (FEC) presents a more viable alternative, as it reduces the network delay. But even after the application of FEC, there are still a number of lost packets. So concealment schemes [4] are used to conceal the lost packets and to help in decoding the error correcting codes.

In section 2 of this paper, a brief description of error concealment techniques is given. The packet loss model used is described in section 3. Section 4 gives the performance analysis used to evaluate the proposed system. Simulation results are given in section 5 and finally section 6 concludes the paper.

2. Error concealment

Audio signals, in particular speech, exhibit large amount of short-term self-similarity. So, a similar packet, which has been successfully received, can replace the lost packets. Error concealment techniques are receiver-based and require no assistance with the sender. However, these techniques work for small packets (4 - 40 ms). When the loss length approaches the length of a phoneme (50-100 ms) these techniques break down, since the listener may miss a whole phoneme. V. Hardman et al [4] give a Taxonomy of various receiver-based recovery techniques.

3. Packet loss model

The Gilbert model, also known as the two-state Markov-chain model, is most commonly used to describe bursty losses often found in the Internet [5]. The lossy network can be modelled as having two states; state 0 represents the state of successful packet arrival and state 1 represents the state of packet lost. Fig. 1 shows the state diagram of the Gilbert Model.

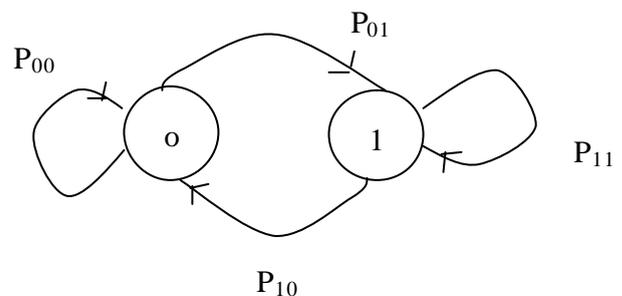


Fig.1 Gilbert model

$P_{10} = P(\text{packet } n+1 \text{ is not lost} | \text{packet } n \text{ was lost})$
 $P_{01} = P(\text{packet } n+1 \text{ is lost} | \text{packet } n \text{ was not lost})$

P_{10} and P_{01} are conditional probabilities such that the steady state probabilities p_0 and p_1 are obtained as follows:

$$p_0 = \frac{P_{10}}{P_{10} + P_{01}} \quad \text{and} \quad p_1 = \frac{P_{01}}{P_{10} + P_{01}}$$

where p_0 is the probability that a packet is not lost, p_1 is the probability that a packet is lost and $p_0 + p_1 = 1$.

4. Performance Analysis

The performance of the proposed system is evaluated by looking at the quality of speech delivered using an objective technique; the signal-to-noise ratio, and a subjective technique; the Mean Opinion Score.

Mean Opinion Score

Mean Opinion Score (MOS) is based on obtaining listener opinion. Listener opinion tests use speech material in the form of sentences, and listeners judge the speech over the system being tested according to a given criterion. Listeners use a five-category rating scale (i.e., excellent, good, fair, poor, and bad) with points assigned for each level. The scores of all the listeners are averaged to provide a score for each source being evaluated.

Prior to the test the concept of MOS was introduced to listeners and the various criterion used to rate speech were explained, as shown in table 1.

Table 1: Mean opinion score ratings

score	Quality	Impairment
5	excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Non-integer scores like 3.5 are possible. Following the evaluation, a numerical MOS score for overall speech quality is calculated. An MOS of 3.0 is generally referred to as ‘toll quality’, meaning ‘good enough to pay for’. Digital voice users may tolerate

MOS levels less than three if they get additional benefits, such as simultaneous voice and data services.

5. Simulation results

Recorded Voice signals are sampled using the Matlab function ‘wavread’ and then it is quantized. The digital samples are then encoded and packetized before transmission. The Internet is simulated using the Gilbert model. The lost packets are replaced by different error concealment techniques before decoding. The results of the simulation are illustrated using SNR v/s network loss rate graphs and MOS v/s network loss rate graphs.

(a) SNR ANALYSIS

(i) SNR analysis of silence substitution combined with FEC schemes

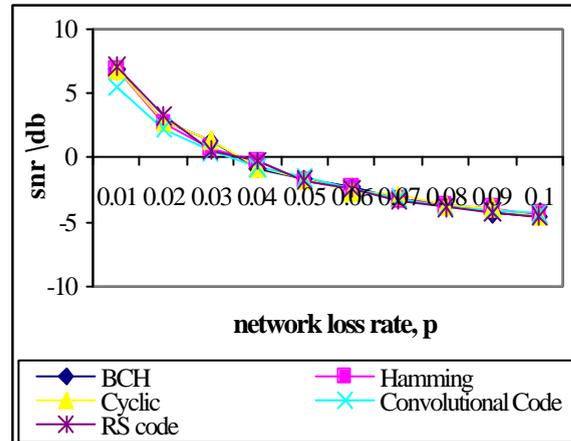


Figure 2: SNR analysis of silence substitution combined with FEC schemes

In this scheme, silence substitution is used to replace lost packets. Figure 2 shows that for small network loss rate (<3%) Reed Solomon coding scheme when combined with silence substitution gives the best quality of speech delivered. But at higher network loss rates, all the schemes give similar results.

(ii) SNR analysis of noise insertion combined with FEC schemes

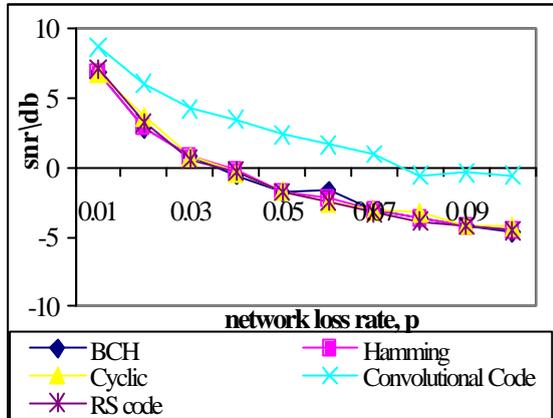


Figure 3: SNR analysis of noise insertion combined with FEC schemes.

From Figure 3 it appears that convolutional coding gives better performance when the lost packets are replaced by noise packets. Having the highest SNR value, this would indicate that the signal suffers little degradation. The schemes based on block codes have similar performances and much lower SNR values.

(iii) SNR analysis of packet repetition combined with FEC schemes

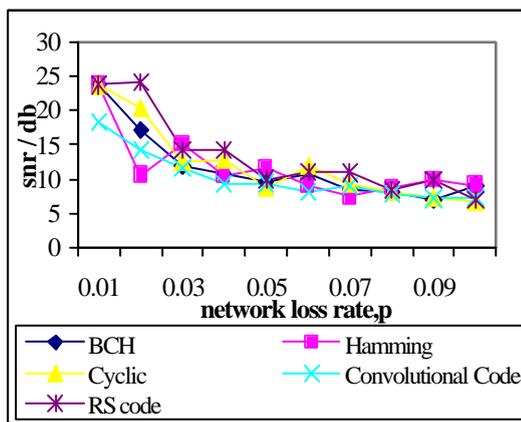


Figure 4: SNR analysis of packet repetition combined with FEC schemes

Figure 4 shows the results when the lost packets are replaced by the previous received packets. All the FEC schemes have higher SNR values than with silence substitution and noise substitution. Reed

Solomon coding scheme gives the best result at lower network loss rates, but at higher network loss rate all the schemes give similar results.

(b) MOS ANALYSIS

(i) MOS analysis of silence substitution combined with FEC Schemes

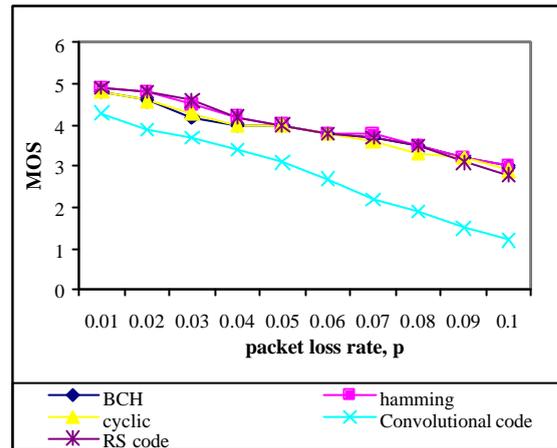


Figure 5: MOS analysis of silence substitution combined with FEC schemes.

Figure 5 shows that the coding scheme employing convolutional code has the worst MOS performance, when compared to FEC schemes employing block codes. The block codes used to encode the digital voice data give similar performance in terms of MOS. For low network loss rate silence substitution combined with Reed Solomon code and Hamming code perform slightly better than other codes.

(ii) MOS analysis of noise insertion combined with FEC schemes

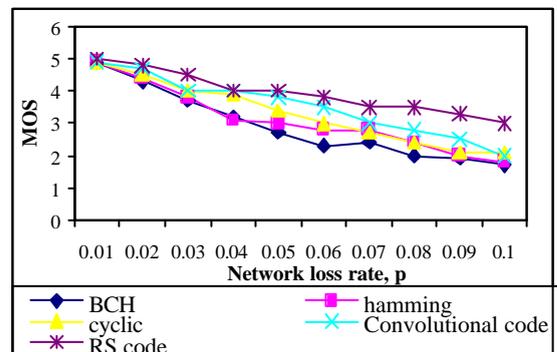


Figure 6: MOS analysis of noise insertion combined with FEC schemes.

The analysis of the MOS of the different FEC schemes used in tandem with noise insertion shows that the Reed Solomon scheme gives best performance compared with the other coding schemes. Even at high network loss rate, the speech remains intelligible, the user does not require much effort to make, so as to understand the message carried by the speech. As compared with silence substitution scheme, the quality of speech is slightly poorer with noise substitution.

(iii) MOS analysis of packet repetition combined with FEC schemes

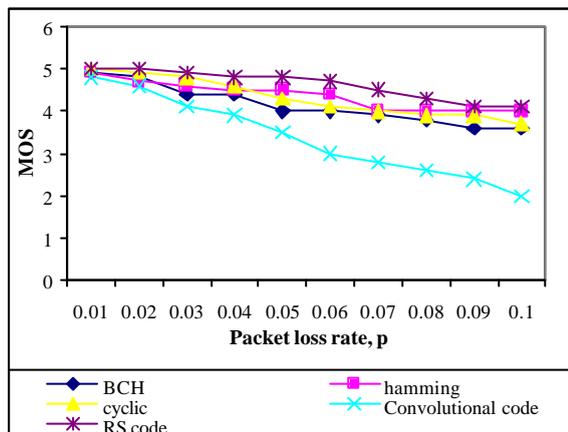


Figure 7: MOS analysis of packet repetition combined with FEC schemes.

Except for the scheme employing convolutional coding, all the other schemes, based on block codes, give very good performance when applied along with packet repetition, **Figure. 7**. The Reed Solomon scheme gives the best performance and the speech delivered is comparable to the original transmitted speech. The listener fails to distinguish the slightest distortion in the speech signal that may occur if the speech is played from a speaker, even for the other block codes.

6. Conclusion and future work

The RS code combined with packet repetition concealment scheme gives the best result compared to the other combinations considered in this paper. However, there is a need to reduce the high decoding delay associated with it. Since, the position of the lost packet is known, only the error correction part of the decoding process can be considered. Thus, the use of RS erasure code combined with packet repetition is being investigated. This will reduce the decoding

delay and also use the self-similarity property of voice packets.

References

1. C. Perkins and O. Hodson, "Options for repair of streaming media", Internet Engineering Task Force, June 1998.
2. J.C Bolot and A. Vega Carcia, "The case for FEC-based error control for packet audio in the Internet", ACM Multimedia systems, 1993.
3. M.G. Podolsky and S. Mc Canne, "Soft ARQ for layered streaming media", Journal of VLSI signal processing 27,pp.81-92, 2001.
4. C. Perkins, O. Hodson and V. Hardman, "A survey of packet loss recovery techniques for streaming audio", IEEE Network Magazine, pp.40-48, Sept/Oct. 1998.
5. M. yajnik, S. Moon, J. Kurose and D.Towsley, "Measurement and modelling of the temporal dependence in packet loss", Proc. Conference on computer communication (IEEE Infocom), New York, March 1999.

A Fast and Accurate Intrusion Detection System

A. Mirzaei, S.H. Davarpanah
{amirzaei, Davarpanah}@iust.ac.ir
Department of Computer Engineering
Iran University of Science and Technology
Narmak, Tehran, Iran (+98-21-8912451)

Abstract

This paper proposes a new method that uses multivariate analysis of the variance algorithm (MANOVA) as a basic technique and makes a hybrid intrusion detection method.

This proposed methodology manipulates the original features to produce a new feature set with two major advantages; The first advantage while this method reduces the similarity between different classes, the similarity between the instances of each class will be increased. As a result of this feature, the resulted classes will have a more convex shape. As a second advantage, this algorithm will produce new orthogonal features each holding a larger amount of information than their ancestor[1].

The nearest neighbour algorithm is used to detect intruders in the next stage. Although this algorithm is fast and less complex than the other classification algorithms in this content, promising results are achieved.

Keywords

Intrusion Detection, Multivariate Analysis of Variance Algorithm, Nearest Neighbour Algorithm.

1. Introduction

With the rapid development of the Information Technology (IT), Computer networks and the Internet, today's modern society has gained lots of benefits; On the other hand, the new era has its own complexities as well. Using the IT technology in every part of the modern life has created tempts which leads some users to commit illegal actions in the system. Therefore in order to secure the integrity of the IT systems, it is necessary to have some strategies and tricks developed

to stand against these actions and to prevent them from compromising system's integrity.

Therefore intrusion and intrusion detection as are described in above are constantly a field of challenge for the researchers in the concept of the computer security. Heady [2] has defined the intrusions as:

"Any set of actions that attempt to compromise the integrity, confidentiality or availability of a resource."

The aim of intrusion detection is to detect attacks against computer systems and networks. Due to the interactive nature of the system and the necessity to provide a provably secure information system, intrusion detection systems have a monitoring system included within them. In particular, intrusion detection systems look for attempts in user/parties abuse of privileges or exploitation of security flaws.

The Intrusion Detection Systems (IDS) are often categorized based on three aspects:

- The Model of intrusion detection (anomaly detection/ misuse detection).
- The data source (host based / network based).
- The response time (offline/online).

With respect to the utilised intrusion detection technique, IDSs can be categorized into two classes:

1. *Anomaly detection*, known and unknown intrusions are detected by analysing changes in the normal pattern of system utilisation or change in the behaviour model of the system. These approaches don't use information about the system behaviour when an intrusion is in progress.
2. *Signature or Misuse Detection*, based on the attack pattern for different known intrusion strategies, looking at the computer system behaviour the known intrusions are detected. This approach works with a predefined attack pattern [3]. Obvious problem with this approach is that IDS will fail in detecting new attack patterns (or those without a known signature).

A host based IDS will monitor all the activities in its host system. Its task is to ensure that none of the systems security policies being violated. A network IDS will monitor activities all over the network and it analyses the network traffic searching for potential security breaches or violations.

IDSs also differ in whether they work on-line or off-line. Off-line IDSs run periodically and their purpose is to detect the intrusion using after-the-fact strategy and the system logs. On-line systems are designed to detect intrusions while they are happening, therefore providing capability for a fast intervention. On-line IDSs are computationally expensive because they require continuous monitoring of the system logs/traffic. In the on-line IDS, decisions have to be made quickly with the minimum data available, therefore they are not reliable enough.

There are different approaches reported in intrusion detection research area[4]. These approaches can be categorized into three groups:

The first group is the statistical-based detection approach. This method is not sensitive to the sequence of the events occurring in the system. It is a difficult task to estimate a threshold for detecting the intrusion. The statistic-based methodology heavily rely on some initial assumptions such as audited data (or user's behaviour) distribution obeying Gaussian distribution. But user behaviour is actually stochastic. These hypotheses may result in high false alarm rate [5]. Most studies on statistical-based anomaly detection techniques are based on a statistical technique developed for IDES/NIDES. Based on a single this technique computes statistics for a test data with a normal distribution (called Q statistic and S statistic). This technique has several drawbacks. First of all, due to the normality assumption this technique is sensitive to the distribution of the test data. On the other hand, this technique is unvaried in such a way that a statistical norm profile is built for only one measure of activities in information system. However, intrusions often affect multiple measures of the activities collectively [6].

Detection in the second group is based on existing rules. In this group using a representation method such as rule-based technique, the behaviour of the system is modelled and explained. Expert systems and fuzzy systems are some examples that might be used in this group. An example of a rule-based intrusion detection system is Snort. Garcia reports that the specificity of the rules in the Snort rule base, limits the effectiveness of the approach. Finding a way to define the rules against potential intruders in an effective way is a general problem with expert system approach [7].

Some researchers have used Evolutionary algorithms as a learning methodology to create rules for

the IDSs [3][7][8][9]. Florez used a set of fuzzy association rules that are mined from audit data to detect anomalous behaviour [10].

If there are no rules to model the behaviour of the system, the only way left will be to model it using training samples. This can be done using different techniques, researchers have reported using Artificial Neural Network and Hidden Markov model are examples for this group [11][12][13][14].

Various feature vectors were extracted to describe the traffic characteristics of the user. It is generally desirable that different vectors in the feature space can be clearly separable from each other. On the other hand similar vectors should be close to each other. From the pattern recognition point of view for good recognition accuracy, the inter-class variances of the vectors should be large and intra-class variances should be small.

Multivariate Analysis Of Variance Algorithm (MANOVA) will search for a linear combination of an original variable set with the largest separation among the groups. These coefficients are computed by measuring the ratio of the between-group variance to the within-group variance. This algorithm produces orthogonal features and therefore the correlation effect on distance measure is reduced [1].

This paper proposes that a combination of MANOVA together with a classification technique can construct a new methodology for intrusion detection. In this algorithm feature reduction can be used to decrease the complexity of the feature space for the problem. In This research, the DARPA data acquired from MIT's Lincoln Labs are used. By using MANOVA algorithm, a linear combination of primary features is produced that includes new distinctive features. Considering the new feature space, the most suitable features were selected. Finally the Nearest Neighbour algorithm is applied to this newly selected data to detect the intrusion. The results out of the implemented method were studied found to be within an acceptable accuracy and speed. Results are reported in section 5.

This paper consists of the following parts: The first part introduces the format of the dataset used as a sample space for the classification system. In the second part the implemented methodology for improving the variances of the inter-classes and intra-classes is presented. The third part deals with the implementation results for the proposed method. The final part is to derive conclusions out of the presented work.

2. DARPA Data for Intrusion Detection¹

¹.This section Information is collected from reference[11].

The data was acquired from the DARPA intrusion detection [12][15] evaluation program. They set up an environment to acquire raw TCP/IP dump data for a local-area network (LAN) simulating a typical U.S. Air Force LAN. They operated the LAN as if it was a true environment, but blasted with multiple attacks. For each TCP/IP connection, 41 various quantitative and qualitative features were extracted. Attacks fall into four main categories:

1. DOS: Denial Of Service
2. R2L: unauthorized access from a remote machine
3. U2R: unauthorized access to local super user (root) privileges
4. Probing: surveillance and other probing

Table 1 shows 32 different exploits that were used in the 1998 DARPA intrusion detection evaluation. This table presents attacks broken up into different categories by their type and operating system.

<i>Attack class</i>	<i>OS: Solaris</i>	<i>OS: SunOS</i>	<i>OS: Linux</i>
Denial of Service	Apache2 Back Mail bomb Neptune Ping of death Process table Smurf Syslogd UDP storm	Apache2 Back Mail bomb Neptune Ping of death Process table Smurf Syslogd UDP storm	Apache2 Back Mail bomb Neptune Ping of death Process table Smurf Syslogd UDP storm
Remote to user	Dictionary Ftp-write Guest Phf Xlock Xnsnoop	Dictionary Ftp-write Guest Phf Xlock Xnsnoop	Dictionary Ftp-write Guest Imap Named Phf Sendmail Xlock Xnsnoop
User to superuser	Eject Ffbconfig Fdformat Ps	Load module Ps	Perl Xterm
Probing	Ip sweep Mscan Nmap Saint Satan	Ip sweep Mscan Nmap Saint Satan	Ip sweep Mscan Nmap Saint Satan

Table 1: Attacks used in DARPA evaluation

A. Denial of Service Attacks

A denial of service attack is a class of attacks in which an attacker makes some computing or memory resource too busy or too full to handle legitimate requests, or denies legitimate users access to a machine. Examples are Apache2, Back, Land, Mailbomb, SYN

Flood, Ping of death, Process table, Smurf, Syslogd, Teardrop, Udpstorm.

B. User to Root Attacks

User to root exploits are a class of attacks in which an attacker starts out with access to a normal user account on the system and is able to exploit vulnerability to gain root access to the system. Examples are Eject, Ffbconfig, Fdformat, Loadmodule, Perl, Ps, Xterm.

C. Remote to User Attacks

A remote to user attack is a class of attacks in which an attacker sends packets to a machine over a network—but who does not have an account on that machine; exploits some vulnerability to gain local access as a user of that machine. Examples are Dictionary, Ftp_write, Guest, Imap, Named, Phf, Sendmail, Xlock, Xsnoop.

D. Probing

Probing is a class of attacks in which an attacker scans a network of computers to gather information or find Known vulnerabilities. An attacker with a map of machines and services that are available on a network can use this information to look for exploits. Examples are Ipsweep Mscan, Nmap, Saint and Satan.

E. List of Features

A complete list of the feature set defined for the connection logging is given in the following tables.

<i>Feature name</i>	<i>Description</i>	<i>Type</i>
Duration	Length (number of seconds) of the Connection	Continuous
Protocol type	Type of the protocol, e.g. tcp, udp, etc.	Discrete
Service	Network service on the destination, e.g., http, telnet, etc.	Discrete
Src_bytes	Number of data bytes from source to destination	Continuous
Dst_bytes	Number of data bytes from destination to source	Continuous
Flag	Normal or error status of the connection	Discrete
Land	1 if connection is from/to the same host/port; 0 otherwise	Discrete
Wrong_fragment	Number of "wrong" Fragments	Continuous
Urgent	Number of urgent Packets	Continuous

Table 2: Basic features of individual TCP connections

Feature name	Description	Type
Hot	Number of ``hot" Indicators	Continuous
Num_failed_logins	Number of failed login attempts	Continuous
Logged in	1 if successfully logged in; 0 otherwise	Discrete
Num_compromised	Number of ``compromised" conditions	Continuous
Root_shell	1 if root shell is obtained; 0 otherwise	Discrete
SU_attempted	1 if ``su root" command attempted; 0 otherwise	Discrete
Num_root	Number of ``root" Accesses	Continuous
Num_file_creations	Number of file creation operations	Continuous
Num_shells	Number of shell Prompts	Continuous
Num_access_files	Number of operations on access control files	continuous
Num_outbound_cmds	Number of outbound commands in an ftp session	Continuous
Is_hot_login	1 if the login belongs to the ``hot" list; 0 otherwise	Discrete
Is_guest_login	1 if the login is a ``guest" login; 0 otherwise	Discrete

Table 3: Content features within a connection suggested by domain Knowledge

Feature name	Description	Type
Count	Number of connections to the same host as the current connection in the past two seconds	Continuous
Error_rate	% Of connections that have ``SYN" errors	Continuous
Rerror_rate	% Of connections that have ``REJ" errors	Continuous
Same_srv_rate	% Of connections to the same service	Continuous
Diff_srv_rate	% Of connections to different services	Continuous
Srv_count	Number of connections to the same service as the current connection in the past two seconds	Continuous
Srv_error_rate	% Of connections that have ``SYN" errors	Continuous
Srv_rerror_rate	% Of connections that have ``REJ" errors	Continuous
Srv_diff_host_rate	% Of connections to different hosts	Continuous

Table 4: Traffic features computed using a two-second time window.

3. Feature Enhancement

Bad features can limit the performance of the system. A feature is bad that can not distinguish between classes so well. An efficient feature extraction can yield to series of good features, however using an effective feature enhancement results can be improved.

As it was mentioned before, DARPA dataset and its features were used as main features in this work, thus only the feature enhancement task has to be performed on the dataset. In other words the extracted features from DARPA dataset will be used as input data for following algorithm. The Output of the MANOVA algorithm is a linear combination of these vectors. The coefficient of each term (feature) is used to show the significance of its value in the classification process.

Once the input dimensions of the data are reduced, a faster classifier can be designed and implemented. The results of experiments in this research show that this reduction will increase the classification accuracy. Therefore, only some selected features resulted from the MANOVA were used for the training.

Algorithm

In the propose approach the MANOVA is used for canonical analysis, where the linear combination of the original variables with the largest separation between groups will be selected. Having the resulted selection from the first run, it can be feed in to the MANOVA for the second and more runs to increase the separation value for the selections. Since this method will produce orthogonal features, calculating the distance between the features it will not be necessary to consider the correlation between them. The proposed algorithm is as follows:

Let's $X_{N \times M}$ be the observation matrix. Rows refer to users typing samples and columns represent the mentioned features.

I. Normalize the features to zero mean.

$$X_m = X_m^{old} - \bar{X}_m \text{ For } m=1 \dots M \quad (1)$$

Where X_m is m^{th} feature (column)

II. Compute Total sum of squares (matrix T)

$$T = X^T * X \quad (2)$$

Where X^T is the transposed of X

III. Compute Within-groups sum of squares(matrix W)

a) Create $W_{M \times M} = 0$

b) If Z be the matrix of instances of class j, normalize its features to zero mean.

$$Z_m = Z_m^{old} - \bar{Z}_m \text{ For } m=1 \dots M \quad (3)$$

Where Z_m is m^{th} feature of class j

$$c) W = W^{old} + Z^T * Z \quad (4)$$

d) Repeat b and c for all classes

IV. Compute Between-groups sum of squares (matrix B)

$$B = T - W \quad (5)$$

V. Calculate the eigenvec and eigenval using (6, 7). eigenvec is a matrix which defines the coefficients of the linear combinations of the original variables. The eigenval is a vector measuring the ratio of the between-group variance to the within-group variance for the corresponding linear combination.

Each value of new features is a linear combination of the mean-centered original variables, using coefficients from the eigenvec matrix.

$$\text{eigenval} = \text{eigenvalues of } W^{-1}B \quad (6)$$

$$\text{eigenvec} = \text{eigenvectors of } W^{-1}B \quad (7)$$

VI. Sort eigenvec with respect to eigenval in descending order and choose K greater of them. So the eigenvec will be a matrix of $M \times K$ dimensional that is selected heuristically through experiments. This selection is such that the first largest K eigenvalues are greater than the rest. Figure 1 illustrates a typical example of an eigenvalue distribution.

VII. Use the eigenvec to project the training sample on a new domain.

$$X = X^{old} * \text{eigenvec}_{M \times K} \quad (8)$$

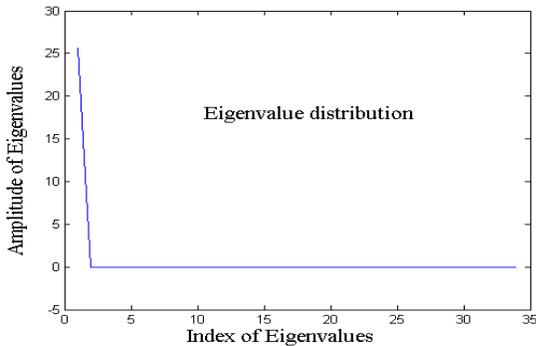


Figure 1: Typical Eigenvalue distribution

4. Intrusion Detection

Output of prior section was an eigenvec matrix that will be used to project new instances with M features on new domain with K dimensions. Let x be the feature vector for a new user, then:

$$x_{1 \times K} = x_{1 \times M} * \text{eigenvec}_{M \times K}$$

Now the class containing this instance should be determined. To do so the nearest neighbour algorithm is

utilised. A new sample is a member of a class if the Euclidean distance between the centre of the class and the sample is calculated to be smaller than a threshold value. Otherwise the user will be identified as unknown.

Figure 2 and Figure 3 show that the first two canonical variables produce more separation between groups than any other pair of original variables.

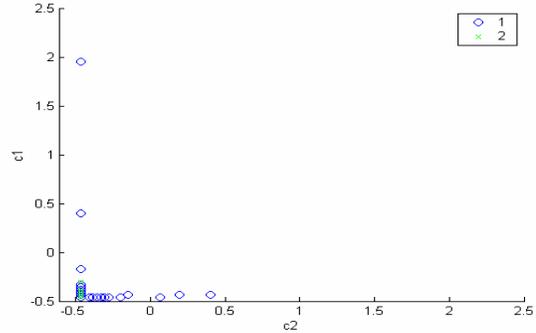


Figure 2: Two original variables

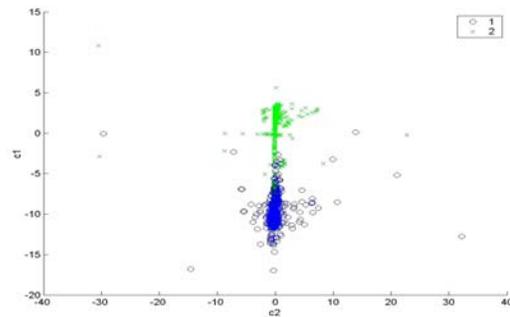


Figure 3: First two canonical variables

5. Experiments and Results

Experiments were performed using randomly selected instances from DARPA dataset. Each one of the instances has 41 features. The feature enhancement phase of the algorithm will affect the final value for the number of the features. Having number of selected features altered, different results will be produced. For classification, the Mahalanobis distance is calculated in a 1-nearest neighbour scheme.

Experiments were run twice each time with a different assumption as follows:

1. 14115 instances are used. 7057 of them for training and 7058 for testing. In this case, there are two classes: normal and intrude samples. Results are presented in Table 5.
2. The classifier determines the type of intrusion. There are five choices can be selected. One of which is the

normal class that presents normal state and four other classes (DOS, R2L, U2R and Probing) show type of the intrusion. 7245 instances were used for both the training and the testing. Table 6 shows the error distribution for this experiment. As it is illustrated in this table False Alarm (FA) and Detection Rate (DR) are respectively equal to 14% and 99.5%.

6. Conclusions

This paper reports a new approach to the design and development of an intrusion detection system. High accuracy and speed are two major advantages in this approach. Using MANOVA algorithm, accuracy of classifier is increased and due to the feature reduction phase the classification time is reduced. The results of this research show that increase in number of features does not increase accuracy. The accuracy of the classifier is determined by the existence of distinguish features in the dataset.

As the correlation between the resulted features is low and the correlation between the different classes is low as well and since the variance between the samples within one class is high, therefore the results have a high quality. This so because they are orthogonal and orthogonal features produce the best results.

References

- [1]. M. Analoui, A. Mirzaei, and S. H. Davarpanah, "Using Multivariate Analysis of Variance Algorithm in keystroke Identification", 2th International Symposium on Telecommunications, August 16-18 2003, Isfahan, Iran.
- [2]. R. Heady, G.Luger, A. Maccabe, and M. Sevilla, "The Architecture of a Network-level Intrusion Detection System", Technical report, CS920-20, Dept. of Computer Science, University of New Mexico, Albuquerque, NM 87131.
- [3]. J. Gomez, and D. Dasgupta, "Evolving Fuzzy Classifiers for Intrusion Detection", Workshop on Information Assurance, United States Military Academy, West Point, NY June 2001.
- [4]. D. Denning, "An intrusion model", IEEE Transaction on Software Engineering, SE-13(2):222-232, 1987.
- [5]. L. Wang, Ge Yu, G. Wang, and D. Wang, "Method of Evolutionary Neural Network-based Intrusion Detection", 2001. Proceedings. ICII 2001 - Beijing. 2001 International Conferences on Info-tech and Info-net, Volume: 5, 2001,Page(s):13-18vol.5
- [6]. Nong Ye, Senior Member, IEEE, Syed Masum Emran, Qiang Chen, and Sean Vilbert, "Multivariate Statistical Analysis of Audit Trails for Host-Based Intrusion Detection", 810-820, IEEE TRANSACTIONS ON COMPUTERS, VOL. 51, NO. 7, JULY 2002
- [7]. R. C. Garcia, and J. Cannady, "Boundry Expansion of EcperT Systems: Incorporating Evolutionary Computation with Intrusion Detection Solutions",IEEE 2001.
- [8]. C. Sinclair, L. Pierce, S. Matzner, "An Application of Machine Learning to Network Intrusion Dtection"
- [9]. F. Neri, "Comparing Local Search with respect to Genetic Evolution to Detect Intrusions in Computer Networks",IEEE 2000.
- [10]. G. Florez, S. M. Bridges, and R. B. Vaughn, "An Improved Algorithm for Fuzzy Data Mining for Intrusion Dtection", IEEE 2002.
- [11]. S. Mukkamala, G. Janoski, and A. Sung, "Intrusion Detection Using Neural Networks and Support Vector Machines", IEEE, 2002.IJCNN '02. Proceedings of the 2002 International Joint Conference on Neural Networks, Volume: 2, 2002 Page(s):1702-1707
- [12]. J.M. Bonifacio, A.M. Cansian, A.C.P.L.F. De Carvalho, E.S. Moreira, "Neural networks applied in intrusion detection systems", Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on , Volume: 1 , 4-8 May 1998, Page(s): 205 -210 vol.1
- [13]. S.C. Lee, D.V. Heinbuch, "Training a neural-network based intrusion detector to recognize novel attacks ", Part A, IEEE Transactions on Systems, Man and Cybernetics, Volume: 31 Issue: 4 ,Jul 2001, Page(s): 294 -299
- [14]. C. Sung-Bae, "Incorporating soft computing techniques into a probabilistic intrusion detection system", Part C, IEEE Transactions on Systems, Man and Cybernetics, Volume: 32 Issue: 2 , May 2002, Page(s):154-160
- [15]. <http://kdd.ics.uci.edu/databases/kddcup99/task.html>

K	2	3	5	9	10	34	41
False No	78	88	108	133	108	58	5717
Accuracy	98.89	98.75	98.47	98.12	98.47	99.17	0.19

Table 5: Effect of Number of selected features

Class no.	1	2	3	4	5	Total Instance No	Accuracy Percent
1	1190	109	1	0	86	1386	85.86%
2	3	206	0	0	0	209	98.56%
3	1	34	5558	0	0	5593	99.37%
4	1	0	0	0	0	1	0%
5	22	0	0	0	34	56	60.71%

Table 6: Accuracy of different classes in KDD-CUP99 data se

A Method for Upper Bounding on Growth of Network Access Speed

Thomas Phillip Knudsen, Jens Myrup Pedersen, Ole Brun Madsen

Aalborg University, Institute of Electronic Systems

Fredrik Bajers Vej 7C DK-9220 Aalborg Øst, Denmark

Thomas@control.auc.dk; Jens@control.auc.dk; obm@control.auc.dk

Abstract: This paper presents a method for calculating an upper bound on network access speed growth and gives guidelines for further research experiments and simulations. The method is aimed at providing a basis for simulation of long term network development and resource management.

Keywords: Network Resource Management, Internet Access, Large Scale Networks, Computer Networks, Network Planning.

Introduction:

In the past few years, the driving force behind growth in Internet access speed has been visual context and increasingly real time streaming video. At the same time access speeds are approaching the limit for what the old copper-based access nets can carry; the MPEG2 standard specifies convergence to usage profiles with speeds up to 100 megabit/s [1], well above the limits for the xDSL technologies. The observed growth rates for Internet access speed have been exponential with rates of around 25% per year; this has been formulated as Nielsen's Law for NT access speed [2], predicting continued growth at this rate. Extending this through the twenty-first century and beyond leads to figures in the environ of:

Table 1 Projected network access speed

Years from now:	Speed:
Now	1 Megabit/s
10	10 Megabit/s
20	87 Megabit/s*
40	7.5 Gigabit/s
50	70 Gigabit/s
75	18 Terabit/s
100	4.9 Petabit/s
150	344 Exabit/s = 344.000.000.000.000 Megabit/s

* suffices for MPEG2 high level-main profile.

In 1032 years there will be 1 Megabit/s for each of the estimated 10^{100} nuclear particles in the universe.

It should be noted that some plans operate with higher figures in the short term, such as the Swedish IT commission, aiming at 5 megabit/s in a new fiber optical access network by 2005, with feasibility for doubling every year up to 100 megabit/s [3].

It is obvious that such growth must be bounded at some point, but there is no reason to assume this bound will be at 100 megabit/s. In resource management, dimensioning and planning for future access nets and actual upgrading of the MAN-WAN, it is highly relevant to have an estimate of how far this development can go on, particularly in view of the life expectancy of such network cabling being counted in decades. Such an estimate can be reached by analysis of the bandwidth needed to provide a human observer with full virtual experience in real time; at this point any further detail introduced into the virtual environment would be impossible for the observer to perceive and thus redundant, providing an end to content driven growth.

This paper presents a method for calculating a bound on growth of network access speed based on analysis of the technical requirements for providing the human sensory apparatus with a full virtual environment. The context for this analysis is the global Internet, and the scenario is full real time telepresence with arbitrary contents viewed. The focus will be on the requirements of a single individual. The analysis proceeds in three main steps, where first the premises are given, the global Internet and the human visual capabilities, second a gross bound is calculated from these premises, third different methods for reducing such bandwidth are introduced, and finally an estimate is given and discussed.

Premises.

An estimate of long term bound on demand must be based on factors that can be assumed to remain constant for the timeframe in question; this excludes any consideration of specifics of content to be transmitted, as this is changing in unpredictable ways. The chief constant factor is the human eyesight, characterised by its ability to

perceive details, motion, and colour. For a long term scenario it is therefore assumed, that the maximum demand defining the bound is visual reproduction in real time of a different arbitrary locale, with quality indistinguishable from viewing it directly. The visual reproduction is characterised by the data that must be presented to the observer; no references to or assumptions are made concerning the screen technology used to present it, except that it is not goggles worn by the observer; developments in screen technologies are to be expected mitigating it as a limiting factor [5], and no assumptions are necessary. Based on these, several scenarios must be considered:

For an observer that is static in relation to the screen surface, the image quality should be indistinguishable from the real. An example could be a room where all surfaces are covered by screens displaying a live image of a distant landscape; this should be indistinguishable from viewing the landscape directly. Given the characteristics of human vision it can be calculated what level of detail is necessary, and any detail beyond this level is redundant - anything less is insufficient. This detail level is independent of the distance between the observer and the screen surface, as it is defined by the angular resolution of the human eye.

For a moving observer the image behaviour when approaching an image (the screen) should appear natural, the image should zoom in without delay while retaining full visual quality. For this scenario it is necessary to include the latency in the image update as response to the changed viewing point. Sufficient extra data must be transmitted - buffered in the image - to allow observer-local computation to adjust the image. Given the Internet as context buffering in the image must be sufficient to compensate for global delay.

In order to adjust the estimate on the bound on access speed, known methods for compression must be considered; these fall in three categories:

Compression based on properties of human vision, removing further redundant information, such as Retinally Reconstructed Images (RRI) [6]; destructive compression methods; non-destructive compression methods. The characteristics of normal human eyesight are [6]:

Table 2 Characteristics of human sight.

Angular resolution near focus	1 minute of arc
Time resolution of motion	50 Hz
Colour depth	24 bit

To display a given resolution image point separation must be half that.

Static gross bound

The most basic estimate is based on an image surrounding the observer with a resolution matching that of the eye. This is modelled as a sphere with matching point resolution. Each point contains 24 bit colour information, see Table 3 and Table 4.

Table 3 Number of points on sphere.

Angular resolution of points 1/2'	No. points
Full Sphere	5.9e+8
160° by 90°	2.1e+8

Table 4 Bits/s at 24 bpp colour and 50 Hz update.

Full sphere	7.1e+11
160° by 90°	2.5e+11

Moving gross bound problem

The static bound can be equated with a moving observer, where the future positions are known. The delay factor becomes significant for access speed calculations when the movement of the observer is not fully predictable. In this case additional information must be transmitted to allow the image to zoom in when approached by the observer. The amount of information is a function of the delay in adjusting the transmission to take account of the new observer location and of the uncertainty in predicting that position. Figure 1 illustrates this by showing a movement resulting in a viewing position with higher angular resolution; the outer circle shows the amount of movement possible before the transmission can be adjusted.

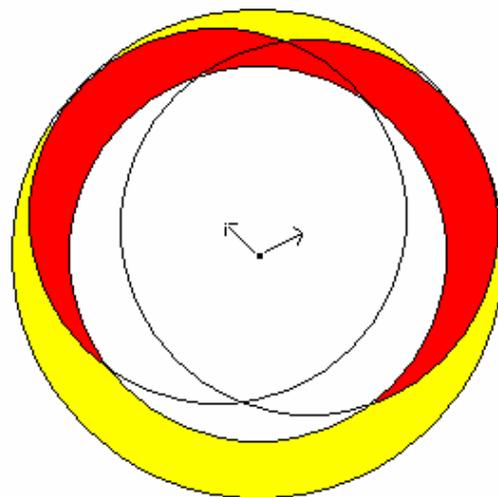


Figure 1 Movement and delay.

Global delay and update times

Factoring in global delay for arbitrary choice of pair of locales presupposes selection and experimentation of transmission technologies and protocol implementations. Some general statistics, though, can be given. If fiber optics is assumed as the basis signal propagation speed is around 200.000 km/s giving a minimum global round trip delay of 0.2 second; if a wireless system, such as Low Earth Orbit (LEO) satellite systems is assumed the round trip delay minimum is 0.13 second. Using LEO for such transmissions is doubtful given the shared bandwidth problem. To these figures must be cumulatively added delays from queuing and routing in network nodes. Advances in optical transmission technologies have made it possible to transmit optical signals over almost global distances without the need for optical-electrical-optical conversion, allowing transmission delay potentially to approach the lower limit set by signal propagation speed in the optical fiber. Routing and switched delays may be minimised using circuit switching.

Compression

Define compression and the main differences destructive and non-destructive; RRI is a grey zone as it prunes away redundant information. It is a modification of the redundant premise from indiscriminate angular resolution to direction specific angular resolution.

Retinally reconstructed images (RRI) is a technique to reduce data in image transmission based on variations in the ability of the human retina to distinguish details [6]. This ability varies greatly, being highest at the point of focus and dropping off fast as the angle from this point exceeds more than a few degrees. Given knowledge of where in an image the eyes focus, resolution in the image can be progressively reduced with angular distance. The degree of compression rises as the image covers a larger part of the field of vision. Up to two orders of magnitude of compression is reported for this technique [6]. Combining this technique with eye tracking equipment and feedback to the transmitter sets up a control loop. The latency of this loop is a limiter on the applicability of this technique as, with growing delay, the uncertainty of where the focus is moving to becomes significant. Here not only transmission delay but also complexity of the reconstruction algorithms must be taken into consideration; these are reported as polynomial to the number of image points processed [6] with figures of tens of MFLOPs for 30 frames/s at 256x256 resolution.

For very short latency this control loop can effectively reduce the amount of data transmitted, but it is reported as doubtful with higher delays, such as 100 ms as well as with saccadic eye movement, where accelerations of $40.000^{\circ}\text{s}^{-2}$ and peak velocities of 500°s^{-1} occur with movement ranging up to 40° [8]. For global distances the moving observer problem greatly reduces the effectiveness of this form of compression, and the feasibility has yet to be demonstrated.

Destructive compression enables potentially unlimited compression, where only the usage tolerance for introduction of a systematic error in the data is a limiting factor. Many effective algorithmic approaches have been developed. The results, whether subjectively evaluated or measured, vary highly with the data characteristics, and choice of, for instance, filter in Wavelet transform based compression is significant [7]. In estimating the bound for arbitrary content it is necessary to take into account worst case scenarios, where no suitable algorithm is found, or where the signal source is unable to determine in real time either the type of data or the receiver tolerance for the systematic error introduced by compression. In such cases defining a maximum bit rate and applying destructive compression compromises quality. As guidelines can be taken the values for MPEG2 compression, Table 5:

Table 5 MPEG2 compression ratios [1].

Profile at Level	Compression ratio
mainP@highL (60Hz)	39.8
highP@highL (60Hz)	31.8
mainP@mainL (30Hz)	19.9
highP@mainL (30Hz)	14.9

Non-destructive compression on the other hand guarantees that full quality is retained. The fundamental principle of this form of compression is to reduce a dataset to a shorter representation while retaining the full information content. The information content of a dataset can be calculated as the entropy of the data [4]. Information theory shows two fundamental approaches to such compression, Huffman compression and Lempel-Ziv compression. For these two the theoretical limit for compression shows that a dataset can be compressed to within one bit of its entropy. For an arbitrary dataset as is the premise here, the problem of this form of compression lies in that it can not be assumed that any compression is achievable; if the dataset is already in a representation with no redundancy, no compression is possible. Therefore any dimensioning of capacity must be based on peak, rather than any other measure of resource consumption; the consequence for upper bounding

with arbitrary content is that non-destructive compression has no effect.

Bound estimate

The upper bound can thus be calculated as a function of the gross bound, movement buffering, RRI and destructive compensation; movement buffering and RRI being dependent on the transmission loop delay. Using these factors with the following values: double angular resolution for movement compensation, two orders of magnitude compression for RRI, and factor 20 for destructive compression, the bounds appear as in Table 6.

Table 6 Bound estimates, values in bits/second

Full sphere 160° by 90°		Without movement compen- sation	With 2x movement compen- sation
R R I	With Destructive compression	3.6e+8 1.3e+8	1.4e+9 5.0e+8
	Without Destructive compression	7.1e+9 2.5e+9	2.8e+10 1.0e+10
No R R I	With Destructive compression	3.6e+10 1.3e+10	1.4e+11 5.0e+10
	Without Destructive compression	7.1e+11 2.5e+11	2.8e+12 1.0e+12

Discussion:

The method presented in this paper defines the necessary components of modelling for an upper bound on content driven growth of network access speed. One key element is missing from the model; this is the added demand stemming from purely technical applications such as distributed control and surveillance. In a WAN context these applications are still only emerging, and their long term properties are less well defined. Also, with such applications distributed data processing gives much greater opportunity for reducing the amount of data transmitted.

The rough figures presented here are, thus, a first indication of an upper bound on network access speed. Given the nature of the global delay problems and quality loss in destructive compression the upper bound on access speed should not be assumed less than 1.0e+11 bits/second. The figures will necessarily undergo some modification as the technological development continues, but the figures have a high degree of robustness to new developments because they rely mainly on factors that can be assumed to

remain constant over time. To achieve a higher reliability for the figures further research in several areas must be carried out. These include determining suitable values for observer acceleration, as well as research on global delay properties and bounding thereof in the long term. Given such further research there is good opportunity for algorithmic formulation of the model and, hence, simulation of long term network growth. But already these figures are significant for resource management in network provisioning, as the figures can be a guideline for dimensioning tube and fiber cable capacity for optical fiber cables when upgrading MAN/WAN, thereby avoiding the cost of digging multiple times. Particularly must be noted the trap of basing an access FTTH (Fiber to the Home) infrastructure on low-capacity multimode optical fibres.

References:

- [1] ISO/IEC 13818-2 "Information technology -- Generic coding of moving pictures and associated audio information: Video"
- [2] J. Nielsen, "Nielsen's Law of Internet Bandwidth," *Jakob Nielsen's Alertbox for April 5, 1998*
<http://www.useit.com/alertbox/980405.html>
- [3] The Swedish ICT Commission, "General guide to a future-proof IT infrastructure" *Report 37/2001, Technical Report*, June 2001 The Swedish ICT Commission, <http://www.itkommissionen.se>
- [4] T.M. Cover, J.A. Thomas "Elements of Information Theory" Wiley-Interscience; 1991 ISBN: 0471062596
- [5] J.T. Klosowski et al., "Deep View: High-resolution Reality" *IEEE Computer Graphics and Applications*. 2002, vol. 22 (3), p. 12-15.
- [6] T. Kuyel et al., "Retinally Reconstructed Images: Digital Images having a Resolution Match with the Human Eye" *IEEE Transactions on Systems, Man & Cybernetics, Part A (Systems & Humans)* Vol. 29 Issue.2 1999
- [7] A. Jensen, A. la Cour-Harbo, "Ripples in Mathematics" Springer Verlag; 2001 ISBN: 3540416625
- [8] P. M. Sharkey, D.W. Murray, "Feasibility of Using Eye Tracking to Increase Resolution for Visual Telepresence." *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*. Vol. 2. p. 1078 -1083, ISBN 0780340531

A Multi-constrained QoS Aware Scheduler for Class-based IP Networks

Pedro Sousa, Paulo Carvalho and Vasco Freitas

Universidade do Minho, Departamento de Informática, 4710-057 Braga, Portugal

email: {pns,pmc,vf}@di.uminho.pt

Abstract: This article presents a novel modular scheduler with powerful semantics able to differentiate simultaneously multiple QoS metrics in class-based IP networks. In opposition to traditional scheduling mechanisms, this scheduler encompasses rate, loss and delay differentiation capabilities in a flexible way. This behaviour stems from new relative and mixed differentiation models able to bound QoS parameters on high sensitive traffic classes.

1 Introduction

The scheduler proposed in this work is able to achieve independent control of delay, loss and rate differentiation, through the use of two priority disciplines acting at distinct points of the proposed scheduler architecture. The delay differentiation modules are based on theoretical schemes [1] and, in particular, proportional differentiation [2, 3] is considered as one of the possible options for delay differentiation. Other differentiation schemes are also supported [4, 5, 6] by the scheduler, including an hybrid model specially devised for real-time differentiation. These delay models aggregate a packet drop mechanism in order to provide (i) loss differentiation or (ii) rate allocation with distinct work conserving behaviour. If required, for specific scenarios, the packet drop mechanism is able to provide simultaneously loss and rate differentiation semantics. The present scheduling proposal can be viewed as a modular traffic control mechanism able to be configured with distinct semantics depending on each class QoS requirements, enhancing the scheduling QoS capabilities of a network node. The proposed model has been implemented and tested in the network simulator (NS-2).

2 Reactive Rate Differentiation

This section focuses on one of the roles of the packet drop mechanism associated with the scheduler. The mechanism is able to induce output rate differentiation among multiple traffic classes by controlling the corresponding loads. Consider that the traffic arriving at a network node, to be forwarded to a specific output link, is classified in N distinct traffic classes contributing with individual loads $R_{in_i}(t)$ with $0 \leq i \leq N - 1$. From queuing theory, the server associated with the corresponding output link enters in an unbalanced

state ($\rho > 1$)¹ when the total traffic class load at the input exceeds the output capacity of the link, C . This situation, illustrated in Eq. (1), leads to packet loss and to different levels of throughput share depending on the service discipline, class load and buffering resources.

$$C < \sum_{i=0}^{N-1} R_{in_i}(t) \quad (1)$$

$$C \geq \sum_{i=0}^{N-1} \min(R_{in_i}(t), R_{max_i}) \quad (2)$$

$$C \geq \sum_{i=0}^{N-1} R_{max_i} \quad (3)$$

The first step in the mechanism design assures that Eq. (1) is not verified, i.e. the total arriving load does not exceed the output capacity of the server. Thus, to each $Class_i$ is assigned a value, R_{max_i} , which is the maximum input rate to be submitted to the server. If $R_{in_i}(t)$ measures $Class_i$ input load at time t then Eq. (2) is valid and assures that the server is always under a balanced state ($\rho \leq 1$)². Assuming N distinct classes, it is clear that the sum of R_{max_i} values should not exceed the output capacity of the server, as denoted by Eq. (3). $R_{in_i}(t)$ is estimated resorting to an adaptive exponential weighted moving average, Eq. (4), where l_i^k is the length of the k^{th} packet of $Class_i$ and $\Delta t_i^k = t_{0_i}^k - t_{0_i}^{k-1}$ is the inter packet arrival time. The parameter T acts as a reference value which should have a similar order of magnitude of the time period for which the estimation module is expected to provide average rate information. In addition, the dropping mechanism was conceived so that the unused share of bandwidth of $Class_i$ is assigned to a variable $credit_i(t)$ (see Eq. (5)) representing the amount of bandwidth provided by $Class_i$ to the differentiation node for subsequent distribution. The sum of all $credit_i(t)$ values is represented by $Credits(t)$ ³.

$$R_{est_i} = (1 - 2^{-\frac{\Delta t_i^k}{T}}) \cdot \frac{l_i^k}{\Delta t_i^k} + 2^{-\frac{\Delta t_i^k}{T}} \cdot R_{est_i}^{old} \quad (4)$$

$$credit_i(t) = \begin{cases} R_{max_i} - R_{in_i}(t) & \text{if } \neg(cong_i) \\ 0 & \text{if } cong_i \end{cases}$$

$$Credits(t) = \sum_{i=0}^{N-1} credit_i(t) \quad (5)$$

¹ $\rho = \lambda \cdot \bar{S}$, λ is the arrival rate and \bar{S} the average service time.

²This means that, assuming enough buffering resources, the server is able to forward all traffic, i.e. on average, the R_{max_i} will also represent the output rate share obtained by the $Class_i$.

³The boolean variable, $cong_i$, is true if $R_{in_i}(t) \geq R_{max_i}$.

Within this work conserving behaviour, Eq. (6) determines the server operating under a balanced state. The function $limit_i(t)$ defines the maximum throughput share for each class. If the traffic class exceeds its R_max_i then $limit_i$ will increase R_max_i of a value given by a given credit distribution function, $dist(t)$ ⁴. The dropping mechanism associated with Eq. (6) is now ruled by Eq. (7) assuring a reactive response to load oscillations and redirecting the unused bandwidth to the congested classes⁵.

$$limit_i(t) = \begin{cases} R_max_i & \text{if } \neg(cong_i) \\ R_max_i + dist(Credits(t)) & \text{if } cong_i \end{cases}$$

$$C \geq \sum_{i=0}^{N-1} \min(R_in_i(t), limit_i(t)) \quad (6)$$

$$drop_prob_i(t) = 1 - \frac{limit_i(t)}{R_in_i(t)} \text{ if } (R_in_i(t) > limit_i(t)) \quad (7)$$

3 Enhanced Delay Differentiation

This section overviews four delay differentiation models included in the proposed scheduler [5, 6]. Lets consider N classes $Class_i (0 \leq i \leq N-1)$ having C_0 the highest priority.

3.1 Proportional Model

Assume that $p_i(t)$ is the priority function associated with the queue i and U_i the corresponding differentiation parameter. In the proportional model this function is given by Eq. (8), with t_{0_i} denoting the arrival time of packet to queue i and $U_0 > U_1 > \dots > U_{N-1}$. Under heavy load conditions, it is expected that Eq. (9) is valid for all classes ($0 \leq i, j < N$) where \bar{d}_i, \bar{d}_j are the mean queuing delays of the classes i and j , i.e. the proportional delay relations are ruled by the U_i parameters.

$$p_i(t) = (t - t_{0_i}) * U_i \quad (8)$$

$$\frac{U_i}{U_j} \approx \frac{\bar{d}_j}{\bar{d}_i} \quad (9)$$

3.2 Additive Model

The additive model differentiates queues by an additive constant as expressed by Eq. (10), with $U_0 > U_1 > \dots > U_{N-1}$. The focus of this model is on the possibility of achieving additive differentiation in class delays, as expressed by Eq. (11),

⁴An example of $dist(t)$ can be a *strict priority* function where credits are allocated to traffic classes according to their priority, i.e. server credits are firstly allocated to high priority classes (see details in Sec. 5.1).

⁵Relaxed versions of the rate differentiation module are possible. For instance, this mechanism may operate only during specific probing periods or the $limit_i$ and $drop_prob_i$ values are only computed for specific time intervals, despite the class rate estimation being continuously updated.

denoting that high priority classes may have a delay gain over low priority classes similar to the difference between the differentiation parameters.

$$p_i(t) = (t - t_{0_i}) + U_i \quad (10)$$

$$[\bar{d}_i - \bar{d}_j] \approx [U_j - U_i] \quad (i > j) \quad (11)$$

3.3 Upper Time Limit Model

This model tries to impose a finite queuing delay, reflected by U_i (see Eq. 12) and, the lower the boundary time is, the higher the priority function slope will be. At the limit ($(t - t_{0_i}) \geq U_i$), the server is *forced*⁶, to dispatch the packet waiting service. This model protects high priority classes, giving that packets remain queued for a maximum value U_i , with $U_0 < U_1 < \dots < U_{N-1}$. This allows to establish delay bounds on the highest priority class and, simultaneously, achieve proportional differentiation between the other classes. For instance, $Class_1$ can be *protected* by a realistic upper time limit, and $Class_2$ and $Class_3$ by *virtual* limits (e.g. $U_2, U_3 \gg U_1$). Proportionality between $Class_2$ and $Class_3$ is obtained as explained by Eq. (8).

$$p_i(t) = \begin{cases} \frac{(t - t_{0_i})}{U_i - t + t_{0_i}} & \text{if } t < t_{0_i} + U_i \\ \infty & \text{if } t \geq t_{0_i} + U_i \end{cases} \quad (12)$$

3.4 An Hybrid Delay Model

This model is useful to distinguish real-time traffic with distinct sensibilities to queuing delays and excess delays. In this model, the priority function assumes values starting from an infinity negative reaching zero when the queuing time of the packet matches the upper time parameter. After that, if the packet is still enqueued, the scheduler switches to a new working region of positive values, where the priority behaviour is ruled by a *congestion parameter*⁷ which determines the slope of the priority function. The final priority function is given by (13).

$$p_i(t) = \begin{cases} \frac{\delta_t - U_i}{\delta_t} & \text{if } \delta_t < U_i \\ (\delta_t - U_i) * C_i & \text{if } \delta_t \geq U_i \end{cases} \quad (13)$$

with $\delta_t = t - t_{0_i}$ and $0 \leq i \leq N - 1$.

$$d_i = d_i^o + d_i^* \quad (14)$$

The total delay (d_i) affecting $Class_i$ can be divided in two components: one induced by the priority function when it assumes negative values ($t < t_{0_i} + U_i$), which we call *upper time delay*, d_i^o , and the other one when the function assumes positive values, which we call *congestion delay*, d_i^* (see Eq. (14)). The magnitude of

⁶When congestion occurs or the load of high priority classes becomes very high, the time limit may be exceeded.

⁷In this context, the term *congestion* is used in a relaxed way as it may reflect heavy load conditions in the server; heavy load conditions in $Class_i$ impairing the expected upper time limit or feasibility problems in the differentiation parameters.

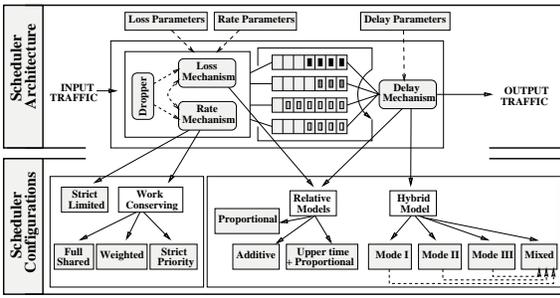


Figure 1: The scheduler architecture implemented in NS-2.

d_i° is controlled by U_i whereas C_i controls the magnitude of d_i° . This means that fundamental differentiation relations among classes, i.e. $d_0 \leq d_1 \leq \dots \leq d_{N-1}$, can be achieved through different combinations of d_i° and d_i^\bullet , and consequently by different combinations of U_i and C_i . In summary, a distinct delay behaviour can be induced depending on the relations between the *upper time* and *congestion* delays of the traffic classes⁸.

4 Enhanced Loss Differentiation

As regards packet loss, the model resorts to Eqs. (8), (10) and (12) to achieve loss differentiation. In this case, instead of using the packets queuing time, i.e. $(t - t_0)$, the models use the ratio l_i/A_i , with l_i being the number of packet drops and A_i the number of packet arrivals for *Class*_{*i*}⁹. Whenever the buffer overflows, the class selected to drop a packet is the one with the lowest $p_i(t)$ value. The traffic classes are configured with loss differentiation parameters $L_0 > L_1 > \dots > L_{N-1}$.

5 A Multi-constrained QoS Engine

The scheduling architecture presented in Fig. 1 aggregates all the previously explained differentiation mechanisms and was implemented in NS-2. This section illustrates that the proposed scheduling architecture is able to decouple the rate, loss and delays differentiation behaviour, i.e. the differentiation mechanisms can act jointly but, simultaneously, can provide independent QoS metric differentiation. Due to the high number of possible differentiation schemes this section only covers examples of specific configuration modes. The selected examples were taken from a scenario where three classes contend for a 4.5 Mbps capacity link, with packet lengths of 500 bytes uniformly distributed over the interval [250, 750]. The scheduler was tested successfully for distinct traffic sources as CBR, exponential, pareto and

⁸For instance, in Sec. 5.1, the classes are configured with distinct upper time parameters, having two of them similar congestion parameters. In the example included in Sec. 5.2, the classes have distinct congestion parameters, having two of them similar upper time parameters.

⁹ l_i and A_i are measured for a specific time interval which can be configured in the differentiation node.

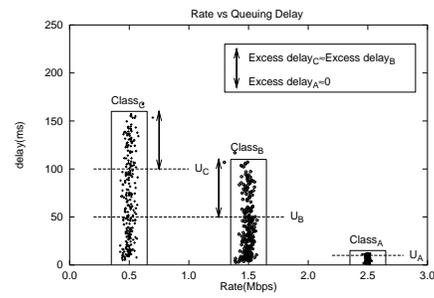


Figure 2: Rate differentiation with hybrid delay model (conf. II+III) for $(R_{maxA}, R_{maxB}, R_{maxC}) = (2.5 Mbps, 1.5 Mbps, 0.5 Mbps)$, $(U_A, U_B, U_C) = (10ms, 50ms, 100ms)$ and $(C_A, C_B, C_C) = (20, 1, 1)$.

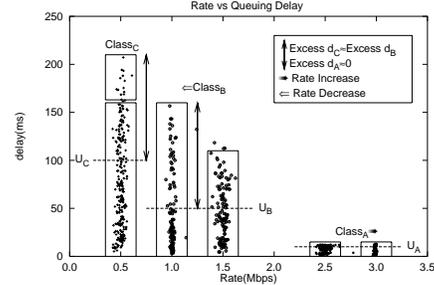


Figure 3: Strict priority rate with hybrid delay model (conf. II+III) for $(R_{maxA}, R_{maxB}, R_{maxC}) = (2.5 Mbps, 1.5 Mbps, 0.5 Mbps)$, $(U_A, U_B, U_C) = (10ms, 50ms, 100ms)$ and $(C_A, C_B, C_C) = (20, 1, 1)$.

combinations thereof¹⁰.

5.1 Rate vs. Delay

Strict Priority Rate Model with Hybrid Delay - This example illustrates the use of the hybrid delay differentiation module and the rate differentiation module for the configuration parameters shown in Fig. 2. In the delay configuration mode, *Class*_A is the highest protected class as regards both rate and delay violations and *Class*_B and *Class*_C have distinct upper time parameters but similar congestion parameters, meaning that they have similar sensibility to absorb excess delays despite having different upper time delays. Fig. 2 shows the average output rate (x-axis) and queuing delays (y-axis) obtained by the classes, clearly corroborating the expected differentiation behaviour. Fig. 3 illustrates this delay differentiation mode and the strict priority rate differentiation. The rate credits of the server are now first distributed to the high priority classes and the remaining credits, if any, are allocated to low priority classes. With this purpose, Fig. 3 plots the differentiation behaviour when *Class*_B decreases its rate to 1 Mbps. As plotted in Fig. 3, only *Class*_A, which has the highest priority, has assigned extra bandwidth, being shifted to the right side of the graph with an offset of 0.5 Mbps, exactly the share provided by *Class*_B. As a consequence, a new delay distribution is achieved by the server and both *Class*_B and *Class*_C delays increase. For *Class*_C, all plots are

¹⁰For a simulation period of 120s with a QoS metric evaluation interval of 1s and the overall class load above the link capacity to force packet loss.

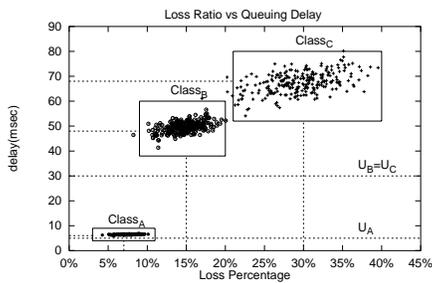


Figure 4: Proportional loss and hybrid delay model for Conf. I+II, with $(U_A, U_B, U_C) = (5ms, 30ms, 30ms)$, $(C_A, C_B, C_C) = (40, 2, 1)$, $(L_A, L_B, L_C) = (4, 2, 1)$.

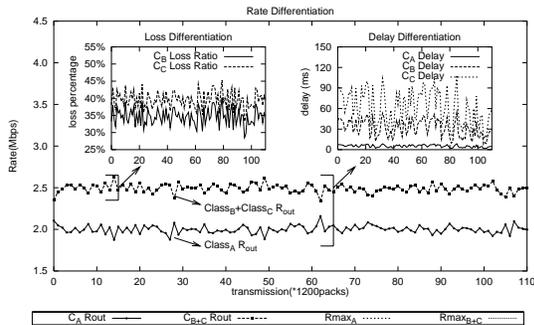


Figure 5: Rate differentiation with additive loss and upper time delay model with $(R_{maxA}, R_{maxB+C}) = (2Mbps, 2.5Mbps)$, $(L_B, L_C) = (0.05, 0)$, $(U_A, U_B, U_C) = (10ms, 100ms, 200ms)$.

still centered on $0.5Mbps$ as this class has not received any extra bandwidth share. The increase in $Class_C$ excess delays is represented by a second box above the previous obtained delays. The magnitude of $Class_B$ and $Class_C$ excess delays is still similar even after the rate sharing, while $Class_A$ delay violations keep a low value due to its high C_A parameter.

5.2 Loss vs. Delay

Proportional Loss and Hybrid Delay - In this example the classes were configured to have proportional loss differentiation with $(L_A, L_B, L_C) = (4, 2, 1)$. They are also configured with the hybrid delay differentiation mechanism in the mixed configuration I+II with $(U_A, U_B, U_C) = (5ms, 30ms, 30ms)$ and $(C_A, C_B, C_C) = (40, 2, 1)$. This means that a proportional packet loss is expected and, due to a very high congestion parameter, $Class_A$ should have queuing delays close to $5ms$. In addition, the congestion delays of $Class_C$, i.e. the difference between the obtained delays and the target delay of $30ms$, should be twice the congestion delay of $Class_B$, which have a similar delay target of $30ms$, but a congestion parameter two times higher than $Class_C$. This behaviour is illustrated in Fig. 5.2 showing the delay vs. loss experienced by the classes during the simulation.

5.3 Rate vs. Loss vs. Delay

Rate Differentiation with Additive Loss and Upper Time Delay - This example illustrates the three differenti-

ation modules acting together. It was assumed that $Class_A$ is used for high loss and time sensitive traffic and the traffic load is controlled at network edges imposing to differentiation nodes a bandwidth allocation of $2Mbps$ for the class. $Class_B$ and $Class_C$ are used for low priority traffic and, depending on the network conditions, packet loss is likely to occur. In this context, the rate parameters were configured as $(R_{maxA}, R_{maxB+C}) = (2Mbps, 2.5Mbps)$. The additive model was used to guide loss differentiation between $Class_B$ and $Class_C$ with $(L_B, L_C) = (0.05, 0)$, meaning that $Class_B$ should experience a loss percentage which is 5% lower than the obtained by $Class_C$. Finally, the upper time model was used to limit the queuing delay of $Class_A$ to a maximum value of $10ms$, with proportional relations between $Class_B$ and $Class_C$. As depicted in Fig. 5, the results show the correctness of this configuration.

6 Conclusions

This article presents a modular scheduler architecture providing enhanced rate, loss and delay differentiation behaviour. The diversity of the configuration modes for the three QoS metrics turns the proposed scheduler in an useful component to be used in network scenarios aiming at QoS differentiation. The proposed scheduler allows to achieve independent QoS metrics differentiation behaviour, avoiding coupling effects which may affect other differentiation mechanisms. Due to the enhanced differentiation semantics, many combinations of rate, loss and delay differentiation behaviour are possible using a small set of simple and intuitive configuration parameters.

References

- [1] G. Bolch et al. *Queueing Networks and Markov Chains - Modeling and Performance Evaluation with Computer Science Applications*. John Wiley and Sons INC., 1998.
- [2] C. Dovrolis et al. A case for relative differentiated services and the proportional differentiation model. *IEEE Network Magazine*, 1999.
- [3] C. Dovrolis et al. Proportional differentiated services: delay differentiation and packet scheduling. *IEEE/ACM Transactions on Networking*, 10(1), Feb. 2002.
- [4] P. Sousa et al. End-to-end delay differentiation of IP traffic aggregates using priority queueing models. In *Proc. of the IEEE Workshop on High Performance Switching and Routing (HPSR2002)*, pages 178–182, Kobe, Japan, May 26-28 2002.
- [5] P. Sousa et al. Tuning delay differentiation in IP networks using priority queueing models. In E. Gregori et al, editor, *Proc. 2nd International IFIP-TC6 Networking Conference*, pages 709–720, Pisa, Italy, 2002. LNCS 2345, Springer-Verlag.
- [6] Pedro Sousa et al. Scheduling Time-Sensitive IP Traffic. In G. Goos et al, editor, *Proc. 6th IFIP/IEEE MMNS International Conference*, pages 368–380, Northern Ireland, Belfast, 2003. LNCS 2839, Springer-Verlag.

A New Architecture and MAC Layer for Quality of Service Support in Wireless Home Networks

Robert Sizeland, Abraham Fapojuwo and Robert Davies
University of Calgary and TRILabs, Calgary, Alberta, Canada
{sizeland,fapojuwo,davies}@cal.trilabs.ca

Abstract—Wireless home networks and high quality digital entertainment have both undergone tremendous growth in recent years. The next logical step is to provide high bandwidth multimedia streams over a wireless network in the home. Lots of research has been done to add Quality of Service for multimedia traffic to the list of features of IEEE 802.11 networks, but this protocol was not designed to carry this type of traffic. This paper proposes a new MAC protocol and network architecture built specifically for the distribution of high quality multimedia over a wireless link in the home. It is shown that the proposed protocol exhibits a quality that rivals that provided by a wired solution.

I. Introduction

As the services home networks provide evolve from printer and file sharing to intelligent home management centres and entertainment sources, the underlying structure of these networks must also evolve. The hardware and protocols must be designed with the environment and traffic type seen in a home in mind. Simple control commands to many already wired fixtures will require one type of physical layer (PHY) and low level protocol, while large, time sensitive multimedia streams to a few mobile devices will require another physical layer and set of protocols.

While some networked services in the home can be provided over existing wiring (power and phone lines), the bandwidth and delay required for high quality digital video and sound requires either new wires or a wireless solution. Since new wiring is expensive and inconvenient, in many cases a wireless solution is necessary. Not only will a wireless home network for media need to cope with the wireless channel, it must provide Quality of Service (QoS) in terms of latency and bandwidth guarantees to the traffic it carries.

This paper proposes a Medium Access Control (MAC) sublayer dubbed MediaMac, designed with QoS for multimedia streams over a high bandwidth wireless link. Also presented is how MediaMac and its corresponding physical layer would integrate into a home network. The protocol is simulated with the ns2 network simulator [1] and compared with a simulation of an IEEE 802.11 [2] network using an Enhanced Distributed Channel Access (EDCA) style of differentiated channel access.

The remainder of this paper is organized as follows. Section two justifies the need for a new MAC layer for multimedia traffic and describes related work. The novel

MAC layer is described in some detail in section three, as well as its place in a home network. Simulation results are shown and discussed in section four and section five concludes the paper.

II. Home Networks

IEEE 802.11 networks have become popular and inexpensive, so they are a likely contender for a wireless home network and are therefore the focus of discussion in this section. At first glance, protocols like 802.11a/g, boasting data rates of up to 54 Mbit/s look like they would be suitable for streaming high bandwidth content around the home. This may not be the case, however, as these protocols evolved as a wireless Ethernet replacement. The medium access scheme used is Carrier Sense Multiple Access with Collision Avoidance, where a station requiring the wireless medium waits a random amount of time before transmitting, depending on how heavily the network is loaded. This style of distributed medium access control where each traffic flow is treated equally does not work well with a mix of traffic that has different QoS requirements.

Research has been done that considers the delivery of multimedia over 802.11 networks, such as [3] and others, but they often propose a change in the standard or vendor specific algorithm. 802.11 networks are best used with delay tolerant applications that operate on the TCP/IP stack [4]. Though some traffic in the home fits this description, much of it will be high quality multimedia.

III. A New Access Scheme

With regards to high bandwidth multimedia in the home, the situation is less likely to resemble a networked office and more likely to consist of a home gateway, or computer, with access to several sources of media, namely video discs and a streaming source from a high bandwidth connection from outside the home. Therefore, a centralized networking protocol operating out of the home gateway is appropriate. A centralized approach gives up some of the flexibility of distributed protocols such as IEEE 802.11, but buys some key features required to distribute multimedia in a small network. First, QoS is much easier to manage out of a central location. Not only can packets be scheduled in a controlled way, but deterministic bounds can be set on parameters such

as latency, jitter and bandwidth. Next, admission control can be applied to keep existing streams safeguarded from other traffic. Finally, the management of nodes in the network is simplified.

A wireless media protocol should address the problem of the highly asymmetric nature of the bit rate of high quality media traffic. Depending on the size and resolution of the video, the downlink stream can have a data rate of as much as 20 Mbit/s that requires tight delay and jitter control. The uplink, on the other hand, might only be a few kbit/s of control information and acknowledgment packets that may or may not require QoS. With this in mind, the protocol designed uses a separate, downlink only channel for the media content, while acknowledgment and control information use an existing shared channel.

At the expense of another physical channel, this scheme provides many advantages. First, the downlink can be made contention-free, as requests for resources or management information can be made on the control channel. Next, the new media channel does not have to conform to an existing standard intended for another purpose - it can be custom designed to meet the needs of media traffic. Finally, a separate channel for multimedia allows easier growth to higher bit rates. By dedicating less of the time in the channel to management tasks such as protocol headers and acknowledgment packets, throughput will scale more linearly with channel data rate.

A. Home Network Node Architecture

Each node in the network architecture is set up as in Figure 1, called enhanced nodes. The node works as follows: all packets to or from the node are directed to the entry point, as labeled in Figure 1. An address demultiplexer directs the packets either to the peer of an application that generated them or to the network if the destination warrants. The source/sink block is responsible for generating the media and non-media traffic and is complete from the application to IP layer. The classifier block filters traffic to either the contention-based control network or the non-contention (multimedia) network. OSI Layers 1 and 2 of the contention half of the node are built with the 802.11b standard. This means the physical layer is a 2.4 GHz wireless channel with a 11 Mbit/s maximum data rate. Not only are these network interfaces readily available and cheap, this would allow any enhanced node to communicate with legacy equipment. OSI Layers 1 and 2 of the non-contention downlink half of the node are built with MediaMac, the MAC developed by the authors specifically for multimedia traffic requiring QoS and an 802.11a physical layer. This layer was chosen because it is designed for the indoor channel and has sufficient data rate for high quality multimedia applications, but any wireless physical layer fitting this description could be used.

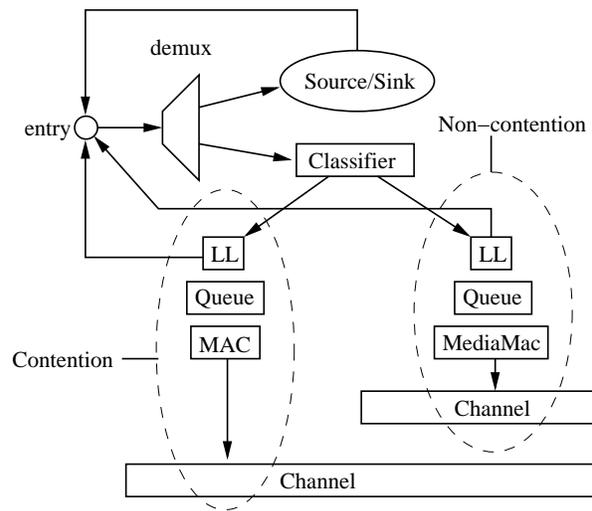


Fig. 1. Structure of a network node enhanced with MediaMac

B. MediaMac Description

A MAC layer designed for the challenges of delivering multimedia over a wireless channel in the configuration specified in Figure 1 must:

- Group incoming packet into logical traffic flows
- Monitor the timing characteristics of packets delivered from each flow
- Schedule packets such that they meet their QoS requirements
- Have a robust acknowledgment/retransmission scheme

1) *General Description:* MediaMac meets these requirements by assigning each flow a separate queue within the MAC and monitoring the status of each queue. Since the uplink and downlink are so asymmetrical, the functions of the server, or home gateway, are different than for a station receiving the media. The basic structure of the MAC at the gateway and station are shown in Figure 2, where N flows are shown in the gateway. Each flow has a set of properties associated with it, including maximum bandwidth and latency, which the scheduler block is aware of. Both the gateway and station are equipped with queues and timers relating to the acknowledgment and retransmission scheme described below.

2) *Scheduler:* The purpose of building a separate channel for multimedia traffic is to provide convenience and scalability to users while preserving the quality of the best wired solution. Since the wireless channel is time-varying and unpredictable, the admission control policy used in a QoS aware high-end entertainment system should be very conservative. This takes many demands off the scheduling algorithm, so a relatively simple scheduler is used. The results generated in this paper are with a scheduler which randomly and fairly selects a flow $n \in [0 \dots N - 1]$, as follows. If x is a

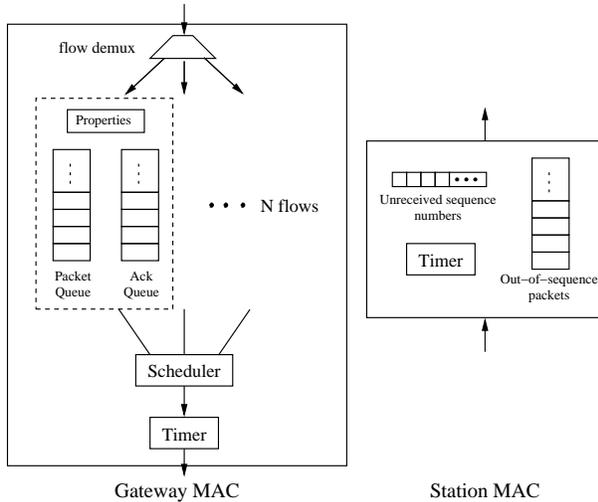


Fig. 2. MediaMac basic structure

random variable uniformly distributed between 0 and 1, and flow i has a maximum bandwidth of BW_i , then flow n is selected if

$$\sum_{i=0}^{n-1} \frac{BW_i}{\sum_{j=0}^{N-1} BW_j} < x < \sum_{i=0}^n \frac{BW_i}{\sum_{j=0}^{N-1} BW_j}$$

and the packet has not exceeded the maximum allowed latency. In that case, the packet is dropped, as further attempts to transmit it would unnecessarily occupy channel time. Although the scheduler can impact how a protocol performs as in [5], our simulation results show even this simplistic scheduler gave favorable results.

3) *Acknowledgment Scheme*: The acknowledgment scheme used must be delay tolerant and reliable, as it is being sent not over a controlled, contention free channel, but over the contention based network which may or may not be loaded with other traffic. A bit vector approach which provides redundancy, as described in [6], was used. For each flow, a monotonically increasing sequence number i is generated and added to the MAC header of data packet D_i . When the receiver successfully demodulates a packet with sequence number M larger than the last acknowledged packet or an acknowledgment timer of T seconds expires, an acknowledgment packet is generated and sent. The packet contains the sequence number i , implying a positive acknowledgment of packet D_i and a bit vector representing the positive or negative acknowledgment of the last A packets, $D_{i-A} \dots D_{i-1}$. The vector is generated by setting bit $(i-1 \dots i-A)$ to 1 for a positive acknowledgment and 0 for a negative acknowledgment. The values used in the simulation were $M = 15$, $T = 60ms$, $A = 128$.

IV. Simulation Results

Two video delivery scenarios were simulated to evaluate the ability of the new MAC protocol to deliver high

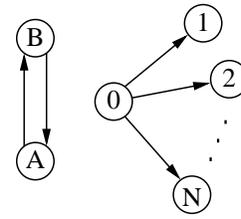


Fig. 3. Network setup used in simulation

quality media in a small network, as well as to compare with a MAC that relies upon distributed control and contention. The IEEE 802.11 protocol with and without EDCA is used for this comparison. Nodes are configured as in Figure 3, with the home gateway being labeled 0, and the stations receiving media are labeled 1 through N. Two additional stations, A and B, transfer traffic on the contention link. The media traffic is generated from a trace file generated from actual video streaming over a loop-back interface from a DVD (Gangs of New York) with a peak rate of 10.0 Mbit/s and an average of 5.50 Mbit/s. Each stream has a random starting position in the file so the traffic is not synchronous.

A. Pure Video Delivery

In this experiment, the background traffic from nodes A and B is non-existent and video streams are established from the gateway to the stations for 300 seconds. There was no difference in the results if more simulation time was used, so 300 seconds was deemed long enough. The enhanced nodes use an 802.11a PHY for the media link and an 802.11b protocol stack for the contention channel. For comparison, an 802.11a PHY with EDCA is used. The delay characteristics are shown in Figure 4. The figure shows how the decreased overhead on the contention free channel allows seven flows when compared to the four flows allowed by 802.11e. These results are important when streaming real-time or near real-time applications. Not only can more streams be delivered at a lower latency, but incrementally adding streams (to a point) does not change the overall delay characteristics drastically.

As data rates go up, the overhead required by the 802.11 protocol causes the MAC efficiency to drop. Using the methods described in [7], the 802.11a protocol at a rate of 54 Mbit/s achieves an efficiency of around 60% at packet size of 1500 bytes. Those same methods yield a maximum efficiency of 90% with MediaMac, the overhead mainly being PHY headers and a 20 byte MAC header. Not only is the MAC more efficient, but the delays are lower for a given number of flows.

B. Video Delivery with Background Traffic

This experiment shows the real value of differentiating the delivery method based on the requirements of the traffic. In all cases, two videos were streamed from the

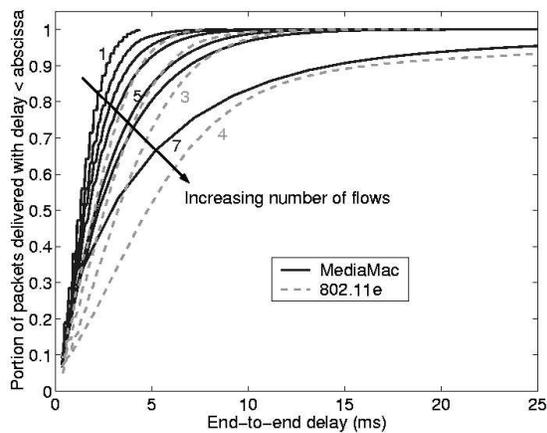


Fig. 4. Delay characteristics with no background traffic

gateway to two stations. Based on the earlier experiment, this does not represent a load that would strain the network. Between 20 and 50 seconds of simulation time, the background traffic is turned on. In order to evaluate the worst case scenario, the background traffic is UDP packets provided at a rate that will keep the queues of nodes A and B full at all times, representing a heavy, bursty user.

A trace that is representative of many simulation runs is shown in Figure 5. Shown in the figure are delay traces of one of the video streams. The delay traces have been low-pass filtered in order to show trends by removing small delay differences in consecutive packets. The filter used is a third order butterworth filter with an impulse response of about one second. Three protocols were used in this experiment, 802.11 with EDCA, MediaMac and 802.11a. The findings are summarized as follows:

1) *802.11 with EDCA*: Here, two traffic classes are created and are differentiated with channel access parameters. Specifically, the video streams have a minimum contention window size of 15 slots, while the background traffic has a minimum contention window of 31 slots. This mechanism seems to be effective at preserving the low delay needed for streaming multimedia.

2) *MediaMac*: The remarkably robust acknowledgment scheme employed in this protocol allows the video to stream with a small delay even when the acknowledgment packets have to traverse a very heavily loaded contention-based channel. The protocol performs at least as well as 802.11 with EDCA.

3) *802.11a*: This experiment is the same as the EDCA experiment except that the background traffic is placed in the same traffic class as the video traffic, with a minimum contention window size of 15 slots. The top of the trace is not shown in the Figure, as it was too high (about two seconds) to show on the same plot as the other protocols. The actual delay depends on the buffer sizes and in a real situation with finite buffer size, many packets would be dropped. This would lead to

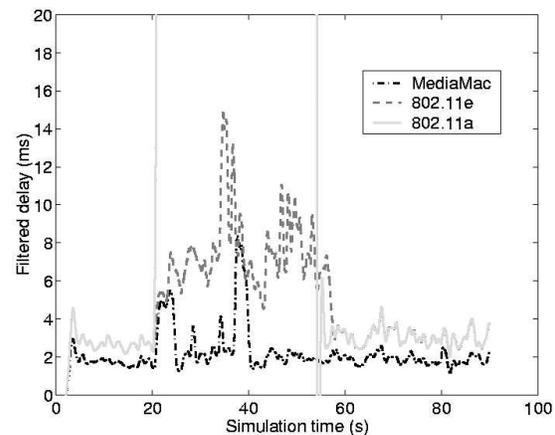


Fig. 5. Delay trace with background traffic from 20-50 s

pausing, popping and clicking or an outright stopping of the video which would be unacceptable in a high quality system. Clearly, this type of protocol does not lend itself to streaming multimedia.

V. Conclusion

While existing wireless data networking protocols, with enough modification, can provide some level of Quality of Service to streaming multimedia, a solution that is built from the ground up is needed for widespread use. As data rates rise, separating network administrative tasks and the delivery of time sensitive media streams is essential in terms of delay and MAC efficiency. MediaMac, a new MAC protocol and network architecture for the home, provides a centralized solution for wireless media while incorporating enough legacy technology to be backward-compatible with existing data applications.

References

- [1] S. McCanne and S. Floyd. ns network simulator. [Online]. Available: <http://www.isi.edu/nsnam/ns/>
- [2] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std., 1999.
- [3] P. Berthou, T. Gayraud, O. Alphand, C. Prudhommeaux, and M. Diaz, "A multimedia architecture for 802.11b networks," in *Proceedings of Wireless Communications and Networking*, March 2003, pp. 1742–1747.
- [4] P. Fowler, "5 GHz goes the distance for home networking," *Microwave Magazine, IEEE*, vol. 3, no. 3, pp. 49–55, September 2002.
- [5] D. E. Wrege, E. W. Knightly, H. Zhang, and J. Liebeherr, "Deterministic delay bounds for VBR video in packet-switching networks: fundamental limits and practical trade-offs," *IEEE/ACM Transactions on Networking*, vol. 4, no. 3, pp. 352–362, 1996.
- [6] H.-S. W. So, Y. Xia, and J. Walrand, "A robust acknowledgement scheme for unreliable flows," in *Proceedings of the IEEE Infocom 2002*, vol. 3, New York, NY, June 2002, pp. 1500–1509.
- [7] J. Jun, P. Peddabachagari, and M. Sichitiu, "Theoretical maximum throughput of IEEE 802.11 and its applications," in *Proceedings of the Second IEEE International Symposium on Network Computing and Applications*, April 2003, pp. 249–256.

A NEW MODEL FOR THE PROBABILITY DISTRIBUTIONS OF CROSSTALK IN TWISTED PAIR CABLES

Nils Holte, Norwegian University of Science and Technology

ABSTRACT

The paper presents a new and more accurate model for calculating the probability distributions of pair-to-pair crosstalk and crosstalk power sum for randomly selected pairs in a twisted pair cable. The calculation is based on averaged measurements of crosstalk in each individual pair combination. Probability distributions are calculated for different cable sizes for one specific type of cable that is based on 10-pair binder groups. The results show that crosstalk power sum is approximately gamma distributed both for near end and far end crosstalk for all cable sizes from 10 to 200 pairs. It is also shown that crosstalk from other 10-pair binder groups gives a significant contribution to crosstalk and should not be neglected.

1. INTRODUCTION

Crosstalk between different pairs is usually the dominating noise mechanism in transmission systems for twisted pair cables. There has recently been large renewed interest in the modelling of crosstalk in twisted pair cables. This is due to the extensive deployment of different types of DSL (Digital Subscriber Line) systems and the use of existing copper cables for broadband access. A further significant improvement of DSL technology is now underway by the introduction of DSM (Dynamic Spectral Management) [1]. The use of DSM strongly increases the importance of accurate crosstalk modelling.

One of the problems in modelling of crosstalk is that the observed level of crosstalk is caused by deviations from nominal cable geometry. Therefore, it is necessary to use statistical models. In 1963, Cravis and Crater [2] published the first successful crosstalk model for twisted pair cables, assuming random coupling along the cable and correctly predicting the average crosstalk as a function of the cable length and frequency.

The planning of twisted pair cable systems is usually based upon the 99% point of the crosstalk power sum; thus the probability distributions of crosstalk are important. Several different models for the probability distributions have been proposed, some based on mathematical modelling in combination with empirical constants [2, 3, 4], and some more empirical [5, 6, 7].

It is worth pointing out that the average coupling between two specific pairs in a cable is in fact a *deterministic* function of the properties of the two pairs. All

previous authors [2, 3, 4] have handled the average coupling statistically, but this is only an approximation. In this paper it is assumed that the average crosstalk is known for all pair combinations in the cable. The probability distributions of pair-to-pair crosstalk and power sum crosstalk for both near end and far end crosstalk are then calculated analytically from the individual probability distributions of the crosstalk in each pair combination. The new method is demonstrated for several cables with different numbers of pairs, all based upon 10-pair binder groups.

2. CROSSTALK MODELS

Near end crosstalk (NEXT) between two pairs in a multi-pair cable will consist of contributions as illustrated in Fig. 2.1. Both pairs are terminated by their characteristic impedances.

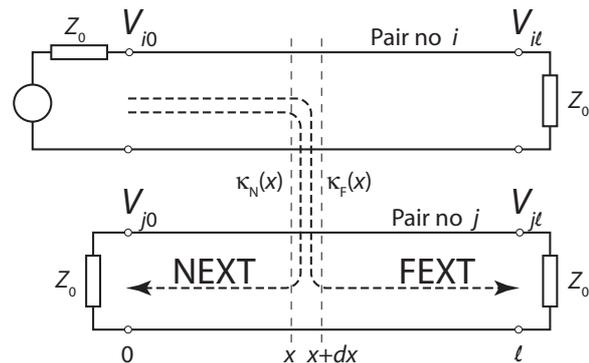


Figure 2.1 Near end and far end crosstalk between two pairs.

The *near end crosstalk transfer function* for high frequencies ($f > 100$ kHz) between pair no. i and pair no. j of a cable is given by Klein [8]:

$$H_{Ni,j}(f) = \frac{V_{j0}}{V_{i0}} = j\beta_0 \int_0^{\ell} \kappa_{Ni,j}(x) \exp(-2\alpha x - 2j\beta x) dx. \quad (2.1)$$

β_0 is the lossless phase constant of the cable in rad/km, α is the attenuation constant of a pair in Neper/km, β is the phase constant of a pair in rad/km, and ℓ is the cable length. The normalised NEXT coupling coefficient between pair i and j at position x is: $\kappa_{Ni,j}(x) = [C_{i,j}(x)/C + L_{i,j}(x)/L]/2$, where $C_{i,j}(x)$ and $L_{i,j}(x)$ are the mutual capacitance and the mutual inductance per unit length between pair i and j , respectively. C and L are the capacitance and the inductance per unit length of the pairs.

The coupling coefficients are, in accordance with [4], assumed to be stationary, white, Gaussian, random processes with autocorrelation functions:

$$R_{Ni,j}(\tau) = E[\kappa_{Ni,j}(x) \cdot \kappa_{Ni,j}(x + \tau)] = k_{Ni,j} \delta(\tau). \quad (2.2)$$

The constants $k_{Ni,j}$ are different for all pair combinations in a cable and can be estimated from crosstalk measurements. For a given cable design, $k_{Ni,j}$ is a deterministic function of the pair combination, and it is mainly determined by the average distance between the two pairs and the difference in twisting periods between them [9].

The *average near end crosstalk power transfer function* for cables of length more than a few hundred metres is found from (2.1) and (2.2):

$$P_{Ni,j}(f) = E[|H_{Ni,j}(f)|^2] = k_{Ni,j} \cdot \beta_0^2 \int_0^\infty \exp(-4\alpha x) \cdot dx = \frac{k_{Ni,j} \cdot \beta_0^2}{4\alpha}. \quad (2.3)$$

For frequencies above 100 kHz, α is proportional to \sqrt{f} and β is proportional to f . Consequently, average NEXT increases 15 dB/decade with frequency.

The *NEXT power transfer function* in one specific pair combination at frequency f , $x_{Ni,j} = |H_{Ni,j}(f)|^2$, will according to Cravis and Crater [2] be gamma distributed [10] with one degree of freedom (equivalent to the exponential distribution).

Far end crosstalk (FEXT) between two pairs in a multi-pair cable is also illustrated in Fig. 2.1. The *far end crosstalk transfer function* for high frequencies ($f > 100$ kHz) between pair no. i and pair no. j of a cable is given by Klein [8]:

$$H_{Fi,j}(f) = \frac{V_{j\ell}}{V_{i\ell}} = j\beta_0 \int_0^\ell \kappa_{Fi,j}(x) dx. \quad (2.4)$$

$\kappa_{Fi,j}(x) = [C_{i,j}(x)/C - L_{i,j}(x)/L]/2$ is the normalised FEXT coupling coefficient between pair i and j at position x .

In the same way as for NEXT, it is assumed that the coupling coefficients are stationary, white, Gaussian random processes with autocorrelation functions:

$$R_{Fi,j}(\tau) = E[\kappa_{Fi,j}(x) \cdot \kappa_{Fi,j}(x + \tau)] = k_{Fi,j} \delta(\tau). \quad (2.5)$$

There is a different constant $k_{Fi,j}$ for each specific pair combination.

The *average FEXT power transfer function* is found from (2.4) and (2.5):

$$P_{Fi,j} = P_{Fi,j}(f) = E[|H_{Fi,j}(f)|^2] = k_{Fi,j} \cdot \beta_0^2 \cdot \ell. \quad (2.6)$$

Hence, average FEXT increases 20 dB/decade with frequency and 10 dB/decade with the cable length for frequencies above 100 kHz.

The *FEXT power transfer function* for one specific pair combination at frequency f , $x_{Fi,j} = |H_{Fi,j}(f)|^2$, will according to Cravis and Crater [2] be gamma distributed [10] with 0.5 degrees of freedom.

3. CABLES USED IN EXAMPLES

The approach presented in the previous chapter may be applied to all types of cable designs; both cables

based upon a layered structure, and cables based upon binder groups of different sizes. PE-insulated twisted pair cables based on 10-pair binder groups are used as an example in this paper. The binder groups are produced by the cross-stranding technique [11]. The conductor diameter is 0.6 mm, and the capacitance is 45 nF/km in the example.

The average near end crosstalk matrix for pairs within a 10-pair binder group (or 10-pair cable), \mathbf{P}_{N10} , has been estimated from an extensive set of measurements [12]. The measurements have been averaged on a power basis over different binder groups and different cables for each specific pair combination within the 10-pair group in order to estimate the different elements of this matrix. An element, $P_{Ni,j}$, denotes the estimated average near end crosstalk power transfer function between pair i and pair j for a single frequency.

The crosstalk matrix is symmetric due to the reciprocity of crosstalk. Corresponding matrices for NEXT in cables containing 30 to 200 pairs have been constructed by means of measurements of crosstalk between pairs in different binder groups. Crosstalk matrices for FEXT have been produced in the same way.

4. PAIR-TO-PAIR CROSSTALK

The crosstalk power transfer function (NEXT or FEXT) for a randomly selected pair combination is denoted:

$$y = y(f) = |H_{i,j}(f)|^2. \quad (4.1)$$

The pair combination (i,j) is uniformly distributed over the $N(N-1)/2$ pair combinations in the cable. Taking the reciprocity into account, the probability density function of NEXT or FEXT for a random pair combination will be:

$$p(y; f) = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N p_{i,j}(y; f), \quad y \geq 0. \quad (4.2)$$

The probability density function, $p_{i,j}(y; f)$, of NEXT or FEXT for pair combination i,j is a gamma distribution as described in Section 2.

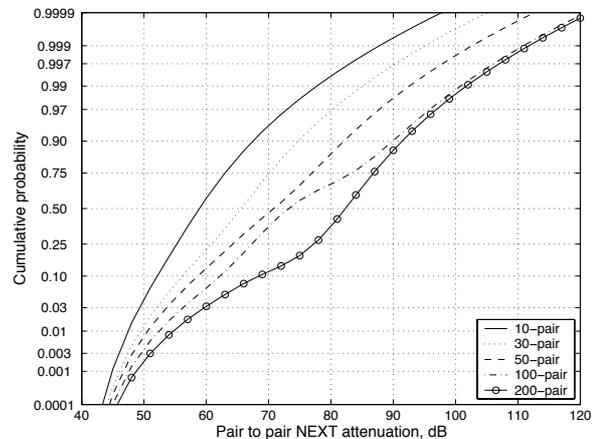


Figure 4.1 Pair-to-pair NEXT attenuation for different cable sizes presented for frequency 1 MHz.

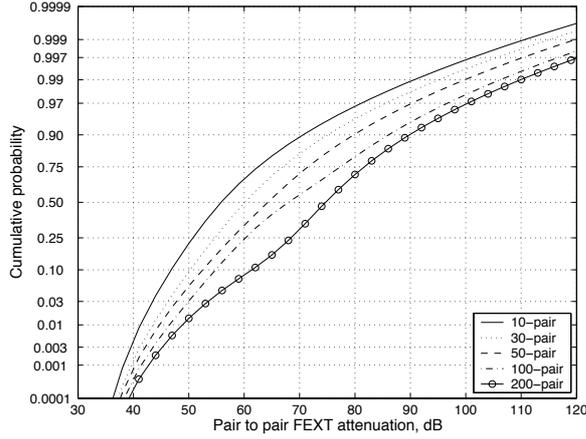


Figure 4.2 Pair-to-pair FEXT attenuation presented for frequency 1 MHz and cable length 1 km.

The probability distributions of pair-to-pair crosstalk for a random pair combination have been calculated according to (4.2) in cables with 10 to 200 pairs. The results are shown in Fig. 4.1 for NEXT and Fig. 4.2 for FEXT. All results are presented for frequency 1 MHz and cable length 1 km in normal probability plots.

It has been common to assume that pair-to-pair crosstalk is log-normal distributed [4], but the results show that this is not true for this type of cable. A log-normal distribution would have been a straight line in the above diagrams.

5. CROSSTALK POWER SUM

For a multi-pair cable, crosstalk from different pairs will add on a power basis because signals in different pair are statistically independent. If the same type of system is used in all pairs of the cable, the effective crosstalk is given by the crosstalk power sum. The *crosstalk power sum* for pair no. i at frequency f is expressed:

$$z_i = z_i(f) = \sum_{\substack{j=1 \\ j \neq i}}^N |H_{i,j}(f)|^2 = \sum_{\substack{j=1 \\ j \neq i}}^N x_{i,j}, \quad (5.1)$$

where $x_{i,j}$ is the crosstalk power transfer function between pairs i and j in one specific cable (NEXT or FEXT). The probability distribution of crosstalk power sum in pair no. i is denoted $p_i(z_i; f)$, and it is found by convolving the probability distributions for all disturbing pairs. This is a convolution of different gamma distributions, and the result is calculated by means of an algorithm given by Moschopoulos [13].

The crosstalk power sum for a randomly selected pair is denoted $z = z(f)$. Exploiting the uniform distribution of i , the probability distribution of z is expressed:

$$p(z; f) = \frac{1}{N} \sum_{i=1}^N p_i(z; f), \quad z \geq 0. \quad (5.2)$$

A gamma distribution is a potential candidate for a simple approximation to $p(z; f)$. The approximated probability distribution of the crosstalk power sum will hence be given by [10]:

$$p_A(z) = \frac{1}{\Gamma(\nu)} \cdot \left(\frac{\nu}{Q}\right)^\nu \cdot z^{\nu-1} \cdot \exp\left(-\frac{\nu \cdot z}{Q}\right), \quad (5.3)$$

where Q is the average power sum and ν is the number of degrees of freedom. The parameters of the approximating distribution are determined by matching the first and second order moments of the crosstalk power sum.

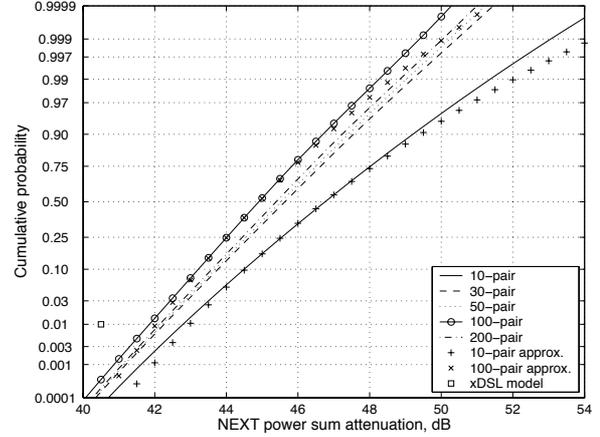


Figure 5.1 Average NEXT power sum attenuation for different cable sizes presented for frequency 1 MHz.

The average power sum for a random pair has been calculated according to (5.1) and (5.2), and the results are shown in normal probability plots in Fig. 5.1 (NEXT) and Fig. 5.2 (FEXT) for frequency 1 MHz and cable length 1 km.

The results for NEXT power sum show that crosstalk within the 10-pair binder groups is the major crosstalk contribution, because there are moderate differences between a 10-pair cable and larger cables. However, the crosstalk contributions from the other binder groups are still significant, and Fig. 5.1 shows that the median (50% point) for NEXT is approximately 2 dB stronger in a 100-pair cable than within a 10-pair group. Hence, the results show that crosstalk between different binder groups may not be neglected for the case of 10-pair binder groups.

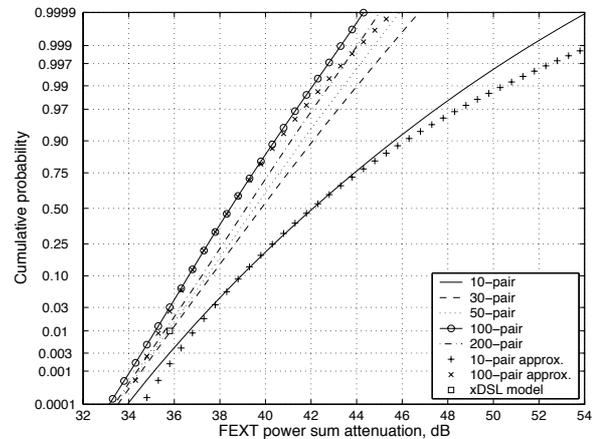


Figure 5.2 Average FEXT power sum attenuation for different cable sizes presented for frequency 1 MHz and cable length 1 km.

The standard model for NEXT power sum used for xDSL systems [6] is also shown in Fig. 5.1. The 99%

point of NEXT power sum is 40.5 dB for a completely filled cable. The differences to the above results are approximately 1.5 to 2.5 dB. This shows that the standard NEXT model used for xDSL is somewhat pessimistic for this type of cable.

The results for FEXT power sum show almost the same basic differences as for NEXT. The median of FEXT is approximately 3.5 dB stronger in a 100-pair cable than within a 10-pair group. Hence, it is even more important not to neglect crosstalk between different binder groups for FEXT.

Figs. 5.1 and 5.2 also show the approximations of NEXT and FEXT power sum by gamma distributions according to (5.3) for 10-pair and 100-pair cables. There are some deviations from the correct distribution for high crosstalk attenuation, but this is of almost no practical importance. The results show that the probability distributions of both NEXT and FEXT power sum may be well approximated by simple gamma distributions.

The standard model for FEXT power sum used for xDSL systems [6] is also shown in Fig. 5.2. The 99% point of NEXT power sum is 35.8 dB for a 100% filled cable. The differences to the above results vary from -0.6 to +0.8 dB. This shows that there is good agreement between FEXT power sum for this type of cable and the standard FEXT model used for xDSL.

The planning of transmission systems for twisted pair cables is usually based on the 99% point of crosstalk power sum. This corresponds to the 1% point in the above figures because crosstalk attenuation in dB is used as abscissa axis. The 99% point (worst-case) of both NEXT and FEXT are shown for two different cable sizes in Tab. 5.1. The table shows that the effect of crosstalk from other 10-pair binder groups is moderate for 99% power sum. FEXT in a 100-pair cable is increased by 1.4 dB in comparison with a 10-pair cable due to the contributions from other 10-pair binder groups.

	Average power sum, dB	ν gamma approx.	99 % power sum, dB	Approx. 99% PS dB	Approx. error dB
NEXT 10-pair	46.5	5.5	42.8	43.0	-0.2
NEXT 100-pair	44.8	10.0	41.8	42.0	-0.2
FEXT 10-pair	41.5	2.7	36.6	36.9	-0.3
FEXT 100-pair	38.3	8.5	35.2	35.4	-0.2

Table 5.1 NEXT and FEXT crosstalk power sum attenuation for different cable sizes at 1 MHz.

The table also presents the 99% point of the approximate gamma distribution. This shows that the 99% point is underestimated by 0.2 dB for NEXT and by 0.2 - 0.3 dB for FEXT. This is a negligible approximation error in comparison with the normal accuracy of crosstalk data.

6. CONCLUSIONS

It is shown how the probability distributions of pair-to-pair crosstalk and crosstalk power sum for randomly selected pairs at a single frequency may be

computed analytically for both near end and far end crosstalk in a twisted pair cable. The calculations are based on measurements of crosstalk in each individual pair combination and the probability distributions of crosstalk between two pairs. Probability distributions have been calculated for cables with 10 to 200 pairs, which are all based on 10-pair binder groups. Furthermore, the results show that crosstalk power sum is approximately gamma distributed, both for near end and far end crosstalk and for all cable sizes from 10 to 200 pairs. The accuracy of this approximation is such that it is suitable for planning of transmission systems for twisted pair cables. The results also show that the crosstalk contributions from other binder groups in a cable are not negligible for cables based on 10-pair binder groups. Finally, it appears that the standard crosstalk models used for xDSL systems are somewhat pessimistic for NEXT, but they are almost completely correct for FEXT in this type of cable.

REFERENCES

- [1] Kee Bong Song et. al., "Dynamic spectral management for next-generation DSL systems," *Commun. Mag.*, vol. 40, pp.101-109, Oct. 2002.
- [2] H. Cravis, T. V. Crater, "Engineering of T1 carrier system repeatered lines," *Bell System Techn. Journ.*, vol.42, pp. 431-486, March 1963.
- [3] S. D. Bradley, "Crosstalk considerations for a 48 channel PCM repeatered line," *IEEE Trans. on Commun.*, vol. 23, pp. 722-728, July 1975.
- [4] A. J. Gibbs, R. Addie, "The covariance of near end crosstalk and its application to PCM system engineering in multi-pair cable," *IEEE Trans. on Commun.*, vol. 27, pp. 469-477, Feb. 1979.
- [5] C. Valenti, "NEXT and FEXT models for twisted-pair North American loop plant," *IEEE Journ. on Sel. Areas in Commun.*, vol. 20, pp. 893-900, June 2002.
- [6] ITU-T Recommendation G.996.1, "Test procedures for digital subscriber line (DSL) transceivers," Geneva, 2001.
- [7] S. H. Lin, "Statistical behaviour of multipair crosstalk," *Bell System Techn. Journal*, vol.59, July-Aug. 1980.
- [8] W. Klein, "Die Theorie des Nebensprechens auf Leitungen," Springer-Verlag, Berlin, 1955.
- [9] N. Holte, "A Crosstalk Model for Cross-Stranded Cables," *Int. Wire and Cable Symp.*, Nov. 1982.
- [10] M. Abramowitz, I. A. Stegun, "Handbook of mathematical functions," Dover Publications, New York, 1970.
- [11] S. Nordblad, "Multi-paired Cable of nonlayer design for low capacitance unbalance telecommunication networks," *Int. Wire and Cable Symp.*, 1971, pp. 156-163.
- [12] N. Holte, "Crosstalk in subscriber cables, Final report," SINTEF report no. STF44 F85001, Trondheim, Jan. 1985, (in Norwegian).
- [13] P. G. Moschopoulos, "The distribution of the sum of independent gamma random variables," *Ann. Inst. Statist. Math. (Part A)*, vol. 37, pp. 541-544, 1985.

A New Proposed Burstness Factor For Real-time Traffic Pricing across the Internet

Hussein Sherif Eissa

Assistant Prof., Computers & Systems Dept., Electronics Research Institute,
El-Tahrir st., Dokki, Cairo, Egypt, Tel.:202-3310503, FAX.: 202-3351631,
hussein@eri.sci.eg, heissa@mcit.gov.eg

Abstract: In this paper, a new factor has been proposed to measure the real-time traffic burstness. Actually, this factor effects the pricing equations for any real-time voice and video traffic that requires a guaranteed quality of service across the Internet. Without having any burstness parameter that indicates the burstness amount in any real-time quality of service session, the ISP (Internet Service Provider) could not estimate the reserved session cost.

Keywords: Burstness Factor, Pricing Equation, Poisson Processes, Real-time Traffic, Internet.

1 Introduction

Normally, at the network traffic-modeling phase, the packets size/inter-arrival are often assumed to be *poisson* processes for analytic simplicity, even though a number of studies have shown that the distribution of packet size/inter-arrival clearly differs from exponential formats [1] especially for the voice and video traffic. The failure of *poisson* modeling is caused by ignoring the burstness occurrence during the real time voice/video session. This leads to unacceptable performance for such guaranteed services. So, the real time voice/video traffic should be analyzed carefully without ignoring the bursty pattern within any studied session. Then, the ISP should be provided with all the factors that effect the obtained pricing equation for the reserved quality of service session.

2 The New Proposed Factor

In this paper, a new factor has been proposed, called Burstness Factor (BF), which measures the unexpected normalized amount of the transfer rate change than the average transfer rate. This new factor has been proposed to provide the ISP with the unexpected amount of the transfer rate change than the reserved transfer rate (extra bandwidth) so, the ISP could reserve sufficient bandwidth for such burstness amounts (i.e. it is a measure for the session burstness). But actually, this factor is insufficient to provide the

ISP with the burstness effects during the session because there is a need to know the burstness time long. So, in this paper the burstness time (S) has been evaluated as well. This time could be defined as the maximum time in which the transfer rate equals to the peak rate. By having both the previous two factors, the ISP could estimate the burstness cost during the pricing phase. Finally, the burstness factor and the burstness time values could be evaluated based on the followings:

$$\text{Burstness Factor (BF)} = \frac{P - R_{avg}}{R_{avg}} \quad (1)$$

$$\text{Burstness Time(S)} = \left(\frac{\gamma_{bursty} \times \max.ps}{P - R_{avg}} \right) - \sum_{i=1}^{i=\gamma_{bursty}-1} t_i \quad (2)$$

Where: R_{avg} is the average rate, P is the peak rate, t_i is the inter-arrival time between two successive packets, $\max.ps$ is the maximum packet size and γ_{bursty} is number of packets during the burstness time.

3 Experiments Setup

In this section, the experiments setup implementation has been presented to study the encoded voice/video traffic characteristics. Actually, many parameters have been measured

and analyzed during these experiments, such as; (maximum and average) packet size, (average and minimum) inter-arrival time, (peak and average) rate, burstness factor, and the burstness time.

The implemented setup consists of; two fully equipped Personal Computers, MS NetMeeting v.3.0 [2] to generate the voice/video traffic, and the NetSUIT analyzer package v.2.0 [3] to capture the voice/video packets that have been generated from the NetMeeting clients. The capture process has been done on an assigned filter (Ethernet/IP/UDP) at the analyzer software. In these experiments, the packets have been captured only from the source to study the voice and video traffic characteristics for the following voice/video studied encoders:

- G.723.1 (6.4), G.723.1 (5.3), G.711 (a-law), G.726 (ADPCM) for voice sessions [4], [5],[6].
- H.263 for video sessions with different image sizes and different video qualities [7].

4 Voice over IP Experiment

This experiment has been implemented to capture the studied voice encoder's traffic. The voice session consists of combination of talk/pause, high/low tone and fast/slow speech verified that no video is being sent or received. Table 1 summaries some of the obtained results from the voice experiment and from equations 1 & 2.

Table 1: Voice encoders' results.

	Max. PS byte	Avg. PS byte	Avg. Int. ms	Avg. rate kbps	P kbps	BF	S ms
G723 (6.4)	138	78	30	21	80	2.8	70
G723 (5.3)	138	74	29.5	20	64	2.2	72
G711	310	310	30.5	81	128	0.6	105
G726	310	310	61	40.5	88	1.2	104
						1.7	88

From the obtained results, it has been concluded that the average voice burstness factor value for the studied types of encoders equals to

1.7. Whereas the average burstness time value equals to 88 ms. During the voice sessions, the bursty pattern with very low probability (ranging from .003 to .03) have been illustrated in figure 1. Whereas the normal patterns, as shown in figure 2, consists of some successive packets with a high occurrence probability (ranging from .96 to 99).

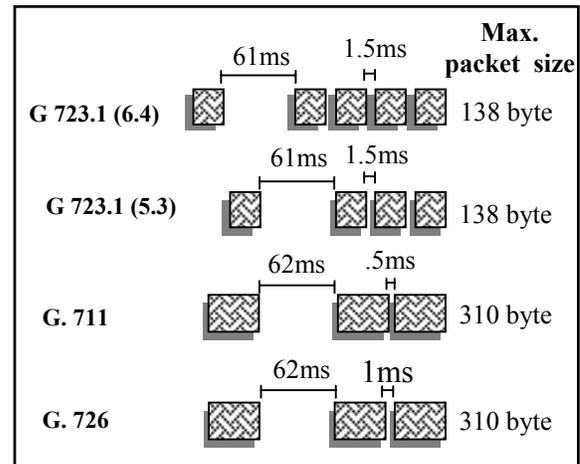


Figure 1: The voice encoders' bursty patterns.

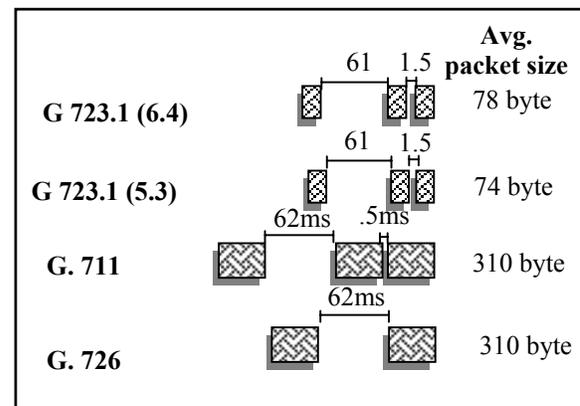


Figure 2: The voice encoders' normal patterns.

5 Video over IP Experiment

This experiment has been implemented to capture the studied video encoders' traffic. The video sessions includes a person moving normally (setting down, standing up, talking, waving his hands/head, and sometimes someone is passing-through as a background). It has been done for small video image with low quality (SQCIF (Fast)), small video image with better

quality (SQCIF (Better)), medium video image with low quality (QCIF (Fast)), and medium video image with better quality (QCIF (Better)). The audio is disabled because this is an only video conference session, verified that no voice is being sent or received. From the obtained results, it is noticed that the video packet size and the inter-arrival time are variables and table 2 summaries the obtained results from the video capturing experiment and from equations 1 and 2.

Table 2: Video encoders' results.

	Max. PS byte	Avg. PS byte	Avg. Int. ms	Avg. rate kbps	P kbps	BF	S Ms
SQCIF (F)	1408	977	34	229	456	1	244
SQCIF (B)	1408	807	27	237	488	1.1	305
QCIF (F)	1408	893	19	378	592	0.6	413
QCIF (B)	1408	863	17.5	394	592	0.5	504
Avg.						0.8	367

From table 2, it could be observed that the average video burstness factor value for the studied types of encoders equals to 0.8. Whereas the average burstness time value equals to 367 ms. Figures 3 and 4 illustrate the video encoders' bursty/normal patterns that have been obtained from the video sessions' experiments with occurrence probability ranging from .0001 to .0002 for the bursty patterns and from .79 to .85 for the normal patterns.

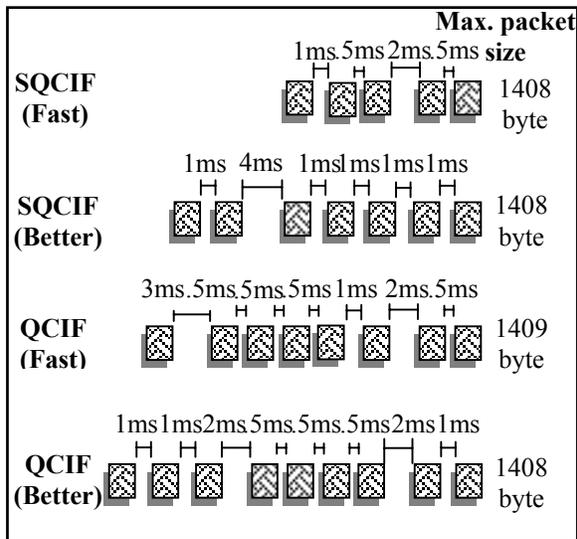


Figure 3: The video encoders' bursty patterns.

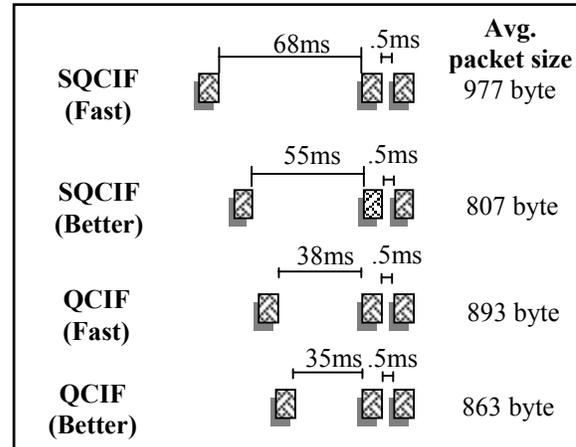


Figure 4: The video encoders' normal patterns.

6 Conclusion

Most of the researchers ignore the burstness pattern of the voice/video traffic because it has a low probability of occurrence. The experimental observations have lead to the fact that; when the burstness pattern occurs, a set of users are most certainly suffer potentially severe performance degradation. However the ignored percentage is low, this value will be enough to face unacceptable performance for the users who will pay for such guaranteed voice/video sessions.

So all of the above mentioned considerations provide a motivation to take a closer look at the behavior of voice and video over IP and experimentally study their pattern shape for different types of encoders.

In this paper, four voice encoders and four video picture formats have been analyzed experimentally without ignoring the burstness patterns. By comparing the obtained voice and the video burstness values, it could be concluded that, although the real-time video traffic need much more bandwidth than voice traffic, but the voice traffic is much more bursty than the video traffic, where the average voice burstness factor is more than the double of the average video burstness factor. But, it should be noted that, the average burstness time of the video is more than four times than the average burstness time of the voice (for only the eight studied encoders). Practically, the burstness factor effects the session cost price more than the average burstness time. This fact could be concluded logically, because the required

unexpected amount of bandwidth is considered one of the most serious problems that face the real-time voice/video traffic. In addition, this paper has concluded by calculations that, in general, the voice traffic is much more bursty than the video traffic.

Finally, it is important to highlight on the proposed burstness factor and burstness time effect on the pricing equations. Without having such factors, the ISP could not estimate both the required network resources for the guaranteed quality of service session, and the session cost. By having both the proposed measurement parameters, the ISP could have an accurate forecasting for the burstness effect during the real-time voice and video session.

References

- [1] V. Paxson & S. Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling," *IEEE/ACM Transactions on Networking*,3(3), pp.226-244, June 1995.
- [2] Microsoft NetMeeting Software, <http://www.microsoft.com/windows/netmeeting>
- [3] NetBoy suite V.2.0 Software, <http://www.ndgsoftware.com>
- [4] ITU-T Recommendation G.723.1, "Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s," Mar. 1996.
- [5] ITU-T Recommendation G.711, "Pulse Code Modulation (PCM) of Voice Frequencies," 1988.
- [6] ITU-T Recommendation G.726, "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," June 1990.
- [7] ITU-T Recommendation H.263, "Video Codec for Low Bit Rate Communication," 1998.

A New Traffic Engineering Approach for IP Networks

Ljiljana Adamovic and Martin Collier

Research Institute for Networks Communications Engineering (RINCE)

Dublin City University, Dublin 9, Ireland

e-mail: {ljiljana, collierm}@eeng.dcu.ie

Abstract

MPLS has received considerable attention as a protocol for transporting IP packets while providing traffic engineering. It requires label distribution, which represents a significant overhead when processing connectionless traffic. This paper describes a new protocol called subIP which efficiently provides connectionless service to IP traffic while remaining the simplicity of MPLS packet forwarding. When used in conjunction with a concept called multiple topology routing, which allows multiple routes to be obtained using standard shortest path routing algorithms, it allows best effort traffic to be flexibly routed across a network domain.

1 Introduction

With the intensive growth of the Internet there is a need for introducing control of traffic routes in IP networks in order both to improve the network utilization and to provide different types of service to the customers. This type of traffic control is called traffic engineering. In general, current proposals for traffic engineering are:

- optimizing link weights in link-state routing protocols in order to cause better traffic distribution with connectionless IP routing [2],

- introducing connection-oriented services with a protocol, such as MPLS [1], and balancing the traffic among established explicit paths based on various criteria.

In this paper we present a new approach. Different virtual topologies are derived from the actual network topology determined by an external entity, such as a network management system or a human operator. They are distributed and used by subIP, a new forwarding protocol proposed in this paper for balancing the connectionless traffic in order to improve network utilization.

subIP adopts ideas from both IP and MPLS. As with MPLS, subIP should be implemented below the IP layer. It adds a 4 byte subIP header between the data link header and the IP header to each IP packet. It may thus be considered as a new version of the MPLS protocol. However, it provides connectionless service similar to IP. A subIP area has no more than 256 routers where each router has a unique 1 byte long subIP address. While traversing the area each packet carries the subIP address of the last area router to be visited on

its way to the destination, the area destination router. Packet forwarding through the area is based on interpretation of this address.

The subIP routing tables are calculated based on the area topology information distributed by the existing routing protocol and the virtual topology information distributed by the subIP Control Message Protocol (sCMP), which is an integral part of subIP. For each topology defined each router calculates shortest paths for the area destination. The value of the *control* field in the subIP header determines which of the topologies should be used on packet forwarding. This concept allows balancing the traffic among shortest paths of different topologies, which may not be the shortest paths in the actual area topology.

subIP is simple. It simplifies and speeds up IP routing and improves network utilization with multiple topology routing. It provides better performance than MPLS connectionless hop-by-hop routing and may be expanded to support MPLS explicit routing. With a new subIP hierarchical addressing scheme and virtual paths established by MPLS there is a prospect of building multiple hierarchy networks with geographically significant addresses, which would significantly simplify traffic engineering.

In this paper we discuss subIP implementation within an autonomous system (AS) implementing the OSPF link-state routing protocol [5]. It may also be applied for networks implementing the IS-IS routing protocol [6], due to the similarity of the two protocols.

Since subIP adopts ideas from both IP and MPLS, a brief overview of the two protocols is given in section 2. The subIP protocol is described in section 3. Multiple topology routing is presented in section 4. Further network developments are discussed in section 5. The paper is summarized in section 6.

2 Existing Protocols

Apart from multiple topology routing which is explained in section 4, the subIP ideas are based on concepts of existing protocols IP and MPLS. We give a short overview of the two protocols in this section. For the MPLS forwarding concept the implementation of a label distribution protocol, such as CR-LDP or RSVP-TE is obligatory. The protocol stack is given in Fig. 1.

As will be discussed in the paper, more flexible protocol

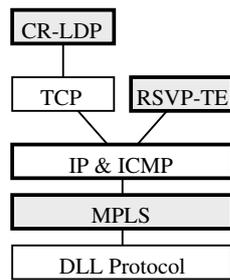


Figure 1: Protocol Relationship

structure can be provided with subIP.

2.1 IP

The Internet Protocol (IP)[4] is the core of the data exchange in the Internet. It provides connectionless, best-effort delivery of datagrams through the network. It also provides fragmentation and reassembly of long packets, if necessary, for transmission through *small packet* networks.

The IP routes packets through the network based on the interpretation of the destination address in the IP header. The information distributed by a routing protocol is used by routers for calculating shortest paths towards the reachable destinations. The results are stored in routing tables and each table entry contains a destination address prefix, next hop to reach the destination and the cost of the path to the destination. On packet forwarding a router finds in its routing table the longest prefix matching the packet's destination address and forward the packet towards the next hop associated with the found longest prefix.

The IP header is at least 20 bytes long. In general, the fields of concern for packet forwarding are: *Source Address* and *Destination Address*, the 4 byte IP addresses of the source and destination of the packet and the 1 byte *Time to Live* (TTL) field, which contains the number of hops a packet may take on its way to a destination. The TTL field is decremented with each hop of the packet. If zero is reached the packet is discarded and a message is sent back to the sender. The purpose of the field is to prevent packets from looping endlessly due to topology changes or in the case of some error.

To provide endnodes with feedback about the problems in the network IP uses the Internet Control Message Protocol (ICMP) [7]. Typically it reports errors in the processing of datagrams, such as *destination unreachable*, *parameter problem*, etc. ICMP is an integral part of IP. Its messages are sent in the data portion of an ordinary IP packet, with the protocol field in its IP header identifying it as an ICMP packet.

IP routing along shortest paths may cause congestion on some network links while longer not used path exists. Current IP implementations do not provide balanc-

ing the traffic among paths of different costs. The traffic may though be balanced among multiple paths of equal cost using ECMP algorithm[3].

2.2 MPLS

Multiprotocol Label Switching (MPLS)[1] is a forwarding protocol implemented below the network layer protocol. Its forwarding procedure is based on labels. The two neighboring nodes negotiate about a number, the *label* to be used on forwarding an IP packet with certain characteristics. The label has local significance. It is stored in the MPLS header added to each packet when it enters an MPLS domain before it is forwarded. At subsequent hops through the domain the label is used as an index into a table which specifies the next hop and a new label to replace the old one before forwarding the packet to its next hop. There is no IP header examination nor longest prefix match table lookup while the packet is traversing the MPLS domain. The scheme is conceptually very similar to ATM cell switching.

The labels are distributed by a separate label distribution protocol, such as CR-LDP or RSVP-TE, which is thus necessary for MPLS operation. Some routing protocols (e.g. BGP) have been extended so that they also distribute the labels, though not OSPF. For establishing label switched paths (LSPs) MPLS defines two types of routing: hop-by-hop and explicit routing. Hop-by-hop routing provides the same paths within the MPLS domain as when IP routing is used. To accomplish it each MPLS router independently assigns a label to each address prefix in its routing table. When the network topology changes the paths are recalculated and new labels need to be assigned and distributed. Since packet forwarding is based on labels, this prolongs the time the network is operating with inconsistent forwarding information and decreases network stability. Alternatively, a router may keep all the labels received from the neighbors assigned for each address prefix, which allows for quicker adaptation to routing changes, but requires many more labels to be maintained [1]. The paths established using explicit routing are determined based on some criteria by the ingress domain routers and labels are distributed along the path on setup.

The MPLS header is shown in Fig. 2. It is 4 bytes long, most commonly encapsulated between the data link header and the network layer header. Beside the la-

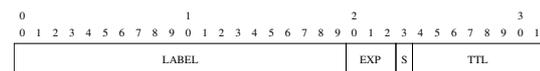


Figure 2: The MPLS Header

bel, it contains 3 experimental bits not yet defined, a 1 bit (S) top of the label stack indication and 8 bits of Time To Live (TTL) field.

The concept of the label stack allows packets to carry more than one label while traversing the network. Each

label is placed in a separate MPLS header and the headers are organized as a *label stack*. Processing of the labelled packet is always based on the label of the top header. This is used when tunnelling packets through a nested MPLS domain [1].

The TTL field in the MPLS header has the same meaning as in the IP header. It is copied from the IP header when the packet enters the MPLS domain and decremented at each hop along the label switched path. If its value reaches zero somewhere along the path, the packet is discarded. When the packet leaves the domain the TTL field is copied back to the packets IP header.

Initially, the main goal of MPLS was to speed up packet forwarding by implementing simple and fast switches within the MPLS domain that forward packets based on short labels placed in a new MPLS packet header. With the emergence of fast longest prefix match algorithms that sped up IP forwarding, the justification for MPLS in IP networks now is its use in balancing the traffic among explicit paths in order to increase network utilization.

3 subIP

subIP provides connectionless packet forwarding in an area with no more than 256 routers. Its routing concept is the same as IP routing. It does not provide packet fragmentation and reassembly, unlike IP. The subIP should be implemented below the IP in the layered protocol architecture (Fig. 3), and it adds a 4 byte header to each

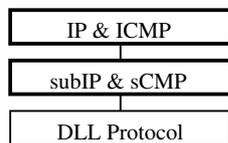


Figure 3: Layered Protocol View

IP packet between the data link and IP header, same as MPLS. However, due to a different routing concept it does not require a label distribution protocol, unlike MPLS.

A unique per subIP area 1 byte subIP router address is assigned to each area router, which bounds the subIP area size to 256 routers. When a packet enters the area the ingress router determines the last area router on the packet's way to the destination, the egress area router, based on the IP destination address in packet's IP header and the information collected by a routing protocol. The ingress and the egress router may be considered as the source and destination of the packet within the subIP area. The ingress router stores the subIP addresses of the source (its own address) and destination (egress) area routers in the subIP header of the packet. The forwarding concept through the area is now the same as in IP routing. Based on the subIP area topology information each area router calculates shortest paths to the routers

within the area using the existing shortest path algorithms. The results are stored in routing tables where each table entry contains area destinations, next hops along the path and the cost of the path to reach those destinations. Since only the area routers are considered as destinations (up to 256 routers) and the destinations are identified with 1 byte subIP addresses, the routing tables are small. At each hop the routing table entry containing the destination subIP address in the packet's subIP header determines the next hop of the packet. There is no IP header examination nor longest prefix match table lookup while the packet is traversing the subIP area.

The subIP area topology information is distributed to each area router by the existing routing protocol implemented in the area. In order to provide some control of the traffic routes, subIP also allows new virtual topologies to be distributed to all the area routers using its subIP Control Message Protocol (sCMP), which is an integral part of the subIP. Its messages are sent in the data portion after the subIP header, with the protocol field in subIP header identifying it as an sCMP packet (similarly to ICMP used with IP). The new topologies are derived from the actual subIP area topology. The area routers calculates shortest paths for all the defined topologies. The ingress area router controls which topology should be used on forwarding with *control* field of subIP header. This allows traffic balancing among shortest paths of different topologies. The multiple topology routing is discussed in section 4.

The subIP depends on the information collected by a routing protocol. We discuss its implementation in an autonomous systems (AS) implementing the OSPF [5] link state-routing protocol. The OSPF defines two hierarchy levels in an AS, areas that communicate over the second level backbone. Our subIP areas will be determined by the areas defined by the OSPF, including one subIP area for the backbone.

3.1 subIP Header

The format of the subIP header is shown in Fig. 4. The *control* field is used for specifying the topology to be used on packet forwarding in multiple topology routing, explained in section 4. It may also be used in hier-

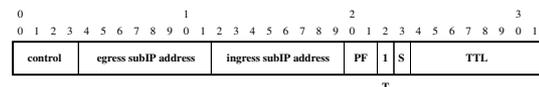


Figure 4: The subIP Header

archical networks to indicate the level of hierarchy. The *egress subIP address* presents the subIP area router address of the egress area router. It is used for determining the next hop on packet forwarding. The *ingress subIP address* presents the subIP area router address of the ingress area router. It may be used on sending an ICMP messages from within the area in order

to avoid determining the egress router (and thus longest prefix table lookup) on sending messages back to the source.

Both *ingress* and *egress subIP addresses* may be used as an ID of the ingress-egress flow aggregation, for balancing the traffic among equally shortest paths within the area, similar as in IP routing.

The *PF* field is the protocol field. It allows only 4 different protocol to be implemented *on the top* of subIP. Two values of this field are reserved for the IP and subIP Control Message Protocol(sCMP).

The T (type of service) bit is set to 1 to indicate subIP connectionless service. If it is set to 1 it may indicate the MPLS connection-oriented service and the MPLS header, as will be discussed in section 5.

The S bit is the top of the label stack indication. It is used to allow sending more than one subIP header with one packet organized as a label stack, same as in MPLS. The TTL field contains the number of hops a packet may take on its way to a destination. As in MPLS, it is copied from the packet's TTL field in the IP header when an IP packet enters the subIP area and decremented at each hop. If zero is reached the packet is discarded, else the field is copied back to the IP header TTL field when the packet leaves the area.

3.2 Addressing the Domain

Each router in a subIP area has a unique 1 byte subIP address. As shown in Fig. 5, the subIP network with new addresses underlies the existing IP network. While the IP addresses are assigned per router's interface, the subIP addresses are assigned per router and present a unique per area prefix of all the IP addresses of a router. In order to identify different subIP areas within a two level hierarchy AS, we assign a 4 byte hierarchical subIP AS router address to each AS router in the form *AS:backbone router:area:area router*, where, for simplicity, each byte represents a subIP address of the AS, backbone router, area and area router, respectively. The length of these addresses may depend on the number of areas and routers, as long as the backbone router address and the area router address are not more than 1 byte long. They are used on subIP forwarding within the backbone and an area, respectively. The new subIP

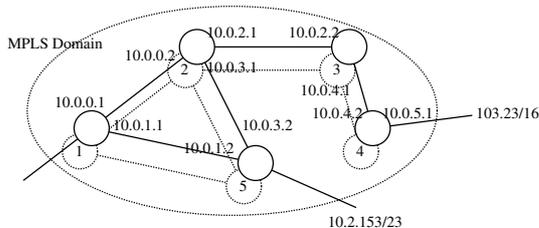


Figure 5: The subIP Addresses

address carried in the subIP header and the subIP routing concept allow exact match table lookup on packet

forwarding through the domain. A similar approach is used in CLNP where a node has unique address per area and level 1 routing uses exact match table lookup [6].

Since the 4 byte subIP address uniquely identifies a router in the AS, the best and the most appropriate way to distribute subIP addresses is in the routing protocol messages as router IDs, if implementation allows. This ID is a 4 byte identifier of a router that sent the routing protocol message. Currently, one of the router IP addresses is set as router ID. Alternatively, a file containing mappings of the router IDs to the subIP addresses can be manually added to each router.

3.3 Determining the Egress Router

In subIP routing the ingress subIP area router needs to determine the last subIP area router on the packet's way to the destination, i.e. the egress area router. This may be done based on the packet's IP destination address in its IP header and the information gathered by the routing protocol. In our case, a subIP area matches an OSPF area and router ID in routing protocol messages is the subIP address of the router. Each link state advertisement (LSA) distributed by OSPF contains the ID of the router that sent it and is kept in the router link state databases [5]. subIP may thus expand the existing IP routing tables for each address prefix with the subIP area router address of the router that advertised it. For the routes generated from within an area (our subIP area) the egress router is the router that advertised the longest prefix matching the destination in the packets IP header. If there is more than one router advertising the prefix, the one that provides the shortest path is chosen. The cost of the path to a particular subIP area router may be found in the subIP forwarding tables, while the OSPF link advertisements contain the cost of the path towards a destination from the router advertising it. Table 1 is a simplified example of the ingress router subIP routing table for the subIP area given in Fig. 5. For a

Table 1: Address Advertisements

subIP area router who advertised	IP address prefix/mask
4	103.23/16
5	10.2.153/23

packet with destination address 10.2.153.178 by performing the longest prefix match lookup of the table it may be found that the subIP address of the router that has advertised it is 5 and that is the egress area router for this packet.

For the paths outside the subIP area the egress subIP router is the area border router (ABR) that provides the shortest path to a particular destination.

3.4 sCMP

The subIP Control Message Protocol (sCMP) is a protocol used with subIP for collecting information from the network. An sCMP message may be very short, placed in the next label stack entry or longer, placed in the data portion sent after the subIP header. However, the messages cannot be longer than the MTU per path since subIP does not provide fragmentation and reassembly. For example, the minimal MTU within the domain may be the maximal packet size.

An advantage of the sCMP messages is that they are simple to send. In order to send a message back to the ingress router of the received packet, a subIP area router may copy the address of the packet's ingress router from the subIP header of the received packet. This address is now the *egress subIP address* in the subIP header of the message, while the *ingress subIP address* is the subIP address of the area router sending the message.

The subIP provides means for balancing the traffic within a subIP area using multiple topology routing, described in section 4. sCMP is used to distribute the required topology and control information. Other sCMP messages may also be defined. For example, a short sCMP message indicating congestion indication on a particular subIP link is useful for better traffic distribution decisions.

4 Multiple Topology Routing

The concept of shortest paths routing used both in subIP and IP may cause congestion on particular links while a longer unused path exists. Currently IP only provides traffic balancing among multiple paths of equal cost. However, subIP also provides traffic balancing among multiple paths of different costs in order to avoid congestion along shortest paths.

The subIP shortest paths routing is based on the topology information collected by a routing protocol (in our case OSPF). Since each router has the same topology information, the calculation of the shortest path between two nodes gives the same result at any node in the same routing domain. We may thus define several virtual topologies based on the physical topology information collected by the routing protocol and distribute it to all the area routers. For example, this information may contain a bit mask matrix or a list of links that should be omitted from the full topology before shortest path calculations for each defined topology. The subIP addressing scheme allows its compact presentation. Each topology will be marked as a different level topology and each area router will have a separate next hop field in its routing table for each topology. The 4 bit *control* field in the subIP header, set by the ingress router, will be used to indicate which of the defined topologies should be used on packet forwarding.

A simplified example is given in Fig. 6. The physical area topology is marked as T0 and its next hop field in

the routing table is NextHop0. In our T1 topology information we indicated that link 3-5 should be pruned before shortest path calculations. Using a shortest path algorithm each area router determines the next hops for the area destinations for this topology and stores them as NextHop1 in the routing table. In the given example the router R3 will find different next hops for destination R6 and R7 for the two topologies. A received packet will be routed according to the NextHop0 if the *control* field in its subIP header is set to 0 and according to the NextHop1 if this field is set to 1. By assigning packets to

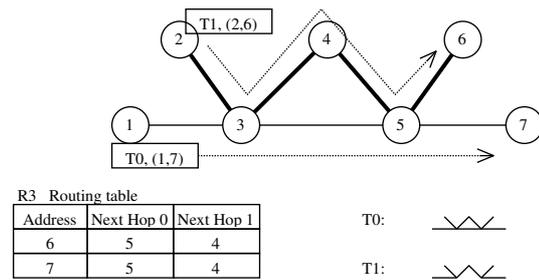


Figure 6: Multiple Topology Routing

different topologies the ingress routers balance the traffic in the domain. In the given example the ingress router R2 forward packets to the egress R6 along the topology T1, while the ingress router R1 sends packets to the egress router R7 along the topology T0. Alternatively, each ingress router may split its traffic to these destinations by performing a finer traffic aggregation based on the address prefix in routing tables.

Different topologies and traffic distribution are determined by an external entity, such as a network management system or a human operator, to achieve certain traffic engineering goals. Generally, a new topology is defined rarely and distributed rarely. However, distribution of the new topology information does not disrupt the current network routing. Once the information is distributed throughout the subIP area, the same external entity may trigger the use of the new topology on packet forwarding by sending a message to one or more edge area routers. This also provides some control of the edge routers in traffic balancing. Within the area a packet received with non-0 *control* field in its subIP header will trigger routing table calculations for a particular virtual topology and it may be timed out when no more subIP headers with non-0 *control* fields are detected. The ingress LER may assign packets to different topology levels based on their egress address or a finer traffic aggregation may be applied based on the address prefix in the table of address advertisements or packet destination address. For the control message exchange cCMP is used as discussed in section 3.4.

Defining different topologies can be kept simple and used only to avoid congestion in the network, but there are also prospectives for more sophisticated traffic engineering techniques, which requires further study. Link

utilization, congestion or failures in a certain period of time, for example in peak hours, are some of the factors that can be considered on defining a new topology. A separate topology may also be defined for the inter-AS traffic and the intra-AS traffic. Since subIP does not provide fragmentation and reassembly, it may be useful to define a topology based on the MTU size of links so that regardless of basic topology changes a long packet cannot be routed to a link that cannot transmit it. The advantages of the multiple topology routing approach are:

- it has **centralized approach** for determining area routes based on an area-wide view of the topology and traffic, rather than the local views at each router
- distributing the network feedback information to a single router instead of a number of edge routers introduces **less protocol overhead** and also simplifies the operation of the edge routers
- it supports **route pinning**, which allows the movement of some traffic from one path to another without disrupting the paths for other traffic
- **backup paths** may be included in a new topology, which allows faster rerouting in the event of a network failure, and also the physical network topology is always available as a backup for the edge routers

The disadvantages of the presented approach are:

- recalculating shortest paths for a number of topologies on topology changes may slow down routing in bigger networks, although the time required for calculations may be reduced by modifying the shortest path algorithms so that the paths are calculated in parallel

The implementation of subIP in IP networks may thus lead to better network utilization. The advantage of this approach comparing to optimizing link weights in routing protocols is in providing route pinning. Comparing to the approach of balancing the traffic among explicit routed paths, multiple topology routing requires simpler management. In the former case the number of explicit routes that needs to be established and maintained is in general, proportional to N^2 , where N is the number of domain routers. The number of new topologies defined in the latter case cannot be greater than 16, which is determined by the 4 bit *control* field in subIP header. This is though a trade off with providing better quality of service (QoS). Still, to achieve satisfying QoS per path similar traffic engineering techniques need to be applied throughout the Internet. Simple multiple topology routing may thus be a good transient solution before the conditions for implementing more sophisticated traffic engineering techniques are globally fulfilled.

5 Further Developments

subIP may be extended with the MPLS connection-oriented service. The change required in the current MPLS header in order for it to inter-operate with subIP is that one of the three experimental bits (T) needs to be reserved to distinguish the two protocols. In the MPLS

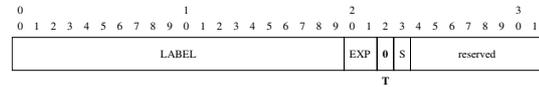


Figure 7: The Modified MPLS Header

header it should be set to 0. Given that the connectionless service is provided by subIP instead of MPLS hop-by-hop routing, this change may also allow the TTL field of the MPLS header to be redefined. In general, the TTL field is used with connectionless service to prevent packets from looping. Since with connection-oriented services loops cannot occur, in this case the ingress router may process the TTL field in the IP header of the packet by reducing it for the hop count of the explicit route (or an estimate of the path length). This will prevent packets that need more hops through the area than are allowed entering the area, there will be no TTL processing at each hop and the TTL field may be used for other purposes, such as defining different types of traffic, which is indicated in Fig. 4 by marking the field as *reserved*.

subIP can support a new protocol, independent of IP, as long as the maximum packet size of the protocol does not exceed the minimum subIP area MTU. This may be a new routing and signaling protocol to be used for explicit route establishment and distribution of the new subIP topology information (sLDP) (Fig. 8). A protocol

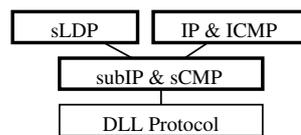


Figure 8: A New Routing and Signaling Protocol

implemented *on top* of subIP introduces less protocol overhead since the IP header is not included in protocol messages and also shorter area addresses may be used on defining paths. It also provides better network modularity.

While in MPLS the importance of explicit routing is emphasized in order to provide traffic balancing in an MPLS domain, we emphasize its importance in establishing virtual paths in order to build an arbitrary network topology. Together with the new addressing scheme that may be introduced by subIP it may be a step towards a multiple hierarchy network with hierarchical addresses with geographical significance which would

significantly simplify traffic engineering. A number of ASs may be grouped into a new subIP hierarchy level where the similar approach of routing applied within an AS may be introduced. However, the routing between the ASs is controlled with an Exterior Gateway Protocol, such as BGP, which is conceptually different from OSPF and additional topology and addressing information needs to be distributed in order to implement subIP, which is for further study.

6 Conclusion

In this paper we present a new approach for optimizing network utilization within an AS. It is based on the subIP protocol proposed in the paper.

subIP is a modification of the MPLS protocol which provides connectionless service similar to IP. It basically divides the Internet into small areas with local short subIP router addresses. The IP routing concept is then applied within the areas based on the new addresses. The new information is distributed in the 4 byte subIP packet header. Since subIP requires smaller forwarding tables and uses fast exact match table lookup on forwarding, it provides simpler and faster packet forwarding than IP.

Similar to MPLS hop-by-hop routing, subIP routes packets along the same path as IP. However, as opposed to MPLS hop-by-hop routing, the subIP concept does not require a label distribution protocol, which reduces protocol overhead and simplifies the protocol. The reaction of the protocol on topology changes is faster since forwarding tables are calculated based on available topology information, while the MPLS approach may introduce latency for redistributing the labels. The subIP also requires smaller forwarding tables and does not require label swapping on forwarding.

The disadvantage of the subIP is that it bounds the area size to 256 routers. Still, in view of future multiple hierarchy networks with small interconnected domains which provides better scalability and simpler management, this area size is acceptable.

The concept of traffic balancing within the area provided by subIP is based on defining different area topologies based on the physical area topology. The calculation of the shortest paths within the area based on the new topology information, but consistent at each area router, may give different results than calculations based on the physical topology. Using the *control* field of subIP, traffic is thus balanced among shortest paths belonging to different topologies. Some of the advantages of the approach are centralized control, route pinning and providing backup routes. It is also simple and may lead to a better network utilization.

The paper discusses subIP implementation in an AS implementing the OSPF routing protocol. Its new addressing scheme and the MPLS explicit routing which may provide virtual paths and thus arbitrary network topology may be a good step towards a new multiple hier-

archy network with addressing scheme with geographical significance. This would significantly simplify traffic engineering. However, the subIP implementation with conceptually different exterior gateway routing protocols requires further study.

References

- [1] R. Callon E. Rosen, A. Viswanathan. RFC 3031: Multiprotocol Label Switching Architecture, January 2001.
- [2] Bernard Fortz, Jennifer Rexford, and Mikkel Thorup. Traffic engineering with traditional IP routing protocols. *IEEE Communications Magazine*, pages 118–124, October 2002.
- [3] C. Hopps. Rfc 2992: Analysis of an equal-cost multi-path algorithm, November 2000.
- [4] Information Sciences Institute. RFC 791: Internet Protocol, September 1981.
- [5] J. Moy. RFC 2328: OSPF Version 2, April 1998.
- [6] Radia Perlman. *Interconnections: Bridges, Routers, Switches and Internetworking Protocols, Second Edition*. Addison Wesley, September 2001.
- [7] J. Postel. RFC 792: Internet Control Message Protocol, September 1981.

A Novel MAC Protocol and Traffic Scheduling for HFC Network with QoS Support

Char-Long Chen
Industrial Technology Research Institute
Computer & Communications Research Laboratories
Hsinchu, Taiwan
Email: along_chen@itri.org.tw

Abstract

In order to support Quality of Service (QoS) for real-time data, HFC (Hybrid Fiber Coax) networks must provide an effective priority scheme. In this paper, a new scheme to support QoS in HFC network is proposed. In this scheme, three priority queues for constant bit rate (CBR), variable bit rate (VBR) and available bit rate (ABR) traffic and different competitive ratios are assigned to priority queues respectively are proposed. On the other hand, a novel ideal microslot is proposed to support CBR traffic and allow continuously transmit data without requesting the upstream bandwidth again. Furthermore, an effective and workable traffic scheduling is proposed to arrange the competitive ratio. Consequently, the proposed scheme offers the guaranteed QoS for high priority traffic in HFC network.

1. Introduction

In HFC network, upstream is separated into discrete base time slots, called minislots, by time division multiple access (TDMA) technologies. The CMTS determines the frame format by setting the number of request minislots and data minislots [2][3] in each frame. Request minislots are one minislot long and used by CMs to send request for bandwidth. Data minislots are several minislot long and used by CMs to transmit data.

To cope with growing multimedia application, DOCSIS defines five upstream service in DOCSIS 2.0 [1]: Unsolicited Grant Service (UGS), Unsolicited Grant Service with Activity Detection (UGS-AD), Real-Time Polling Service (rtPS), Non-Real-Time Polling Service (nrtPS), and Best Effort (BE) service. In general, we can divide those services into three parts, CBR, VBR, and ABR by their character. UGS and UGS-AD are CBR like, rtPS and nrtPS are VBR like, and BE is ABR like. According to the classification, we assign those services into three priority queues by their traffic character respectively. And an algorithm, Contention-Scheduling Matrix (CS_MATRIX), is proposed to dynamically schedule competitive ratio. We also proposed a novel ideal, microslot, which attaches before request minislots in upstream. By the method, we can realize UGS and UGS-AD (CBR) traffic in DOCSIS HFC network and arise the utilization of bandwidth.

The remainder of the paper is organized as

follows. Section 2 explains the contents of DOCSIS MAC protocol. Section 3 gives a detail description of our scheme. Section 4 presents several simulations and analyzes the result. Finally, the conclusion and summary are presented in section 5.

2. DOCSIS MAC protocol

The MAC protocol controls the access on upstream channel and uses a request/grant mechanism for bandwidth allocation. CMTS controls most of the bandwidth allocation and traffic-scheduling operations. It regularly sends a control signal through downstream, which is called MAP (MAC management message)[4], to describe the allocation of shared upstream bandwidth. According to the instructions within the MAP message, CMs can perform collision-free transmission.

In DOCSIS2.0, there are five upstream services defined to support QoS, they are UGS, UGS-AD, rtPS, nrtPS and BE. UGS is defined to support CBR data transmission, which guarantees bandwidth and access delay. UGS-AD almost resembles UGS, which is designed to support UGS flows that may become inactive for substantial portions of time, such as voice over IP with silence suppression. RtPS is designed for supporting VBR traffic. CMTS periodically polls the specific CM and CM can perform transmission through these opportunities. NrtPS closely resembles rtps but with longer polling interval. The last one service is BE. For the service, CMs generally use conformed request opportunities for bandwidth demand.

3. Proposed DOCSIS QoS Enhancement

3.1 Priority Queues

As shown in figure 1, the priority queues consist of CBR queue, VBR queue and ABR queue to furnish multiple QoS levels. When traffic generates, the traffic classifier distinguishes them into different priority queues by the service flow type. Then the Contention Ratio Management (CRM) assigns different competitive ratio parameters R_c , R_v and R_a to CBR, VBR and ABR queue, respectively. The relation of those parameters is as below:

$$R_c > R_v > R_a, \quad R_c + R_v + R_a = 100\%$$

According to these parameters, the traffic scheduler arranges different frequencies for CBR, VBR and ABR traffic to contend request minislots.

If CBR traffic requests bandwidth successful, it

transmits by marking microslot and does not contend request minislots with other traffic. As the figure 1 shown, the transmission will change from path 1 to path 2. At this time, CRM changes R_c to zero and raises R_v, R_a value. At this situation, the relation of parameters is as below:

$$R_c=0\%, \quad R_v>R_a, \quad R_v+R_a=100\%$$

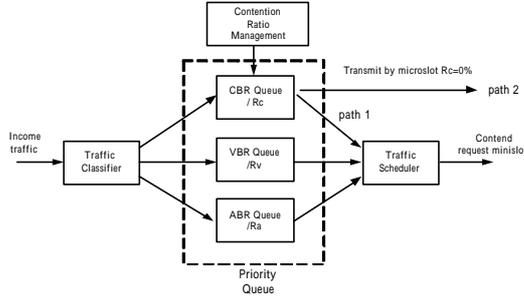


Figure 1. Priority queue architecture

3.2 A Novel traffic Schedule

3.2.1 Contention Scheduling Matrix

In order to implement traffic scheduler, Contention Scheduling Matrix (CS_Matrix) is proposed. The concept of CS_Matrix is to use different contention ratios to describe the arranging condition of priority queues. The CS_Matrix is $n*m$ matrix as shown in figure 2. Its elements consist of "0" and "1".

$$CS_Matrix = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 1 & 0 & 1 & 0 & 1 & \dots \\ 0 & 1 & 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Figure 2. An example of CS_Matrix

The n (row) indicates the total numbers of priority queues. The m (column) is a significant parameter that indicates a period of contention time, it also denotes as scheduling sub-cycle. The other important parameter is M , which denotes total scheduling cycle. The m_j ($j \in M$) indicates the contention sequence of one scheduling sub-cycle; m_1 indicates sequence one and m_2 is sequence two. During the same m_i , there is only one queue can contend the request minislot. For instance, $CS_Matrix(i, j)$ denotes the number i (n_i) queue can contend request minislot at sequence j (m_j).

The traffic scheduler arranges the CS_Matrix by competitive ratio parameters. For example, when there are three priority queues ($n=3$), CBR queue, VBR queue and ABR queue, total scheduling cycle is ten ($M=10$) and CRM assigns R_c, R_v , and R_a ($R_c>R_v>R_a$) to the three queues respectively, the traffic scheduler will generate a $3*10$ ($n*M$) CS_Matrix and arrange the elements in the CS_Matrix, as shown in figure 3. The m_1 assigns to CBR queue, m_2 assigns to VBR queue, m_3 assign to ABR queue, m_4 assigns to CBR queue, and so on, up to m_{10} assigns to CBR queue.

$$\begin{matrix} CBR \\ VBR \\ ABR \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 1 & \dots & \dots & 1 \\ 0 & 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & 0 \end{bmatrix} \begin{matrix} 50\% \\ 30\% \\ 20\% \end{matrix}$$

$$m_1 \ m_2 \ m_3 \ m_4 \ \dots \ \dots \ \dots \ m_{10}$$

Figure 3. Content of CS_Matrix

Because the CS_Matrix size is fixed and the competitive ratio parameters are arranged in turn, there may be a critical problem to decrease the performance in a special condition. The figure 4 presents the unideal condition. The sequence $m_1, m_2, m_3, m_4, m_5, m_8, m_9, m_{10}$ have "1" element to contend request minislot, but m_5 and m_6 don't. This is because that at m_5 and m_6 sequence the corresponding queues have no data to transmit, they needn't contend request minislot. Therefore, there are two contention sequences wasted in unideal condition. As the result, an algorithm is developed to predict the empty sequence and omit it during the sub-cycle time. The predictable algorithm will be described in next section.

$$\begin{matrix} 50\% \\ 30\% \\ 20\% \end{matrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$m_1 \ m_2 \ m_3 \ m_4 \ m_5 \ m_6 \ m_7 \ m_8 \ m_9 \ m_{10}$$

Figure 4. An unideal condition of scheduling

3.2.2 Predictable Algorithm

In order to solve the empty sequence problem, a predictable algorithm is proposed. The pseudocode for the algorithm is as follows.

```
//Global variable
Empty_Counter=1 // Count the empty sequence

//Variable definition
//m is the number of sequence
for (j=1;j<=m;j=j+Empty_Counter) {
    Empty_Counter=1
    Contend_CS(mj) // mj to contend request minislot
    If (j<m)
        Check_Empty(j) //Check the empty sequence
}
Check_Empty(j){
    index=j+1
    while(m_index is empty and index<=M)
        //M is total scheduling cycle
        {
            index=index+1
            Empty_Counter=Empty_Counter+1
        }
    index=0 }
}
```

The variable number Empty_Counter is a global Contend_CS() is responsible to contend request minislot, Check_Empty() is to check the empty sequence and increase the Empty_Counter value.

3.3 Microslot for CBR Traffic

The new MAC frame format with microslot is shown at figure 5. The microslots are attached in front of request minislots and the CM marks its own microslot to transmit CBR data.

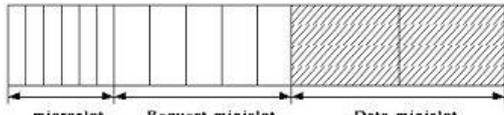


Figure 5. New DOCSIS MAC format

The state machine of microslot is shown in figure 6. When a CM power on, CMTS assigns a microslot to it and records the relation between microslots and CMs. After initiation, if the CM contends successful and system can provide enough bandwidth to its CBR traffic, CMTS keeps this CBR traffic information (used bandwidth, delay jitter... etc.) in resource table and informs the CBR Traffic Grant Message to the CM by downstream. On the contrary, if system cannot provide enough bandwidth, CMTS sends rejection message to the CM and terminates the procedure. After the CM receives the grant message, it will be aware that CMTS has reserved the resource for it. At this time, the CM can transmit CBR traffic immediately by marking its own microslot or suspend to transmit without marking it. When CMTS receives the marked microslot from the CM, it looks up the resource table and sends the Data Slot Grant Message to the CM by downstream. Then the CM begins to transmit CBR data after receiving the grant message. Therefore, the CM can transmit CBR traffic at a stable time and acquires enough bandwidth by microslot. When the CBR data is the last one, the CM adds Microslot Termination Message (MTM) in DS. After CMTS receives the MTM, it removes the information from resource table and releases the bandwidth.

As mentioned above, we can satisfy the character of UGS and UGS-AD by microslot. By marking it or not, the resource for CBR traffic can be temporary released to increase the utilization of bandwidth. It also reduces the overhead of MAC frame, because the size of microslot can be set very small just for marking. As a result, the upstream bandwidth increased and it provides a method to realize the UGS and UGS-AD.

4. Simulation Result and Analysis

We have built a simulation program to evaluate our scheme. The common simulation parameters are listed in table 1. In this simulation, timestamp, i.e. the duration of one minislot, is the elementary time unit, and all simulation results are measured in the unit of timestamp. There are two experimental models in the simulation, non-priority model and priority model. The non-priority model is the original MAC protocol without our scheme. On the contrary, the priority model supports QoS by priority queues and microslot.

As follows, we will use these two experimental models to accomplish the experiment.

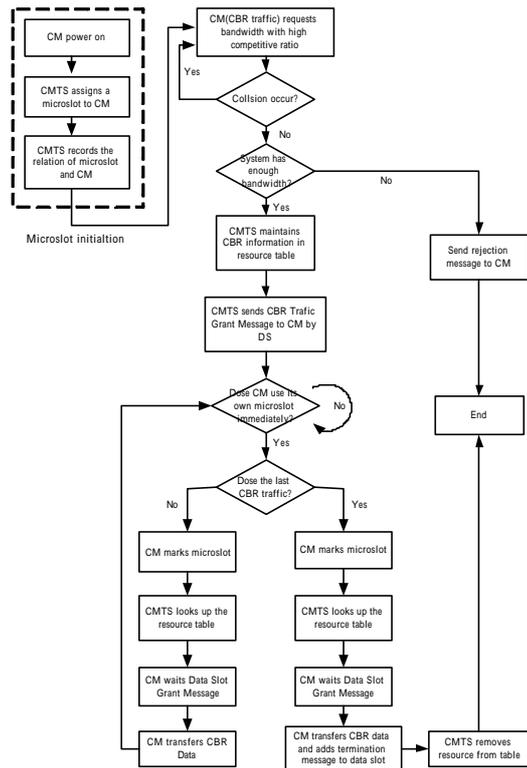


Figure 6. The state machine of microslot

Simulation Parameter	Value description
The number of CM	20
Upstream bandwidth	3 Mbits/sec
Timestamp	0.02 micro second
Simulation Time	1000000 timestamp
Average length of CBR	500 timestamp
Traffic Ratio	CBR(40%),VBR(30%),ABR(30%)
Competitive ratio parameters (before CBR contends successful)	Rc=70%, Rv=30%, Ra=10%
Competitive ratio parameters (after CBR contends successful)	Rc=0%, Rv=70%, Ra=30%

Table 1. Simulation Parameter

4.1 Experiment 1: Average Buffer Delay of CBR, VBR and ABR traffic

In experiment 1, average buffer delay of CBR, VBR and ABR traffic are quantified. Average buffer delay is the average time period of one packet arrives at the queue to the time it is transmitted. Figure 7 shows the average buffer delay of CBR traffic. For non-priority model the average buffer delay grows conspicuously due to the CBR traffic without any guarantee. On contrary, the curve of priority model is very gently regardless of traffic load. The result explains that the scheme keeps the CBR traffic

transmits at the same rate and conform to the characteristics of CBR traffic.

Figure 8 presents the average buffer delay of VBR traffic. By the scheme, the VBR traffic acquires higher competitive ratio and also raises the request percentage successfully. Therefore, the average buffer delay of priority model is better than non-priority one. Nevertheless, when traffic load reaches 0.7, the average buffer delay of priority model increases substantially and higher than non-priority. This is because that when the traffic load is heavy, the load of CBR traffic also increases and it occupies the most bandwidth by microslot. As the result, the VBR traffic acquires fewer bandwidth and increases the average buffer delay a lot. In contrast with non-priority model, all traffic are the same priority and CBR traffic cannot occupy the bandwidth for a long time. So VBR traffic acquires bandwidth more easily in heavy traffic load.

The average buffer delay of ABR traffic is illustrated in figure 9. By employing the scheme, CBR and VBR traffic occupies most bandwidth while ABR traffic loses the competitive predominance. Therefore, the average buffer delay of ABR traffic increases violently. By contrast, because all traffic is at the same level and without any guarantee for high priority traffic in non-priority model. ABR traffic contends the bandwidth fairly with other traffic. So, the average buffer delay of ABR traffic increases moderately.

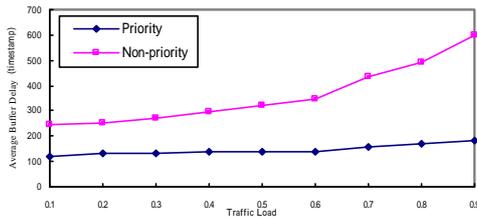


Figure 7. Average buffer delay of CBR traffic

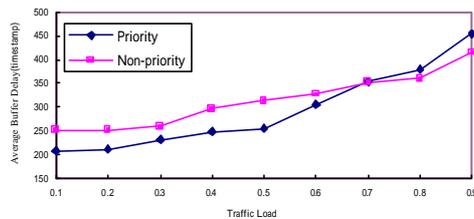


Figure 8. Average buffer delay of VBR traffic

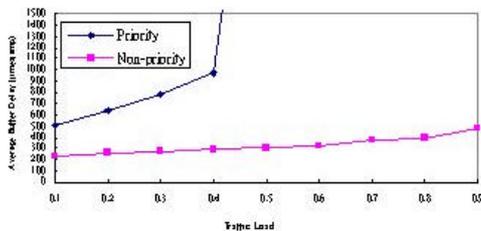


Figure 9. Average buffer delay of ABR traffic

4.2 Experiment 2: Lost rate of CBR Traffic

In this experiment, the lost rate in priority model and in non-priority model are compared. The lost rate is the percentage of CBR traffic in and out the CBR queue. According to figure 10, the CBR traffic with microslot retains the same lost rate regardless of traffic load. On the contrary, the lost rate of CBR traffic without microslot is seriously affected by traffic load. When the traffic load reaches 0.5, the curve falls in a sharp rate. Therefore, this experiment shows that CBR traffic with microslot can retain the stable delay jitter to conform to the characteristics of CBR.

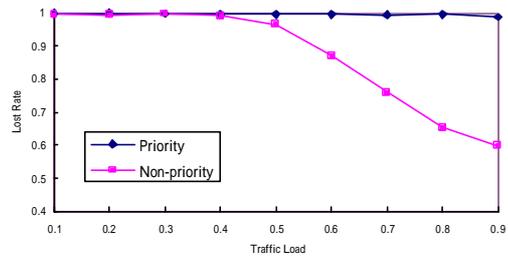


Figure 10. Lost rate of CBR traffic

5. Conclusion

In this paper, priority queue and microslot are proposed to support QoS over HFC network. The CRM assigns different competitive ratios to priority queues. CS_Matrix with predictable algorithm is proposed to schedule the competitive ratio. Consequently, high priority traffic acquires more probability to contend request minislots. Finally, a novel ideal microslot only used for CBR traffic to support absolute priority. Therefore, the CBR traffic can transmit in good performance and conform to the characteristics of CBR traffic.

Reference

- [1] Data-over-cable service interface specification (DOCSIS v2.0)-Radio frequency specification-SP-RFI v2.0-I01-011231.
- [2] N. Glomie, F. Mouveaux, and D.Su, "A comparison of MAC protocols for hybrid fiber coax networks: "IEEE802.14 vs. MCNS", IEEE communication, Apr. 1999.
- [3] W. M. Yin and Y. D. Lin, "Statistically optimized minislot allocation for initial and collision resolution in hybrid fiber coaxial networks", IEEE Journal on Selected Areas in Communications, Vol.18, No. 9, Sep. 2000.
- [4] Hue-jiun ju and Wanjiun Liao "Fast Request Transmission in DOCSIS-Based CATV Networks," to appear in Proc. IEEE International Conference in Multimedia and Expo(ICME), Aug, 2002.

A Prediction Method of Self-Similar Network Traffic Based on the Covariation Orthogonality Criterion

Ge Xiaohu, *Won-Sik Yoon, Zhu Guangxi, *Shaokai Yu, *Seung-Wook Min

Department of E & I Engineering,
Huazhong University of Science &
Technology, Wuhan, P.R. China

*Department of Electronics & Computer
Engineering,

Ajou University, Suwon, 442-749, Korea

Abstract

In this article, based on the covariation orthogonality criterion an unbiased linear prediction method is developed for self-similar network traffic. The covariation orthogonality criterion ensures that the procedure of the prediction is efficient and simple. The simulation experiments demonstrate that the new prediction method can predict the changes of the self-similar network traffic, especially in forecasting the burstiness changes, which can be used for avoiding network congestion.

Key words: Prediction, alpha-stable process, covariation orthogonal, self-similar

1. Introduction

A number of recent empirical studies have shown that the actual network traffic is self-similar or fractal in nature [1]. Based on the alpha-stable processes a new self-similar network traffic model has been proposed [2]. Recent research conclusions provide much evidence that the alpha-stable network traffic model is better than any other self-similar network traffic models in simulations [3].

Developing the self-similar network traffic model is only the first step. The next step is to find a prediction method for self-similar network traffic, which can be used for avoiding the network congestion and improving the performance of network. Because the variance of self-similar network traffic processes is infinite, the classical minimum mean-squared-error criterion cannot be use for these processes. Very little research exists in the area of forecasting the network traffic which has the

property of the infinite variance. In the present paper, a new prediction method is proposed based on the alpha-stable network traffic model.

The contents of the paper are organized as follows: Section 2 the covariation and the relations between the conditional expectation and the covariation orthogonal are described. According to the covariation orthogonality criterion an explicit method used for obtaining the prediction value is proposed in Section 3. The simulation experimentations have demonstrated that the prediction method is available and efficient in Section 4. Finally, in the conclusion we summarize the contributions of the paper and refer to the future work.

2. Conditional Expectation and Covariation Orthogonal

2.1 Covariation

The covariance function is an extremely powerful tool in the study of Gaussian random elements, but it is not defined in the alpha-stable processes when $\alpha < 2$. However, the covariation is designed to replace the covariance when $\alpha < 2$ [4].

Definition 1: Let X and Y be stable random vectors with $\alpha > 1$ and spectral measure Γ , and (u, v) is the processes (X, Y) denoted the polar coordinates on the \mathbb{R}^2 unit circle S_2 ,

the stable covariation of X on Y is denoted as

$$[X, Y]_{\alpha} \equiv \int_{S_2} uv \langle \alpha^{-1} \rangle \Gamma(du, dv) \quad (1)$$

2.2 covariation orthogonal

Denote by \hat{X}_{n+k} any prediction value

of the alpha-stable random variable X_{n+k} .

Prediction *unbiasedness* is defined by the

$$\text{property } E[X_{n+k} - \hat{X}_{n+k} | \mathbf{X}] = 0,$$

where \mathbf{X} is the sample vector. Next, denote by

$$\Psi_m^\alpha \equiv \left([X_{n+1-j}, X_{n+1-i}]_\alpha \right)_{i,j=1}^m$$

the cross-covariation matrix.

Definition 2: Let S be a normed vector space.

A vector $X \in S$ is said to be covariation orthogonal to a vector $Y \in S$ ($X \perp_\alpha Y$) if and only if

$$[X, Y]_\alpha = 0 \quad (2)$$

Proposition 1 [5]: Let the vector

$$(X_\omega, \dots, X_n; X_{n+k}), \quad n \geq 1,$$

be jointly alpha-stable with spectral measure Γ , and characteristic exponent $\alpha \in (1, 2)$. Consider

any arbitrary $-\infty \leq \omega \leq 1$, and let (v, \mathbf{u})

denote the spectral representation of (X_{n+k}, \mathbf{X}) .

$$\text{A) } E[X_{n+k} | \mathbf{X}] = \sum_{i=1}^m a_i X_{n+1-i} \text{ for some}$$

sequence $\{a_i\}_{i=1}^m$ if and only if $\forall \lambda \in \mathfrak{R}^m$,

$$\int_{S_{m+1}} (v - \mathbf{a}'\mathbf{u})(\lambda'\mathbf{u})^{(\alpha-1)} \Gamma(dv, d\mathbf{u}) = 0 \quad (3)$$

$$\int_{S_{m+1}} (v - \mathbf{a}'\mathbf{u})|\lambda'\mathbf{u}|^{\alpha-1} \Gamma(dv, d\mathbf{u}) = 0 \quad (4)$$

B) Consider the following set of equalities:

- i. $E[X_{n+k} | \mathbf{X}] = \mathbf{a}'\mathbf{X}$
- ii. $E[X_{n+k} - \mathbf{a}'\mathbf{X} | \mathbf{X}] = 0$
- iii. $[X_{n+k} - \mathbf{a}'\mathbf{X}, X_{n+1-t}]_\alpha = 0, \quad \forall t=1, \dots, m$

Then, (i) \Rightarrow (iii). Moreover, the m -vector \mathbf{a} in

(3), (4), and (i)—(iii) is unique for every $\alpha \in (1, 2)$, if and only if the associated stable

scale is bounded, and the kernel of Ψ_m^α is

identical the singleton $\{0\}$: $\ker \Psi_m^\alpha = \{0\}$.

3. Prediction Method

Based on the alpha-stable processes, in this paper a linear prediction formula is designed. We explicitly treat the value of traffic carried by the network element X_n which is governed by the alpha-stable network traffic model. If we assume the equations (3) and (4) are true, there exists an unbiased linear prediction for a k -ahead forecast. The linear prediction formula is follows:

$$\hat{X}_{n+k} = \sum_{i=1}^n a_i X_{n+1-i} \quad (5)$$

where \hat{X}_{n+k} is the prediction value of traffic

after the k time units, $\{X_1, \dots, X_n\}$ are the n

sample values. In order to work out a prediction value, a coefficient set

$\{a_i\}, i=1, \dots, n$ must be obtained. If we can find a coefficient set $\{a_i\}, i=1, \dots, n$ such

that the error of the prediction value is covariation orthogonal to the sample values,

and according to the conclusion of the reference [5] which proves that the

unbiasedness linear prediction for alpha-stable processes exists if and only if the covariations

of alpha-stable random variables are orthogonal, cf. Proposition 1.B.iii, then the

unbiased prediction exists uniquely, and is necessarily linear.

Because we assume the linear predictor is unbiased, the following equality can be obtained,

$$E[X_{n+k} | \mathbf{X}] = \mathbf{a}'\mathbf{X} \quad (6)$$

According to the Proposition 1, the equation (equ 6.) can be translated into the following

equation.

$$[X_{n+k} - \mathbf{a}'\mathbf{X}, X_{n+1-t}]_\alpha = 0, \quad \forall t=1, \dots, m \quad (7)$$

and then it can be translated into the equ.8

according to the decomposability of covariation and the covariation stationarity,

$$A \cdot B = C \quad (8)$$

$$A = \begin{bmatrix} [X_n, X_n]_\alpha & [X_{n+1}, X_n]_\alpha & \cdots & [X_{n+m}, X_n]_\alpha \\ [X_n, X_{n-1}]_\alpha & [X_{n+1}, X_{n-1}]_\alpha & \cdots & [X_{n+m}, X_{n-1}]_\alpha \\ \vdots & \vdots & \vdots & \vdots \\ [X_n, X_{n+m}]_\alpha & [X_{n+1}, X_{n+m}]_\alpha & \cdots & [X_{n+m}, X_{n+m}]_\alpha \end{bmatrix}$$

$$B = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}; C = \begin{bmatrix} [X_n, X_{n-k}]_\alpha \\ [X_n, X_{n-1-k}]_\alpha \\ \vdots \\ [X_n, X_{n+1-m-k}]_\alpha \end{bmatrix}$$

Proposition 2: Let (X, Y) be alpha-stable with $\alpha > 1$. Then for all $1 < p < \alpha$

$$\frac{EXY \langle p-1 \rangle}{E|Y|^p} = \frac{[X, Y]_\alpha}{\|Y\|_\alpha^\alpha} \quad (9)$$

where $\|Y\|_\alpha$ denotes the scale parameter of Y .

According to the Proposition 2, the covariation can be express as the moment. Because the moment $EX_t X_{t-1}^{(p-1)}$ is just the population average of $X_t X_{t-1}^{(p-1)}$ and we assume that the vaule of n is large enough, the new matrix function is denoted as

$$A' = B' \cdot C' \quad (10)$$

$$A' = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}; C' = \begin{bmatrix} \frac{1}{n} \sum_{t=k+1}^n X_t X_{t-k}^{(p-1)} \\ \frac{1}{n} \sum_{t=k+2}^n X_t X_{t-k-1}^{(p-1)} \\ \vdots \\ \frac{1}{n} \sum_{t=k+m}^n X_t X_{t+1-k-m}^{(p-1)} \end{bmatrix}$$

$$B = \begin{bmatrix} \frac{1}{n} \sum_{t=1}^n X_t X_t^{(p-1)} & \frac{1}{n} \sum_{t=2}^n X_{t-1} X_t^{(p-1)} & \cdots & \frac{1}{n} \sum_{t=m}^n X_{t+m-1} X_t^{(p-1)} \\ \frac{1}{n} \sum_{t=2}^n X_t X_{t-1}^{(p-1)} & \frac{1}{n} \sum_{t=3}^n X_{t-2} X_t^{(p-1)} & \cdots & \frac{1}{n} \sum_{t=m}^n X_{t+m-2} X_{t-1}^{(p-1)} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{n} \sum_{t=m}^n X_t X_{t+m}^{(p-1)} & \frac{1}{n} \sum_{t=m}^n X_{t-1} X_{t+m}^{(p-1)} & \cdots & \frac{1}{n} \sum_{t=m}^n X_{t+m-1} X_{t+m}^{(p-1)} \end{bmatrix}$$

where $m+k \leq n$ must be assumed. After solving the matrix equation, the coefficient set $\{\mathbf{a} | a_1, a_2, \dots, a_m\}$ can be obtained. If the coefficient set is substituted into the prediction

$$\widehat{X}_{n+k} = \sum_{i=1}^m a_i X_{n+1-i},$$

the prediction value can be estimated.

4. Simulation Experimentation

In order to validate the prediction method, the actual network traffic data (file Oct89Ext.TL) which was collected by Leland at Bellcore Morristown Research and Engineering facility are used for simulation experimentation. In the reference [3] it is found that the actual network data (file Oct89Ext.TL) is governed by the alpha-stable distribution.

The data file contains a million network packets. All packets of the network data (file Oct89Ext.TL) are divided into ten data sets each has 100,000 packets, and then the number of packets is counted in every data set passing in a range of ten seconds. So ten new data sets are obtained, in each of which every element expresses the number of passing packets in 10s time scale. We will analyze every new data set.

First, the 1-step ahead prediction method is used for those data sets (i.e. $k=1$). The sample bound n is set to 10. Because of the requirement $m+k < n$, the number of prediction coefficient m is set to 5. The algorithm of prediction is as follows:

- 1) The 1~10 samples are picked up;
- 2) Let $n=10, m=5, k=1$, and the ten samples are substituted into the equ.10, the prediction coefficients are obtained after calculation;
- 3) The sample values and the prediction coefficients are substituted into the equ.5, a prediction value can be obtained by the calculation;
- 4) The 2~11 samples are picked up again, and then go to **step** (2) for computing the next prediction value, until the samples are picked up to the end.

As a result, all prediction network data are obtained. We optionally select one data set, and it is the fifth data set. Because the α

parameter of the fifth data set is estimated as 1.5 (the detailed procedure of estimation can be referred to the [3]), in order to predigest the calculation the value of the p parameter is set to 1.5. The prediction result of the fifth data set can be obtained after all parameters and samples are substituted into the algorithm of prediction. The prediction trace and the actual network trace can be compared in the following Fig 1.

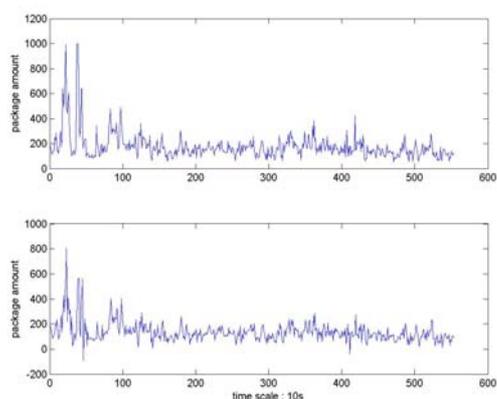


Fig.1 The sample trace plot and prediction trace plot. The top plot is the sample trace which is picked up from the fifth data set, the bottom plot is the prediction trace which is 1-step ahead.

In the Fig.1, the prediction trace can capture most burstinesses in the sample trace. The mean value of sample trace is 177.3, and the mean value of prediction trace is 136.2. Comparing with two traces, it is shown that the change scale of the prediction trace is same with the change scale of the sample trace. Therefore, it is said that the result of the prediction can efficiently describe the changes of the fifth data set. And all other results of experiments are consistent with the fifth result of the experiment, so the 1-step ahead prediction method can be thought to be very efficient and actual.

Many simulation experiments have been done for n, m, k equaling to different values, and the results of all the other experiments are similar with the results of the above experiment. According to the conclusion of the

simulation experiments, it is shown that this prediction method can efficiently predict the burstiness of actual self-similar network traffic.

5. Conclusion

We presented a new prediction method for self-similar network traffic, which is based on the covariation orthogonal criterion. The new prediction method overcomes the problem of the unbounded variance which is caused by the self-similarity. The results of the simulation experiment show that the new prediction method can accurately predict the burstiness changes of the actual self-similar network traffic. The new method implications along with more analytical work, such as selection of the sample bound and the number of prediction coefficients, are the objectives of our future research in this exciting field.

Reference

- [1] W. Leland, M. Taquu., W.Willingerand D. Wilson.,(1994) "On the Self-Similar Nature of Ethernet Traffic (Extended Version)", IEEE/ACM Transactions on Networking, Vol. 2, No. 1, February 1994.
- [2]Ge Xiaohu, Zhu Guangxi, Zhu Yaoting,(2003) "An Improved Modeling for Network Traffic Based on Alpha-stable Self-Similar Processes", Chinese Journal of Electronics, Vol.12, No.4, October 2003, pp.494-498.
- [3]Ge Xiaohu, Zhu Guangxi, Zhu Yaoting, "On the testing for alpha-stable distributions of network traffic", Computer Communications, Vol 27/5, Feb, 2004, pp 447-457.
- [4] G. Samorodnitsky and M. S. Taquu,(1994). Stable Non-Gaussian Random Processes. London, U.K.: Chapman & Hall.
- [5]Jonathan Hill.(2000) "Minimum Dispersion and Unbiasedness: Linear Predictors for Stationary Linear alpha-Stable Processes", working paper 12, University of Colorado, Boulder.

A Size-limited Cluster Maintenance Technique for Mobile Ad-hoc Networks

Gayathri Venkataraman*, Sabu Emmanuel**, Srikanthan Thambipillai**

School of Computer Engineering, Block N4, Nanyang Avenue,

Nanyang Technological University, Singapore-639798

*gayathrivenkat@pmail.ntu.edu.sg, **asemmanuel, astsrikan@ntu.edu.sg

ABSTRACT

As networks grow in scale, hierarchical structures are employed to address the scalability issues of the network. Algorithms for organizing nodes into clusters build the foundation for the development of a hierarchical structure. In this paper, we propose a distributed technique for cluster maintenance for large mobile ad-hoc networks. The proposed scheme maintains the clusters by limiting the cluster size, in terms of number of nodes in a cluster. The proposed method is shown to perform well in large mobile ad-hoc networks. Extensive simulations and comparisons were carried out to verify its effectiveness on cluster maintenance.

Key words: Mobile ad-hoc network, Distributed algorithm, Cluster maintenance.

1. INTRODUCTION

Mobile ad-hoc networks (MANET) are formed by mobile nodes, which can freely move around in the network, leave the network and join the network. Hence the topology of the network can change arbitrarily. Each mobile node in the MANET acts as a router as well as a host. Each of these mobile devices has varied capabilities in terms of memory (storage) space, processing speed, battery power etc. Supporting varied capability and different types of devices in MANET is a major challenge especially when the devices are to perform routing and host functions. It is the dynamic topology, varied/limited mobile node capability and limited link bandwidth [5, 7] that are posing a challenge for the scalability of MANET. In order to deal with the scalability issue of MANET, a hierarchical approach based on partitioning the network into clusters can be employed. Since the network is partitioned into smaller logically separate clusters, it is easy to keep track of mobile nodes in a cluster, nodes joining/leaving a cluster, topology of the cluster and node capabilities.

Clustering can be seen as two phases, cluster formation phase and cluster maintenance phase. The formed clusters need to be maintained and updated in a mobile scenario such as MANET. In this paper, we assume the existence of clusters and propose a distributed technique for maintaining clusters. Cluster-based hierarchical approach also reduces the storage requirement at each node for storing the routing table, communication over

head for the route computations. However in this paper we focus on the cluster maintenance technique and would not be discussing the routing issues further.

In a resource limited network such as MANET (especially with nodes of varied capability), it is worthwhile to control the number of nodes per cluster for good planning and management of resources (such as memory, link bandwidth, battery power etc) in the cluster. Our proposed technique limits the number of nodes in a cluster. There are other cluster maintenance strategies where they use number of hops as the limiting parameter [1, 2, 5, 8]. The hop-limited clusters do not guarantee the number of nodes in a cluster to be limited and hence it is possible to have large number of nodes per cluster which may make it tedious to plan and manage the resources in that cluster.

Clustering algorithms use distributed or centralized approach for cluster maintenance. Clustering algorithms that use centralized approach for cluster maintenance first elect a cluster head and then the cluster head is made in charge of the cluster maintenance [1, 4]. Some typical cluster maintenance functions are admission of new node in the cluster, releasing a leaving node from the cluster etc. Certain other clustering algorithms use distributed clustering technique for cluster maintenance [2, 5]. In distributed clustering algorithm there is no special node that is in charge of cluster maintenance, therefore, they do not suffer from single point (cluster head) of bottle neck especially in a highly mobile environment where many nodes join and leave clusters frequently. Servicing time for a cluster maintenance request (such as joining, leaving a cluster) is less in case of distributed algorithm than in case of centralized algorithms, since the cluster head need not be consulted for every cluster maintenance decision. Thus a distributed algorithm is preferred for cluster maintenance. Our cluster maintenance strategy employs a distributed approach.

The rest of the paper is organized as follows. The related work in this area is discussed in the next section. The characteristics of the proposed algorithm are described in section 3. In section 4 we discuss the simulation study and results. Some concluding remarks are given in section 5.

2. LITERATURE REVIEW

Many algorithms have been proposed for dealing with clustering. We discuss some of these in this section.

Gerla et.al. [1] discusses two algorithms, one based on lowest ID and the other based on highest degree (degree is defined as number of neighbours each node has). In the lowest ID algorithm, node with the lowest ID becomes the cluster head, where as in the highest degree algorithm, highest degree node becomes the cluster head. Since the degree of a node changes with changes in topology, the cluster heads are not likely to stay as cluster heads for very long [6]. On the other hand, using the Lowest-ID algorithm, the nodes with low ID stay as cluster heads most of the time. In the lowest ID algorithm, nodes with low ID tend to become cluster heads often, which is unfair to the low ID nodes. This unfairness does not exist in our algorithm. In both the algorithms, cluster head relinquishes its position as and when another prospective (lowest ID/highest degree) node joins the cluster. Therefore there is a likelihood of more cluster head changes with changes in topology. In our algorithm, the cluster head retains its position till it opts to leave the cluster.

Lin R. et.al. [2], is a modified version of lowest ID algorithm. It employs a distributed technique with number of hops as limiting parameter. Within each cluster, nodes must be able to communicate with each other in at most two hops. This algorithm may generate single node clusters [5] which will increase the number of clusters formed. Thus it seems that additional procedures for merging or rearranging clusters may be desirable. Our distributed cluster maintenance scheme has tackled this problem by employing merging of clusters based on a threshold size.

In the paper proposed by Ramanathan et.al. [4], the clustering procedure autonomously groups the switches into clusters, clusters into super clusters, and so on, forming a hierarchical control structure. In this method, the cluster maintenance decisions like splitting and merging is done by a cluster leader, which is not a distributed approach. The cluster leader is the lowest ID entity in the cluster. This algorithm may generate clusters with more than two hops in diameter. Cluster maintenance is complicated and difficult in this algorithm especially under high mobility [3].

Clustering technique by Lin C.H et.al. [7], uses hop distance to control the cluster structure. One of the nodes in each cluster takes the role of the cluster head. The distance between a cluster head and any member of the cluster is within a predetermined maximum number of hops. Restricting the number of hops does not restrict

the number of nodes in a cluster. Therefore this technique will lead to formation of clusters with larger size.

3. PROPOSED CLUSTER MAINTENANCE ALGORITHM

This is a distributed cluster maintenance algorithm which limits the number of nodes inside a cluster. The state of a node can be identified by six tuples, cluster ID c_{id} , node ID n_{id} , cluster head ID ch_{id} , node count in a cluster n_c , time stamp τ , cluster size limit in terms of number of nodes in the cluster N .

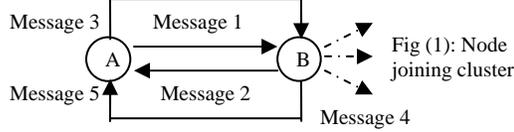
State of a node: $\{c_{id}, n_{id}, ch_{id}, n_c, \tau, N\}$

The cluster ID is unique to a cluster and is used to identify a cluster. The node ID is unique within a cluster. The cluster ID appended with the node ID forms a unique identifier for every node in the network, which is dynamically allocated when a node joins a cluster. The dynamic allocation of ID is more preferred than globally unique fixed ID, as dynamically allocated ID length is dependent on the number of active devices in the network. The fixed ID length is governed by the total number of devices (active and inactive). The cluster head ID is the node ID of the cluster head in the cluster. Every node in the cluster has information about its cluster head so that it can communicate across the cluster. Every cluster head maintains a cluster head information table wherein the information about other cluster heads is stored. Node count n_c is maintained by each node so that it does not exceed the cluster size limit in terms of number of nodes N . The time stamp is used to find the time of entry of each node in a cluster. The cluster size limit in terms of number of nodes in a cluster is used to limit the number of nodes inside a cluster. Further, we will consider node joining a cluster, node leaving a cluster and special conditions pertaining to it.

3.1 Node joining a cluster

When a new node A wants to join a cluster, it broadcasts a join request message. The nearby nodes to A, which are within its transmission range, respond to A. These nearby nodes can be from one or more clusters. The response can be an acceptance or a reject. A reject occurs when the nearby node is not able to accommodate the new node A due to its node count limit. Acceptance/reject messages can be from one or more nearby nodes. When new node A receives one acceptance reply from a nearby node, it joins that cluster. If it receives more than one reply, it compares the node count of each cluster and joins the one with lower node count and rejects the join reply of other clusters. This joining uses a protocol as shown in Fig (1). In Fig (1), node A is the new node, which wants to join a cluster. Node B is the nearest node, which is already a part of the cluster, which A wants to join. When A receives a reject from all the nearby nodes and still wants to communicate, A forms a cluster of its own. It

communicates with the nearby cluster head, which allocates A, a non-allotted cluster ID. This cluster head then informs the other clusters of a new cluster with A as cluster head. Node A can also communicate with a nearby cluster head which then contacts the central authority for assigning a cluster ID to A.



- Message 1: A → B The new node A broadcasts a join request
- Message 2: B → A The nearby node B sends a join reply message with 6 tuples to A: $\{c_{id}, n_{Bid}, ch_{id}, n_c, \tau, N\}$.
- Message 3: A → B The new node A sends join accept
- Message 4: B → Its members B broadcasts a message with 3 tuples to its members informing them of a new node arrival: $\{n_{Bid}, n_{Aid}, \tau\}$.
- Message 5: B → A B sends a join confirm message to A with the allocated node ID, n_{Aid} .

3.1.1 Cluster ID allotment

Every cluster head is capable of allocating the cluster ID. The node that wants to form the cluster, communicates with nearby cluster head, which allocates the new cluster ID. This cluster head then informs the other cluster heads, the arrival of a new cluster with the cluster ID allocated.

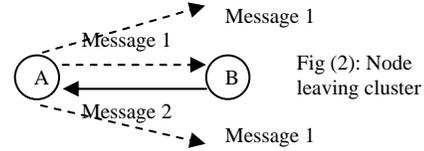
3.1.2 Node ID allotment

Every node is capable of allocating a node ID for a new member. The nearby node that admitted it in the cluster allocates the new node ID. Every node maintains a table of node IDs of its cluster members. This helps to keep track of the non-allocated node IDs. When a new node arrives then the lowest of the non-allocated node IDs is allocated. When two nodes want to join but they communicate with two distinct nodes of the same cluster, it is possible that same node ID be allocated to both of them. To solve this issue every node broadcasts a message with its own node ID, newly allocated ID and time stamp (Fig (1), Message 4). The lowest time stamp node wins and a reject is offered to the other which then selects another non-allocated node ID.

3.2 Node leaving a cluster

When a node A wants to leave a cluster, it broadcasts to its members that it wants to leave the cluster. It releases its node ID and waits for an ACK message from atleast one of its members. In Fig (2), node A wants to leave the cluster. Node B is a nearby node and is a part of the same cluster to which A belongs. The node leaving the cluster ensures that the ACK message from atleast one of its members is received. If not the leaving message has to be resent. When A is in the periphery of its own cluster and wants to leave its current cluster and join another nearby cluster, it uses soft /hard handoff technique to leave the current cluster and join the nearby cluster. When a

cluster head wants to leave the cluster, it broadcasts a message to its members. Every member maintains a list of the entered nodes with time stamp in a table. This helps them to identify the last entered node. This last entered node then becomes the cluster head.



- Message 1: A → its members Node A wants to leave the cluster. It broadcasts the message with 6 tuples to its members: $\{c_{id}, n_{Aid}, ch_{id}, n_c, \tau, N\}$. It also releases its node ID.
- Message 2: B → A The nearby member B sends an ACK message. It marks the node ID of the leaving node vacant in its table and decrements the node count by 1

There can be special cases where the leaving member does not inform the members while leaving. As a result there is a possibility that the members do not know of its leaving. To circumvent this, the cluster head periodically initiates a node count inside the cluster and broadcasts information about lost nodes.

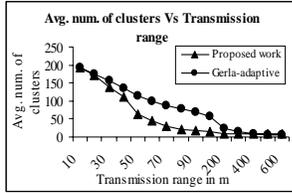
3.2.1 Election of cluster head

When a cluster head wants to leave the cluster, a new cluster head is elected. The most recently entered node in a cluster is made the cluster head. However, we can also employ a weight based election of the cluster head [5]. For mobile environments such as inter-vehicular communication the nodes often tend to move in one direction. Therefore for inter-vehicular communications, the recently entered nodes have higher likelihood to stay in the cluster for a longer duration when compared to the other nodes. In the weight based method, each node has a weight that will indicate its suitability for cluster head role, and weight is decided by a generalized formula [5].

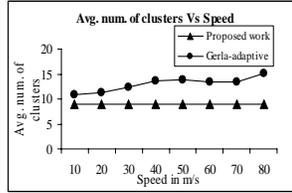
4. EXPERIMENTAL RESULTS

4.1 Network model

The terrain dimensions used in these simulations were 1000m by 1000m and the number of nodes taken for simulation was 200. The transmission range taken for study was in the range 10 to 600 metres. The speed range taken for study was between 10 m/s to 80m/s. We study the stability of cluster in terms of number of cluster head changes and the number of clusters formed for different transmission ranges and speed. Fig (3) plots the average number of clusters for varying transmission ranges. The average computed here is the time average over a period of 120 seconds. With 200 nodes in terrain, the proposed cluster maintenance strategy shows the formation of lesser clusters when compared to the algorithm by Gerla [2]. This is because, the nodes in the cluster are restricted by the hops leading to more number of clusters with varying number of nodes in each cluster. This restriction of number of hops is not existent in our algorithm. We limit the number of nodes in the cluster.



Fig(3) Avg. num. of clusters with tx range



Fig(4) Avg. num. of clusters with speed

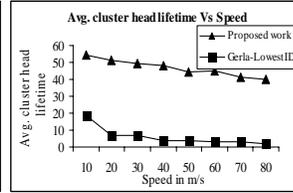


Fig (5) Avg. cluster head lifetime with speed

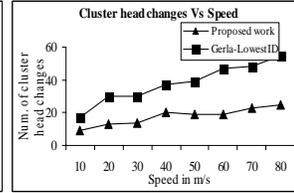


Fig (6) Cluster head changes with speed

We also observed that there were a few clusters which did not reach its cluster limit. We therefore introduced merging of clusters to solve this problem. This has not been discussed in this paper due to space constraints.

Fig (4) depicts the time average number of clusters for different speeds. With increase in speed, cluster members change clusters frequently however the average number of cluster remains the same. The cluster heads in our algorithm maintain a cluster head information table. This table size will not grow to larger sizes as the number of clusters is found to be relatively constant. In the algorithm by Gerla [2], we find that the average number of clusters increase with increase in speed. This may not be desirable because, maintaining a large table will be tedious, especially in high mobility environments. With a low variation in number of clusters, with fixed number of nodes in a cluster, the table size would be easily maintainable.

In Fig (5) the average lifetime of a cluster decreases for high speeds. Average lifetime is computed as a time average over 120 secs. The lifetime of a cluster head in our proposed strategy is more than the lowest ID algorithm [1]. One of the important criteria in clustering is stability. The frequent changes of cluster head may adversely affect the performance of the clustering algorithm. In the lowest ID algorithm the lowest ID will keep changing with changes in topology. The cluster head of the lowest ID algorithm relinquishes its position when another lower ID joins the cluster. In our algorithm, the cluster head relinquishes its position only when it leaves the cluster. This results in more lifetime of the cluster head when compared to lowest ID. More life time of a cluster head implies that there are lesser changes in the cluster head. This is shown in graph Fig (6). Lowest ID algorithm shows more cluster head changes when compared to our strategy.

5. CONCLUSION

The distributed cluster maintenance strategy limits the number of nodes per cluster. In addition, it also leads to better cluster maintenance. The cluster maintenance algorithm shows the formation of lesser clusters for varying speeds and transmission ranges when compared to the algorithms taken for comparison study

The variation in number of clusters with speed is less and is stable when compared to the adaptive algorithm by Gerla [2]. Our simulations show that this strategy leads to better stability of clusters. It was also found that the cluster head lifetime is more in our algorithm compared to the lowest ID algorithm [1]. Our investigations show that limiting the maximum number of nodes per cluster is highly effective when both the speed and the transmission range are varied. The proposed work was simulated and compared with the other algorithms.

6. REFERENCES

- [1] Gerla M. and Tsai T.C., "Multicluster, Mobile, Multimedia Radio Network" ACM Baltzer J. Wireless Networks, Vol. 1, No. 3, pp. 255-265, 1995.
- [2] Lin C.R. and Gerla M., "Adaptive Clustering for Mobile Wireless Networks," IEEE Journal on Selected Areas in Communications, Vol. 15, No.7, pp. 1265-1275, September 1997.
- [3] Basu.P, Khan.N and Little.T.D.C," A mobility based metric for clustering in mobile ad hoc networks", In Proceedings of Distributed Computing Systems Workshop 2001.
- [4] Ramanathan R. and Steenstrup M., "Hierarchically-organized, Multihop Mobile Wireless Networks for Quality-of-service Support," ACM/Baltzer Mobile Networks and Applications, Vol. 3, no. 1, pp. 101-119, Jun. 1998.
- [5] Chen.G, Nocetti.F, Gonzalez.J, Stojmenovic.I, "Connectivity-based K-hop clustering in wireless networks", Proceedings of the 35th Annual Hawaii International Conference on System Sciences - HICSS'02
- [6] Johansson T., Carr-Motyckova L.,"On Clustering in ad-hoc networks", Proceedings of the First Swedish National computer networking workshop, Sept 2003.
- [7] Lin H.C. and Chu Y.H., "A Clustering Technique for Large Multihop Mobile Wireless Networks," Proceedings of the IEEE Vehicular Technology Conference, May 15 - 18, 2000.
- [8] Krishna P., Vaidya N.H., Chatterjee M., and Pradhan D.K., "A cluster-based approach for routing in dynamic networks", ACM SIGCOMM Computer Communication Review, pp. 49-65, April 1997.

An Energy-Efficient Clusterhead Assignment Scheme for Hierarchical Wireless Sensor Networks

(An Extended Abstract)

Dajin Wang

Department of Computer Science
Montclair State University
Upper Montclair, NJ 07043, USA

1 Introduction

A Wireless Sensor Network (WSN) is composed of a large number of sensor nodes, and a few (at least one) “central” node(s). The sensors are embedded into various physical environments mainly for the collection of physical world data. The data are transmitted to, or gathered by, the central nodes for aggregation, analysis, and processing. The central nodes also play the role of manager of the WSN. The communication among nodes is all via wireless means. Therefore all nodes are equipped with radio transceivers/receivers. WSNs have very promising prospect in many applications, such as environment monitoring, traffic monitoring, target tracking, and fire detection.

Different models of WSN have been proposed. However some basic characteristics can be observed that are common in most proposed models.

- They are all composed of a large number of sensor nodes, and a small number of (in some models just one) master nodes (central nodes);
- All sensor nodes are relatively low cost, perform relatively limited computational operation. Their main job in the whole system is to collect raw data, and render it to the master nodes, with or without some primitive pre-processing;
- The master nodes collect the data from all sensors, and analyze/process them. They are much more powerful, costlier processors than ordinary sensors. The master nodes are also the managers of the network.

A WSN can have either just one master node or a group of master nodes, depending on the network’s

scale of geographical coverage and/or cost effectiveness consideration. In a single-master WSN, the master node (also called base station) collects and processes data from all sensors. It is also the sole manager of the entire network system. In a multi-master WSN, the tasks of data collection, aggregation, processing, and network management are distributed among a group of nodes working collaboratively. The organization of these master nodes is one of the essential issues in the design of WSN architecture.

In [2], a cluster-based, hierarchical model for WSN, named COSMOS, was proposed. COSMOS takes up a hierarchical network architecture comprising of a large number of low power, low cost sensors. The sensors are organized into spatial clusters. For each sensor cluster, there is a *clusterhead*. Sensors within a cluster communicate in a time synchronized manner, using single hop communication. The clusterheads form a mesh-like topology and communicate asynchronously. Algorithms basic to sensor networks, such as sorting and summing, are addressed using COSMOS as the underlying architecture.

As a matter of fact, the topologies of many distributed systems are more or less hierarchical. If distributed functions are performed in such a way as to reflect the underlying hierarchical topology, the algorithm design can be greatly simplified. A hierarchical architecture may also help improve scalability of the distributed functions, or even the scalability of the network itself. The hierarchical approach has been used in solving many different problems of distributed nature, such as distributed monitoring, resource scheduling, and network routing, either to effectively coordinate the local control activities or to enhance the overall system performance.

In this paper, we study the problem of adequately

assigning clusterheads (master nodes) in a hierarchical WSN. That is, for a given WSN, how many clusterhead nodes the WSN should have, and where they should be positioned. There are numerous factors affecting the assignment of clusterheads. A solution optimizing all performance metrics, such as time, memory space, and energy consumption, is impossible to obtain. A solution to the optimization problem is both application and network topology dependent. In this work, we will focus on the important issue of energy efficiency of the WSN. A widely accepted convention of WSN is that nodes are running on batteries. Therefore power saving is an especially important goal in architecture design. We will propose a hierarchical WSN architecture toward the end of saving energy of both sensor nodes and clusterheads.

2 The sensor network model

A wireless sensor network resembles a conventional parallel and distributed systems in many ways. However, several unique characteristics stand out to call for redefinition, or modification, of the network model. Those characteristics include energy efficiency consideration, communication reliability, and global awareness of individual nodes, among others. Because of the wide diversity of sensor applications, it is hard to capture all characteristics in a single model.

In this paper, we will adopt, with slight modification, the WSN model called COSMOS (standing for Cluster-based heterOgeneous MODEL for Sensor networks), proposed by Singh and Prasanna [2]. A WSN model aiming at large size and scalability, COSMOS features a cluster-based, hierarchical network architecture. It comprises of a large number of low power, low cost sensors, presumably distributed in a large physical environment. The distribution of sensors is close to uniform. That is, in each unit area there is a sensor with high likelihood. Sensors are organized into equal-sized, square-shaped clusters according to their spatial proximity. For each sensor cluster, there is a clusterhead, which is costlier, more powerful in computational capability and radio transmission range. The clusterheads of the whole WSN form a mesh-like topology. The sensors' main job is to collect first-hand, raw data, with or without some initial processing. The clusterheads perform more intensive, more complex tasks. It is at the clusterheads that the data of the sensor network get processed in a collaborative manner. Fig-

ure 1 illustrates the basic structure of the hierarchical WSN.

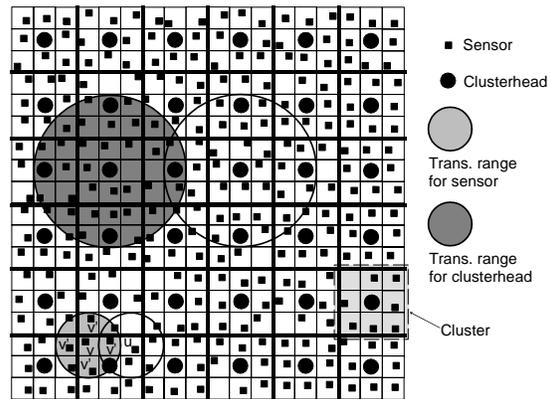


Figure 1: A clustered, hierarchical wireless sensor network.

In Figure 1, the sensors (represented by black squares) are almost evenly distributed in a two-dimensional terrain. Each unit area (or cell) contains one sensor with a high probability. There may be a few unit areas that have no sensors. A group of nearby sensors are organized into square-shaped clusters. At the center unit area of a cluster is stationed a clusterhead (represented by a bigger, black circle). For the purpose of energy saving, an ordinary sensor's communication capability is presumably very limited, consuming as low as possible power in radio transmission. We assume it can only communicate directly with its four immediate neighboring sensors. For example in Figure 1, sensor v can only guarantee reliable transmission to the four v' sensors. To communicate to sensor u , the message has to be relayed by one v' sensor.

On the other hand, a clusterhead is equipped with more powerful transceiver that can communicate with any node within the cluster. However, again for the purpose of energy saving, we do not assume a limitless, super powerful clusterhead that can send/receive radio messages to all sensors/clusterheads in the system. Beyond all nodes in its own cluster, the transmission range of a clusterhead is such that it can only guarantee reliable communication with the clusterheads of its four neighboring clusters. See Figure 1.

For the communication mechanism, since a sensor has very limited radio range, if it wants to "actively" send message/data to its clusterhead, it can only do so by relaying through intermediate sensors (routing scheme in this context is another issue, which will

not be addressed in this work). Most of the time the sensor data are “passively” picked up by its clusterhead. For analysis purpose, we quantify the energy dissipated by one round of sensor transmission to a simplified, normalized unit. Refer to the example in Figure 1 again: If sensor v wants to send one unit of data to sensor u , 2 units of energy will be consumed — 1 for transmission from v to v' , 1 from v' to u . The clusterhead, we assume, can support multiple power states to transmit to sensors/clusterheads of different distances. We also assume that the clusterhead fetches a unit of data in one unit time.

We demonstrate the communication model in Figure 2. In Figure 2 (a), the darkest node repre-

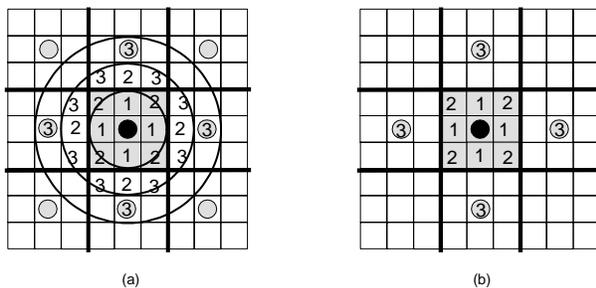


Figure 2: (a) The center clusterhead uses different power states to transmit to sensors/clusterheads of different distances; (b) Cells the center clusterhead will transmit to with different power states.

sents the clusterhead for the central cluster, drawn in grey (for clear viewing, ordinary sensors are not drawn). In this particular example, a clusterhead is equipped with three different transmission power states, represented by the three circles in (a). The largest transmission power allows a clusterhead to transmit to the four clusterheads of its neighboring clusters. The number in a cell stands for the clusterhead’s power state level needed to reach that cell. Note that a rather strict standard is adopted here to guarantee reliable transmission: In the grey cluster, the four corner cells are not covered by a state-1 transmission because the circle does not completely cover the cell area. They will be properly covered by a state-2 transmission, incurring more power. The grey nodes in (a) are all clusterheads, of which four will be reached by the central clusterhead with a one-hop, state-3 transmission. Figure 2 (b) summarizes the cells covered by the central clusterhead with various power states.

3 A power-efficient cluster-head assignment scheme

A “super” cluster that contains all sensors would be desirable. However it is not feasible as the size of WSN grows larger. Considerations such as energy limitation, cost, and scalability make a single-centered WSN not only unfavorable, but also difficult to implement. The proposal of hierarchical organization of WSN [1, 2, 3] is to distribute the computational and managerial tasks to a group of clusterheads. The approach will reduce the communication traffic in network, and will allow the deployment of less powerful, lower priced processors to do the WSN’s computational and managerial jobs in a collaborative manner. One basic question in constructing a hierarchical WSN is how to cluster the sensor network, or equivalently, how many clusterheads are to be used and where to position them.

Due to the diversified nature of WSN’s data processing tasks and concerns like cost effectiveness and system robustness, it is impossible even to define a comprehensively “optimal” architecture, let alone achieve it. An optimal, or asymptotically optimization is practical only in terms of narrowed optimization target. In this work, we will consider such a narrowly defined optimization target, i.e., we are trying to find a clusterhead assignment so that the WSN’s overall battery power dissipation is minimized. The scenario we use to justify our assignment scheme is an operation that presumably requires most amount of energy: The central processor of the WSN needs to process data collected from all sensor nodes. We assume that the nature of the application allows “partial preprocessing” of data before they reach the central processor. There are many such data in both computational and managerial tasks. For instance, the aggregatable sensor data is of such nature. Another example is to get the sum of certain value from all sensors: It is not necessary for the central processor of the WSN to collect all addends before it performs the addition — partial sums can be obtained by clusterheads, and sent to the central processor. That will prevent the central processor from collecting all data from afar, reducing the energy use. It is in this context that we propose an optimal scheme for clustering the WSN.

The target is to find a hierarchical clustering, so that the data collection/processing task by the center station dissipates the minimum amount of energy overall. To formulate problems quantitatively for analysis, we assume a simple model for calculating

battery consumption. It should be pointed out that the model is a normalized abstraction from vastly variable real scenarios. Refer to Figure 2 (b) again. Firstly, we use the battery state level to represent needed power to transmit to a cell. This representation is characteristic of the reality: The farther the sensor, the more power the clusterhead needs for transmission. Secondly, we assume each transmission round will collect data from one sensor. That is a simplification of real situations where multi-channel transmission/reception may be supported. However the results obtained from the single-channel model can be applied to multi-channel models with minor adaptation. With the above assumption, in Figure 2 (b), if the center clusterhead wants to collect one round of data from its own cluster, a total of $1 \times 4 + 2 \times 4 = 12$ units of power will be consumed.

In a hierarchical WSN design, the whole square terrain is divided into a set of smaller, square-shaped clusters. There is still a center station for the whole WSN, located at the center of the mesh composed of all clusterheads. The data collection of central station is performed in two phases. In the first phase, all clusterheads collect data from sensors in their own clusters. The data is aggregated and/or preliminarily processed in clusterheads. In the second phase, the WSN's center station collects data from all clusterheads. Figure 3 illustrates the structure of hierarchical WSN.

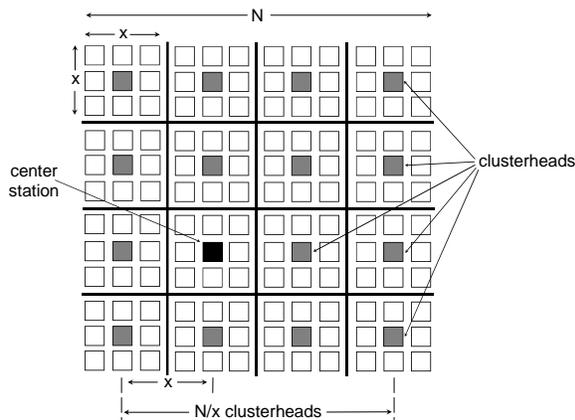


Figure 3: The hierarchical WSN. The $N \times N$ -cell terrain is divided into $(\frac{N}{x})^2$ $x \times x$ -cell clusters. The grey nodes are clusterheads, the darkest node is the center station of the WSN.

In Figure 3, each square-shaped cluster consists of $x \times x$ cells. The whole WSN is divided into $(\frac{N}{x})^2$ clusters. Each cluster has a clusterhead located at the center. The $(\frac{N}{x})^2$ clusterheads form a mesh, and

at the center of the mesh is the WSN's center station. But what is x , the size of the clusters? The choice of this size can affect many aspects of the WSN. However, as stated earlier, we will narrow our target and figure out an appropriate cluster size x favoring minimizing the power consumption incurred by communication. The main result of our work can be summarized in the following theorem.

Theorem 1 *In a hierarchical $N \times N$ -cell WSN, if the clusters are of dimension $x \times x$, so that*

1. x is as close to $\sqrt[3]{2N}$ as possible
2. x divides N

then the system's total power consumption for one round of data collection/aggregation is minimum.

The saving of power gained by this hierarchical scheme is quite substantial. The ratio of minimal hierarchical power cost versus non-hierarchical cost is (assuming even N , even x) given by $\frac{3}{2} \sqrt[3]{\frac{2}{Nx}}$, which is ever decreasing as N grows. When $N = 10$, the min-hierarchical-cost/non-hierarchical-cost ratio is about 40%; when $N = 100$, less than 9%; when $N = 200$, less than 6%.

With the analytically optimal cluster size, we conducted simulation experiments for more realistic situations. The simulation results showed power reduction rate similar to the ideal situation, based which the cluster size was determined.

References

- [1] P. Ramanathan, K. C. Wang, K. K. Saluja, and T. Clouqueur, "Communication Support for Location-Centric Collaborative Signal Processing in Sensor Networks," *DIMACS Workshop on Pervasive Networks*, Piscataway, NJ, May 2001.
- [2] Mitali Singh and Viktor K. Prasanna, "A Hierarchical Model for Distributed Collaborative Computation in Wireless Sensor Networks," to appear in *International Journal of Foundations of Computer Science*, 2004.
- [3] M. Younis, M. Youssef, and K. Arisha, "Energy-Aware Routing in Cluster-Based Sensor Networks," *Tenth ACM Int'l Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, Fort Worth, TX, Oct. 2002.

An Improved Jitter Time-Stamp Method for Source Timing Recovery in AAL2

J. Walker

Department of Electronic and Computer Engineering
University of Limerick, Limerick, Ireland
jacqueline.walker@ul.ie

Abstract-The AAL2 (ATM Adaptation Layer 2) does not currently provide a method for source rate recovery. Recently, the jitter time-stamp (JTS) method was proposed for this purpose. An analysis of the method is presented which shows that its calculation of playout timing introduces an error term which can disrupt playout. An improved JTS method which does not suffer from this problem is proposed and its performance tested by simulation.

I. INTRODUCTION

Transport of voice by carriers is being progressively migrated to packet transport [1] and there are now many competing voice packet transport technologies such as MPLS (multiprotocol label switching), FR (frame relay), VoIP (voice over IP) and ATM [2]. ATM has the advantage of a widely installed base in the backbone network, whereas in the access network it has been standardized for voice over DSL [1]. ATM also offers integrated transport of voice, data and video traffic, QoS guarantees and proven performance in existing implementations [3]. Originally, voice over ATM was only available for constant bit rate voice [4]. More recently, recommendations for AAL2 (ATM Adaptation Layer 2) have been made [5], [6]. The new AAL2 is designed for low bit-rate delay sensitive traffic such as voice, which is inherently VBR (variable bit rate) due to silent periods. To achieve greater efficiencies, AAL2 allows multiple VBR voice channels to share a single ATM connection while still guaranteeing delay requirements [7]. Because the CPS (Common Part Sublayer) allows variable size packets, less than 45 bytes, from different channels to be assembled into a single ATM cell, the multiplexing gain of AAL2 over the use of AAL1 and AAL5 (which only allows one packet per ATM cell) is significant [8], [9].

In packet voice networks, a number of different approaches are used to reconstruct the timing of the source voice stream at the destination. When clocks at the source and destination are synchronized, it is possible to use an absolute timing approach such as a time stamp which encodes the source clock in terms of a reference clock. The AAL1 [4] applies this approach relying on synchronous network clocks and information at the destination to fully reconstruct the source clock [10]. If synchronous network clocks are not available, slips may occur.

A periodic packet stream, although jittered by variable network delay, can be smoothed by adding a

constant delay, and read back from a data buffer at the same rate using relative timing information, such as sequence numbers, to order the packet stream [11], [12]. As the source and destination clocks are only nominally the same frequency, buffer overflow and underflow may occur. For particular traffic types, such as voice, controlled buffer slips may be allowed as they can be disguised by expanding or compressing silent periods [11], [12]. Another approach is to monitor the destination buffer level and use this to adjust the destination clock frequency [12]. Extension of this approach has led to more complex adaptive clock recovery algorithms [12].

The current AAL2 does not specifically provide for source timing recovery and the JTS (jitter time-stamp) method has been recently proposed for this purpose [13]. According to its proposers, source rate recovery at the AAL2 layer reduces the computational load on the application layer. As it allows VBR sources to be sent as VBR, it removes the need for traffic smoothing and allows multiplexing of VBR sources while still satisfying QoS requirements [13].

The paper is organized as follows: in Section II., the JTS method is described and analysed and some problems with the original proposal are uncovered. In Section III., an improved JTS method is introduced and analysed and the results of a preliminary simulation of the performance of the original JTS method and the improved JTS method are presented.

II. THE JITTER TIME-STAMP (JTS) METHOD

In the JTS method, every Nth packet generated by the source is a timing packet and carries a number known as the TI (timing indication). The TI is generated every N reference clock cycles by a b -bit counter and is therefore $\text{mod}(2^b)$. The lower bound of b is $b_{lb} = \lceil \log_2(2J_{max}/T_r N) \rceil$, which depends on the reference clock period, T_r , the maximum jitter expected in the network, J_{max} , and the value of N. In the JTS method, the reference clock period at the source is synchronous with the source clock [13] so that latching by the timing packet arrival does not cause issues of either metastable failure or synchronization jitter as in SRTS [10].

The arrival time of the packet at the receiver will depend upon both a constant network delay and the

variable network delay or jitter. The aim of the JTS method is to calculate the jitter component and then remove it. The recalculated arrival time which results is known as the AdAT (adjusted arrival time) and is used to calculate the timing of the playout pulses. As shown in Fig. 1: at the receiver, a timing packet arrival latches a pair of counters. One counter, C_r , has b bits and is clocked at the rate of the reference clock divided down by N . The other counter, C_{tc} , is clocked at the rate of the reference clock and has b_{1C} bits where

$$(b_{1C})_{lb} = \lceil \log_2(2J_{max}/T_r) \rceil = b_{lb} + \lceil \log_2 N \rceil. \quad (1)$$

The arrival of a timing packet is captured in two numbers: τ_{1C} is $\text{mod}(2^{b_{1C}})$ and expresses the arrival to the nearest reference clock cycle; τ_{NC} is $\text{mod}(2^b)$ and expresses the arrival to the nearest N th reference clock cycle. Finally, counter C_{tc} is also latched on every N th reference clock cycle giving a value τ_{1C_N} which is also $\text{mod}(2^{b_{1C}})$. The algorithm then estimates the packet's expected arrival time (EAT) as

$$EAT = \left(TI + \left\lceil \frac{D_{ref}}{N} \right\rceil \right) \text{mod} 2^b. \quad (2)$$

An initial value for the delay for a given session, D_{ref} , is measured at setup of the connection [13]. EAT could be used for playout calculation but instead the i th value of the network jitter is calculated as

$$J_i = J_{main} + J_{tail}. \quad (3)$$

J_i is made up of a main component, J_{main} (the variable network delay) and a tail component, J_{tail} . The main component of the jitter is calculated as

$$J_{main} = N[(EAT - \tau_{NC}) \text{mod} 2^b]. \quad (4)$$

Both EAT and τ_{NC} are positive b -bit numbers which represent instants in time. Now, as b_{lb} is at least one bit greater than the number of bits necessary to represent the maximum jitter expected in the network (in units of N reference clock cycles), then, for values of $|J_{main}| \leq |J_{max}|$, the result of the subtraction in (4) will be within $[-2^{b_{lb}-1}, 2^{b_{lb}-1} - 1]$. The final size of J_{main} when $|J_{main}| \leq |J_{max}|$ is thus $b_{lb} + \lceil \log_2 N \rceil$. The 'tail' component of the jitter J_{tail} provides a measure of the fine detail of the jitter in the packet arrival due to the difference in resolution of the clocking of the counters and is calculated as [13]

$$J_{tail} = (\tau_{1C} - \tau_{1C_N}) \text{mod} 2^{b_{1C}}. \quad (5)$$

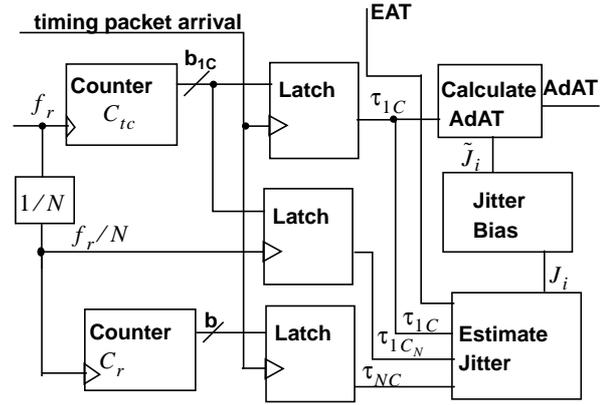


Fig. 1: Dejittering and AdAT calculation at the receiver.

Now since, it is always the case that $\tau_{1C_N} < \tau_{1C}$, J_{tail} is always positive. The range of possible values for J_{tail} is $\{0, 1, 2, \dots, N-1\}$ and the significant number of bits in J_{tail} is $\lceil \log_2 N \rceil$. Thus, $|J_{tail}|_{max} \leq |J_{main}|_{max}$ since $b_{lb} > 0$. The maximum value of J_i is $J_{i,max} = N2^{b_{lb}-1} - 1$ where only the case of $J_{main}, J_{tail} > 0$ need be considered as J_{tail} is always positive. The maximum value of J_i can be represented using $b_{lb} + \lceil \log_2 N \rceil$ bits, which from (1) will be possible using b_{1C} bits.

The final step in the algorithm is the calculation of the adjusted arrival time AdAT as

$$AdAT = (\tau_{1C} - \tilde{J}_i) \text{mod} 2^{b_{1C}} \quad (6)$$

where \tilde{J}_i is the bias adjusted jitter value. Bias adjustment of the jitter is done using the mean value of the jitter over some estimation period [13]. The AdAT is used to generate the timing pulses which, after smoothing by a phase-lock loop, will provide the de-jittered playout clock. Calculation of the playout pulses proceeds from timing packet to timing packet. The duration of a single playout interval is given by

$$\Delta = (AdAT_2 - AdAT_1) \text{mod} (2^{b_{1C}}) \quad (7)$$

where $AdAT_2$ and $AdAT_1$ correspond to the first two timing packets in the data buffer and the number of pulses to be played out is given by

$$\lambda = \frac{SN_2 - SN_1}{N} \quad (8)$$

where SN_2 and SN_1 likewise correspond to the first two timing packets.

It can be shown that, for a given bias adjusted jitter value, the mean jitter values only affect the playout time interval once at the time the bias adjusted jitter value is updated. In between these changes, it is valid to consider a simplified form of (6) as

$$AdAT = (\tau_{1C} - J_i) \text{mod} 2^{b_{1c}} \quad (9)$$

Equation (9) can be simplified as shown in Appendix A to give

$$AdAT = (\tau_{1C_N} - J_{main}) \text{mod} 2^{b_{1c}} \quad (10)$$

Generating AdAT by subtracting the estimated variable network delay should be effective. However, J_{main} is defined as [(ideal time) - (actual time)] and will be negative when the packet is delayed and positive when the packet arrives early. Thus, when the packet is delayed, the value of J_{main} is negative and the effect of subtracting a negative value from the actual arrival time is to make the adjusted arrival time more positive - which corresponds to more delay. Conversely, when the packet arrives early, the adjusted arrival time is adjusted to be even earlier. Hence, (10) tends to *exacerbate* the effect of the network jitter, rather than improve on it.

As shown in Appendix A, (10) can be further simplified to

$$AdAT = (2\tau_{1C_N} - N(EAT)) \text{mod} 2^{b_{1c}} \quad (11)$$

Substituting (11) in (7) and gives

$$\Delta = (2((\tau_{1C_N})_2 - (\tau_{1C_N})_1) + NT'_r) \text{mod} 2^{b_{1c}} \quad (12)$$

where from (2) $N((EAT)_2 - (EAT)_1)$ is proportional to N reference clock periods and the notation T'_r indicates that the reference clock at the destination is not identical to that at the source. Note however the extra term in (12). Until wrap-around of counter C_{tc} , $(\tau_{1C_N})_2 > (\tau_{1C_N})_1$, so this term will add an error term to the playout clock signal.

III. AN IMPROVED JTS METHOD

The observed problems with the JTS method can be overcome by reversing the sign convention of J_{main} and calculating it as

$$J_{main} = N[(\tau_{NC} - EAT) \text{mod} 2^b] \quad (13)$$

and it can be shown by applying the approach of Appendix A that

$$AdAT = (N(EAT)) \text{mod} 2^{b_{1c}} \quad (14)$$

Substituting for the new value of AdAT in (7) gives

$$\Delta = (NT'_r) \text{mod} 2^{b_{1c}} \quad (15)$$

The playout interval should then track the source rate within the constraint that the reference clocks at source and destination are not synchronous.

A simulation was used to compare the performance of the original and improved JTS algorithms. A 64 kbps CBR source was assumed which, with 20 octets per packet, has $T_r = 2.5$ ms. J_{max} is set at 100ms and $N = 8$. To simulate network jitter, as in [13], a geometrically distributed jitter distribution with probability of success equal to 0.3 is used. A comparison of the performance of the standard JTS algorithm and the improved JTS algorithm are plotted in Fig. 2. The plot shows that the standard algorithm has encountered an erroneous offset value near the start which has effectively reset the playout frequency to a very low value. The improved algorithm does not display this behaviour.

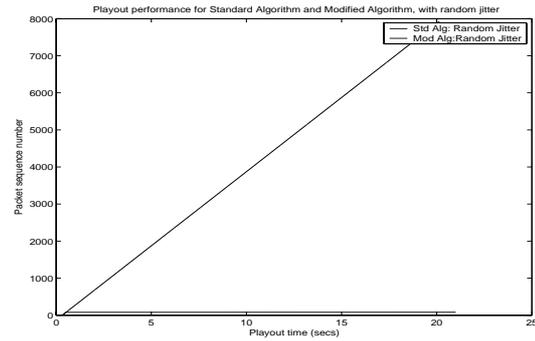


Fig. 2: Packet playout for standard JTS algorithm and improved JTS algorithm with random network jitter.

In this paper, analysis of the JTS method for AAL2, as originally presented in [13], shows that the method is flawed in that it may result in a frequency offset on the playout clock. It has also been shown that an improved JTS method can be obtained by changing the way in which the AdAT is calculated. Preliminary simulation results suggest that the error term encountered in the standard JTS algorithm does prevent correct packet playout and that the improved JTS algorithm does not show this behaviour.

However, due to this analysis, it is also shown that calculating the AdAT by removing the jitter from the actual arrival time of the packet is effectively regenerating the EAT or expected arrival time. A simpler approach would be to calculate playout times directly based on the EAT as shown in (2). The network jitter, J_{main} , may still be measured and can be used to calculate the bias adjusted jitter bias term, which is used

to periodically adjust the playout interval and may be important because the JTS method does not require synchronized clocks at source and destination. Further study is required to fully analyse the effectiveness of the jitter bias estimate adjustments.

The JTS method as originally proposed [13], requires significant overhead in the transport of packets. Compared with the SRTS used in AAL1, which uses 4 bits for every 8 ATM cells [4], JTS, with $N = 8$ and one AAL2 packet per ATM cell, uses 8 bits for every 8 ATM cells [13]. In addition, it assumes a sequence number of 7 bits in every packet [13]. According to [14], a sequence number of at least 4 bits is required to prevent packet slips arising when variable network delay is significant. These sequence numbers are provided to AAL2 voice packets at the AAL2/SSCS (service specific convergence sublayer). In addition, the current AAL2 standard provides for 5 bits of UUI (User-to-User Indication) in the 3 byte CPS Packet header which can be used to indicate voice coding type and sequence numbers of AAL2 packets [14].

APPENDIX A

Let $2^{b_{1c}} = M$, then applying the identity $(x + y) \bmod M = (x \bmod M + y \bmod M) \bmod M$ twice and substituting for J_i from (3) in (9) gives

$$AdAT = (\tau_{1c} \bmod M - (J_{main} \bmod M + J_{tail} \bmod M) \bmod M) \bmod M$$

Substituting for J_{tail} using (5), removing the extra $\bmod M$ and then applying the identity in reverse twice allows simplification to yield (10). Re-apply the identity and substitute for J_{main} from (4) in (10), rewritten as $(N(EAT) - N(\tau_{NC})) \bmod 2^{b_{1b} + \lceil \log_2 N \rceil}$ (where the modulo operation has been adjusted to reflect the significant number of bits following the multiplication operation as discussed in Section II.) Since $b_{1c} \geq b_{1b} + \lceil \log_2 N \rceil$, apply the identity in reverse to give $AdAT = (\tau_{1c_N} - N(EAT) + N(\tau_{NC})) \bmod M$ which can be simplified to (11) by arranging that $N(\tau_{NC}) = \tau_{1c_N}$ by an appropriate choice of the latching cycle of counter C_{tc} .

REFERENCES

[1] D. Wright, "Voice over MPLS Compared to Voice over Other Packet Transport Technologies," *IEEE Comms. Magazine*, vol. 40, pp. 124-132, Nov. 2002.

[2] ITU-T Recommendation I.363, "B-ISDN ATM Adaptation Layer (AAL) Specification", Mar. 1993.

[3] B. Subbiah and S. Dixit, "Low-bit-rate voice and telephony over ATM in cellular/mobile networks," *IEEE Personal Communications*, pp. 37-43, Dec. 1999.

[4] ITU-T Recommendation I.363.1, "B-ISDN ATM Adaptation Layer 2 specification: Type 1 AAL," Aug. 1996.

[5] ITU-T Recommendation I.363.2, "B-ISDN ATM Adaptation Layer 2 specification: Type 2 AAL," Nov. 2000.

[6] ITU-T Recommendation I.366, "AAL type 2 service specific convergence sublayer for narrowband services" Nov. 2000.

[7] B. Subbiah and Y. Raivio, "Transport architecture for UMTS/IMT-2000 cellular networks," in *Proc. IEEE International Performance, Computing and Communications Conference 2000*, vol. 1, pp. 208-214, 2000.

[8] D.W. Petr, R. R. Vatte, P. Sampath and Y. Lu, "Efficiency of AAL2 for voice transport: simulation comparison with AAL1 and AAL5," *Proc. IEEE ICC*, vol. 2, pp. 896-901, June, 1999.

[9] H. Saito, "Performance evaluation and dimensioning for AAL2 CLAD," in *Proc. IEEE Infocom*, vol. 1, pp. 153-160, Mar. 1999.

[10] J. Walker and A. Cantoni, "Determining parameters to minimize jitter generation in the SRTS method," *IEEE Trans. Commun.*, vol. 46, pp. 82-90, Jan. 1998.

[11] W. A. Montgomery, "Techniques for packet voice synchronization," *IEEE J. Selected Areas Commun.*, vol. 1, pp. 1022-1028, Dec. 1983.

[12] R. P. Singh, S. H. Lee and C. K. Kim, "Jitter and clock recovery for periodic traffic in broadband packet networks," *IEEE Trans. Commun.*, vol. 42, pp. 2189-2196, May 1994.

[13] W. Su and I. F. Akyildiz, "The jitter time-stamp approach for clock recovery of real-time variable bit rate traffic," *IEEE/ACM Trans. on Networking*, vol. 9, pp. 746-754, Dec. 2001.

[14] K. Sriram, T. G. Lyons and Y. T. Wang, "Anomalies due to delay and loss in AAL2 packet voice systems: performance models and methods of mitigation," *IEEE J. Selected Areas Commun.*, vol. 17, pp. 4-17, Jan. 1999.

Analysis and Evaluation of Traceback Schemes for Distributed Denial of Service Attacks

Fawad Abbas, Usman Bukhari
University of Southern California, Los Angeles CA
(fabbas, bukhari @ usc.edu)

Abstract:

IP provides us with no mechanism for tracing the source of a received packet. The recent increase in the DDoS attacks has increased the importance of IP trace back. In this paper we have analyzed and evaluated seven different traceback schemes. The emphasis in the comparison has been on the computational overhead; traffic overhead and storage related issues. This paper does not prove one scheme as being superior to other; rather it brings out different pros and cons of these schemes. In the end, we categorize these schemes on the basis of approaches they have taken and discuss the deployment issues related to these approaches.

1. Introduction

DDoS attacks on Yahoo and Ebay [10] are the example of network attacks that we can predict to take place with increased frequency in future. One of the major contributors to the difficulty of mitigating this kind of attacks is the simplistic approach of IP where it does not put any validation check on the source address. Realizing the importance of tracing the actual source, various traceback schemes have been proposed. In this paper we analyze and compare seven of these schemes. The schemes are briefly described and the emphasis has been on the chosen figure of merits for the comparison. These are mainly computational overhead (router and endsystem), storage overhead (router and endsystem), traffic overhead, deployment ease and post mortem though other areas have also been addressed (e.g. scalability and precision).

The rest of the paper is arranged as: Section 2 gives definitions terms used in this paper. Section 3 describes these seven traceback schemes along the figure of merit described above. Section 4 provides the conclusion.

2. Definitions

Following are the definitions of some of the terms used in this paper. *Router Computational Overhead* is defined as the overhead incurred by router during its computation when implementing a traceback

scheme as compared to without it. *End System Computational Overhead* is defined as the overhead incurred by the end system during the implementation of a traceback scheme as compared to without it. *Router Storage Overhead* is defined as the extra storage on the routers while End system storage is the extra storage at the victim when the scheme is implemented. *Traffic Overhead* includes both the increase in the number of packets due to the implementation of this scheme and the increase in the of the normal packet size as a result of corresponding trace back scheme implementation. *Deployment Ease* is composed of three factors. They are compatibility to the existing protocols/infrastructure, amount of ISP intervention required and the incrementally deployment. *Scalability* is defined as the ability of the system to scale in the presence of multiple attacks. *False Positive* is defined as including of a router/node in a reconstructed trace back graph where it actually is not part of that actual attack graph [1]. *False Negative* is defined as not including a router/node in a reconstructed trace back graph where it is part of actual attach graph [1].

3. Traceback Schemes

Following are the seven traceback schemes that have been recently proposed to track down the origin of an IP packet and their respective merits.

3.1 Probabilistic Packet Marking [8]

In this scheme, as described by Savage, routers probabilistically mark packets with their address information before forwarding them. Edge sampling is performed by two adjacent routers on a packet chosen with some probability (1/20,000) and the new distance field is set to zero. If the packet has already been marked, routers increment distance field. Savage proposes to use the 16bit IP header identification field for this purpose. Traffic overhead is not experienced in this scheme. Computational overhead experience by the routers is low. Routers have to select the packets with a certain probability and append the edge information in it. Computational overhead experienced by the end system in calculating the traceback though is very high as demonstrated by Dawn [1]. The storage overhead at the routers is low because they do not need to store

any information. The storage overhead on the end system is very high. According to [1] the number of packets required to trace a single source at 30 hops away with the marking probability of 1% is around 7500. And this number increases with the increase in the number of sources to be traced back. This scheme is incrementally deployable and does not require ISP management intervention for deployment. The issue of backward compatibility however exists. Fragmented traffic is an example, where we actually require the identification field to be used. Overall we can rate the development ease of this scheme as medium. This scheme can easily be used for post mortem analysis because we generate the traceback graph using the marked packets which are stored at victim's location. This scheme has very bad scalability issues as described by [1]. Increase in the number of attackers exponentially increases the time for the trace back graph generation and the number of false positives. This scheme also has compatibility issue with the deployment of IPv6. The precision of this scheme strongly depends upon the probability of marking. This though is inversely proportional to the storage issues at the end system.

3.2 Advanced and Authenticated Marking Schemes [1]

This paper proposes three new schemes which are improvements on top of the Savage's [8] idea. The advanced marking schemes introduces the idea of using hash values of the router IP addresses to be stored in the IP identification field. The second scheme uses multiple hash functions to encode the edge. Authenticated Marking Scheme introduces authentication information about the routers either by introducing a shared key or by using the public key encryption used by each router. The paper also proposes using a dynamically changing key concept. The problem with this approach though is that the routers and endsystems need to be synchronized. Traffic overhead in Advanced Marking schemes is very low. All the traceback information is embedded in the data packet itself. Authentication Marking Scheme though requires a little bit of extra traffic for key information but that very low. The computational overhead of the routers is more than Savage's scheme. Some extra computational overhead comes in while calculating hash value and also the authentication (encryption/ decryption). Overall this overhead can be termed as low. The end system overhead of this scheme is medium. As shown by the simulations in [1], that this overhead is much less than the Savage's [8] scheme. The routers have almost negligible storage overhead in this case, though the overhead of the end system is high. This scheme has the same issues as that of Savage's

scheme. Additionally, in case of Authenticated Marking Scheme, the handling of Keys adds another factor that causes extra overhead. Overall we rate the Deployment Ease of this scheme to be medium. Just like [8], this scheme can perform post mortem analysis, as all the packets marked probabilistically are stored at the end system. This scheme offers more scalability as compared to Savage with the time taken to trace back 2000 attackers well below 100 seconds.

3.3 ICMP Traceback Scheme [3]

This scheme proposes a new ICMP message that is emitted by the router with a certain probability when forwarding a packet. The destination of this ICMP message is same as that of the original packet. This packet contains the information about the router that emitted it, the time of generation, some part of the original packet and authentication information. Traffic overhead of this scheme is medium. The ICMP packet is generated with a very low probability. Computational overhead on routers is low. The end system has a high computational overhead because it computes the traceback from the logged ICMP packets. Storage overhead of this scheme is very low on the routers. The end system storage overhead of this scheme is very high. The end system needs to store enough marked packets from all the routers, which are in the path from the attacking node to the destination. This scheme needs almost no intervention from the ISP and is incrementally deployable. However a new ICMP protocol update has to be done on each and every router, which makes the deployment difficult. Another issue is that a lot of ISPs and Network administrators block the ICMP messages because of security concerns. We would rate the deployment ease of this scheme as medium. Post mortem is easy for the victim because of the logged packets. The precision though is highly dependent on the packet authentication and ICMP packet generation probability. This scheme introduces the idea of using key encryption for router identification but this cause scalability issues.

3.4 Hash based IP Traceback [6]

This scheme proposes SPIE (Source Path Isolation Engine) architecture for traceback. It is composed of three parts: DGA (Data Generation Agent), SCAR (SPIE collection and reduction agent) and STM (SPIE traceback manager). Every router has a DGA agent attached which calculates the digest of each packet leaving the routers using bloom filters and stores it in digest tables. These tables are regularly stored on SCAR for permanent storage. STM receives the information about the packet to be traced back and uses this information from the SCAR and digest tables to generate the trace back within its

administrative domain. Traffic overhead for this scheme is minimal. End system computational overhead for this scheme is low. Most of the overhead in the network is handled by a separate STM and SCARs and in common case the overhead for normal routers involves calculating hash values of every passing packet using bloom filters and paging the digest table out on SCAR once it is filled. This overhead can be termed as medium. The storage overhead on the end systems is low. The storage overhead on routers is high both on its memory and attached SCAR agent. This scheme is built on top of the existing IP architecture and is incrementally deployable. The ISP intervention is required for installation but then the whole process can be automated to make it easily manageable. Overall the deployment ease of this scheme is medium. Post mortem is possible because the hashed value of the packets in the form of bloom filter is stored in the SCAR database. This scheme scales under one administrative domain. The precision of this scheme depends upon the hash and bloom filter to be used.

3.5 Control Flooding [9]

This is an interesting scheme proposed by Hal [9]. The basic idea of this scheme is to load all the links of the first hop routers one by one. When we see a disruption in the attacking packet stream, then the flooded link is added to the attack graph. Then we flood the links going out of the router at the other end of the previously added links one by one and continue until we trace down the originator for that specific packet stream. Traffic overhead of this scheme is very high. Router computational overhead is almost negligible. The end system overhead is also low. It is only involved in the creation of the attacking graph once the disruption in the stream has been identified. Storage overhead on both the end system and the routers is very low in this scheme. It is very difficult to find a favorable argument for deploying this kind of scheme, which can cause denial of service on all the networks through which it is being deployed. This scheme does not allow for the post mortem analysis. This scheme has high scalability issues. In case of multiple hosts attacking with the same spoofed source address this scheme will not be able to distinguish one attacking source from another. This scheme is very precise if only one source is spoofing the packets but in case of multiple sources, the precision of this scheme is degraded.

3.6 CenterTrack [4]

CenterTrack is an IP overlay network having virtual adjacency to all the edge routers and physical connection to the backbone. Hop by hop tracking is used to traceback in this system where all the traffic to the victim is passed through this overlay network

using tunnels at the edge routers. When tracing back a source, all of its traffic is to be routed through overlay network, which can be source of congestion on some links. This overhead is medium. The computational overhead on the end system is negligible. The computational overhead on a router is low (except the edge routers). The storage overhead on both the end system and routers is low. This scheme is compatible with the existing protocols and also has the potential of being incrementally deployed. However ISP intervention is required and this scheme can only locate the ingress point of the attack packets. Overall the deployment ease is medium. This scheme cannot be used for post mortem analysis because it does not keep any state about the passing traffic. The precision of this scheme is high in terms of finding the ingress or the origin of the attacking packets in the ISP network.

3.7 Deciduous (Decentralized Source Identification for Network-Based Intrusions) [5]

Deciduous uses the concept of security associations which are established dynamically through IPSEC/ISAKMP. The component of Deciduous framework associated with tracing the source of the packet is called ASIS (Attack source identification system). This scheme transforms the mesh topology to linear [5] and then establishes security association with the middle router and finding the segment responsible for the attacking traffic. This procedure takes place iteratively until the source is identified. Packet size increases when we use IPsec for a specific security association. So we can term the traffic overhead as medium. The computation overhead on the end system is negligible. The total computational overhead on the routers is in using IPsec in a Security Association, which can be termed as medium low. The storage overhead on both the end system and the routers is low. This scheme can be incrementally deployed. However the scheme will be implemented in one administrative domain. The Deployment Ease of this scheme is medium. This scheme does not support post mortem. It fails to address the issue of tracing back multiple attackers because of high control traffic for dynamically making security associations. However this scheme is very precise.

4. Conclusions

We have analyzed and evaluated seven different traceback schemes in this paper. These schemes can be categorized as following.

1. Probability Packet Marking: Marking the packet or sending its information to the destination with some defined probability [1,2,3]

2. Network Based Storage and Processing: Stores information on Routers [6]
3. Administrative Layout for Real time Analysis: Creating an additional administrative architecture on top of the current network [4,5].

Control Flooding does not fit in any of the categories and offers a unique idea which is difficult to scale because of extra ordinary high traffic load.

Following are some observations about distributed denial of service attacks.

1. The frequency of actually doing a traceback is very low as compared to the daily load traffic.
2. Tracing back the attacker while the attack is still going on is more important than the post mortem.
3. The ISPs will be interested in a traceback approach which does not affect the normal traffic behavior and which can be used in tracing the route or origin of the streaming packets in their own network or else the egress point in their network.
4. Because tracing back will be rare, a small overhead while the tracing is going on is acceptable.
5. The traceback scheme should not have any high computational or storage overhead on the routers.

Based on these five observations and my evaluation of the schemes, we conclude the following about the three trace back approaches.

The first approach either requires changes in the IP packet format or tries to incorporate path information in the already present fields. Introducing complexity in IP does not scale because of two major reasons. First, it requires universal deployment to be successful. Second, using IP fields for a purpose different than what they were designed for is not a good idea and it affects the portion of the traffic which does use these fields (e.g. fragmented traffic in our case). IP is the middle point in the hour glass approach and should be kept simple. The second approach (network based storage and processing scheme) has the problem of putting too much load work on routers in terms of storage and processing. This makes it very difficult to scale. We believe that the third approach is the right direction to move forward. With the internet divided into various Autonomous Systems, a practical approach is to let each administrative domain handle its own network for DDoS attack tracebacks. Having traceback administration separate from the normal routing network helps in keeping the flow of the normal traffic unaffected and is scalable. ISP can charge the customer for this service. A further improvement is possible where each ISP tracks down the intruder in its own domain using any of these administrative schemes and then we can develop a protocol for the edge routers to communicate the information to the neighbor Autonomous Systems. In the worst case we can limit the source to the ingress point of our

network. If the source is originating from within the network, it can be tracked down using the same administrative approach and appropriate measures can then be taken.

The schemes which are described as an example of third approach need improvements in a lot of different areas e.g. automation, ISP to ISP communication, security issues of nodes performing the administrative task, introduction of a hierarchical structure to the administrative framework and making the administration distributed.

References:

- [1]. Dawn X. Song and Adrian Perrig, "Advanced and Authenticated Marking Schemes for IP Traceback", Proceedings of InfoCom 2001
- [2]. Drew Dean, Matt Franklin, and Adam Stubblefield, "An algebraic approach to ip traceback," Proceedings of Network and Distributed System Security Symposium, NDSS '01, February 2001
- [3]. Steven M. Bellovin, Editor, "ICMP Traceback Messages", Internet Draft draftbellovin-itrace-00.txt, submitted Mar. 2000
- [4]. Robert Stone, "Centertrack: An IP overlay network for tracking DoS floods," Proceedings of 9th USENIX Security Symposium, Aug. 2000
- [5]. H.Y. Chang, R. Narayan, S.F. Wu, B.M. Vetter, M. Brown, J.J. Yuill, X. Wang, C. Sargor, F. Jou, F. Gong, "Deciduous: Decentralized Source Identification for Network-based Intrusions", 6th IFIP/IEEE International Symposium on Integrated Network Management, IEEE Communications Society Press, May 1999
- [6]. Alex C. Snoeren, Craig Partridge, Luis A. Sanchez, Christine E. Jones, Fabrice Tchakountio, Stephen T. Kent, and W. Timothy Strayer, "Hash-Based IP Traceback", Proceedings of the ACM SIGCOMM 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, August 2001
- [7]. John Ioannidis, Steven M. Bellovin, "Pushback: Router-Based Defense Against DDoS Attacks", In Network and Distributed System Security Symposium, NDSS'02, Feb 2002
- [8]. Stefan Savage, David Wetherall, Anna Karlin and Tom Anderson, "Practical Network Support for IP Traceback", Proceedings of the 2000 ACM SIGCOMM Conference, pp. 295-306, Stockholm, Sweden, August 2000.
- [9]. Hal Burch and Bill Cheswick. "Tracing Anonymous Packets to their Approximate Source", Usenix LISA, December 2000.
- [10] http://www.businessweek.com/2000/00_08/b3669001.htm

Bluetooth Interference Suppression in IEEE 802.11b

Erik Björnemo

Signals and Systems Group, Uppsala University¹
Box 528, SE-751 20, Uppsala, Sweden
eb@signal.uu.se

Abstract— We study the coexistence of IEEE 802.11b and Bluetooth, and more specifically the performance gains obtained using a frequency domain Bluetooth interference suppression scheme at the IEEE 802.11b physical layer. Focus is on the reduction of packet error rates in heavily interfered environments.

The frequency domain structure of the signals is utilised to separate multiple interferers, thereby simplifying the frequency domain suppression. We develop a minimum mean square optimal suppression method for batch-wise processing of baseband data, and a reduced complexity sub-optimal implementation. The latter turns out to be equivalent to a Wiener smoother based on estimated signal energies.

Simulations indicate that the Wiener based smoother drastically reduces the packet error rate in interfered transmissions, especially for the lower transmission rates. The gains are substantial for single as well as multiple interferers.

1 Introduction

As the license exempt use of the 2.4 GHz band increases, so does the need for communication systems to efficiently handle interference. Several theoretical and empirical studies concerning coexistence of IEEE 802.11b and Bluetooth have been published, showing that substantial performance losses may be caused in both systems (see for example [1] and [2]). In this paper we investigate the suppression of Bluetooth interference in an IEEE 802.11b receiver. We propose a method which tracks and suppresses the interfering Bluetooth signals at the IEEE 802.11b physical layer.

Previously proposed coexistence approaches include collaborative methods for interference avoidance by coordination of Bluetooth and IEEE 802.11b packet transmissions. While effective, these coordinated actions require collocated systems, or changes in the standards, to enable inter system communication. Also non-collaborative solutions aimed at interference avoidance have been investigated. Typically, they adapt packet scheduling – in time or frequency – based on measured channel quality [3]. An adaptive frequency hopping scheme of this kind was recently adopted in the Bluetooth specification [4].

A complement to interference avoidance is interference suppression at the physical layer, which is the approach we adopt in this work. In [5], a recursive least squares filter is used to remove Bluetooth signals from IEEE 802.11b baseband data. It is shown to perform well for the 1 Mbit/s rate with only one interferer, but performs worse for the 11 Mbit/s rate and fails to handle more than one interferer.

2 Overall Goal and Proposed Approach

We consider an IEEE 802.11b device in receive mode. The aim is a non-collaborative, physical layer method that relies

on received baseband data only, and does not impose any changes in the Bluetooth or IEEE 802.11b standards. The overall goal is to minimise the packet error rate. We pursue this goal indirectly by minimising the mean square error in the data.

The suppression algorithm handles multiple simultaneous interferers, and we use the following frequency domain structure of the signals to simplify this. First, compared to the bandwidth of the IEEE 802.11b receiver, roughly 22 MHz, Bluetooth signals are narrow-band and cover approximately 1 MHz. Second, the Bluetooth transmitters hop between 79 known carrier frequencies, spaced 1 MHz apart. We therefore apply the fast Fourier transform (FFT) as an efficient method for separation of simultaneous interferers present within the receiver bandwidth. The FFT approach implies batch-wise processing of data, which in turn introduces a delay. This delay, and the tracking of hopping interferers, require the batches to be much shorter than the IEEE 802.11b packet length.

In summary, we batch-wise minimise the mean square error in discrete Fourier transformed baseband data, with the goal of minimising the packet error rate.

3 Signal Models and Parameter Uncertainties

In this section we present our models of the signals, along with all relevant prior information I about their parameters. Based on these models and the parameter uncertainties, we assign the probability distributions relevant to the suppression scheme. We will throughout this paper use probability theory as extended logic, thus using it to do plausible reasoning based on our state of knowledge. The product and sum rules of probability theory constitute the only internally consistent way of reasoning under uncertainty, and they process our incomplete information with full efficiency. All prior probabilities are assigned using transformation groups and the principle of maximum entropy (MaxEnt) [6].

In our model of the received baseband data r_n , we include the IEEE 802.11b signal s_n , thermal noise v_n , and Bluetooth interferers $x_{m,n}$, where n denotes the sample number and m the channel number. Sampling is at 802.11b symbol rate. We model the frequency domain data as

$$R_k = S_k + \sum_{m=1}^M X_{mk} + V_k, \quad (1)$$

where $k=0, 1, \dots, N-1$ is the frequency bin number, N is the batch size and M is the number of non-overlapping Bluetooth channels within the receiver bandwidth. We use the subscript mk when bin k is within channel m . All

¹This work was partly funded by the graduate school PCC++.

propagation channels are modelled as slowly flat fading, i.e. as a constant attenuation over each batch.

3.1 IEEE 802.11b

These systems use differential BPSK or QPSK combined with a spreading sequence or complementary codes [7]. Assuming perfect symbol synchronisation and no inter symbol interference, we use the received transformed signal model

$$S_k = \frac{\sqrt{E_s}}{\sqrt{N}} e^{j\theta_s} \sum_{n=0}^{N-1} C_n e^{-j2\pi \frac{kn}{N}} \quad (2)$$

where $\sqrt{E_s} e^{j\theta_s} C_n = s_n$. Here, E_s is the received energy per symbol, θ_s is a phase rotation, and C_n are the transmitted complex symbols.

E_s At the start of the reception, the energy is unknown but bounded. From above by regulations and receiver efficiency, and from below by the sensitivity of the receiver. Since E_s is a scale parameter, this uncertainty prescribes $p(E_s|I) \propto 1/E_s$ over $[E_{s-}, E_{s+}]$.

θ_s The differential demodulator does not track absolute phase, and therefore θ_s is unknown. We assign a uniform distribution $p(\theta_s|I)$ over $[0, 2\pi]$.

C_n The received symbols are from the alphabet $\{\exp(j(1/2 + d)\pi/2)\}$, where d is the encoded unknown transmitted information (BPSK: $d = 0, 2$ QPSK: $d = 0, 1, 2, 3$). All symbols are independently equally probable.

We conclude from (2) that, given E_s , all S_k are uncorrelated and have probability distributions $p(S_k|E_s I)$ with zero mean and variance E_s . By virtue of the central limit theorem these distributions are Gaussian for large N , and independent.

3.2 Bluetooth

A Bluetooth signal is well described as a continuously phase modulated signal [4], and its transform is

$$X_{mk} = \frac{\sqrt{E_m}}{\sqrt{N}} e^{j\theta_m} \sum_{n=0}^{N-1} e^{j\phi(nT_s - t_0, \mathbf{Y}) - j2\pi(f_m T_s + \frac{k}{N})n}. \quad (3)$$

Here, E_m is the received energy per receiver sample, $\phi(t, \mathbf{Y})$ is the continuous phase modulation of the binary information sequence \mathbf{Y} (modulation index h , GMSK pulse of bandwidth 0.5 MHz), f_m is the frequency offset between the receiver's and interferer's carriers, t_0 is the timing offset, θ_m is the phase difference, and T_s is the receiver sampling time:

E_m Bounded in the same manner as E_s , provided an interferer actually is present. Assign $p(E_m|B_m I) \propto 1/E_m$ over $[E_{xmin}, E_{xmax}]$, where B_m denotes interferer presence in channel m .

θ_m Unknown: $p(\theta_m|I)$ is uniform over $[0, 2\pi]$.

t_0 Unknown, but maximally one Bluetooth bit period: $p(t_0|I)$ is uniform over $[0, 1] \mu s$.

\mathbf{Y} Unknown information sequence: ± 1 equally probable.

f_m The allowed deviation according to the Bluetooth specification is roughly ± 100 kHz. We assign a uniform distribution $p(f_m|I)$ over this range.

h Required to be within $[0.28, 0.35]$, over which $p(h|I)$ is uniform.

We treat the parameters E_m and θ_m analytically, while resorting to numeric calculations for the others. We do this by computing the variance σ_{mk}^2 for $p(X_{mk}|E_m \theta_m I)$ using Monte Carlo simulation of (3). The uniform $p(\theta|I)$ assures further that $p(X_{mk}|E_m I)$ has zero mean. Now, by use of only the mean and variance as prior information, this latter probability distribution is Gaussian according to the maximum entropy principle. The numerical procedure is an approximate marginalisation over the nuisance parameters t_0, \mathbf{Y}, f_m and h . In this way we include our initial uncertainty about their values in $p(X_{mk}|E_m I)$, but we have also permanently removed them from subsequent calculations. This simplification of our model in (3) will degrade the suppression performance, but we believe that it is a very small degradation, especially compared to the reduction in complexity.

3.3 Thermal Noise

The noise power is assumed to be known with good accuracy. We assign a Gaussian zero mean probability distribution $p(V_k|I)$ with variance E_v corresponding to the noise power.

3.4 Interferer Presence

Let B stand for Bluetooth presence within the receiver's range, and let \bar{B} stand for an interference free situation. Denote the presence of exactly i interferers B^i . We make a conservative probability assignment $P(B|I) = P(\bar{B}|I) = 0.5$ based on the fact that we do not know in which type of environment – heavily disturbed or undisturbed – the IEEE 802.11b receiver is operating. Furthermore, Bluetooth networks operate independently of each other, so we assign independent probabilities $P(B^i|I) = P(B^1|I)^i$ for $i \geq 1$. In summary, we get

$$P(B^i|I) = \begin{cases} \frac{1}{2} & , i = 0 \\ (\frac{1}{3})^i & , i > 0 \end{cases} \quad (4)$$

There are $L = 79$ Bluetooth channels, which leads to $P(B_m|B^i I) = 1 - (78/79)^i$. Marginalisation over B^i gives

$$P(B_m|I) = \frac{P(B^1|I)}{1 - P(B^1|I)} - \frac{(L-1)P(B^1|I)}{L - (L-1)P(B^1|I)}, \quad (5)$$

which numerically is $P(B_m|I) = 1/106$.

4 Optimal Interference Suppression Scheme

Our goal is to make a minimum mean square estimate of $\{s_n\}$ for each batch $\{r_n\}$ in order to reduce the packet error rate. This estimate is given by the mean value of the probability distribution $p(\{s_n\}|DI)$, where D denotes the batch of data. An equivalent procedure is to find the mean of $p(\{S_k\}|DI)$ for the frequency domain data. For brevity, let E denote $E_s, \{E_m\}$ and E_v collectively. Then

$$p(\{S_k\}|DI) = \int p(\{S_k\}|EDI) p(E|DI) dE \quad (6)$$

in which

$$p(\{S_k\}|EDI) = \prod_k p(S_k|EDI). \quad (7)$$

Proceeding with the individual distributions for S_k and Bayes' theorem, we get

$$\begin{aligned} p(S_k|EDI) &= p(S_k|R_kEI) \\ &= \frac{p(R_k|S_kEI)p(S_k|EI)}{p(R_k|EI)}. \end{aligned} \quad (8)$$

All distributions on the right hand side in (8) are Gaussian: $p(R_k|S_kEI)$ has mean S_k and variance $E_m\sigma_{mk}^2 + E_v$; $p(S_k|EI)$ has zero mean and variance E_s ; $p(R_k|EI)$ has zero mean and variance $E_s + E_m\sigma_{mk}^2 + E_v$. It can be shown that

$$\langle S_k|EDI \rangle = \frac{E_s}{E_s + E_m\sigma_{mk}^2 + E_v} R_k. \quad (9)$$

Insertion of (9) in (6) leads to the minimum mean square estimate

$$\langle S_k|DI \rangle = R_k \int \frac{E_s}{E_s + E_m\sigma_{mk}^2 + E_v} p(E|DI) dE. \quad (10)$$

This is a marginalisation over uncertain energies in a Wiener smoother – weighting the smoother for each set of energies with their respective probability density gives the optimal result.

The evaluation of $p(E|DI)$ in (10) does not lead to a closed-form solution. We work out everything needed for the implementation of a suppression scheme, but leave some details out due to space limitations. Let D_m denote data from channel m , and let E_{sv} denote E_s and E_v collectively. The posterior probability for E is

$$\begin{aligned} p(E|DI) &= p(\{E_m\} | E_{sv}DI) p(E_{sv}|DI) \\ &= p(E_{sv}|DI) \prod_m p(E_m | E_{sv}D_mI). \end{aligned} \quad (11)$$

We proceed by studying $p(E_m | E_{sv}D_mI)$ and $p(E_{sv}|DI)$ separately. The posterior probability distribution for E_m depends on the probability for B_m :

$$\begin{aligned} p(E_m | E_{sv}D_mI) &= p(E_m | B_m E_{sv}D_mI) P(B_m | E_{sv}D_mI) \\ &\quad + \delta(E_m) P(\bar{B}_m | E_{sv}D_mI) \end{aligned} \quad (12)$$

where

$$p(E_m | B_m E_{sv}D_mI) \propto p(D_m | B_m E_m E_{sv}I) p(E_m | B_m I) \quad (13)$$

and

$$\frac{P(B_m | E_{sv}D_mI)}{P(\bar{B}_m | E_{sv}D_mI)} = \frac{P(B_m|I) \int p(D_m | B_m E_m E_{sv}I) p(E_m | B_m I) dE_m}{P(\bar{B}_m|I) p(D_m | \bar{B}_m E_m E_{sv}I) p(E_m | \bar{B}_m I)}. \quad (14)$$

The posterior for E_{sv} depends on the probability for B^i as

$$\begin{aligned} p(E_{sv}|DI) &\propto p(D|E_{sv}I) p(E_{sv}|I) \\ &= p(E_{sv}|I) \sum_i p(D|B^i E_{sv}I) P(B^i|I). \end{aligned} \quad (15)$$

All the terms in the sum in (15) are straightforward to find by application of the product and sum rules of probability theory, but are not displayed here. Except for this last part, we have now broken up $p(E|DI)$ into already quantified probabilities and probability distributions. Inserting them in (10), we have the mean square optimal solution. Unfortunately, there is no analytically tractable solution. The optimal scheme is however the basis for the approximate

solution we develop in the next section, and could be implemented numerically provided enough computing power is available.

5 Reduced Complexity Implementation

We consider the complexity of a direct numerical implementation of the optimal scheme prohibitive. This is mainly due to the marginalisation over E and B^i in (10) and (15). In this section we develop a reduced complexity scheme based on the optimal scheme. We go backwards from (15) to (9) in the following steps:

1) Signal energy estimation assuming no interferers

Using only the first term in the sum in (15), we make an initial estimate $\hat{E}_{sv} = \langle E_{sv} | B^0 E_v DI \rangle$ from

$$\begin{aligned} p(E_{sv} | B^0 DI) &\propto p(E_{sv} | I) p(D | B^0 E_{sv} DI) \\ &= p(E_{sv} | I) \prod_k p(R_k | B^0 E_{sv} I) \end{aligned} \quad (16)$$

where the distributions $p(R_k | B^0 E_{sv} I)$ are Gaussian with variance $E_s + E_v$, while $p(E_{sv} | I) \propto (E_s + E_v)^{-1}$. This leads to

$$p(E_{sv} | B^0 DI) \propto \frac{1}{(E_s + E_v)^{N+1}} \exp\left(-\frac{\sum |R_k|^2}{E_s + E_v}\right). \quad (17)$$

This distribution is zero outside $[E_v + E_{s-}, E_v + E_{s+}]$, and its mean value is

$$\hat{E}_{sv} = \frac{\Gamma(N-1, b) - \Gamma(N-1, a)}{\Gamma(N, b) - \Gamma(N, a)} \sum_{k=0}^{N-1} |R_k|^2. \quad (18)$$

Here $a = \sum |R_k|^2 / (E_v + E_{s-})$ and $b = \sum |R_k|^2 / (E_v + E_{s+})$. Within the range of possible signal energies, the first factor in (18) is very close to $N-1$. This suggests

$$\hat{E}_{sv} = \frac{1}{N-1} \sum |R_k|^2 \quad (19)$$

as a reasonable approximation. Note that this estimate is based on the truth of B^0 , so we are likely to be in large error if B^0 turns out to be false. This is the price we pay for reduced complexity.

2) Approximate marginalisation over all E_m

We perform the marginalisation involved in (14), for all m , by approximating the integral with a sum over discrete E_m , and replacing E_{sv} with \hat{E}_{sv} . Here we have a clear tradeoff between precision and complexity.

3) Channel classification and energy re-estimation

The posterior odds in (14) are computed and we introduce a threshold to classify channels as clear or interfered. We use this classification to re-estimate E_s without the contaminated data. For the first batch in each packet we also redo the marginalisation in step 2).

4) Interference energy estimation

By application of (12), (13), and the results from steps 2) and 3), we compute minimum mean square estimates $\hat{E}_m = \langle E_m | \hat{E}_{sv} D_m I \rangle$ for all channels classified as interfered.

5) Wiener smoothing with estimated signal energies

Finally, we use a Wiener smoother based on \hat{E}_m , \hat{E}_s and E_v to estimate S_k . The integral in (10) collapses to

$$\hat{S}_k = \frac{\hat{E}_s}{\hat{E}_s + \hat{E}_m \sigma_{mk}^2 + E_v} R_k, \quad (20)$$

which basically is the smoother in (9). Go back to step 2) for the next batch in the packet.

For further reduction in computational complexity, we can use a subset of the data from each channel. For example, when estimating E_m we may use data from the frequency bins where most of the interferer energy is concentrated.

6 Suppression Performance

Simulations indicate that large performance improvements are obtainable with the proposed suppression scheme. In Figure 1 and Figure 2 we present results for the IEEE 802.11b rates 2 Mbit/s and 11 Mbit/s. Note that the packet error rates shown pertain to interfered transmissions. The actual rates are therefore not representative for a real situation, where not all packet transmissions are interfered.

The simulation code is available at www.signal.uu.se.

6.1 Simulation Parameters

The results reflect a constant flat fading environment and a signal-to-noise ratio of 15 dB. Batches are 256 samples ($\approx 23 \mu\text{s}$).

A) *IEEE 802.11b*: Only the physical layer demodulation and symbol decoding is simulated, using perfect symbol synchronisation. The packets are 1500 bytes for 11 Mbit/s, 1024 bytes for 2 Mbit/s, containing pseudo-random data.

B) *Bluetooth*: The simulations include one, two or four interfering Bluetooth networks, hopping at 1600 Hz. In order to provide a severe test, all interferers are restricted to hop only within the receiver bandwidth. The given carrier-to-interference (CIR) ratios are with respect to the total interference power, which is distributed unevenly between the interferers. As suggested by [5], the transmit power is ramped up/down over $2 \mu\text{s}$.

C) *Reduced Complexity Implementation*: The interferer energy E_m is quantised in steps of 1 dB. Out of the 46 bins per Bluetooth channel, only 10 are used in steps 5)-5) in the algorithm.

7 Conclusion

We have shown that a frequency domain approach using a Wiener smoother based on estimated signal energies is very effective in reducing Bluetooth induced packet errors. The gains are especially evident for the 2 Mbit/s rate.

The results, obtained with the reduced complexity approach without any tuning of the parameters, indicate that *i*) we can do even better by implementing of the full solution *ii*) we can achieve significant performance gains also when using more heavily reduced complexity. It should be kept in mind, however, that the results pertain to a flat fading environment and perfect synchronisation. The effects of interference on the synchronisation have not been taken into account, neither have the effects of fast fading.

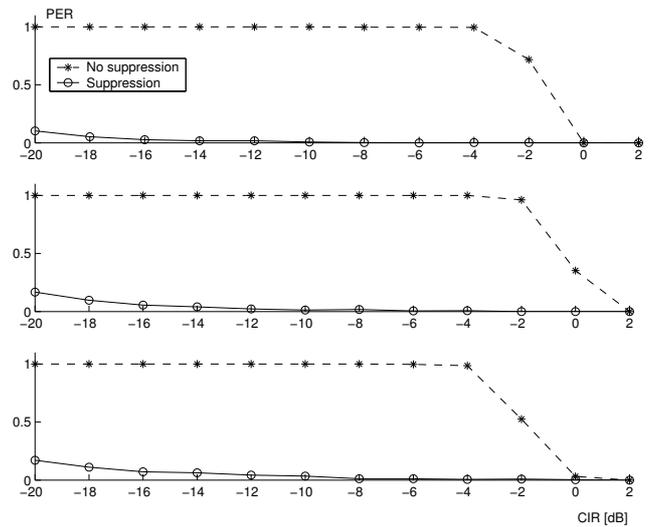


Fig. 1. Packet error rates (PER) for 2 Mbit/s rate. One interferer (top), two interferers (middle) and four interferers (bottom). Solid lines represent results with suppression, dashed lines without suppression.

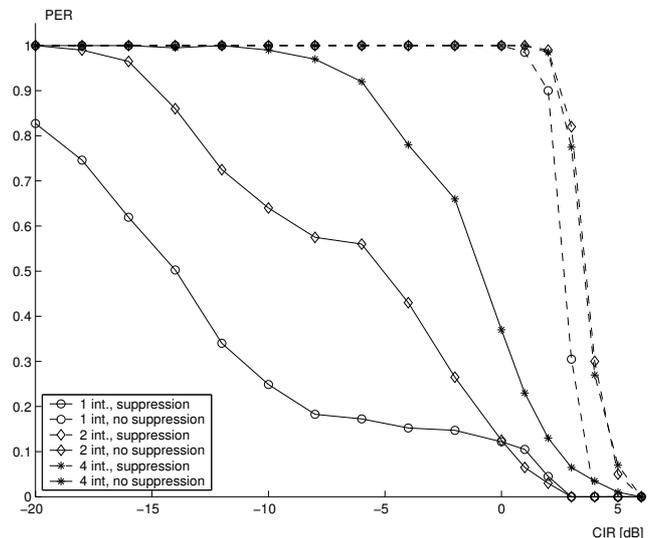


Fig. 2. Packet error rates (PER) for 11 Mbit/s rate. One interferer (circle), two interferers (diamond) and four interferers (star). Solid lines represent results with suppression, dashed lines without suppression.

References

- [1] J. Lansford, A. Stephens and R. Nevo, "Wi-Fi (802.11b) and Bluetooth: Enabling Coexistence", *IEEE Network Mag.*, vol. 15, pp. 20-27, September-October 2001
- [2] M. Fainberg and D. Goodman, "Analysis of the Interference Between IEEE 802.11b and Bluetooth Systems" *Proc. IEEE Vehicular Technology Conf.*, vol 2, pp. 967-971, October 2001
- [3] C.F. Chiasserini and R.R. Rao, "Coexistence Mechanisms for Interference Mitigation in the 2.4-GHz ISM Band" *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 964-975, September 2003
- [4] Bluetooth Special Interest Group, *Specification of the Bluetooth System, Vol. 2, ver.1.2*, November 2003.
- [5] A. Soltanian, R.E. van Dyck and O. Rejala, "Rejection of Bluetooth interference in 802.11 WLANs", *Proc. IEEE Vehicular Technology Conf.*, vol 2, pp. 932-936, September 2002
- [6] E. T. Jaynes, *Probability Theory – The Logic of Science*, Cambridge University Press, April 2003.
- [7] IEEE Std 802.11, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 1999 Edition

Congestion Control of AQM Routers by An Improved PID Controller

Yongji Wang¹, Xiuli Wang¹, Jinhui Zhou^{2,1} and Xiaoxi Wang¹

¹ Laboratory for Internet Software Technologies, Institute of Software, Chinese Academy of Sciences, Beijing 100080, P. R. China

² Xidian University, Xi'an 710071, P. R. China

{ywang, wxl, jinhui, xiaoxi}@itechs.iscas.ac.cn

Abstract- This paper applies a novel PID and PID-like controller design method to AQM control that permits the designer to control the desired dynamic performance of a closed-loop system. A set of desired D -stable regions in the complex plane is first specified and then a numerical optimization algorithm is run to find the controller parameters such that all the roots of the closed-loop system are within the specified regions. The smaller queuing delay and higher throughput are achieved. A comparative study has been carried out and the experimental simulation results show that this algorithm obviously outperforms the RED and PI algorithms.

Index terms- congestion control, active queue management, PID and PID-like controller, D -stability.

I. Introduction

Active Queue Management (AQM) is a very active research area in networking. RED accords with the objective of the AQM, but the subsequent investigations have revealed its two limitations: the tradeoff between the speed of response and stability, and the direct coupling between queue length and loss probability [1]. Thus researchers have developed many AQM algorithms, such as ARED [2], SRED [3], BLUE [4], REM [5], PI [1], AVQ [6], but there are still limitations of each algorithm. The network congestion problem has not been resolved up to now.

The feed back control strategy is crucial for congestion control. [7] has come up with a novel PID and PID-like controller design method. In this paper we apply the method to design a controller that is better suited for AQM than RED and PI. We present guidelines to the stable linear controller. The PID controller is shown via simulations to be a robust controller that outperforms the RED and PI controllers under almost all scenarios considered.

II. TCP/AQM Model

A non-linear dynamic model for TCP/AQM is developed in [8], which can be depicted in (1), while the linearized model is shown in Figure 1, see [9] for more details of linearization.

$$\begin{cases} \frac{dW(t)}{dt} = \frac{1}{R(t)} - \frac{W(t)W(t-R(t))}{2R(t)} p(t-R(t)) \\ \frac{dq(t)}{dt} = \frac{N(t)}{R(t)} W(t) - C(t) \end{cases} \quad (1)$$

where W , R , C , N , q and p denote the TCP window size, the round trip time, the link capacity, the load factor, the queue length of the router and the packet discard ratio.

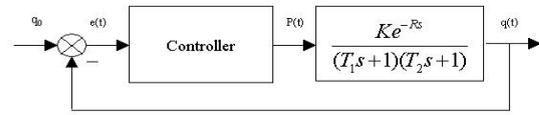


Figure 1. Block diagram of a linearized AQM control system

where $K = \frac{(RC)^3}{4N^2}$, $T_1 = R$, $T_2 = \frac{R^2 C}{2N}$.

III. New PID Controller Design Method

Methods developed for a PID or PID-like controller design can be classified in numerous ways according to the plant models. A novel PID and PID-like controller design method by pole assignment within D -stable regions has been presented [7]. This method can be used for stable and unstable plants with high order degree, for plants with time delay, for controller with more than three design parameters, and for various controller configurations. It also allows a unified treatment of the controller design for both continuous and discrete systems. See [10,11,12] for more details.

It is assumed that the plant has the following general transfer function with time delay

$$G_p(s) = \frac{N_p(s)}{D_p(s)} = \frac{a_0 + a_1 s + \dots + a_{m-1} s^{m-1} + a_m s^m}{b_0 + b_1 s + \dots + b_{n-1} s^{n-1} + s^n} e^{-ts} \quad (2)$$

where $nb \geq na$, the coefficients a_i ($i=0,1,\dots,na$) and b_j ($j=0,1,\dots,nb-1$) are all real, and t is the pure time delay.

The delay may be approximated by a first order Pade approximation

$$e^{-ts} = \frac{1 - \frac{t}{2}s}{1 + \frac{t}{2}s} \approx \frac{1 - Ts}{1 + Ts} \quad (3)$$

where $T=0.5t$. Substituting (3) into (2) yields

This work is supported by the One-Hundred-People Program of Chinese Academy of Sciences under Grant No. BCK35873 and by the National Natural Science Foundation of China under Grant No. 60373053.

Authors gratefully acknowledge the support of K. C. Wong Education Foundation, Hong Kong.

$$G_p(s) = \frac{c_0 + c_1s + \dots + c_m s^m + c_{na+1} s^{na+1}}{d_0 + d_1s + \dots + d_n s^n + d_{nb+1} s^{nb+1}} = \frac{N_p(s)}{D_p(s)} \quad (4)$$

where $c_0=a_0$, $c_i=a_i-Ta_{i-1}$, ($i=1,2,\dots,na$), $c_{na+1}=-Ta_{na}$, $d_0=b_0$, $d_j=b_j+Tb_{j-1}$, ($j=1,2,\dots,nb-1$), $d_{nb}=1+Tb_{nb-1}$, $d_{nb+1}=T$.

Note that if there is a large time delay, a higher order Pade approximation may be used to increase the approximation accuracy. The resulting plant model is still a rational transfer function.

Let $G_c(s)$ denote the PID controller

$$G_c(s) = \frac{N_c(s)}{D_c(s)} = \frac{K_d s^2 + K_p s + K_i}{s} \quad (5)$$

where K_p is the proportional gain, K_i is the integral gain and K_d is the differential gain. Then the transfer function of the closed-loop system is given as

$$G(s) = \frac{N(s)}{D(s)} = \frac{G_c(s)G_p(s)}{1+G_c(s)G_p(s)} \quad (6)$$

and its characteristic polynomial is obtained as

$$D(s) = N_c(s)N_p(s) + D_c(s)D_p(s) = [c_0K_i + (c_0K_p + c_1K_i)s + (c_0K_d + c_1K_p + c_2K_i)s^2 + \dots + c_{na+1}K_d s^{na+3}] + [d_0s + d_1s^2 + \dots + d_{nb+1}s^{nb+2}] = c_0K_i + (c_0K_p + c_1K_i + d_0)s + \dots + (c_{na+1}K_d + d_{nb+1})s^{nb+2} \quad (7)$$

A n -th order characteristic polynomial $D(s, \mathbf{u})$ for a closed-loop system is given as

$$D(s, \mathbf{u}) = a_0(\mathbf{u}) + a_1(\mathbf{u})s + a_2(\mathbf{u})s^2 + \dots + a_n(\mathbf{u})s^n \quad (8)$$

where the unknown vector $\mathbf{u} = [u_1, u_2, \dots, u_m]^T$ representing the m design parameters in the controller, develop an approach to find \mathbf{u} such that the n characteristic roots of the characteristic polynomial (8) are located within a set of specified D -stability regions or on their boundaries.

An n -th order polynomial $D(s)$ with real coefficients has n complex characteristic roots (a real root is a special case). Let

$$s_i = \mathbf{s}_i + j\mathbf{w}_i, \quad i=1,2,\dots,n \quad (9)$$

denote its n roots. To simplify notation in the sequel, let $\mathbf{S}_i = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_i]$, $\mathbf{W}_i = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_i]$, $i=1,2,\dots,n$. Therefore, $D(s)$ can be written as

$$D(s) = \prod_{i=1}^n (s - s_i) = [\mathbf{a}_0(\mathbf{S}_n, \mathbf{W}_n) + j\mathbf{b}_0(\mathbf{S}_n, \mathbf{W}_n)] + [\mathbf{a}_1(\mathbf{S}_n, \mathbf{W}_n) + j\mathbf{b}_1(\mathbf{S}_n, \mathbf{W}_n)]s + \dots + [\mathbf{a}_{n-1}(\mathbf{S}_n, \mathbf{W}_n) + j\mathbf{b}_{n-1}(\mathbf{S}_n, \mathbf{W}_n)]s^{n-1} + s^n \quad (10)$$

where $\mathbf{a}_i(\mathbf{S}_n, \mathbf{W}_n)$ and $\mathbf{b}_i(\mathbf{S}_n, \mathbf{W}_n)$, $i=0,1,2,\dots,n-1$, which are all functions of $\mathbf{S}_n = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n]$, $\mathbf{W}_n = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n]$.

Leads to the following recursive formula:

For $n=1$,

$$\mathbf{a}_0(\mathbf{S}_1, \mathbf{W}_1) = -\mathbf{s}_1, \quad \mathbf{b}_0(\mathbf{S}_1, \mathbf{W}_1) = -\mathbf{w}_1, \quad (11)$$

for $n>1$,

$$\begin{aligned} \mathbf{a}_0(\mathbf{S}_n, \mathbf{W}_n) &= -\mathbf{a}_0(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{s}_n + \mathbf{b}_0(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{w}_n \\ \mathbf{b}_0(\mathbf{S}_n, \mathbf{W}_n) &= -\mathbf{b}_0(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{s}_n - \mathbf{a}_0(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{w}_n \\ \mathbf{a}_j(\mathbf{S}_n, \mathbf{W}_n) &= \mathbf{a}_{j-1}(\mathbf{S}_{n-1}, \mathbf{W}_{n-1}) - \mathbf{a}_j(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{s}_n \\ &+ \mathbf{b}_j(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{w}_n, \quad j=1,\dots,n-2 \\ \mathbf{b}_j(\mathbf{S}_n, \mathbf{W}_n) &= \mathbf{b}_{j-1}(\mathbf{S}_{n-1}, \mathbf{W}_{n-1}) - \mathbf{b}_j(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{s}_n \\ &- \mathbf{a}_j(\mathbf{S}_{n-1}, \mathbf{W}_{n-1})\mathbf{w}_n, \quad j=1,\dots,n-2 \\ \mathbf{a}_{n-1}(\mathbf{S}_n, \mathbf{W}_n) &= \mathbf{a}_{n-2}(\mathbf{S}_{n-1}, \mathbf{W}_{n-1}) - \mathbf{s}_n \\ \mathbf{b}_{n-1}(\mathbf{S}_n, \mathbf{W}_n) &= \mathbf{b}_{n-2}(\mathbf{S}_{n-1}, \mathbf{W}_{n-1}) - \mathbf{w}_n \end{aligned} \quad (12)$$

Recall that the characteristic polynomial (8) has the same roots as (10), we thus have

$$\begin{aligned} a_n(\mathbf{u})\mathbf{a}_j(\mathbf{S}_n, \mathbf{W}_n) &= a_j(\mathbf{u}), \quad \mathbf{b}_j(\mathbf{S}_n, \mathbf{W}_n) = 0, \\ j &= 0, 1, \dots, n-1 \end{aligned} \quad (13)$$

Suppose that a general desired D -stability region is given in the form of $g(\mathbf{s}, \mathbf{w}) \leq 0$, thus

$$g(\mathbf{s}_i, \mathbf{w}_i) \leq 0, \quad i=1,2,\dots,n \quad (14)$$

Note that inequalities (14) contain all the unknown roots $(\mathbf{S}_n, \mathbf{W}_n)$. Let us introduce a vector design variable $\mathbf{x} = [\mathbf{u}, \mathbf{S}_n, \mathbf{W}_n]$. From (13) and (14) we can observe that if a design vector \mathbf{x} can be found such that the equalities (13) and the inequalities (14) are simultaneously satisfied, then a controller satisfying the D stability requirement is found. Thus, it becomes a numerical optimization problem to find the controller parameters. The theory concerning numerical optimization is presented in [13,14].

IV. The PID Controller

From $G_p(s) = \frac{Ke^{-Rs}}{(T_1s+1)(T_2s+1)}$, (3) and (5), the

characteristic polynomial of the closed-loop system can be derived as

$$D(s, \mathbf{u}) = T_1T_2R's^4 + (T_1T_2 + T_1R' + T_2R' - KK_dR')s^3 + (T_1 + T_2 + R' + KK_d - KK_pR')s^2 + (1 + KK_p - KK_iR')s + KK_i \quad (15)$$

where $R'=R/2$. Polynomial (15) has four roots: $s_i = \mathbf{s}_i + j\mathbf{w}_i$, $i=1,2,3,4$. According to (12) and (13), \mathbf{s}_i , \mathbf{w}_i and K_p, K_i must satisfy equality constraints

$$\begin{aligned} T_1T_2R'(\mathbf{s}_1\mathbf{s}_2\mathbf{s}_3\mathbf{s}_4 + \mathbf{w}_1\mathbf{w}_2\mathbf{w}_3\mathbf{w}_4 - \mathbf{s}_1\mathbf{s}_2\mathbf{w}_3\mathbf{w}_4 - \mathbf{s}_1\mathbf{s}_3\mathbf{w}_2\mathbf{w}_4 - \mathbf{s}_1\mathbf{s}_4\mathbf{w}_2\mathbf{w}_3 - \mathbf{s}_2\mathbf{s}_3\mathbf{w}_1\mathbf{w}_4 - \mathbf{s}_2\mathbf{s}_4\mathbf{w}_1\mathbf{w}_3 - \mathbf{s}_3\mathbf{s}_4\mathbf{w}_1\mathbf{w}_2) &= KK_i \\ T_1T_2R'(\mathbf{s}_1\mathbf{s}_2\mathbf{s}_3 + \mathbf{s}_1\mathbf{s}_2\mathbf{s}_4 + \mathbf{s}_1\mathbf{s}_3\mathbf{s}_4 + \mathbf{s}_2\mathbf{s}_3\mathbf{s}_4 - \mathbf{s}_3\mathbf{w}_1\mathbf{w}_2 - \mathbf{s}_4\mathbf{w}_1\mathbf{w}_2 - \mathbf{s}_4\mathbf{w}_1\mathbf{w}_3 - \mathbf{s}_4\mathbf{w}_2\mathbf{w}_3 - \mathbf{s}_1\mathbf{w}_2\mathbf{w}_3 - \mathbf{s}_1\mathbf{w}_2\mathbf{w}_4 - \mathbf{s}_1\mathbf{w}_3\mathbf{w}_4 - \mathbf{s}_2\mathbf{w}_3\mathbf{w}_4 - \mathbf{s}_2\mathbf{w}_1\mathbf{w}_3 - \mathbf{s}_2\mathbf{w}_1\mathbf{w}_4 - \mathbf{s}_3\mathbf{w}_1\mathbf{w}_4 - \mathbf{s}_3\mathbf{w}_2\mathbf{w}_4) &= KK_iR' - KK_p - 1 \end{aligned}$$

$$\begin{aligned}
& T_1 T_2 R' (S_1 S_2 + S_1 S_3 + S_1 S_4 + S_2 S_3 + S_2 S_4 + S_3 S_4 - \\
& W_1 W_2 - W_1 W_3 - W_1 W_4 - W_2 W_3 - W_2 W_4 - W_3 W_4) \\
& = T_1 + T_2 + R + K K_d - K K_p R' \\
& T_1 T_2 R' (S_1 + S_2 + S_3 + S_4) = K K_d R' - T_1 T_2 - T_1 R' - T_2 R' \\
& S_1 S_2 S_3 W_4 + S_1 S_2 S_4 W_3 + S_1 S_3 S_4 W_2 - S_1 W_2 W_3 W_4 \\
& + S_2 S_3 S_4 W_1 - S_2 W_1 W_3 W_4 - S_3 W_1 W_2 W_4 - S_4 W_1 W_2 W_3 = 0 \\
& S_1 S_2 W_3 + S_1 S_2 W_4 + S_1 S_3 W_4 + S_2 S_3 W_4 - W_1 W_2 W_3 \\
& - W_1 W_2 W_4 - W_1 W_3 W_4 - W_2 W_3 W_4 + S_1 S_3 W_2 + S_1 S_4 W_2 + S_1 S_4 W_3 \\
& + S_2 S_4 W_3 + S_2 S_3 W_1 + S_2 S_4 W_1 + S_3 S_4 W_1 + S_3 S_4 W_2 = 0 \\
& S_1 W_2 + S_2 W_1 + S_1 W_3 + S_3 W_1 + S_1 W_4 + S_4 W_1 \\
& + S_2 W_4 + S_4 W_2 + S_2 W_3 + S_3 W_2 + S_3 W_4 + S_4 W_3 = 0 \\
& W_1 + W_2 + W_3 + W_4 = 0 \tag{16}
\end{aligned}$$

Now we specify the D -stability regions is described by a parabola of the following form

$$D \cup \mathbb{D}: g(s, w) = 4s + w^2 + e \leq 0 \tag{17}$$

The physical meaning for the change of e is that when e is bigger, the boundary curve for (17) will move leftwards, and thus the allowed region for the roots is smaller (see Figure 2).

All the roots of within the $D \cup \mathbb{D}$ described by (17) must satisfy

$$4s_i + w_i^2 + e \leq 0 \quad i=1,2,3,4 \tag{18}$$

The new design variables $x = [K_p, K_i, K_d, s_1, s_2, s_3, s_4, w_1, w_2, w_3, w_4]^T$ are then produced in (16) and (18).

Suppose, C is 15Mb/s, N is 60, R is 400ms. From the numerical optimization algorithm with constrains, for different e , the results for K_p, K_i, K_d are given in Table 1 and the D regions and the corresponding characteristic roots are shown in Figure 2.

e	0	1	2	3	4	5	6	6.5	7
$K_p(10^{-5})$	2.4999	2.5836	2.8843	3.1044	3.1059	2.9459	2.8841	2.8708	na
$K_i(10^{-6})$	4.4999	5.7025	10.052	13.341	14.417	13.394	12.704	12.564	na
$K_d(10^{-6})$	8.4999	8.6730	9.3080	9.7860	9.8192	9.6196	9.5991	9.5729	na

Table 1. PID controller parameters for different e values

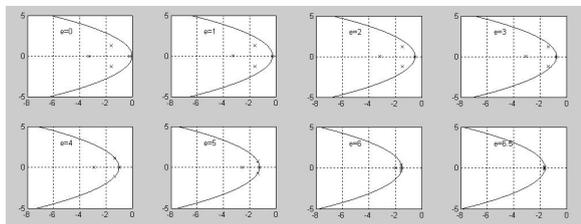


Figure 2. Parabola with different e values and the corresponding characteristic roots generated

Note that the D regions may be a set of disjoint regions, the different controller parameters can be generated by different D regions.

Now we get the PID compensator for AQM routers.

V. Experiments

To validate the performance of the PID controller, throughout this Section we will compare the results obtained from our PID controller to those obtained from the RED and PI controllers using the well known NS2 (Network Simulator Version 2) [15]. The network topology is shown in Figure 3. Because of the limit space of this paper, we only describe three representative simulation results in this Section.

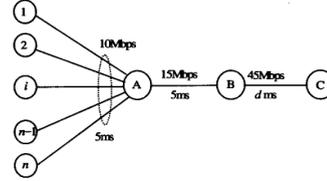


Figure 3. A simple network topology and simulation scenario

The only bottleneck links are the queues A and B , which has 15Mb/s (3750 packets/s, average packet size 500bytes) of capacity and 5ms of delay. There are n ftp flows whose capacity is 10Mb/s and delay is 5ms uniformly. Others use DropTail queue except that node A uses RED, PI or PID. The router's buffer size is 300packets. The maximum threshold is 200packets and the minimum threshold is 100packets of RED respectively. The referenced queue length (q_{ref}) for PI and PID is chosen to be 150packets. The link between the node B and C has d ms of delay.

Experiment 1:

The experiment is designed for $n=60, d=30$ ms. The instantaneous queue length plots for the RED, PI and PID controllers are depicted in Figure 4. The PI controller takes a long time to settle down to the referenced queue length, but the RED and PID controllers have a fast response speed. Then, we keep n invariable and increase d to 70ms. As observed in Figure 5, there are large queue oscillations for the RED controller. It frequently goes down to zero, thereby under-utilizing the link. The PI controller is

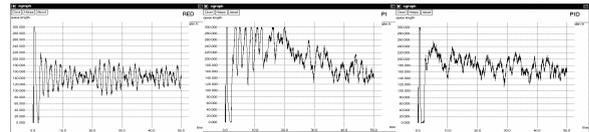


Figure 4. The instantaneous queue length for the RED, PI and PID

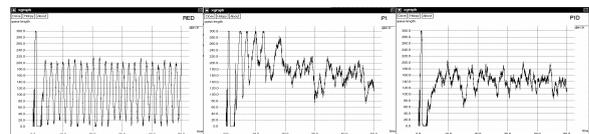


Figure 5. The instantaneous queue length for the RED, PI and PID

still sluggish, while the PID controller has a much fast response speed and small oscillations.

Experiment 2:

In this experiment, we exercise the controllers by setting d to 30ms and increasing the load level n to 90. The simulation results are shown in Figure 6. Then, we decrease n to 30 and keep d constant (see Figure 7). From Figure 6 and 7 we can observe, the RED controller has large oscillations under heavy load or light load. The PI controller loses packets besides its slow response under heavy load. However, the PID controller still shows good performance in both cases.

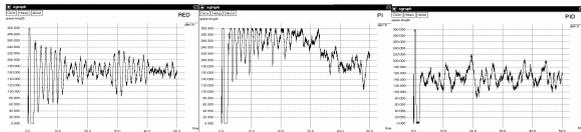


Figure 6. The instantaneous queue length for RED, PI and PID

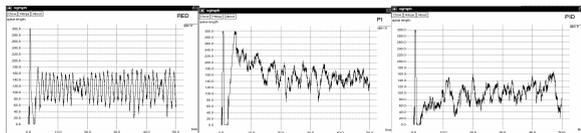


Figure 7. The instantaneous queue length for RED, PI and PID

Experiment 3:

In the last experiment, we set $n=60$, $d=30$ ms, q_{ref} ranged from 10packets to 150packets. The utilization of the bottleneck link versus q_{ref} curve for the RED, PI and PID controllers is illustrated as Figure 8. Then, we plot the utilization versus delay curve in Figure 9 ($n=60$, $q_{ref}=150$ packets, d ranged from 0ms to 100ms). Both Figures indicate that the PID controller shows much higher utilization than the RED and PI controllers. Due to its differential coefficient, the PID controller can estimate and respond to the network congestion in advance, which takes a shorter time to the q_{ref} than the PI controller. The RED controller shows sensitivity to the network parameters. There are large oscillations under small q_{ref} or large delay,

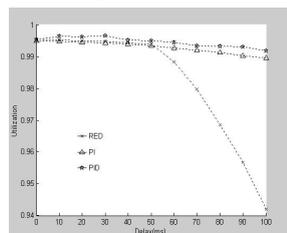
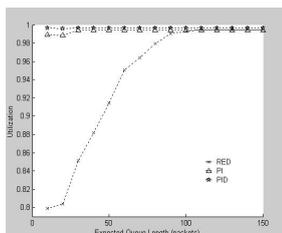


Figure 8. Utilization vs. q_{ref} Figure 9. Utilization vs. delay which increase the probability of the empty queue and lead to the low utilization of the bottleneck link.

The results indicate that the PID controller shows

better performance than the RED and PI controllers.

VI. Conclusions

In this paper we have designed a PID controller for AQM routers. As in any feedback control system design, our approach is driven by closed-loop performance objectives. For AQM performance, we focus on objectives including queue usage and latency control. First, we have specified a set of desired D -stable regions in the complex plane, then, we have run a numerical optimization algorithm in order to find the controller parameters such that all the roots of the closed-loop system are within the specified regions. A comparative study has shown that the PID controller exhibits much better performance and demonstrates its ability to operate the network at a higher link utilization, a better average queue length, a faster response speed and smaller oscillations than the RED and PI controllers in most cases.

References

- [1] Hollot, C.V., Misra, V., et al. On designing improved controllers for AQM routers supporting TCP flows. In *Proceedings of the IEEE INFOCOM*. Anchorage, Alaska: IEEE Computer Society, 2001. 1726~1734.
- [2] Feng, W., Kandlur, D., Saha, D., Shin, K. A self-configuring RED gateway. In *Proceedings of the INFOCOM '99*. New York: IEEE Computer Society, 1999. 1320~1328.
- [3] Ott, T.J., Lakshman, T.V., Wong, L.H. SRED: stabilized RED. In *Proceedings of the INFOCOM '99*. New York: IEEE computer Society, 1999. 1346~1355.
- [4] Feng, W., Kandlur, D., Saha, D., Shin, K. Blue: a new class of active queue management algorithms. *Technical Report, CSE-TR-387-99, University of Michigan*, 1999.
- [5] Sanjeeva, A., Victor, H.L., Steven, H.L., Qinghe, Y. REM: active queue management. *IEEE Network*, 2001, 15(3): 48~53.
- [6] Kunniyur, S., Srikant, R. Analysis and design of an adaptive virtual queue (AVQ) algorithm for active queue management. *ACM Computer Communication Review*, 2001, 31(4): 123~134.
- [7] Wang, Y.J., Schinkel, M., Schmitt-Hartmann, T., Hunt, K.J. PID and PID-like controller design by pole assignment within D -stable regions. *Asian Journal of Control*, 2002, 4(4): 423~432.
- [8] Misra, V., Gong, W.B., Towsley, D. Fluid-Based analysis of a network of AQM routers supporting TCP flows with an application to RED. In *Proceedings of the ACM/SIGCOMM 2000*. Stockholm, 2000.
- [9] Hollot, C.V., Misra, V., Towsley, D., et al. A control theoretic analysis of RED. In *Proceedings of the INFOCOM 2001*. Tel Aviv, 2001.
- [10] Wang, Y.J., Schinkel, M., Hunt, K.J. An engineering approach to simultaneous stabilization and strong simultaneous stabilization with D -stability. In *Proceedings of 40th IEEE Conference on Decision and Control*, Florida, USA, 2001: 967~973.
- [11] Schinkel, M., Wang, Y.J., Hunt, K.J. A robust and performance orientated controller design approach for nondeterministic hybrid systems. *International Journal of Control*, 2002: 1368~1375.
- [12] Hunt, K.J., Wang, Y.J., Schinkel, M., Schmitt-Hartmann, T. Controller design for hybrid systems using simultaneous D -stabilization technique and its application to Anti-lock Braking systems (ABS). In the book on *Nonlinear and Hybrid Control in Automotive Applications*. Edited by Anders Rantzer, Springer-Verlag, pp.97~124.
- [13] Dennis, J.R., J.E., Schanabel, R.B. Numerical methods for unconstrained optimization and nonlinear equations. *Prentice-Hall*, Amsterdam, 1989.
- [14] Leva, L., Colombo, A.M. Method for optimizing set-point weights in ISA-PID autotuners. *IEE Proc.- Control Theory Appl.*, 1999, 146(2): 137~146.
- [15] Fall, K., Varadhan, K. The ns Manual. *The VINT Project*. May 28, 2003.

Dynamic Resource Control for QoS-Aware Multimedia Wireless Networks

Shaokai YU, Seung-Wook Min, Won-Sik Yoon, Yongdeak Kim, and Ge Xiaohu
Dept. of Electrical and Computer Engineering, Ajou University
San 5, Wonchon-dong, Paldal-gu, Suwon, 442-749, Korea
syu@ajou.ac.kr

Abstract -Radio resource is an important issue for current multimedia wireless networks. Efficient traffic control strategies can be explored to optimize resource allocation. Dynamic control and multi-level control schemes are proposed to satisfy quality of service (QoS) constraints in the inference-dominated multimedia wireless environment. The performance of the Markov process is derived by the efficient matrix-analytic methods (MAMs). Results show that the proposed schemes are able to flexibly control system capacity without degrading the required service quality for multimedia services. These results are also important for the design of distributed multimedia wireless systems as well as network operation.

1. INTRODUCTION

The increasing use of multimedia services, such as video services and gaming services, makes the wireless access become the bottleneck of system capacity in broadband networks. In order to meet the stringent QoS for multimedia wireless services, the limited radio resource needs to be efficiently utilized through certain control strategies. These control schemes are not only important to the overall system efficiency, but also significant to the cost saving of network infrastructure. For current CDMA-based 3G networks, there does not exist a constant value for maximum available capacity. The efficient use of radio resource management (RRM) strategies enables to maintain the planned coverage area with required QoS in terms of signal-to-interference ratio (SIR) or bit error rate (BER) constraint. After wireless networks have been implemented, control strategies are still regarded as an effective solution in the RRM algorithms to yield the best performance/cost (P/C) ratio. In addition, the multimedia traffic characteristics with terminal mobility and traffic priority are of importance to the design of the RRM algorithms since the key resource management problems lie in how to exploit the leftover capacity efficiently in a real-time fashion. As a result, effective control schemes for multimedia services, which are sensible to the required QoS and also easy to be

implemented, are still required in the multimedia wireless networks.

For the control strategies, there are two types of call admission control (CAC) schemes. One is known as static policies, and the other is referred to dynamic policies [1]. Both the fixed channel assignment-based CAC (FCAC) and the measurement-based CAC (MCAC) schemes are classified as the static policies [2] and [1]. On the other hand, the soft-decision CAC (SCAC) in [3] and [4] belongs to the dynamic policies. Because the SCAC scheme does not require mutual exchange of network states between adjacent cells, it is seen as a simple measurement and promising control mechanism.

For multimedia services, a plural research has been conducted over the past few years. For example, an admission control scheme for voice and data mixed traffic to share the full bandwidth is proposed in [5]. A SIR-based admission control policy is analyzed for voice and data traffic in [6]. The CAC problem is formulated as a semi-Markov decision process in [7]. Moreover, an effective bandwidth-based admission control procedure is analyzed for a loss-only network and a product-form solution is derived in [8]. However, all these schemes use the centralized control solutions and they are not best suitable for the environment wherein the distributed solutions are required. A 2-D Markov model is numerically solved from the balance equation in [9]. In addition, although mixed traffic capacity in the forward link is analyzed by Zhang in [10], no control policy is enforced and also no mobility study is included. Taking these into consideration, the basic idea of this paper is to design QoS sensible control solutions suitable for multimedia wireless services by using distributed control schemes.

The rest of this paper is organized as follows. Section 2 describes the system model. Section 3 presents the dynamic control schemes and the multi-level control scheme (MLCS) for multimedia wireless traffic classes. Subsequently, performance measures and analyses are detailed in Section 4. Results are discussed in Section 5. Finally, conclusions of this paper are given.

2. SYSTEM MODEL

According to the UMTS QoS traffic classes, there exist real-time traffic and non-real-time traffic. The non-real-time traffic is able to withhold more delay than the real-time traffic. In order to formulate our model, the following assumptions are held in the forward link.

Firstly, real-time traffic has on and off two states. The real-time traffic arrives with a rate of λ_v which is exponential distributed. The handoff real-time traffic from adjacent cells is represented by another arrival rate λ_h . A maximum number of active users c_u with N state spaces can be accommodated while meeting a minimum requirement of QoS. In particular, the handoff real-time traffic has more priority than the new real-time traffic.

Secondly, in this study, the streaming non-real-time traffic is assumed and approximated by a Poisson process with a mean arrival rate of λ_d . Once the arriving real-time traffic finds the system full, the real-time traffic is able to preempt the non-real-time traffic. The preempted streaming non-real-time traffic can be queued in an infinite waiting buffer.

In addition, a uniform distribution mobility model is employed for the real-time traffic. The average cell crossing rate is denoted by $R_s = (VL)/(\pi S)$, where V , L , and S represent average terminal speed, cell perimeter, and cell area with a radius of R_c . Because channel holding time in a cell is the minimum of call holding time and cell sojourn time, the average channel holding time becomes: $\mu_1 = \mu_v + (2V/\pi R_c)$.

Finally, for the real-time traffic, the parameters are represented by the targeted SIR to meet the required QoS γ_v with the bit rate R_v and activity factor α_v . Similarly, for the streaming non-real-time traffic, they correspond to γ_d , R_d , and α_d respectively.

3. THE DYNAMIC CONTROL SCHEMES AND THE MLCS

In this study, identically multiple cells are taken into consideration. There are i active real-time users in the cell. For real-time traffic, the total interference consists of intercell interference I_{in} , intracell interference I_{out} , and thermal noise density N_0 . That is, $I_{in} + I_{out} + N_0 \leq I_{req}$. Similar to [6], taking the traffic activity into consideration, we can define an admittance function for real-time traffic as follows. If $c_{vt} \leq i < c_{vt}$,

$$\pi(i) = 1 - \sum_{q=0}^i B_i(q; i, \rho_v) \frac{1}{2} \operatorname{erfc} \left(\frac{C_{\max} - i - f\bar{i}}{\sqrt{2f\sigma_i}} \right). \quad (1)$$

When $i < c_{vt}$, $\pi(i) = 1$. Otherwise, $\pi(i) = 0$. As it is mentioned previously, the non-real-time traffic is more delay insensitive than the real-time traffic and both traffic classes also have different traffic characteristics. In order to increase the efficacy of resource utilization, the challenge for the non-real-time traffic is to exploit the leftover capacity by the real-time traffic while maintaining optimal throughput-delay tradeoff. The interference limitation for the non-real-time services can be shown as:

$$\sum_{j=2}^{k_j} \rho_{dj} + \sum_{h=1}^{f_k} \rho_{dh} \leq \frac{PG_d(1-\eta_d)}{E_{db}/I_{dreq}}. \quad (2)$$

Due to the effect of fading distributions, f_d can be modeled as a random variable following a Gaussian distribution while the cell loading is low. When the cell loading becomes high, log-normal distribution can be approximately used [10]. In this study, f_d is treated as a random variable followed by a lognormal distribution. Let Z_d represent the equivalent number of users, hence:

$$P(Z_d \leq K_{\max}) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\ln(K_{\max}) - \mu_d}{\sqrt{2}\sigma_d} \right) \right] \quad (3)$$

where the μ_d and σ_d^2 are the mean and variance in the distribution respectively. Therefore, if the total number of channels equals to C , we can define an admittance function for the non-real-time traffic if $c_{dt} \leq j < c_{dt}$,

$$\pi(j) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\ln(\lfloor (C - \alpha_d i) / \alpha_d \rfloor) - \mu_d}{\sqrt{2}\sigma_d} \right) \right]. \quad (4)$$

When $j < c_{dt}$, $\pi(j) = 1$. Otherwise, $\pi(j) = 0$. For both real-time services and non-real-time services, the joint outage probability becomes: $P_{out} = P[(Z_v + Z_d) > (C_{\max} + K_{\max})]$. Using the admittance functions in (1) and (4), the analysis of multimedia traffic in wireless networks can be derived by using the efficient algorithm in next Section.

In a live network, the number of users within the own cell and other cells is rapidly varying and only the statistical data can be obtained. Therefore the predefined admittance functions are hard to reflect the dynamics of an operating network. As a result, a controlled multi-level scheme is proposed for this purpose. The basic idea of the MLCS is to select a suitable level so as to control network traffic in a real-time fashion. In general, a coefficient $\pi(\lfloor (M-l)N/M \rfloor)$ can be used for the control scheme, where $0 \leq l < M$. The number of levels can be actually determined by the operators based on network operation experience or customers' need within cells. For example, if a two-level control scheme is selected, a coefficient $\pi(\lfloor N/2 \rfloor)$ can be used. Similarly, if a four-level control scheme is adopted, coefficients $\pi(\lfloor 3N/4 \rfloor)$, $\pi(\lfloor N/2 \rfloor)$, and $\pi(\lfloor N/4 \rfloor)$ are to be utilized. From an operation viewpoint, the MLCS can be flexibly adopted in

different time slots per day, for example, in the morning, in the afternoon, or at night. As a result, this can dynamically characterize the traffic peak and non-peak periods based on user profiles.

4. PERFORMANCE MEASURES

Based on the description of the system model and the use of aforementioned dynamic control schemes, the equilibrium probability π can be calculated if the 2-D Markov process is irreducible: $\pi T = 0$, $\pi e = 1$, where e is a one column vector and T represents the transition rate matrix.

The sub-matrices are all with the dimension of $N \times N$. A_{12} and A_{01} are all scalar matrices. Therefore, we can easily have $A_{12} = A_{01} = \text{Diag}[\pi(1)\lambda_1, \pi(2)\lambda_2, \dots, \pi(N)\lambda_N]$. Similarly, A_{10} and A_{21} are the diagonal matrices with $A_{10} = A_{21} = \text{Diag}[N\mu_d, (N-1)\mu_d, \dots, 0]$. Then we can obtain $A_{11} = T_{11} - A_{10} - A_{21}$ and $A_{00} = T_{00} - A_{01} - A_{10}$ respectively.

In order to compute the recurrent states in the process, an efficient logarithmic reduction algorithm (LRA) of MAMs is applied [11]. It is known that the rate matrix R has a spectral radius less than 1 if the quasi-birth-death process is recurrent. Hence, we can have: $R^2 A_2 + R A_1 + A_0 = 0$. As a result, the equilibrium probability can be obtained from solving the matrix R iteratively. Hence, $\pi_i = \pi_{c_i} R^{k-c_i}$, while $k \geq c_i$. Subsequently, the average number of data packets sojourning in the system can be computed by:

$$S = \sum_{i=0}^{c_1} \sum_{j=0}^{\infty} j \pi_j$$

If there is no admittance control, the

blocking probability for real-time traffic is:

$$P_B = \sum_{j=0}^{\infty} \pi_{c_j}$$

After the admittance control is taken

into account, the blocking probability for the

$$\text{real-time traffic becomes: } P_B = \sum_{i=c_1}^{\infty} \sum_{j=0}^{\infty} \pi_j (1 - \pi(i)).$$

5. RESULTS DISCUSSION

Our study pertains to the other third generation wireless networks. However, we only take the WCDMA-based UMTS systems into consideration. For current WCDMA networks, the CAC algorithms are implemented in the Radio Network Controller (RNC) and there exist different admission control

algorithms for both forward link and reverse link. Our forward control schemes are able to be operated at each Node B. For implementation, the actual mean and variance of forward link interference from other cells can be measured according to field experiment or simulation results.

In this study, the data rate of real-time traffic is assumed to be 64 kbps and the required SIR target for BER is $10e-05$. The arrival rate of the handoff real-time traffic is equivalent to 10% of the new real-time traffic. As the simulation shown in [6], the lower bound capacity is 14 and the upper bound capacity is 22. In WCDMA systems, transmit power control (TPC) is used to maintain the QoS requirements for each user. For the real-time traffic, due to the inaccuracy of TPC, the impact of varying SIR from the own cell on the intercell interference is shown in Fig. 1. Note that the offset probability is referred to the complementary value of intercell interference's proportion while the requirements of QoS in the own cell is met.

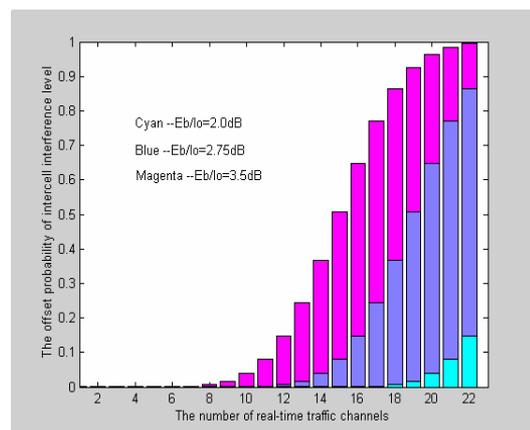


Fig. 1. The offset probability of intercell interference level.

From Fig. 1, we observe that, in order to maintain the required QoS for real-time traffic, the number of active users in the other cells needs to be reduced proportionally according to the increasing active users in the own cell. In addition, it also indicates that the higher the transmit power is required in the own cell, the fewer active users must be accommodated in the other cells.

Taking the admission control for real-time traffic into account, the impact of TPC error for real-time traffic on non-real-time traffic performance is shown in Fig. 2. It shows the correlation relationships between the real-time traffic and non-real-traffic. We observe that, with the increased SIR for real-time traffic, the average number of non-real-traffic become increasing. The reason for this is that the increase of SIR for real-time traffic results in the reduction of average number of real-time traffic in the system. As a result, this leads to the increase of average number of non-real-traffic in the system.

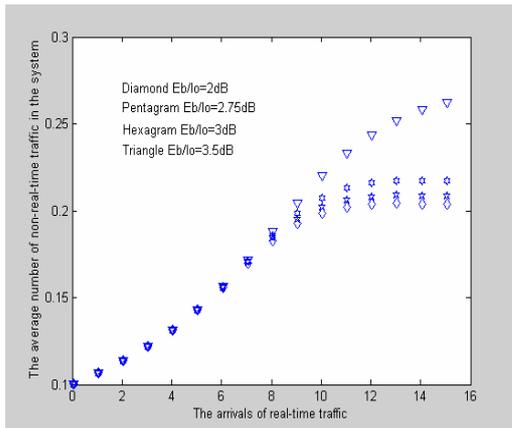


Fig. 2. The impact of real-time traffic on non-real-time traffic performance.

Taking the admittance function for non-real-time traffic into consideration, the use of MLCS can be shown in Fig. 3. Without using the MLCS, a high throughput of non-real-time traffic can be obtained. On the contrary, through the use of the multiple control levels, the average number of non-real-traffic can be reduced proportionally according to the flexible selection of control levels. For example, the use of 1-level, 2-level and 4-level control schemes can reduce average data throughput by 23%, 45% and 64% respectively. Therefore, this strategy can be used to control the non-real-time traffic congestion and relieve the overloading condition of multimedia systems.

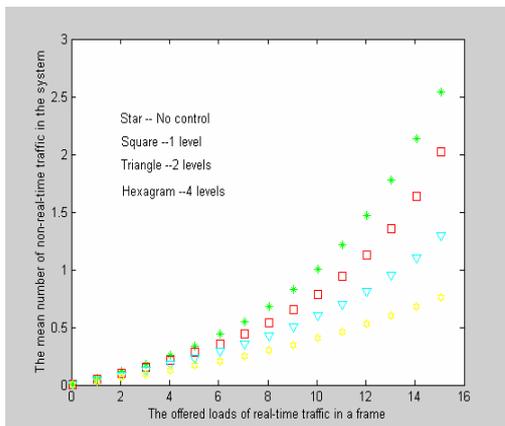


Fig. 3. The impact of MLCS on non-real-time traffic.

6. CONCLUSIONS

In this study, we propose dynamical control schemes for the UMTS-based wireless networks in order to guarantee the sensible QoS requirements. A generic analytical method is developed to study the performance of multimedia traffic in wireless networks. We show that multimedia traffic has an inter-twined relationship for network performance

and the dynamic control schemes are flexible and easy to be implemented for the multimedia wireless systems. In addition, the congestion of non-real-time traffic can be effectively abated through the use of multi-level control solutions.

ACKNOWLEDGMENT

The first author wants to acknowledge the support of IITA, Korea.

REFERENCES

- [1] H. Tong and T. X. Brown, "Adaptive CAC under QoS constraints: a reinforcement learning solutions," *IEEE JSAC*, vol. 18, No. 2, pp. 209-221, Feb. 2000.
- [2] Z. Liu and M. E. Zarki, "SIR-based call admission control for DS-CDMA cellular systems," *IEEE JSAC*, vol. 12, no. 4, pp. 638-644, May 1994.
- [3] P.-V. Vinh and D. D. Luong, "Capacity enhancement with simple and robust soft-decision call admission control for WCDMA mobile cellular PCN's," In Proc. *IEEE VTC* 2001, vol.3, pp. 1349-1353, 7-11 Oct. 2001.
- [4] P.-V. Vinh, "Soft decision call Admission control versus static fixed channel assignment and interference measurement based policies for WCDMA cellular PCN's," In Proc. *IEEE PIMRC*, vol. 2, pp. E-27-E-31, 30 Sept.-3 Oct. 2001.
- [5] J. W. Chang, J. H. Chung and D. K. Sung, "Admission control scheme for soft handoff in DS-CDMA cellular systems supporting voice and stream-type data services," *IEEE Trans. Veh. Technol.*, vol. 51, no. 6, pp. 1445-1459, Nov. 2002.
- [6] S. Anand and A. Chochalingam, "Performance analysis of voice/data cellular CDMA with SIR-based admission control," *IEEE JSAC*, vol. 21, no. 10, pp. 1674-1684, Dec. 2003.
- [7] S. Singh, C. Krishnamurthy and H. V. Poor, "Integrated voice/data call admission control for wireless DS-CDMA systems," *IEEE Trans. on Signal Processing*, vol. 50, no.6, pp. 1483-1495, June 2002.
- [8] J. S. Evans and D. Everitt, "Effective bandwidth-based admission control for multiservice CDMA cellular networks," *IEEE Trans. Veh. Technol.*, vol. 48, pp. 36-46, Jan. 1999.
- [9] C. W. Leong, W. Zhuang, Y. Cheng and L. Wang, "Call admission control for integrated on/off voice and best-effort data services in mobile cellular communications," *IEEE Trans. Commun.*, vol. 52, no. 5, pp. 778-790, May 2004.
- [10] Q. Zhang and O-C Yue, "UMTS air interface voice/data capacity - part 1: reverse link analysis," *IEEE VTC* 2001, Spring, vol. 4, pp. 2725-29, 6-9 May, 2001.
- [11] G. Latouche, "A logarithmic reduction algorithm for quasi-birth-death processes," *J. App. Prob.*, vol. 30, pp. 650-674, 1993.

e-learning promotes students' deep learning: a case study

Dr. M J Lopez and Dr. H R Ponce
University of Santiago of Chile
Av. LB O'Higgins 3363, Santiago, Chile
Tel: + 56 2 776 22 60
Fax: + 56 2 779 97 23
e-mail: mlopez@lauca.usach.cl

Introduction

This paper reports on a second part of a research and development project being conducted at the Department of Industrial Engineering. The first part of the project researched and developed an e-infrastructure to support students' learning, which was qualitatively evaluated, reported at the CSNDSP' 2002 and awarded a Best Paper Prize by the IEEE UK/Republic of Ireland Communications Chapters.

The second part of the project, reported here, has been concerned with a more quantitative evaluation of the impact of the developed e-infrastructure upon students' study styles. Therefore the paper, after this introduction summarises the first part of the project. Then it presents the theoretical background to the second part. Next the course organisation and the instrument used for the quantitative evaluation are presented.

Results of students' responses are then presented and commented. Finally, some remarks are made as conclusions.

Summary of previous work

The paper presented at the CSNDSP' 2002 (Lopez et al, 2002) reported a research and development to create an e-infrastructure to support the new approach to teaching and learning. Firstly, it described the research activities and operational model of students' learning used to build the e-infrastructure. Then, it presented the organisation of courses, which included the modelling of contents, activities and evaluations.

Secondly and in more detailed, the paper described the e-infrastructure developed. This was done in terms of the design of a synchronous/asynchronous web portal to support new courses contents, activities and evaluations.

Finally, it presented some of the reactions of students regarding the use of the e-infrastructure.

Main findings were the strength of course organisation and its structure and the usefulness of the e-infrastructure developed.

Students said that the course organisation was good for them because the teaching team was made up of five people (one lecturer and four assistants). They pointed out that such a team is not usual in departmental practices and represented a good opportunity to have an enriched interaction. This allowed them to a better group work and in an ease manner. Students stated that the way in which the materials were prepared facilitated going through them, gave them the opportunity to better organise their time and plan activities of this and other courses. The organisation of the assessment, it was said by several students, with so many marks made it that individual marks had low impact in the final mark, which decreased pressure on individual marked activities.

The e-infrastructure was the item most mentioned by students in the interviews. For all of them, this was their first course to be based on an e-learning experience, which had extensive support of information technologies integrated to Internet. Students valued not to 'have to' go to the lecture theatre, and 'to go to class' at whatever time they could or had the opportunity, with the "materials there, as if they were waiting for me", as one student put it. It was also stressed that this gave them much autonomy and that this was not always possible to adequately manage. It was also repeatedly mentioned that the e-infrastructure had novelty of use, which produced a new sort of encouragement to revise the materials and that it was fairly easy to access its learning resources.

The major conclusion was that a robust e-infrastructure for distant and distributed learning was developed and, a formative evaluation conducted showed that students had an enthusiastic reaction, that the e-infrastructure was interesting, useful, allowing them a deeper knowledge and that they would very much appreciate more courses in a similar format.

Theoretical background

The development of our e-learning environment has been strongly influenced by both our experiences and two main conceptual learning frameworks. Our experiences had showed us that our traditional approach to teaching made students who recall by memory and, with time passing by, a risk of failing to remember. The conceptual learning frameworks are the “experiential learning approach” and the distinction between “deep learning” approach and “superficial learning” approach. As a result of the use of these frameworks, our own ideas of what learning meant also changed.

The first theoretical element that came to contribute to our development of an e-learning environment was the experiential approach to learning (Kolb, 1984). The crucial question that this approach addresses is how students learn. Although originally formulated to address the question of adult education, it has made an important contribution to understand how students learn in general. In this approach, learning is understood as a process in which “people generate from their experience the concepts, rules and principles that guide their behaviour in new situation” (Kolb, et al., 1991, p. 60). The effectiveness of their behaviour depends on how they adapt their concepts, change their rules or discover new guiding principles.

Learning takes place through a continuous and recurrent sequence of actual experiences and, as experiences by themselves are insufficient, they must be accompanied by thought, observation, abstract concept construction and trying out these concepts in new experiences. Thus, the learning process is conceived as a four-stage cycle. (1) concrete experiences are followed by (2) observation and reflection of such experience, which leads to the (3) formation of abstract concepts and the constructions of principles or generalisations which follows (4) the testing of such concepts in new situations.

Thus, we realised that following this model we could incorporate new learning activities to emphasise each phase of the cycle. Thus, we thought of introducing seminars with small groups, with a view to have debates or discussions on new concepts. Assessing students through the development of a case study applied to a real world situation, which was conducted as a course project. Assigning minor research work on the e-learning environment and developing a more personalised instruction.

The second theoretical source that influenced our development was the distinction between the *superficial approach* and the *deep approach* to learning. With these important concepts, we realized that

some students have different ways of confronting the learning process. While some students take learning simply as a matter of memorizing concepts and reproducing knowledge; other students approach learning with interest in ideas and understanding and with a clear intention to transform such ideas based essentially on their previous experiences and knowledge (Entwistle, 1981).

Having in consideration these distinctions, we rapidly asked ourselves how to promote a deep approach to learning among our students (Entwistle, 2000, Gibbs, 1999). That is to say, we started looking for an e-learning environment that fostered the interest in ideas and understanding in contrast with the prevalent approach observed in our Department principally focus on memorising concepts. Our objective was to conceive a series of related activities supported with e-learning technology that helped us to stimulate and develop, in our students, the ability to seek meaning, relate concepts and make sense of their experiences within and beyond the frontiers of our courses.

Such a challenge required a different teaching model. We found it on what Entwistle (2000) calls a learning oriented approach to teaching. In this model, the teaching and learning process changes from being centred in teaching to be centred in learning. Educators became facilitators and learners are much more active (Gibbs, 1999, Cervera & Gonzalez, 1997).

So, in the process of developing the e-learning technology, our own conception of teaching changed—from a content oriented to learning oriented. From imparting information and structuring what knowledge students learned to an approach that facilitates understanding of concepts and encourage conceptual change. This process was naturally not easy and we learned as we developed the e-learning environment and experimented by supporting our students with such an environment. In this regard, Entwistle (2000) indicates that “the approaches to teaching adopted by teachers also influences their students’ approaches to studying and through those the learning outcomes” (p. 4).

Our own change of conception of how to approach the teaching and learning can be summarised as follows. The lecturer in the traditional approach is a deliverer, unique assessor and decides what and how students learn. Students are dependent, individualist and receptive. On the other hand, in the new approach lecturers are managers, planners, designers, facilitators and guides. Students are autonomous, collaborative and engaged. It is this latter approach that we have attempted to recreate through the development and use of an e-learning environment to support student’s learning.

e-learning courses

The organisation of courses was based on the experiential learning cycle. Consequently, courses were organised to provide students with opportunities for experiences, thoughts, observations, abstract concept construction and probe of new experiences. For abstract conceptualisation there was an initial conference per course unit as well as reading of articles, papers and book chapters. For experiences and probe of new ones, students developed, throughout the course, a case study. For thought and deliberation there were two seminars per unit. Contents, activities and evaluations were available in the e-learning platform.

The Inventory

To investigate if the e-learning infrastructure (synchronous and asynchronous Web portal plus modelled contents, activities and evaluations) had any impact on students' study styles, the Approaches and Study Skills Inventory for Students (ASSIST) was used (ASSIST, 1997).

The inventory was built to discriminate students learning styles. It has three parts: conceptions of learning; approaches to study; preferences for different types of course and teaching.

The part on conceptions of learning is made up of nine questions. The part on approaches to study is divided in three sections: deep approach; strategic approach; surface apathetic approach. The deep approach is divided in three subsections (seeking meaning; relating ideas; use of evidence) with four questions each. The strategic approach is divided in four subsections (organised studying; time management; alertness to assessment demands; achieving; monitoring effectiveness) with four questions each. The surface apathetic approach is divided in 4 subsections (lack of purpose; unrelated memorising; syllabus-boundness; fear of failure) with four questions each. The preferences for different types of course and teaching scored as the sum of the four items (questions) for two sections (supporting understanding and transmitting information). Students respond to items on a 1 - 5 scale. The part on preferences for different types of course and teaching is made up of two sections to set apart if their liking is for the deep or surface approach; each section has four items to be responded.

817 first-year university students drawn from ten contrasting departments in six British universities completed ASSIST. A correlation analysis between the variables deep, strategic and surface apathetic approaches was conducted, as illustrated in the following table.

Table 1: ASSIST - Correlations between factors

	I	II	III
Factor I (Deep)	1.00		
Factor II (Surface Apathetic)	- 0.20	1.00	
Factor III (Strategic)	0.35	- 0.22	1.00

The correlations clearly show that:

- Surface apathetic and deep approaches (-0.20) are almost not related.
- Strategic and deep approaches (0,35) are feebly related.
- Strategic and surface apathetic (-0,22) are not related.

Hence the questionnaire is well designed to discriminate among study styles.

Results

Being a good inventory, ASSIST was duly translated to Spanish and the 82 students in the course were asked to fill it in. Answers were coded into the suggested 1 to 5 scales and processed with the Statistical Package for Social Sciences (SPSS).

The initial statistical analysis was to ensure that the 82 students' responses preserved the no relation between the factors. Therefore the same statistical test was carried out and table 2 shows the results.

Table 2: COURSE - Correlations between factors

	I	II	III
Factor I (Deep)	1.00		
Factor II (Surface Apathetic)	- 0.20	1.00	
Factor III (Strategic)	0.52	- 0.12	1.00

In general the 82 responses had a similar behaviour that those of the 817 British students. The main divergence with ASSIST is that in this case the deep and strategic approaches present some relation.

To segregate responses, some descriptive statistics are presented. Descriptive statistics are used to describe the basic features of the data in the study. These basic characteristics are in the following tables.

Table 3: Factors - Descriptive statistics

	N	Mean	StDv
Factor I (Deep)	81	3.92	0.80
Factor II (Surface Apathetic)	82	3.29	0.92
Factor III (Strategic)	81	3.79	1.09

Regarding students' study approach, the central tendency of the distribution shows that students in the course tended to be deep learners (3.99) rather than superficial (3.29) or strategic (3.79).

The somehow weak relation found between deep and strategic approaches (correlation of 0.52) is reflected by the closeness of their means (3.92 and 3.79 respectively).

Table 4: Factor I (Deep) - Descriptive statistics

	N	Mean	StDv
Seeking meaning	82	3,99	0,73
Relating ideas	82	3,92	0,83
Use of evidence	81	3,95	0,75
Interest in ideas	81	3,84	0,87

Within the deep approach to learning it is interesting to note the nearness of responses for the four items (means between 3.84 and 3.99). It is also notable that their standard deviations are relatively low.

Table 5: Factor II (Surface) - Descriptive statistics

	N	Mean	StDv
Lack of purpose	82	3,42	1,02
Unrelated memorising	81	2,87	1,12
Syllabus-boundness	82	3,27	1,11
Fear of failure	82	3,58	1,12

The means of surface apathetic approach items are markedly lower than the deep items, they range from 2.87 to 3.58 and their standard deviations are over one. This is a more detailed indication that students in the course tended to be deep learners rather than superficial.

Table 6: Factor III (Strategic) - Descriptive statistics

	N	Mean	StDv
Organised studying	82	3,62	0,92
Time management	82	3,66	0,94
Alertness to assessment demands	82	3,79	0,96
Achieving	82	4,08	0,89
Monitoring effectiveness	81	4,02	0,74

The strategic approach ended up to be somehow related to the deep approach (correlation 0.52 between them); hence it follows that strategic items have means in the neighbourhood to those of the deep approach.

Table 7: Preferences for different types of course and teaching

	N	Mean	StDv
Deep approach	79	4,12	0,82
Surface approach	79	3,46	0,97

The last section of the inventory intends to draw students' preference for an approach. Table 7 clearly shows that the students on the course had an inclination for the deep approach, since its mean is 4.12 with a standard deviation below one.

Conclusions

For the case of the students of the Department of Industrial Engineering that used the e-infrastructure, ASSIST was a good instrument to discriminate among the three learning styles (deep, surface and strategic).

Results from ASSIST showed that students not only appeared to have developed a deep approach to learning but also they prefer this approach for courses in general.

The research question whether the e-learning infrastructure (synchronous and asynchronous Web portal plus modelled contents, activities and evaluations) had an impact on students' study styles was positively responded: e-learning promoted, in this case, a deep approach to learning.

References

- ASSIST (1997) Centre for Research on Learning and Instruction, University of Edinburgh
- Cervera, M. & Gonzalez, A. (1997) El docente y los entornos virtuales de enseñanza aprendizaje. **Edutec 97** (Universidad Rovira i Virgili).
- Entwistle, N. (2000) Promoting deep learning through teaching and Assessment: conceptual frameworks and educational contexts. TLRP Conference, Leicester, November.
- Gibbs, G. (1999) *Teaching in Higher Education: theory and practice, how students differ as learners*, Milton Keynes: Open University.
- Kolb, D. (1984) *Experiential Learning: experience as the source of learning and development*, New Jersey: Prentice-Hall.
- Kolb, D., Rubin, I. and Osland, J. (1991) *Organizational Behaviour: an experiential approach*, New Jersey: Prentice-Hall.
- Lopez, M, Bascuñan, C, Bravo, J. & Paredes, V. (2002) Research and development of e-infrastructure to support students' learning, in Carrasco, R (ed). *Proceedings of the Communication Systems, Networks and Digital Signal Processing (CSNDSP' 2002)*, Stafford, England, 15 – 17 Julio.

Acknowledgements

The authors are grateful to A E Arias, J E Bravo and V F Paredes for their enthusiastic support to the ongoing project on e-learning. To F Moreno for most useful information on statistics.

Failure of Adaptive RED in Dynamic Networks

Martin Fricker (PhD Student)
Communications Research Group
University of Wales Swansea
m.l.fricker@swan.ac.uk

Introduction

The Random Early Detection (RED) congestion avoidance algorithm was first proposed in 1993 by Sally Floyd and Van Jacobson [Floyd93]. The main design goals of RED are as follows:

- Ensure congestion avoidance by detecting incipient congestion and notifying sources early through packet loss
- Maximise throughput / utilisation
- Minimise delay by controlling average buffer occupancy
- Accommodate bursty traffic and transient congestion
- Avoid synchronisation of traffic sources

RED aims to fulfil these design goals using an average queue measurement and a series of thresholds. Basing the drop decision on an averaged as opposed to an instantaneous measurement helps to accommodate bursty traffic and transient congestion. In order to perform the congestion avoidance function the RED mechanism drops packets according to an increasing drop probability as the queue size measurement increases. This also helps to improve the distribution of packet drops and hence guard against synchronisation of traffic sources.

Random Early Detection (RED)

The RED algorithm can essentially be subdivided into two main functions, the estimation of the average queue size and the packet drop decision mechanism.

Estimation of Average Queue Size:

The RED algorithm uses an exponentially weighted moving average (EWMA) in order to calculate the average queue size. This average queue size smoothes the instantaneous queue size over time and can be derived using the following equation:

$$\text{Avg}_k = (1 - \text{Weight}_q) \text{Avg}_{k-1} + \text{Weight}_q Q$$

Where:

Weight_q controls the speed at which the algorithm responds to congestion in the network. That is,

weight_q determines how closely the average queue size follows the instantaneous queue size.

Avg_{k-1} is the value of the average queue size the last time it was calculated. (Note for the very first calculation of average queue size *Avg_{k-1}* is set to zero).

Q is the current instantaneous queue size

Packet Dropping Decision:

The RED packet dropping decision is based on a series of thresholds and a calculated drop probability. RED uses two threshold parameters *Min_{th}* and *Max_{th}*, these two thresholds define the boundaries of the transient congestion interval. *Min_{th}* specifies the minimum threshold whilst *Max_{th}* specifies the maximum threshold. When the average queue size is below the minimum threshold (*Min_{th}*) no packets will be dropped. When the average queue size is above the maximum threshold (*Max_{th}*) congestion is persistent, RED then reverts to a congestion control algorithm and drops all incoming packets (as per drop tail). When the average queue size is between the *Min_{th}* and *Max_{th}* thresholds, RED is in the congestion avoidance phase and drops packets with a probability (*Pa*). *Pa* is calculated using an intermediate probability *Pb* which is derived using the following equation:

$$P_b = \frac{\text{Max}P (\text{Avg}_k - \text{Min}_{th})}{(\text{Max}_{th} - \text{Min}_{th})}$$

Where:

MaxP defines the maximum drop probability

The drop probability *Pb* increases linearly with average queue size from zero to *MaxP*. As stated previously *Pb* is an intermediate step in the calculation of the final drop probability *Pa*. *Pb* is not used as the final drop probability because if it were then the number of newly arriving packets between dropped packets is a geometric random variable [Stern97]. This is undesirable as it can lead to global synchronisation. It is possible to transform this geometric random variable into an approximately uniform random variable via the probability *Pa* which is derived using the following equation:

$$P_a = \frac{P_b}{(1 - \text{Count} * P_b)}$$

Where:

$Count$ is the number of new arrivals since the last dropped packet

The result of this equation is that Pa increases slowly as $Count$ increases leading to a roughly uniform distribution of inter-arrival times between dropped packets.

The performance of RED has been shown to be very dependent on the correct parameterisation of its algorithm. However it has proven difficult to achieve optimal parameter selection across a range of network environments [Floyd97].

Adaptive Red

Adaptive RED represents an attempt to dynamically tune the operation of the RED algorithm to the load and characteristics of the network traffic [Feng99]. This is achieved by updating the value of the $MaxP$ parameter when the average queue length crosses the Min_{th} or Max_{th} thresholds. If the average queue size crosses the Min_{th} threshold (in a downward direction) then RED is dropping too aggressively and should be tuned to admit more of the incoming packets to the buffer by decreasing the $MaxP$ parameter by a predetermined factor. If however the average queue size crosses the Max_{th} threshold (in an upward direction) then RED is dropping too conservatively and should be tuned to drop more of the incoming packets. In order to accomplish this the $MaxP$ parameter is increased by a predetermined factor. The precise values of these predetermined factors determine the granularity with which the aggressiveness of the dropping mechanism can be modified.

Experimentation with adaptive RED in papers such as [Feng99] has yielded potentially promising results however as the authors themselves state this experimentation has only been preliminary.

Simulation Configuration

The results presented in this research are derived from simulations conducted using two configurations. The first configuration represents a stable network environment with relatively light congestion as shown in figure 1 below.

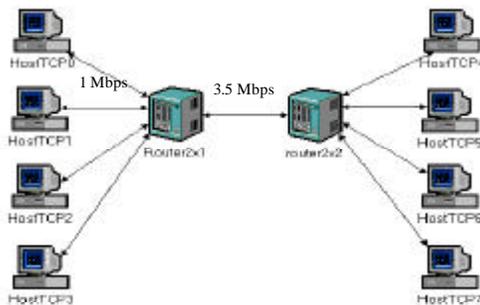


Figure 1 – Stable Network Configuration

The configuration consists of four TCP sources transmitting to four TCP receivers. The startup times for each source are staggered to provide a gradual build up of congestion. RED parameter settings under this configuration are as follows, $MaxP=0.1$, $Weightq=0.002$, $Min_{th} = 15$ packets and $Max_{th} = 45$ packets. The adaptive RED multiplier and divisor parameters are set to 2 and 3 respectively.

The second simulation configuration represents a more dynamic network operating at higher speed with greater numbers of more volatile sources (and the addition of unresponsive UDP sources) as shown in figure 2 below.

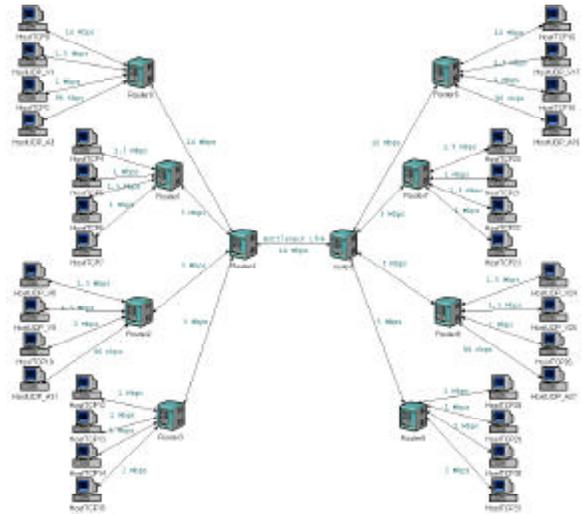


Figure 2 – Dynamic Network Configuration

RED parameter settings under this configuration are as follows, $MaxP=0.03$, $Weightq=0.002$, $Min_{th} = 20$ packets and $Max_{th} = 60$ packets. The adaptive RED multiplier and divisor parameters are again set to 2 and 3 respectively.

Results - Stable Network

Figure 3 below shows the effect of the adaptive algorithm (right) on bottleneck link throughput compared to the standard RED algorithm (left).

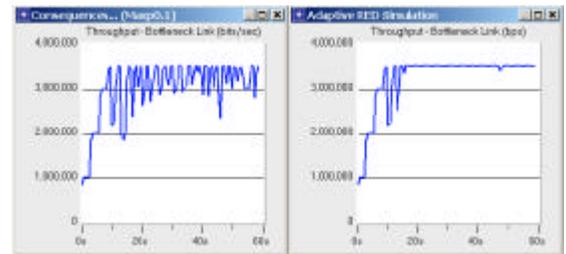


Figure 3 – Throughput (Stable Network)

Clearly the throughput at the bottleneck link is significantly higher and more stable when adaptive RED is used. The $MaxP$ setting of 0.1 which is too aggressive under these network conditions for the

standard RED algorithm is reduced by the adaptive algorithm resulting in more stable buffer occupancy and increased throughput at the bottleneck link.

The more optimal setting of the $MaxP$ parameter under the adaptive algorithm also ensures improved packet drop characteristics.

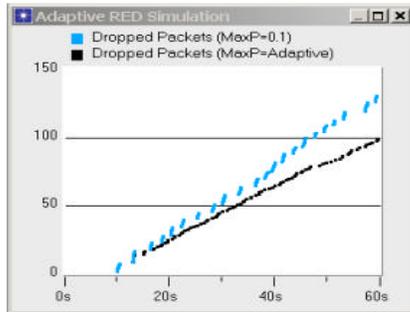


Figure 4 – Packet Drop Characteristics

Figure 4 above shows that packet drops under the adaptive algorithm are less in number and more evenly spaced than under the standard RED algorithm which is causing bursts of packet loss.

Results - Dynamic Network

Figure 5 below shows a comparison of bottleneck link throughput using the standard and adaptive RED algorithms in a dynamic network environment.

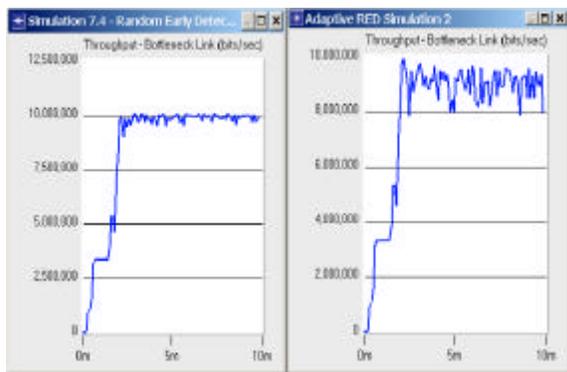


Figure 5 – Throughput (Dynamic Network)

The initial value of $MaxP$ in this simulation is better suited to the network conditions and hence the standard RED algorithm is able to maintain relatively high and stable throughput. Under these conditions however the adaptive algorithm causes large fluctuations in queue occupancy resulting in reduced (and less stable) throughput at the bottleneck link.

Further investigation shows that the adaptive algorithm can exacerbate queue oscillations present in a network with rapidly changing traffic conditions. Under these conditions the adaptive algorithm can cause a pendulum effect due to the

delayed nature of the average queue occupancy calculation. For example the averaged queue occupancy can be crossing the minimum threshold (causing $MaxP$ to be reduced) whilst the actual queue occupancy is already increasing rapidly.

Figure 6 shows the value of the $MaxP$ parameter over the course of the simulation.

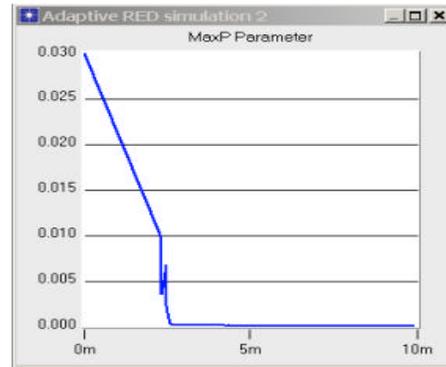


Figure 6 – MaxP Parameter Adjustment

It is clear from figure 6 that the value of the $MaxP$ parameter is rapidly reduced until it reaches (almost) zero. This is due to an interaction between the adaptive algorithm, the actual queue occupancy, the averaged queue occupancy and the physical queue limit.

When the actual queue size rises rapidly it exceeds the maximum threshold and reaches the physical queue limit before the average queue size is able to reach the maximum threshold. When the physical queue limit is reached packets are dropped (in a drop tail manner) causing the TCP sources to overreact and slow their transmission rate. Consequently the actual queue occupancy drops sharply to zero for a short time the average queue occupancy crosses the minimum threshold and the $MaxP$ parameter is reduced. This causes the algorithm to become less aggressive and this in conjunction with the recovery of the TCP sources causes the actual queue occupancy to once again rise rapidly to the physical queue limit.

In essence the problem occurs because the actual queue size is able to remain below the minimum threshold long enough for the averaged queue size to cross that threshold, whereas it is not allowed to remain above the maximum threshold. Consequently the $MaxP$ parameter is decreased far more frequently than it is increased exacerbating the problem by making the algorithm far too conservative in its packet dropping strategy and therefore causing it to revert to drop tail behaviour.

There is also another potential drawback of the adaptive RED algorithm concerning the choice of scaling factors (multiplier / divisor). If the scaling factor is relatively small a more optimal value of *MaxP* will be obtained, however the algorithm will take much longer to adapt to conditions and reach that optimal *MaxP* value, therefore in a dynamic network environment the algorithm may not be able to adjust in a suitable time frame. If the scaling factor is larger, much of the granularity of *MaxP* selection is lost and the selected *MaxP* value will not be as optimal, however the larger scaling factor does cause the algorithm to adapt much more quickly to changes in network status. Ultimately the selection of scaling factors should be decided according to the expected traffic patterns. However the aim of the adaptive RED algorithm is to allow RED to dynamically tune itself to network conditions, if the adaptive algorithm itself requires careful selection of parameters according to traffic conditions then the parameter selection problems with standard RED are simply relocated not solved.

Summary

In conclusion, whilst the adaptive RED algorithm can dynamically optimise the MaxP parameter (over a period of time) in a relatively stable network environment, it does not perform so favourably in a rapidly changing environment. In some circumstances the adaptive algorithm can actually degrade performance to that representative of Drop Tail.

In addition, the behaviour of the adaptive algorithm itself is dependent on careful consideration of the scaling parameters that are used. Finally it should also be noted that the adaptive algorithm only operates on the *MaxP* parameter and so consequently will not improve the performance of RED with poor *WeightQ* or threshold parameterisation.

References

- [Feng99] W. Feng, D. Kandlur, et al, "A Self-Configuring RED Gateway", Proceedings of INFOCOM '99, March 1999.
- [Floyd93] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking August 1993.
- [Floyd97] S. Floyd, "RED: Discussions of Setting Parameters", from an Email Message, November 1997.

- [Stern97] G. Sternberg and J. Boone, "Network Congestion Control", University of Pennsylvania, April 1997.

Acknowledgments

The author would like to acknowledge the financial support of British Telecom Laboratories, Ipswich, UK during the course of this research.

The Generalized Max-Min Fairness Policy for Elastic Traffic based on Linear Programming

I. Moscholios, M. Logothetis and G. Kokkinakis

WCL, Dept. of Electrical & Computer Engineering,
University of Patras, 265 00 Patras, Greece.
E-mail: m-logo@wcl.ee.upatras.gr

Abstract Most of elastic traffic calls require a rate between certain limits, while it is essential for a bandwidth allocation mechanism to allocate the available bandwidth in a fair manner among the elastic calls. Therefore, the min and max bandwidth requirements of calls have to be taken into account in the bandwidth allocation process. Such a process is the Generalized Max-Min Fairness policy (GMM). It has been described without mathematical support in a five-step procedure, in which the fairness criteria are given without coming into details. In this paper, we propose a new algorithm for the GMM policy, in a clear mathematical way, based on Linear Programming (LP), and it is directly convertible into computer software. To clarify the proposed new algorithm, we present selected numerical examples on a well-established testbed network.

1. Introduction

Elastic traffic sources, such as Internet traffic sources using TCP and sources using the Available Bit Rate (ABR) service in ATM networks, can adjust their rates according to a flow/congestion control mechanism, which allocates to them the available bandwidth left over from non-elastic services. This bandwidth needs to be allocated fairly among competing elastic traffic sources (calls), thus guaranteeing that the quality of service in the network remains in an acceptable level [1].

The most popular Fair Bandwidth Allocation (FBA) policy is the Max-Min Fairness (MMF) policy [2], initially adapted by the ATM Forum to allocate available network bandwidth among ABR connections. In the MMF policy, a bandwidth allocation is said to be max-min fair, if the bandwidth allocated to a connection cannot be increased without decreasing at the same time the rate of a connection having a smaller or equal bandwidth allocated (MMF criterion). The main “drawback” of the MMF policy is that it does not take into account the Minimum Cell Rate (MCR) and Peak Cell Rate (PCR) elastic traffic description parameters. This drawback has been well faced by an extension of the MMF policy, the Generalized Max-Min (GMM) [3], which supports MCR and PCR for each call connection.

The GMM policy has been described in procedural steps, by expressing the philosophy of the

GMM policy, without coming into details. This has the advantage of an easy presentation, but fails to describe the policy in a clear mathematical way. The latter can be achieved with the aid of the Linear Programming (LP), which shows clearly the optimization function and the set of constraints. A simple LP model has already been proposed ([4],[5]) for the description of the MMF policy (without the MCR, PCR or weights). Unfortunately, the GMM policy cannot be applied through the MMF LP model, simply by adding the traffic parameters (MCR, PCR) constraint-set. For the description of the GMM policy through LP, we propose a new algorithm. First, we present it in procedural steps (without coming into details), by expressing the key ideas (philosophy of our algorithm) of how the LP model is formulated. This is in order to show the basic differences between the new and the initial GMM algorithm. On the other hand this comparison shows that the new algorithm describes in fact the GMM policy. Then, we give the detailed algorithm in the form of flow-chart. Although the new algorithm is more complicated than the initial GMM procedure ([3]), it has the advantage of the direct computer implementation.

The structure of this paper is as follows. In section 2, we review the GMM policy and include the initial step-by-step procedure whereby this policy can be described. In section 3 we propose the description of the GMM policy through an LP model. We present a step-by-step procedure describing our methodology in a similar way to the initial step-by-step procedure. We discuss the main differences between them and afterwards we give the exact algorithm (in the form of flow-chart) for the LP model formulation. Section 4 is the numerical section, where we give application examples considering a well-known testbed network. We conclude in section 5.

2. Overview of the GMM policy

The GMM policy satisfies the MCR of each connection and then tries to maximize the rate of the smallest connection among all connections, while satisfying this connection’s PCR. Given the best smallest (max) rate allocation, it continues by maximizing the rate of the connection with the second smallest rate and so forth. The GMM rate allocation algorithm can be described in 5 steps [3]:

1. Start the rate of each connection with its MCR.
2. Sort all connections in an increasing rate order.
3. Increase the rate of the connection with the smallest rate among all connections until one of the following events takes place:
 - The rate of this connection reaches the second smallest rate among all connections;
 - A link saturates;
 - The connection's rate reaches its PCR.
4. If a link saturates or the connection's rate reaches its PCR in the previous step, remove from the network: a) the connections that either traverse the saturated link or reach their PCRs (respectively) and b) the network capacity associated with these (the removed) connections.
5. If there is no connection left, the algorithm terminates; otherwise it goes back to step 3 for the remaining connections and network capacity.

3. LP description of the GMM policy - Proposed new algorithm

In order to describe mathematically the GMM policy through an LP model, the following notation is necessary: Let L be the set of network links, S be the set of calls accommodated in the network and S_l be the set of calls which use link $l \in L$. Calls are conveyed by a fixed routing scheme. A call $s \in S$ may require a number of connections, n_s , of the same bandwidth requirements; that is, in this paper, a call is considered as a group of connections with the same bandwidth requirements. We denote by C_l the bandwidth-capacity of link $l \in L$, and by r_s , the bandwidth (rate) which is allocated to each connection of call $s \in S$ (the decision variable of the LP model). The final value of r_s (allocated rate) is denoted by r_s^* . The value r_s^* should be between the minimum, $r_{s,min}$, and the maximum, $r_{s,max}$, bandwidth requirements of call $s \in S$, whereas the vector $\mathbf{r}^* = (r_1^*, r_2^*, \dots, r_s^*)$ should satisfy the MMF criterion. Furthermore, the vectors $\mathbf{r}_{min} = (r_{1,min}, r_{2,min}, \dots, r_{s,min})$ and $\mathbf{r}_{max} = (r_{1,max}, r_{2,max}, \dots, r_{s,max})$ contain the minimum and the maximum permissible rates of all network calls, respectively.

The LP model formulation is distinguished in two phases, initial and final. In the initial phase, the LP model consists of two constraint sets, which are readily defined: the Link Capacity constraints (Set 1), and the Traffic Parameters constraints (Set 2). In the final phase, the minimum permissible rates in Set 2 are modified, while the LP model includes one more set of constraints, the Fairness constraints (Set 3). It is a set of equality constraints between the rates of the connections (in the initial phase of the LP model, Set 3 is empty) and is defined by the proposed new algorithm together with the minimum rates modifications in Set 2. The modifications in Set 2 are done so that the finally allocated rates satisfy the MMF criterion. The objective function of the LP model is to maximize the sum of connections rates.

The constraint sets:

$$\text{Link Capacity Set: } \sum_{s \in S_l} n_s r_s \leq C_l, \text{ for all } l \in L \quad (1)$$

$$\text{Traffic Parameters Set: } r_{s,min} \leq r_s \leq r_{s,max} \text{ for all } s \in S \quad (2)$$

$$\text{Fairness Set: } \text{null} \quad (3)$$

$$\text{Objective function: } \sum_{s \in S} r_s \Rightarrow \max \quad (4)$$

The proposed new GMM rate allocation algorithm is described in 8 steps:

0: Define the **Initial** phase of the (above) **LP model**.

1: Set $\mathbf{r} = \mathbf{r}_{min}$

2: In a vector R , sort the rates of all connections, in the order of increasing rate, including:

- **Minimum** rates
- **Maximum** rates
- **Quantized** values of the available **link bandwidth**, according to the number of connections that traverse the link (see Fig. 1)

3: **Increase** the rate of all connections to the next level (value) of R , while satisfying the minimum and maximum rates of each connection, until the rates reach the upper level of R , R_{max} , or some link saturates. R_{max} is not defined as the maximum value of R , but as: a) the maximum value of minimum rates when the current rates are not equal to one another, or b) the maximum value of the quantized values of the available bandwidth, when the current rates are equal to one another.

4: **If** some **link saturates then remove** from the network the connections that traverse that link as well as the network capacity occupied by such connections and return to step 1 via step 5.

5: **If** the connections of two different calls have the **same rate** (less than their maximum one) and traverse the same link **then add** an **equality constraint** between them in the Fairness set (Set 3 of the LP model) and return to step 1 after updating the minimum rates with the current rates ($\mathbf{r}_{min} = \mathbf{r}$).

6: Having reached the rates up to the level R_{max} , **solve** the initial LP model (i.e. find \mathbf{r}^*), as it has been modified up to this step (basically in Traffic Parameters set and Fairness set).

7: **Give the excess link bandwidth** to the \mathbf{r}^* . This is done by solving the LP model of the initial phase (Set 3 = *null*) while substituting the \mathbf{r}_{min} in Set 2 by the \mathbf{r}^* (defined in Step 6).

To solve the LP models, we use the classical Simplex method or the primal cutting-plane algorithm (when the bandwidth is quantized) [6].

The main difference in philosophy of this procedure from the previous one (section 2), is the proper selection of bandwidth levels and the increase of all connection rates simultaneously, level by level, instead of increasing, level by level, the rates of that calls (in priority) with the smallest bandwidth requirement. The proposed new algorithm is described in detail in the flow chart of Fig. 1.

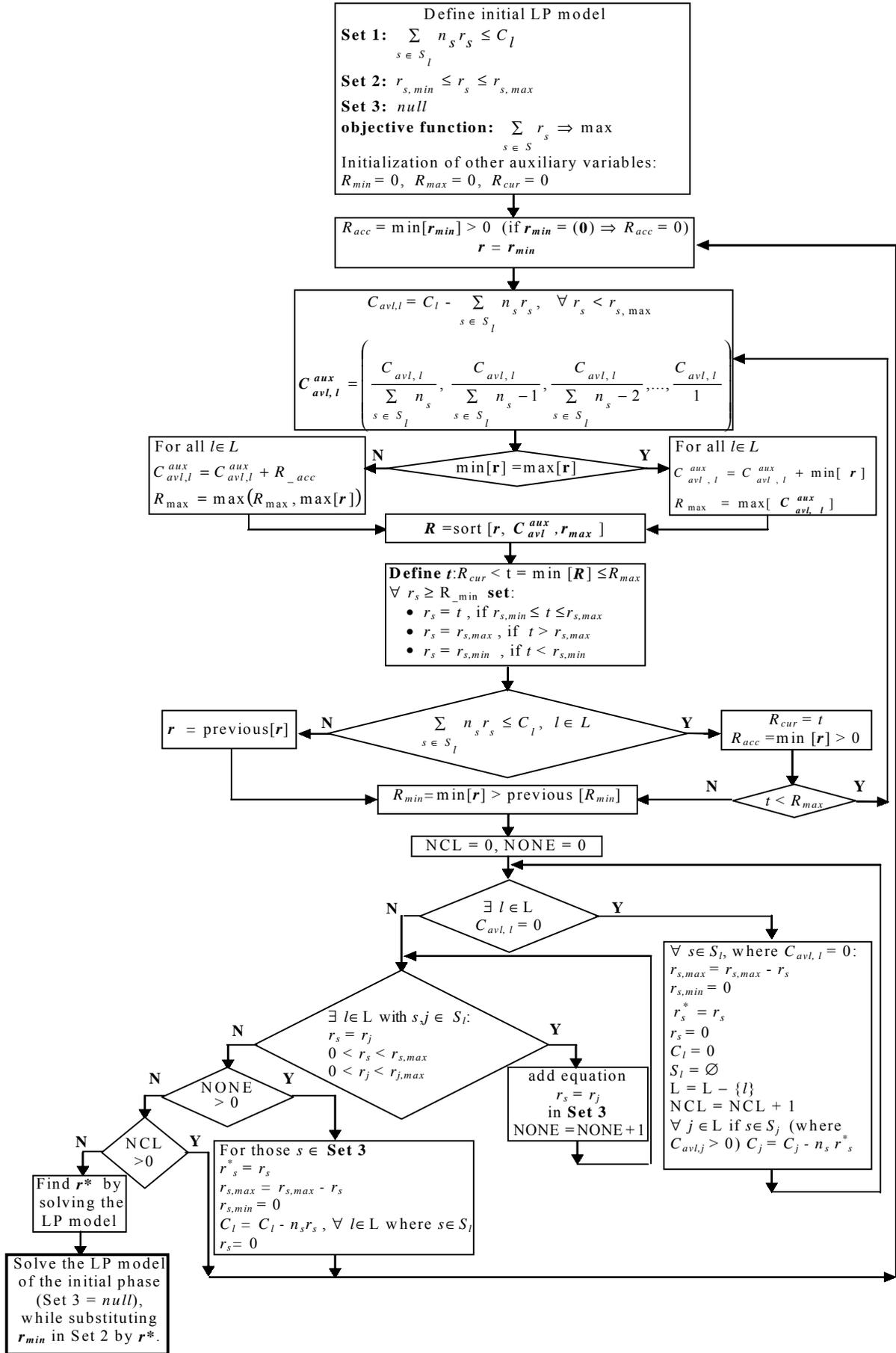


Fig. 1: The flow chart for the LP implementation of the GMM policy.

4. Numerical Examples

We consider the testbed network of Fig. 2, which has been adopted by the ATM Forum for performance comparison of various fairness criteria [7], and apply on it the proposed new algorithm, in order to define the GMM bandwidth allocation among the call connections. As Fig. 2 shows, the testbed network accommodates six calls: A, B, ..., F. Table 1 shows the number of connections, n_s , that each call consists of, as well as the minimum and maximum bandwidth requirements of each calls. More precisely, we consider three pairs of vectors (r_{min} , r_{max}), which are shown in Table 1, as cases 1, 2 and 3.

The constraint sets of the Initial phase of LP models:

$$\left. \begin{array}{l} 3r_A + 6r_D \leq 50 \\ 3r_A + 3r_B + 2r_F \leq 150 \\ 3r_A + 3r_B + 3r_C \leq 150 \\ 3r_B + 6r_E \leq 100 \end{array} \right\} \text{Set 1}$$

$$\left. \begin{array}{l} \text{Case 1} \quad \text{Case 2} \quad \text{Case 3} \\ 4 \leq r_A \leq 34 \quad 2 \leq r_A \leq 12 \quad 2 \leq r_A \leq 7 \\ 7 \leq r_B \leq 22 \quad 4 \leq r_B \leq 18 \quad 4 \leq r_B \leq 10 \\ 24 \leq r_C \leq 35 \quad 12 \leq r_C \leq 20 \quad 12 \leq r_C \leq 20 \\ 5 \leq r_D \leq 9 \quad 1 \leq r_D \leq 5 \quad 1 \leq r_D \leq 5 \\ 8 \leq r_E \leq 24 \quad 3 \leq r_E \leq 14 \quad 3 \leq r_E \leq 14 \\ 20 \leq r_F \leq 100 \quad 10 \leq r_F \leq 100 \quad 10 \leq r_F \leq 100 \end{array} \right\} \text{Set 2}$$

$$\text{null } \left. \right\} \text{Set 3}$$

Sets 1, 3 are common to all LP models (cases 1, 2, 3).

Running the new GMM algorithm for each case, the following LP models (Final phase) are formulated.

The constraint sets of the Final phase of LP models:

Set 1 is the same with that of the Initial phase.

$$\left. \begin{array}{l} \text{Case 1} \quad \text{Case 2} \quad \text{Case 3} \\ 5.56 \leq r_A \leq 34 \quad 6.67 \leq r_A \leq 12 \quad 6.67 \leq r_A \leq 7 \\ 11.11 \leq r_B \leq 22 \quad 6.67 \leq r_B \leq 18 \quad 6.67 \leq r_B \leq 10 \\ 24 \leq r_C \leq 35 \quad 12 \leq r_C \leq 20 \quad 12 \leq r_C \leq 20 \\ 5.56 \leq r_D \leq 9 \quad 5 \leq r_D \leq 5 \quad 5 \leq r_D \leq 5 \\ 11.11 \leq r_E \leq 24 \quad 6.67 \leq r_E \leq 14 \quad 6.67 \leq r_E \leq 14 \\ 24 \leq r_F \leq 100 \quad 12 \leq r_F \leq 100 \quad 12 \leq r_F \leq 100 \end{array} \right\} \text{Set 2}$$

$$\text{null } \left. \begin{array}{l} r_B = r_E \\ r_B = r_E \end{array} \right\} \text{Set 3}$$

Since for some $s \in S$, the $r_{s,min}$ shown in Set 2 of the Final phase of the LP models is the final allocated rate to the call s , because of link saturation (e.g. case 1: $r_{D,min} = 5.56$) or because of $r_{s,min} = r_{s,max}$ (e.g. case 2: $r_{D,min} = 5$), the above LP models can be simplified. This simplification is taken in to consideration by the proposed new algorithm. However, we keep their above form for presentation purposes. Table 2 gives the GMM solution r^* of the LP model for each case, as well as the final allocated bandwidth r^* . The LP solution results (in all cases) by the well-known Simplex method.

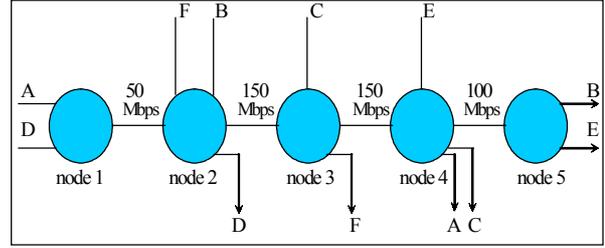


Fig. 2: Generic fairness configuration

Table 1: Elastic traffic parameters for calls A to F.

Call	n_s	r_{min} (Mbps)			r_{max} (Mbps)		
		Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
A	3	4	2	2	34	12	7
B	3	7	4	4	22	18	10
C	3	24	12	12	35	20	20
D	6	5	1	1	9	5	5
E	6	8	3	3	24	14	14
F	2	20	10	10	100	100	100

Table 2: GMM bandwidth allocation (the solution)

Call	Case 1		Case 2		Case 2	
	r^*	final r^*	r^*	final r^*	r^*	final r^*
A	5.56	5.56	6.67	6.67	6.67	6.67
B	11.11	11.11	11.11	11.11	10.0	10.0
C	33.33	33.33	20.0	20.0	20.0	20.0
D	5.56	5.56	5.00	5.00	5.00	5.00
E	11.11	11.11	11.11	11.11	10.0	11.67
F	50.00	50.00	48.33	48.33	50.0	50.0

5. Conclusion

We present the new GMM algorithm in procedural steps to describe its philosophy and in a flow chart to be directly convertible into software. Moreover we support it mathematically based on LP models.

References

- [1] I. Moscholios, M. Logothetis, "Call-level QoS assessment in ATM networks supporting elastic traffic", Proc. ICC2001, Helsinki, June 2001.
- [2] D. Bertsekas, R.Gallagher, "Data Networks", New Jersey, Prentice Hall, 1992.
- [3] Y.T. Hou, Bo Li, S.S. Panwar and H. Tzeng, "On network bandwidth allocation policies and feedback control algorithms for packet networks", Computer Networks 34 (2000), pp. 481-501.
- [4] I. Moscholios, M. Logothetis and G. Kokkinakis, "Evaluation of fair bandwidth allocation policies for elastic traffic", ACTA Tehnica Napocensis Electronics and Telecommunications Journal Vol. 43, No. 2, pp. 12-19, November 2002.
- [5] W. Ogryczak, T. Sliwinski and A. Wierzbicki, "Fair resource allocation schemes and network dimensioning problems", Journal of Telecommunications & Information Technology, pp. 34-42, 3/2003.
- [6] H.M.Wagner. "Principles of Operations Research", Prentice-Hall: Englewood Cliffs, New Jersey 1969.
- [7] R. Simcoe, "Test Configurations for fairness and other tests", ATM Forum/ 94-0557, July 1994.

Gigabit Ethernet over OC-48 Transponder for Metro-Optical Access

Po-Chun Chiang, Kevin Huang, Chih-Feng Cheng

Optical Communication and Networking Technology Dept.,

CCL/ITRI, 195-11 Sec. 4, Chung Hsing Rd., Chutung, Hsingchu, Taiwan 310, R.O.C.

e-mail : {pcchang, jeffyjeng, kevinh}@itri.org.tw

Abstract--For the transmission quality and convenience of the network between MAN and WAN, we design the Gigabit Ethernet Transponder system. The Gigabit Ethernet optical signal transmitted and received on Gigabit Ethernet Transponder is up to an OC-48 SDH/SONET signal, so the Gigabit Ethernet Transponder system was a very complex and novel system. Our Gigabit Ethernet Transponder system includes hardware and software design. This paper will present the function and architecture of each unit in our Gigabit Ethernet Transponder system.

Keywords--Dense Wavelength Division Multiplexing, Synchronous Digital Hierarchy, System Packet Interface Level 3.

I . INTRODUCTION

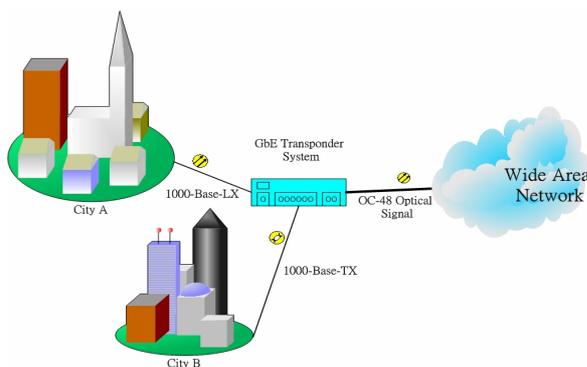


Figure 1. The application of Gigabit Ethernet Transponder.

In local area networks, Ethernet has become the most popular technology which has the benefits of easy management, cost effective, etc. Because of these benefits, many service providers have started to offer Ethernet for network access by extending Ethernet into metro network. So how to provide the package-based traffic on the traditional TDM trunk

network is a big challenge. Ethernet over SDH/SONET [4], [5] using the PPP/HDLC [1] is one of the most popular solutions to reduce the complexity of current trunk network. It can keep the important feature of traditional SDH/SONET network. Figure 1 showed the application of the Gigabit Ethernet Transponder system.

In this paper, we introduced the technology, which can transmit Gigabit Ethernet signal to the OC-48 SDH/SONET signal.

II . SYSTEM ARCHITECTURE

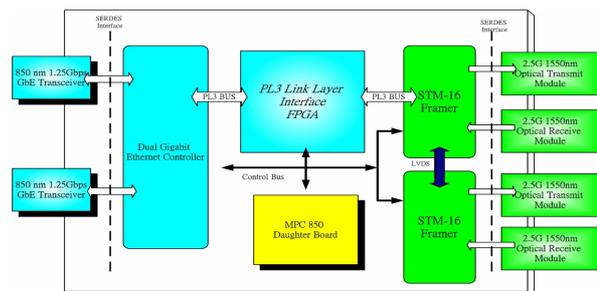


Figure 2. The architecture of the Gigabit Ethernet Transponder.



Figure 3. The picture of the Gigabit Ethernet Transponder.

The Gigabit Ethernet Transponder system can be separated into three parts, which are the Gigabit Ethernet traffic process unit, OC-48 traffic process unit and the micro-controller unit [2], [3]. Two commercial chips, Gigabit Ethernet Controller and STM-16 SDH/SONET Framer, are used in our system. The architecture of the Gigabit Ethernet Transponder is showed in figure 2 and the figure 3 was the picture of Gigabit Ethernet Transponder system.

A. Gigabit Ethernet Traffic Process Unit

The Gigabit Ethernet traffic was received and transmitted by the 850nm Gigabit Ethernet optical transceiver. The serial bit streams from the optical transceiver will be sent to the Gigabit Ethernet Controller. There are two major functional blocks in the controller, the SERDES and the GMAC.

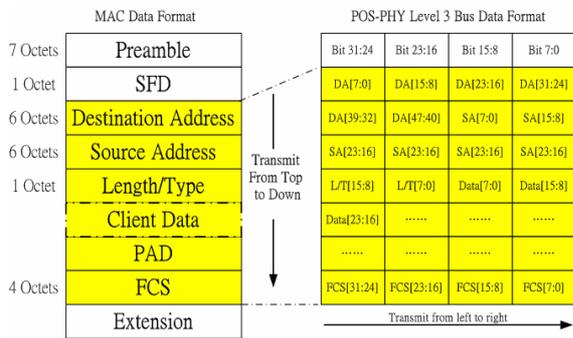


Figure 4. The mapping of the MAC frames to the SPI-3 interface.

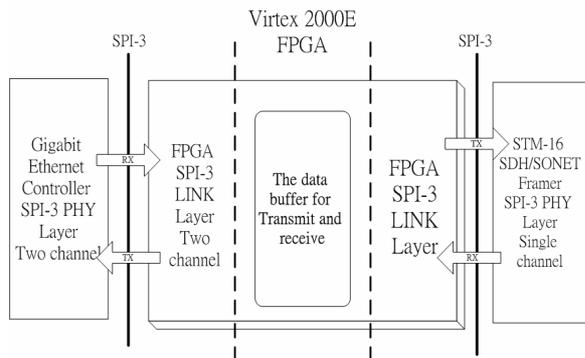


Figure 5. The internal architecture of SPI-3 FPGA.

The SERDES (Serializer-Deserializer) function block was IEEE 802.3-2000 [6] Gigabit Ethernet compatible, and supported Gigabit Ethernet data transfer flows. After the SERDES block, the parallel data coding by 8b/10b will be transferred to the 8b data, and then be sent to the GMAC (Gigabit Media Access Controller). In the GMAC block, there were two functions, flow control and address flitting which can be controlled by user. The format of Gigabit

Ethernet data that will be transferred to the MAC data format is showed in the figure 4. As the Gigabit Ethernet Controller provide a big endian OIF SPI-3 standard [7] interface, figure 4 also showed the data format on the SPI-3 interface.

Both the SPI-3 interface of the Gigabit Ethernet Controller and the STM-16 SDH/SONET Framer are the physical layer interface. Therefore those two chips can not connect and communicate directly. One SPI-3 link layer interface device is necessary to control the data stream between them. We use a Xilinx Virtex 2000E FPGA to connect with SPI-3 interface of these two chips. The internal architecture block was showed in the Figure 5.

The figure 6 showed below was the FSM (Finite State Machine) of the SPI-3 interface operation. This FSM worked according to the SPI-3 standard. There were two buffers in the upstream and downstream direction. Each of them was 64k bytes. If the buffer was underflow or overflow, then the Ethernet package in the buffer will be dropped. As the transmission rate on each signal bit of SPI-3 was above 100MHz, therefore the PCB (Printed Circuit Board) of SPI-3 interface design had to concern about the path length on each signal bit.

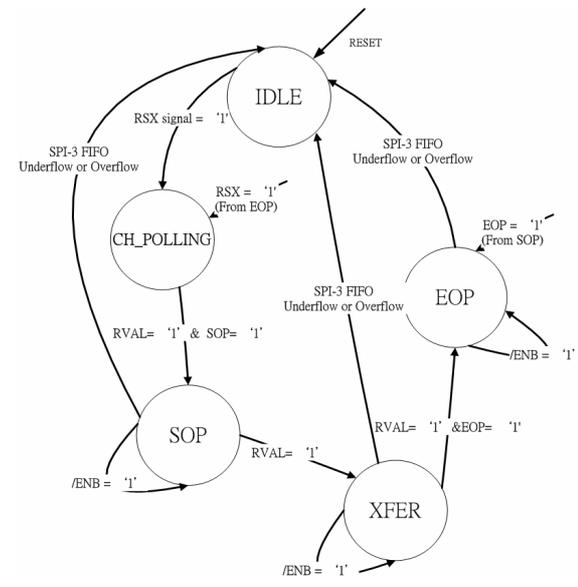


Figure 6. The FSM of SPI-3 interface operation.

B. OC-48 Traffic Process Unit

This unit was used to provide the 2488.32 Mb/s STM-1-16c data stream that made the packets to be aggregated by SPI-3 interface, and implemented Point-to-Point Protocol (PPP) over SDH. Figure 7 presented how the HDLC frame mapping into the STM-16 SDH/SONET frame. The APS (Automatic Protection Switch) port was also implemented for the optical fiber protection.

In the OC-48 traffic process unit, IP traffic was carried over PPP in the HDLC-like frame. The

boundary of each frame was found by searching for the flag character. Flags were also used to fill inter-packet spacing. This octet stream was mapped into the SDH VC-4 with the octet boundaries aligned with the SDH VC-4 octet boundaries. Figure 8 showed the STM-16 signal that carried the HDLC frame data.

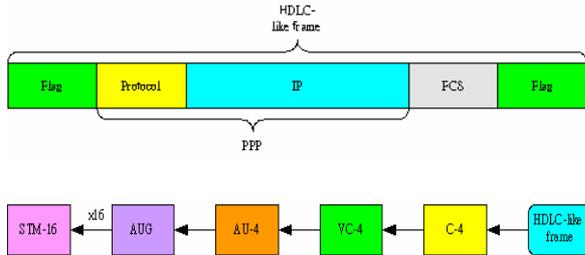


Figure 7. The HDLC frame format and how it mapping into the STM-16 SDH/SONET frame.

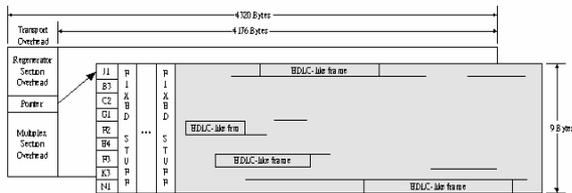


Figure 8. The HDLC frame data on the STM-16 frame payload.

The system needed a 155.52 MHz differential reference clock for both the clock recovery and clock synthesis circuits. The jitter in our system requirement must be less than 1 Pico-sec RMS in band of 12 KHz to 20 MHz in order to comply with Bellcore GR-253 intrinsic jitter specification. For the reason above, a VCXO with PLL circuit was implemented. The figure 9 was the PLL circuit of 155.52MHz on the Gigabit Ethernet Transponder.

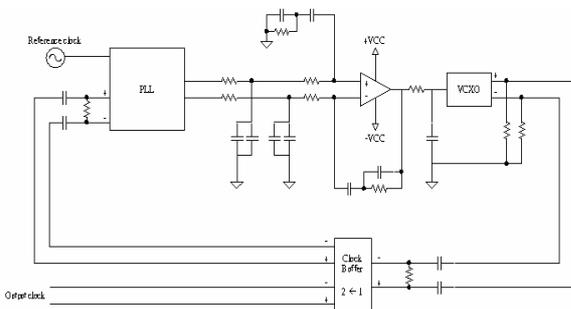


Figure 9. The PLL circuit on Gigabit Ethernet Transponder that worked at 155.52MHz.

A pair of OC-48 optical transmitter and receiver was used in the Gigabit Ethernet Transponder system. The optical transmitter and receiver modules work at 1549.32 nm, so this system can be aggregated to the

DWDM system. Figure 10 and 11 were the optical spectrum and the eye diagram of the OC-48 transmitter respectively. The eye diagram conformed to the standard specification.

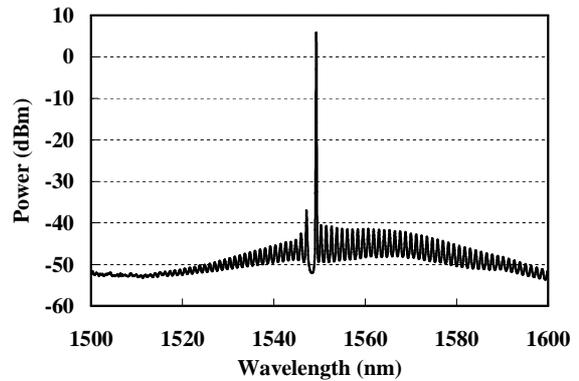


Figure 10. The optical spectrum diagram of OC-48 DWDM transmitter.

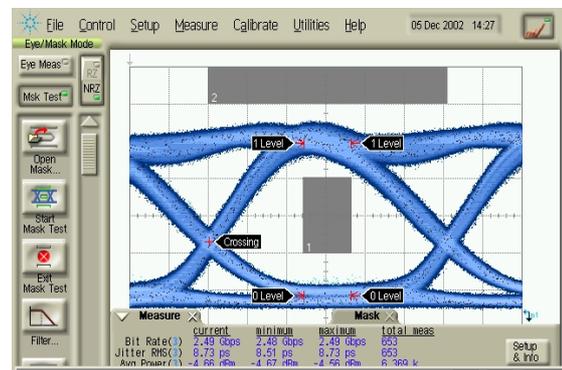


Figure 11. The eye diagram measurement of OC-48 DWDM transmitter.

C. The Software Control Unit

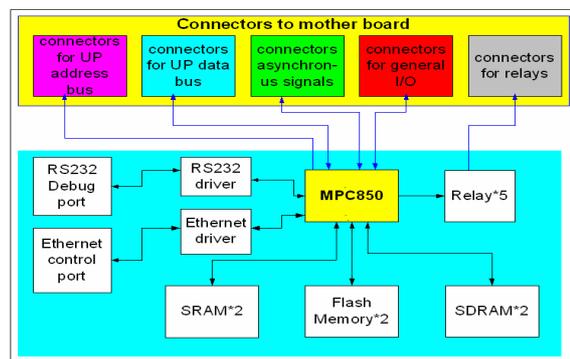


Figure 12. The architecture of MPC850 control unit.

We used Hard Hat Linux as our software platform. It was a standard, off-the-shelf binary distribution of Linux, designed for embedded computer systems, published by MontaVista Software Company.

There is a MPC850 control board on our system for software platform. The functional block showed in figure 12. We use Motorola Power-PC MPC850 to be the core controller of this unit. MPC850 was an integrated communication microprocessor, and provides configurable SCC (Serial Communications Controllers) unit. Each SCC unit can be configured to any kind communication protocol, like HDLC, UART, USB, I2C, and Ethernet. For our system, we configured one of the SCC to be UART port, another to be 10Mbps Ethernet port.

The Linux RTOS (Real Time Operation System) was stored in the Intel flash memory on the MPC850 control board, and communicated with the Gigabit Ethernet Controller and STM-16 SDH/SONET Framer by the micro-controller interface.

III. TEST ENVIRONMENT

The diagram of the test environment is showed as the figure 13. The Nortel OM5200 is a SDH/SONET testbed that can provide OC-3, OC-48, OC-192 and Gigabit Ethernet services. There are three sites in our laboratory, so we can use them to simulate the environment of MAN. First, we used the STM-16 tester to verify the STM-16 signal generated by Gigabit Ethernet Transponder system. Then, we connected the Gigabit Ethernet Transponder to the testbed. A VoD (Video on Demand) server was connected to the testbed by the Gigabit Ethernet switch. There are two VoD client PCs. Each of them installed a Gigabit Ethernet NIC (Network Interface Card) and connected with the Gigabit Ethernet Transponder. Both of the client PCs can request a Video service to the VoD server through our Gigabit Ethernet Transponder. Immediately, the video signal will show up on the PC. So, we can verify each functional unit of the Gigabit Ethernet Transponder in this test environment.

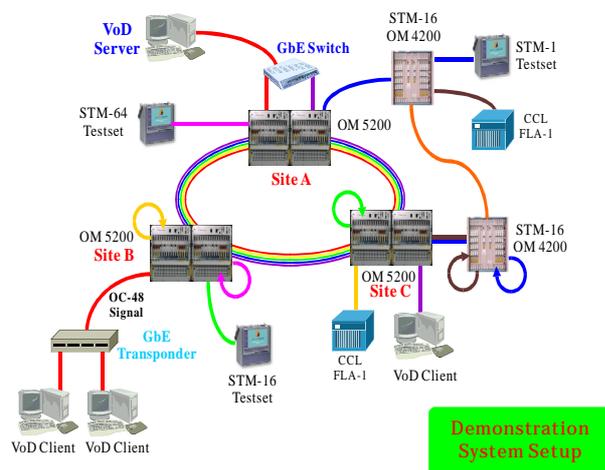


Figure 13. The test environment of the Gigabit Ethernet Transponder

IV. CONCLUSIONS

The new ideas that presented in this paper, included the MPC850 daughter board, POS-PHY Level 3 link layer FPGA, firmware and the system control software. All of these functions were implemented correctly on the Gigabit Ethernet Transponder system. In the Gigabit Ethernet Transponder, two Gigabit Ethernet traffic were combined then translated to OC-48 SDH/SONET traffic. So, the Gigabit Ethernet Transponder system is a good and non-expensive solution for providing the connection between MAN and WAN.

REFERENCES

- [1] A. Malis and W. Simpson, "PPP over SONET/SDH", *RFC 2615*, June 1999.
- [2] COBNET consortium, "Physical and networking Layer issues of COBNET", pp. 1-56, March 1996.
- [3] COBNET consortium, "Hardware specifications and node design for the COBNET demonstrator", pp. 1-65, Sep. 1996.
- [4] ITU-T Rec. G.707, "Network Node Interface for the Synchronous Digital Hierarchy (SDH)", Oct. 2000.
- [5] ITU-T Rec. G.783, "Characteristics of synchronous digital hierarchy (SDH) equipment functional blocks", Oct. 2000.
- [6] IEEE 802.3-2000 Gigabit Ethernet, 2000 Edition
- [7] OIF-SPI3-01.0, System Packet Interface Level 3 (SPI-3): OC-48 System Interface for Physical and Link Layer Devices.

Graphical Simulation Tool for Mobile Networks Routing

Kamlesh Dutta, Manik Karol, Prakul Sharma
kd@recham.ernet.in
{manikkarol|prakulsharma}@hotmail.com
National Institute of Technology,
Hamirpur 177005 (HP) INDIA

1. Introduction

In this paper we present a graphical simulation tool for mobile network routing. We chose to design this framework in such a way that allows easy modifications of the main parameters and, at the same time, the implementation details are transparent to the user. The structure of the cells and other details that are considered are similar to the actual mobile network so that the modeling is as close to real mobile networks as possible[1]. The simulator also maintains a datasheet of database called home-location register (HLR) and visitor-location register (VLR) of the visited-system, the new location. The code is still under development. This paper gives an outline of what the simulation tool presently offers and what features can be further added.

2. Related Work

In [2] low cost mobile network simulation approaches are presented. Terminode routing scheme is presented in [3]. This routing scheme is designed for wide area networks, where a large part or all the nodes are mobile. Terminode routing is a combination of two protocols called Terminode Local Routing (TLR) and Terminode Remote Routing (TRR). TLR is used to route packets to close destinations. TRR is used to route to remote destinations and is composed of the following elements: Anchored Geodesic Packet Forwarding (AGPF), Anchored Path Discovery (APD), multipath routing and path maintenance. The simulations of the TLR and TRR protocols using the GloMoSim simulator are performed. Network simulator ns-2 presented in [5] is essentially a wireless model consisting of the MobileNode at the core, with additional supporting features that allows

simulations of multi-hop ad-hoc networks, wireless LANs etc.

3. Simulator for Routing in Mobile Networks

In this section we will briefly introduce the simulator and its main characteristics that we are developing.

The simulator shows a small map where we have divided the area into various sub-areas with each having its own mobile base station shown as a bigger circle. The mobile users are shown by smaller circles. Each base station has its perimeter i.e. a hexagonal cell. If the mobile user comes within a perimeter range of a base station, the color of the user becomes same as that of the base station. This has been defined by using the clustering algorithm. Each station has its own color and thus it is easy to detect the position of the mobile user in the area.

The mobile users are in continuous motion and the random movements of these users in the graph show this effect. The clustering algorithm is applied and the mobile users keep changing their region denoted by change in colors according to the base station perimeters.

When the simulator is started we have a small option window (Fig. 1) that is displayed. In this window we have certain options like maximum hop length, time interval, size of the cells etc. These are chosen by the user. Some of them can also be varied even when the simulation is in progress.

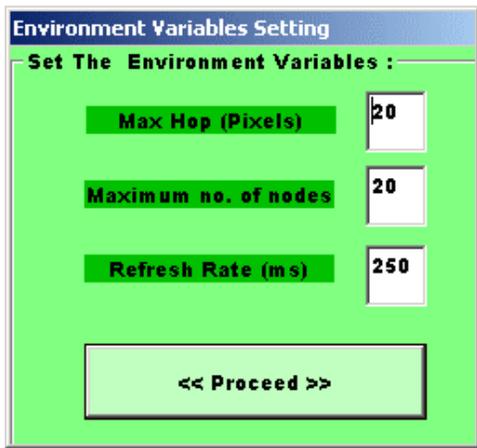


Fig 1 Environment Variable Setting Window.

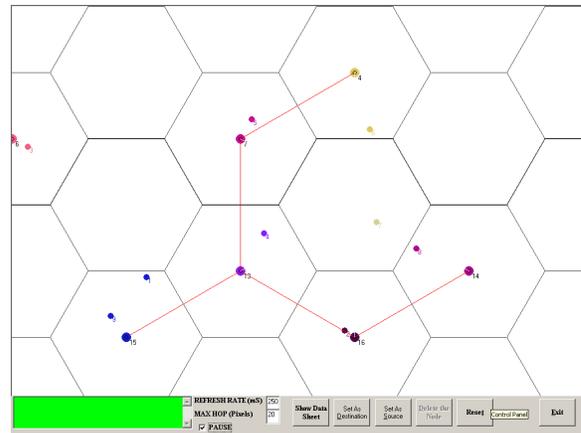


Figure 2.2 Screen shot after a small time interval (shows dynamism)

Initially there are three mobile users. We can increase the number to a maximum. Simple left clicking adds one node (at a time) at the location of the mouse pointer on the field graph. A graph connecting the adjacent base station changes dynamically as the positions of the mobile users changes. This graph is used for the routing in the system

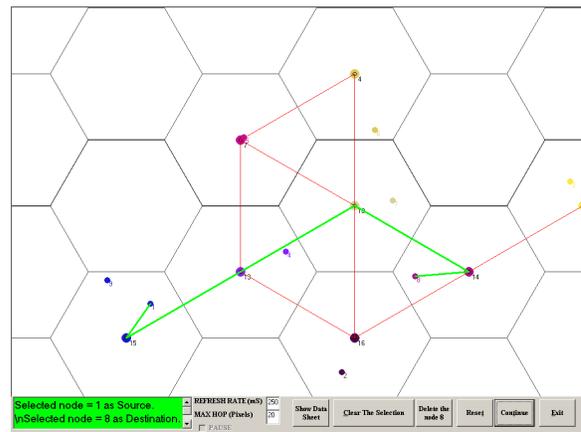


Figure 2.3 Shows the route in different color.

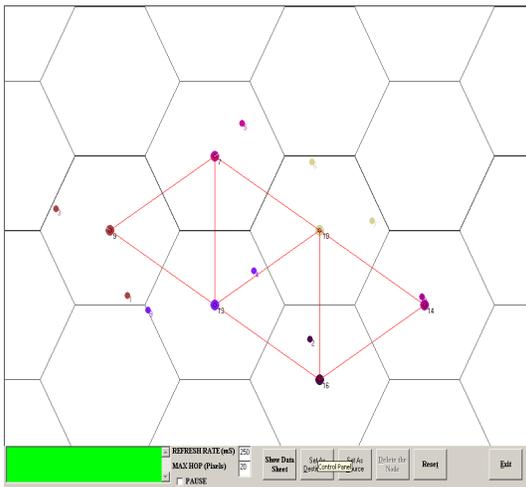


Figure 2.1 Clustering of Mobile Users

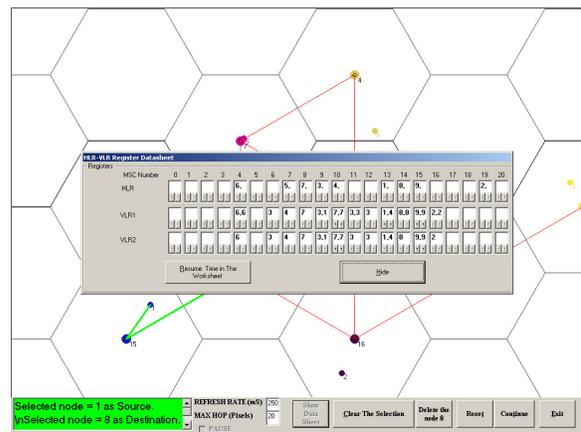


Figure 2.4 Shows the Data Sheet of the network.

A Control Panel at the bottom of the screen has following buttons

Set Destination is used to set the selected node (selected by clicking on the node) as destination node for communication on the mobile network.

Set Source is used to set the selected node as source node for communication.

Pause is used to halt the system so that it is easy to select the source and destination nodes in the system. Then the system could be started again.

Show Data Sheet is used to show the values of the home-location register (HLR) and visitor-location register (VLR) for the various mobile hosts in the network (fig. 2.4). (When a user subscribes the services, a record is created in the system's database called home-location register (HLR). When the mobile user visits a new registration area (RA), a temporary record for the mobile user is created in the visitor-location register (VLR) of the visited-system (the new location). The VLR then sends a registration message to the HLR[2]).

Delete node is used to delete the selected node from the mobile network.

Show route is used to show the route between the source and destination.

Reset is used for resetting the whole system to initial stage with only three nodes.

Clear is used to clear the route that is drawn between the two nodes.

Exit is used for exiting from the system.

The text window at the left hand corner of the screen shows the current status of the nodes i.e. which node is source & which one is destination.

Drag-N-Drop feature helps us simulate the disconnection phenomenon in mobile networks i.e., a user can be selected and using the mouse, can be dragged and dropped into a new cell after disconnecting from its current cell. It is then registered in the new cell as a visitor.

Routing is carried out with the help of a graph that has been generated. If communication is to be carried out between the two nodes, then this is possible only if there exists a path between two

nodes. If there is no path then no routing is possible. Since the graph is dynamic so it may happen that the path that the packet was traversing may change, so to encounter this problem the packet may be lost or may take a new path if so possible. Graphical Display shows the addition of new nodes (fig 2.1), and a snapshot after some time when the position of mobile nodes has changed resulting in new graph is shown in fig 2.2. The path between source and destination nodes is depicted by different colour scheme (fig. 2.3).

This simulator and routing technique used to demonstrate is in the initial stage and changes will be made so as to overcome the shortcoming of the present techniques.

4. Simulator Design

The Simulator designed has been divided into many functions each performing a specific task. Some of the most important are described below.

Segmentation: This function is used to divide the available canvas into various cells. It basically segments the region into hexagonal segments. One of the nodes in this segment is the base station. The idea behind this is to simulate the cell structure that is used in the mobile networks. We consider the centre of the hexagon as the base station. We use it as a reference to draw the hexagonal cell. A mobile host lies in the cell if its distance from the centre of that cell is less than or equal to the boundary of the hexagon.

Timer1_Timer: Control the timing parameter in our simulator. It adjusts the delay for the movement of the mobile hosts in the simulator. This gives dynamism to the mobile network. We are able to fix the value of the timer initially as well as when the process of simulation is in progress. A special menu is provided of this.

Check Cell: This function is used to find out the current cell of a particular mobile host. Here we check this by comparing the distances of that particular mobile host with all the base stations and the one with minimum distance gives the cell. This is an important function of the simulator. For each mobile host the process of finding its parent cell is as under. We check for its distance with centre of every cell. Then

compare those distances. It is obvious that the mobile host will lie in the cell, which has minimum distance from the centre to that host.

Picture1_Click: This is used to generate a new mobile host by clicking on the location in the region. The mobile host is associated with its cell's base station by calling the function *check cell*. The mobile host is then given the color of its cell, which changes if the host changes its position. The number of mobile hosts in the system is fixed at the starting of the simulation.

Assncolor: This is used to assign different colors to the various cells. When a cell is assigned a color, all mobile hosts in that cell and the base station are given that color. The color is assigned randomly. The color scheme makes it possible to easily visualize the whole scenario of the mobile networking.

Drawgraph: This is used to draw the dynamic graph that represents that circuit. The graph keeps on changing as the mobile hosts change their positions. Packets can only be transferred if and only if the two base stations are adjacent to each other. Otherwise the packets are routed through the graph to reach the destination. The graph is represented as an adjacency matrix. This matrix is then used to find the path between the two mobile hosts. We check of adjacency of each cell with its neighboring cells using the distance between their respective centers.

Erasegraph: This is used to erase the graph for updating. After this function the *drawgraph* function is called that updates the graph. This is important for incorporating dynamism in the simulator. If this is not used the canvas will not be refreshed and number of graphs would be drawn on it.

Adjacent: This is used to generate the adjacency matrix for the graph representing the network. This matrix is then used for routing the packets from one host to another. As discussed in section 1.4 that in adjacency matrix we have rows and columns representing the vertices. We put 1 if the two vertices are adjacent and 0 if not in case of simple graph. In case of graphs with multiple edges between vertices we put the number of edges in the respective element of the matrix. This scheme is followed here.

Hexa: This function is used to plot the hexagons on the canvas area, which shows the basic design

model of the simulator. This function takes the centers of the hexagon as the reference for drawing the hexagons. First the centers are computed and then we use this function to draw the sides of the hexagon according to the size that the user has input at the start of the simulator.

Route: This function is used to show the route from the source to the destination. The source and the destination are selected. We have the adjacency graph for the mobile network. We use the adjacency graph recursively to find the route from source to destination. The process is similar to that of finding the minimum path in a graph. First start from the source and then find its most appropriate neighbor and then take that as the source and repeat the process till we reach the destination. Then we highlight those edges in that graph.

5. Conclusion

The simulator developed shows a general overview of communication in a mobile network. The structure of the cells and other details that are considered are similar to the actual mobile network so that the modeling is as close to real mobile networks as possible. This entire abstract model is applicable to general mobile network communication. We shall implement and test other routing algorithms using this software. As future work we plan to add some more features. The most important one is related to making a library with the algorithms we intend to implement for example fault tolerance, check pointing etc.

References

- [1] Andrew S. Tanenbaum, "Computer Networks", 4th Edition, pages 154-156,372-376.
- [2] L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, and H. Yu, "Advances in Network Simulation," *IEEE Computer* 33(5), 59-67, May 2000.
- [3] Ljubica Blažević, Silvia Giordano, Jean-Yves Le Boudec, "Self organized terminode

routing Simulation” Proceedings of the 4th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems Rome, Italy Pages: 81 - 88 ,2001

[4] Yi-Bing Lin, and Wen-Nung Tsai., “Location Tracking with Distributed HLR’s and Pointer Forwarding” *IEEE Transactions on Vehicular Technology*. Vol. 47, No. 1, pp. 58-64, February 1998.

[5] <http://www.isi.edu/nsnam/ns/>

Graphical User Interface to access RMON probe in Java

Peter James Vial
University of Wollongong
School of Electrical
Computer and
Telecommunications
Engineering
Email:
Peter_Vial@uow.edu.au

Ibrahim Raad
University of Wollongong
School of Electrical
Computer and
Telecommunications
Engineering
Email:
Ibrahim@uow.edu.au

Paul Gallon
University of Wollongong
School of Electrical
Computer and
Telecommunications
Engineering

Abstract

Simple Network Management Protocol (SNMP) is considered the current standard for managing an TCP/IP network, such as a local area network. In the early part of the last decade a Management Information Base (MIB) was defined called the Remote Monitoring MIB (RMON). The School of Electrical, Computer and Telecommunications Engineering has two RMON probes available to monitor and manage a simple TCP/IP network as part of its undergraduate network management laboratory. A network management GUI (Graphical User Interface) has been designed and tested to allow these probes to be managed and accessed in training engineering students. This study outlines the modules used and shows the Java GUI developed for this laboratory.

Keywords: Java, RMON, Network Management, SNMP, LANB/250 Remote EtherMeter

1. Introduction

Network management is considered an important area of undergraduate telecommunications engineering learning given the increasing size and scope of networked systems [1]. This is particularly the case with the growth of the telecommunications network based on the TCP/IP standards called generically the Internet. Standards have been put in place in the late 1980's and early 1990's to allow the managing of TCP/IP based networks using the datagram connectionless transport protocol UDP (User Datagram Protocol). One of the Management Information Bases (MIB) defined in the early 1990's was the Remote Monitoring MIB also called the RMON 1 standard (since another standard called RMON 2 was released latter in the 1990's). The

school has been able to recently obtain two RMON 1 probes which can be used in a simple network management laboratory for the use of undergraduate students. The RMON probes used are NAT LANB/250 Remote EtherMeter probes which operate on Thin Ethernet. As a result a method of communicating and tabulating the statistics collected by the RMON probe was required. With this in mind it was proposed that a Java based program should be developed which was able to collect the data available in the RMON probes MIB and then display some of this data graphically in time based format. This paper examines the developed program and is a further extension to material originally described in [2], building on its use of a set of freely available Java methods for RMON 1 and 2 functionality called JoeSNMP [3].

2. Design of RMON GUI

A project initially undertaken in 2002 and described in [2] was to develop a GUI for SNMPv1 and SNMPv2. It used the routines in Java provided freely by JoeSNMP [3]. This project used the work performed in [2] as a basis for the developed GUI interface to the RMON probe. This involved populating the MIB specification for RMON 1 into the software developed in [2]. A simple graphical plotting feature was programmed using the JSwing package available with Java 2.

The RMON probes used in the laboratory were configured using the Windows HyperTerminal program which is standard with most Windows platforms. The associated installation manuals were used to configure the probes and as this is likely different for different RMON probes, it is not covered here. It should, however, be noted that one reason for

using the available probes rather than purchasing an off the shelf package was the cost of available probes in 2002. This was of the order of thousands of dollars (one quote was \$10,000), which also did not include any source code or graphing ability of trends which is traditional with vendor supplied software (such as Aprisma's Spectrum Software [4]).

The design of the GUI management system for a RMON probe, named RMON/SNMP Interface, was developed in Java using the JSwing package. We used the JBuilder v6 [5] integrated development environment. The program extended [2] by the addition of modules and descriptions associated with the RMON MIB. It contained detailed information regarding each MIB element. This involved a large amount of time and effort populating the JoeSNMP associated MIB on RMON with the data and descriptive detail

To add the data to the JoeSNMP framework two approaches were possible. One was to use Abstract Syntax, another was to use the JTree method available in Java. It was decided to use the JTree method as it provides a mechanism to produce a collapsible list of nodes which is suitable for the Object ID tree nature of the Object identification methodology associated with Abstract Syntax Number 1 and SNMP. Each node in the JTree represents an RMON group or element of that group. The initial few lines of code of the method used is shown in Figure 1.

As shown in Figure 1, the JTree MIB works in conjunction with the module ipdesname.java which uses the toString() method to return the relevant node data to the user. Figure 2 shows the GUI provided using this software. A key feature developed, since this software is intended to be used in a network management undergraduate laboratory, is the expanded details function. This is available under the Details|About menu selection and operates on the current Object Identification (OID) of the MIB tree that is selected in the JTree viewing list area to the left section of the RMON/SNMP interface. It contains details about the name, its OID, syntax, access control, status and a description (taken from the standard). This part of the interface was written using the JDialog class in Java. However, every dialog is dependent on the frame, thus when a frame is destroyed so are its dependent dialogs. The program uses the main root frame of the RMON/SNMP interface which means

that when the frame ceases to be displayed on the screen its dependent dialog will also be removed from the screen [6]. Figure 3 shows the dialog with the *etherHistoryIndex* MIB element.

The data from the RMON probe is returned by a series of Get PDU (Protocol Data Units). The data returned is not in a very human readable format (an example is shown in Figure 2 with the main viewing area full of returned data from the RMON probe). The example in the figure uses the OID .1.3.6.1.2.1.16.2.2.1.6.2.6595 through to .1.3.6.1.2.1.16.2.2.1.6.2.6617 with each being a Counter 32 (32 bit counter element) with a number in brackets representing the detail. In the RMON probe the values .6595 through to .6617 represent index increments or samples and the data in square brackets, for example [19768], is the data. Originally the sample index starts at 1 and is incremented each sample. The data is thus a historical data set that forms a time plot of a 32 bit counter. The counter here represents *etherHistoryPkts*, which is the number of good and bad packets received during the sampling interval. Some MIB elements are not plottable, as they represent a single value, whereas other MIB elements are a series of historical or statistical values that can be plotted. The program used flags to indicate the difference between these two types of MIB elements, and determine if plotting is possible.

If a MIB element is plottable, then pressing the Graph button allows this to occur. Data is then collected from the probe as occurs with the Get button, this time it is used to plot the data. This retrieved data is then read using the getText() command into a string construct. It was then necessary to search the string and find the numerical values. The data was displayed in this manner:

```
.1.3.6.1.2.1.16.1.1.3.1 ----->IntegerRange[10]
.1.3.6.1.2.1.16.1.1.3.2 ----->IntegerRange[2]
.1.3.6.1.2.1.16.1.1.3.3 ----->IntegerRange[50]
```

As a result a search of the string was required to find the opening and closing square brackets('[' and ']') so that the data could be extracted. Here this was done by finding the position in the string where the last opening square bracket was located using:

```
int end=line.lastIndexOf("[")
```

```

private void createNodes ( DefaultMutableTreeNode top) {
    DefaultMutableTreeNode Org = null;
    DefaultMutableTreeNode dod = null;
    DefaultMutableTreeNode internet = null;
    Org = new DefaultMutableTreeNode(new
ipdesname(".1.3","original node", "Org","","","TRUE"));
    top.add(Org);
    dod = new DefaultMutableTreeNode(new ipdesname(".1.3.6","department of
defence","dod","","","FALSE"));
    Org.add(dod);
    internet = new DefaultMutableTreeNode(new
ipdesname(".1.3.6.1","internet","internet","","","FALSE"));
    dod.add(internet);
    DefaultMutableTreeNode directory = new DefaultMutableTreeNode(new
ipdesname(".1.3.6.1.1","directory","directory","","","FALSE"));
    internet.add(directory);
}

```

Figure 1: The first few lines of the Method createNodes of JTree used in GUI

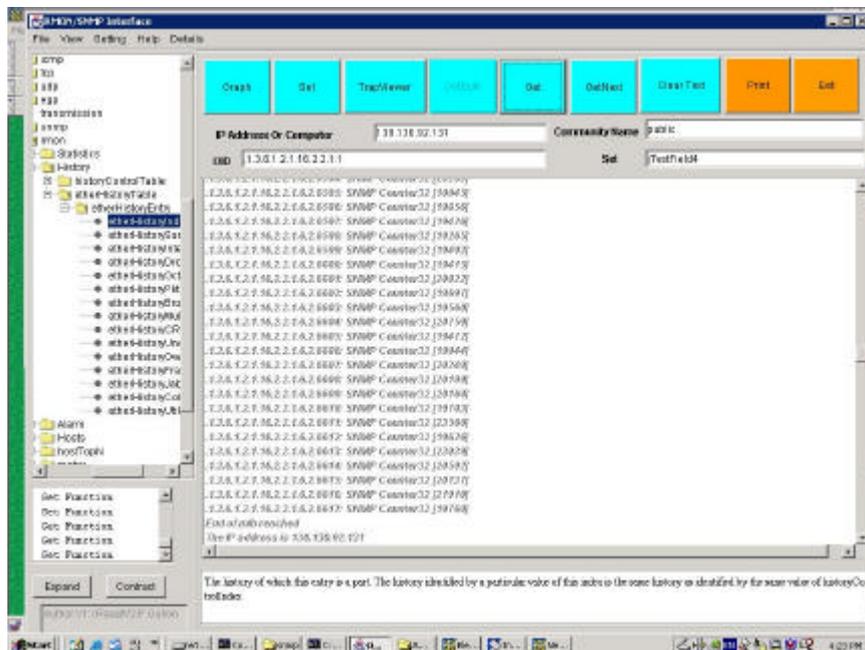


Figure 2: RMON/SNMP Interface using JoeSNMP

Having found the last entry, it was necessary to find out how many data points were in the list by using the code:

```

for(int p=0;p<end;p++){
    place = line.indexOf("[",d );
    d=place+1;
    pos1++; }

```

Using the value stored in the variable pos1, the array which will contain the graph data when the Graph button is pressed can be created using:

```

String[] data = new String[pos1];
int[] dex1 = new int[pos1];
int[] dex2 = new int[pos1];

```

Then the string is searched and the location of each occurrence of the opening bracket is stored in an array called dex1:

```

for(d=0;d<end;d++){
    place1 = line.indexOf("[",d);
    dex1[pos2]=place1;
    d=place1+1;
    pos2++; }

```

Similarly, the string is searched for the location of all closing brackets which is stored in dex2:

```

for(a=0;a<end;a++){
    place2 = line.indexOf("]",a);
    dex2[pos3]=place2;
    a=place2+1;
    pos3++; }

```

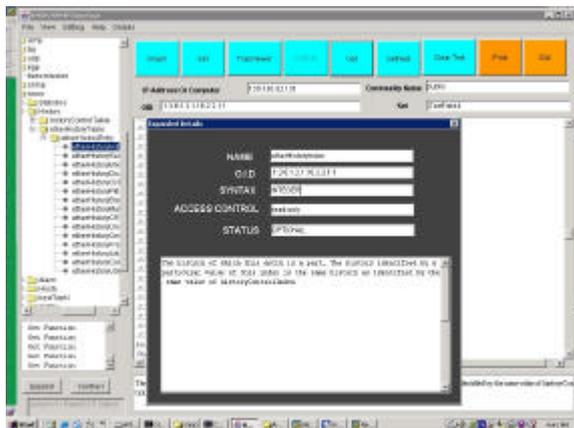


Figure 3: Expanded detail example

The data can then be extracted using the code:

```
for(int j=0;j<pos2;j++)
    data[j]=line.substring((dex1[j]+1),(dex2[j]));
```

Then finally the data can be converted from a string array to an integer array using:

```
for(int k=0;k<37;k++)
    y1[k] = Integer.valueOf(data[k]).intValue();
```

Some code was written to scale the data, though this still needs some development. In order to plot the array of variables obtained from the *Get_Request* it was necessary to use the packages associated with Java2D, Line2D and Rectangle2D. The code to do this was:

```
while (j<sizeofArray) {
    big.drawLine(x[j-1],y1[k-1],x[j],y1[k]);
    big.drawLine(x[j],y1[k],x[j+1],y1[k+1]);
    j=j+2;
    k=k+2; }
```

The image is then buffered using:

```
bimg=(BufferedImage) createImage(width,height)
```

It is then painted onto the screen by using:

```
big=bimg.createGraphics()
```

Followed by the use of the *repaint()* method. Figure 4 shows an example of the plot of some historical data obtained from the RMON probe.

4. Observations and Conclusions

The project was deemed successful in that a GUI was designed and implemented which successfully retrieved the data from the RMON probe. It also provided useful descriptions and OID associated with each RMON MIB element. It was able to plot and

show trends associated with the network which the RMON probe was monitoring (a simple laboratory network intended only for teaching network management principles to undergraduate students).

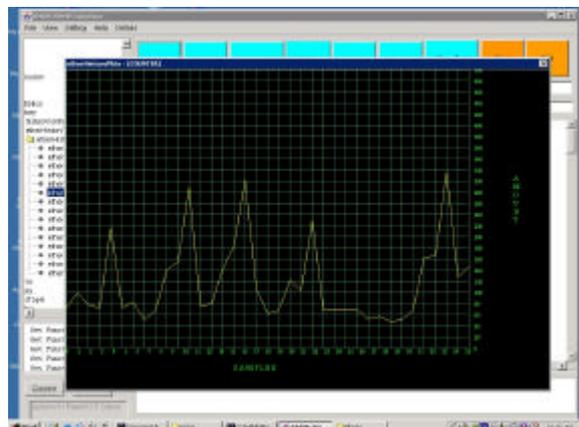


Figure 4: Plot of *etherHistoryPkts* from RMON probe using RMON/SNMP interface

An example of this is shown in Figure 4. It is intended that this software will be incorporated into the laboratory in Spring 2004 when the associated subject is next offered. It will then allow students to experiment with an RMON probe, and provide graphing facilities which are currently not provided within the schools laboratory offering.

5. References

- [1] K V Nagarajan, P Vial, G Awyzio, "Simulation of SNMPv3 Traffic Flow Meter MIB Using ARENA Simulation Modelling Software", *MS2002, IASTED*, Marina del Rey, California, USA, May 2002, pp. 85-90
- [2] I Raad, P Vial, "Telecommunications Network Management Laboratory", Paper submitted to ICTTA'04, Damascus Syria, IEEE French Section
- [3] JoeSNMP is available at <http://www.opennms.org>, visited Feb. 2004, accessed free Java source code in 2002
- [4] Aprisma, Network Management Software called Spectrum, <http://www.aprisma.com>, visited Feb. 2004
- [5] JBuilder 6 from Borland, <http://www.borland.com>, visited Feb. 2004 (latest version is now 9)
- [6] <http://java.sun.com>, visited Feb. 2004
- [7] W Stallings, "SNMP, SNMPv2, SNMPv3, and RMON 1 and 2", 3rd Edition, Addison Wesley, January 2000
- [8] J Knudsen, P Niemeyer, "Learning Java", O'Reilly, First Edition, 2000
- [9] C Cavers, G Friesens, B Koeton, "Special Edition: Using Java 2 Standard Edition", QUE, 2001

IEEE 802.11 WLANs: Performance analysis in presence of bit errors

P. Chatzimisios¹, A.C. Boucouvalas¹ and V. Vitsas²

¹Multimedia Communications Research Group, Bournemouth University, Design, Engineering and Computing, Fern Barrow, Poole, Dorset, BH12 5BB,UK, {pchatzimisios, tboucouv}@bournemouth.ac.uk

² Department of Information Technology, Technological Educational Institution, Thessaloniki, Greece, vitsas@it.teithe.gr

ABSTRACT – IEEE 802.11 is worldwide established and the most used protocol for Wireless Local Area Networks (WLANs). As the assumption of an error-free channel is not always true in a realistic environment, we extend a mathematical model previously presented in the literature that calculates IEEE 802.11 DCF performance to take into account transmission errors for the IEEE 802.11a protocol. Our results take into account all the protocol parameters and packet overheads introduced by both the Medium Access Control (MAC) and the physical (PHY) layers as specified in 802.11a. Finally, we explore the effect of transmission errors, data rate and network size on the performance of the basic access and the RTS/CTS schemes, in terms of throughput and packet delay.

I. Introduction

Continuing advances in wireless communications provide users with a lot of convenience such as mobility, installation speed and simplicity. As a result, wireless technologies and especially the IEEE 802.11 protocol [1][2][3] have gradually become a research focus by many researchers. The IEEE 802.11 Distributed Coordination Function (DCF) medium access control (MAC) method defines two mechanisms to employ packet transmission namely the basic access and the Request-To-Send/Clear-To-Send (RTS/CTS) reservation scheme.

The last few years much research has been done on the performance modeling and analysis of IEEE 802.11 DCF. Bianchi in [4] developed a mathematical model for the 802.11 DCF throughput performance, utilizing a Markov chain model but without considering the impact of bit errors on performance. In [5] and [6], we extended Bianchi's model to calculate average packet delay, under the assumption of ideal channel conditions (no transmission errors) and we produced results for the IEEE 802.11b. Crow in [7] and [8] first studied the effect of errors on performance by means of simulation. Authors in [9] and [10] also considered transmission errors by means of a Markov chain model but investigated only saturation throughput.

The goal of this paper is to derive formulae for the throughput and packet delay of IEEE 802.11 WLANs under an error-prone environment; in a realistic environment, the assumption of an error-free channel is not always true and accurate. We extend the approach in [4] and [6], taking into account

transmission errors¹, therefore, a more realistic model is proposed. Our paper uses the same Markov chain with [4] and [6] under the assumption that packet retransmissions are unlimited. The proposed approach is simple and provides an intuitive understanding of the effect of bit errors on DCF performance and has been validated by OPNET simulation results in [6]. Our paper presents throughput and packet delay results for both the basic access and RTS/CTS schemes that illustrate the dependence of performance on transmission errors. Finally, we study what takes place in an average slot time and we derive simple expressions for the time utilized during collisions or transmission errors per successful packet transmission.

The rest of the paper is organized as follows. Section II briefly provides the essential details about the DCF mechanism. Following that, Section III presents the mathematical analysis and modeling. Section IV presents and discusses some of the derived numerical results of DCF performance in an error-prone environment and the paper concludes with Section V.

II. Description of DCF mechanism

We will briefly introduce the basic components of the binary exponential backoff mechanism employed in DCF, in order to understand the mathematical analysis that follows. Readers can refer to [4] [5] [6] or the IEEE 802.11 standards [1][2][3] for further details.

In DCF, a station with a packet to transmit first senses the medium activity to ascertain whether it is in use. If the medium is sensed to be idle for a time interval greater than the Distributed Inter-Frame Space (DIFS), the station initiates a packet transmission (transmits the data packet in basic access or a short RTS packet first in the RTS/CTS scheme). If the medium is sensed busy, the station defers transmission and initialises its random backoff timer². Note that each station is allowed to transmit only when its backoff timer reaches zero and at the beginning of each slot time.

The value of the backoff timer value for each station is uniformly chosen in the interval $[0, W_i - 1]$, where W_i is the current contention window (CW) size, i is the backoff stage, $i \in [0, m]$ and m represents the number of backoff stages. At the first transmission

¹ Note that bit errors over wireless channels can occur either randomly or in bursts. This paper focuses in random errors.

² The backoff timer is decremented when the medium is idle, is frozen when the medium is sensed busy and resumes again only after the medium has been idle for longer than DIFS.

attempt, CW is equal to the minimum backoff window size $W=CW_{min}$. If two or more stations start a packet transmission simultaneously in the same slot, a collision takes place. After each unsuccessful transmission due to a packet collision or error, W_i is doubled until a maximum backoff window size value is reached. After the successful reception of a data packet, the receiver sends back an acknowledgment (ACK) packet after a Short Inter-Frame Space (SIFS) interval. If the source station does not receive an ACK, the data packet is assumed to have been lost and a retransmission is scheduled according to the previous backoff rules.

III. Mathematical modeling and analysis

Our paper utilizes the same discrete-time Markov chain model with [4] and [6] (is not shown due to limited space). We assume that the network consists of n contending stations, each one always having a packet available for transmission. The key assumption is that the collision-error probability of a transmitted packet is constant and independent of the retransmissions that this packet has suffered in the past.

Our analysis³ considers transmission errors, with p the probability that a transmitted packet encounters a collision or is received in error and is given by:

$$p = 1 - (1 - \tau)^{n-1} \cdot (1 - BER)^{l+H} \quad (1)$$

where BER is the link bit error rate, l is the packet payload size, H is the packet header length and τ is the probability that a station transmits a packet in a randomly chosen slot time. The transmission probability τ is equal to [6]:

$$\tau = \frac{2 \cdot (1 - 2p) \cdot (1 - p^{m+1})}{W \cdot (1 - (2p)^{m+1}) \cdot (1 - p) + (1 - 2p) \cdot (1 - p^{m+1})} \quad (2)$$

Equations (1) and (2) form a non-linear system with two unknowns p and τ . Note that $p \in (0,1)$ and $\tau \in (0,1)$. This non-linear system can be solved using numerical methods and has a unique solution.

A. Saturation throughput

Let P_{tr} be the probability that at least one transmission occurs in a randomly chosen slot time, P_s the conditional probability that this transmission is successful and $PER = 1 - (1 - BER)^{l+H}$ the packet error rate, therefore:

$$P_{tr} = 1 - (1 - \tau)^n, \quad P_s = \frac{n \cdot \tau \cdot (1 - \tau)^{n-1}}{1 - (1 - \tau)^n} \cdot (1 - PER) \quad (3)$$

The probability P_c that an occurring transmission collides (due to the fact that two or more stations transmit at the same time) and the probability P_{er} that a packet is received in error are given by:

$$P_c = 1 - \frac{n \cdot \tau \cdot (1 - \tau)^{n-1}}{1 - (1 - \tau)^n}, \quad P_{er} = \frac{n \cdot \tau \cdot (1 - \tau)^{n-1}}{1 - (1 - \tau)^n} \cdot PER \quad (4)$$

Consequently, the saturation throughput S can be derived as:

$$S = \frac{P_{tr} \cdot P_s \cdot l}{(1 - P_{tr}) \cdot \sigma + P_{tr} \cdot P_s \cdot T_s + P_{tr} \cdot P_c \cdot T_c + P_{tr} \cdot P_{er} \cdot T_{er}} \quad (5)$$

where the denominator of equation (5) denotes the average length of a slot time $E[slot]$, σ is the duration of an empty slot time, T_s , T_c and T_{er} are the average durations the medium is sensed busy due a successful transmission, a collision and a transmission error, respectively.

The values of T_s and T_c depend on the medium access mechanism and for the basic access scheme are given by:

$$T_s^{bas} = T_c^{bas} = T_{er}^{bas} = DIFS + T_{DATA} + SIFS + T_{ACK} \quad (6)$$

and for the RTS/CTS scheme:

$$\begin{cases} T_s^{RTS} = T_{er}^{RTS} = DIFS + T_{RTS} + SIFS + T_{CTS} + SIFS + T_{DATA} + SIFS + T_{ACK} \\ T_c^{RTS} = DIFS + T_{RTS} + SIFS + T_{CTS} \end{cases} \quad (7)$$

where T_{DATA} , T_{ACK} , T_{RTS} and T_{CTS} is the transmission time for a data, acknowledgement, RTS and CTS packet, respectively. According to the IEEE 802.11a standard [3]:

$$T_{DATA} = 20us + 4us \cdot \left[\frac{294+l}{4 \cdot C} \right] \quad (8)$$

$$T_{RTS} = 20us + 4us \cdot \left[\frac{182}{4 \cdot C_{con}} \right] \quad (9)$$

$$T_{ACK} = T_{CTS} = 20us + 4us \cdot \left[\frac{134}{4 \cdot C_{con}} \right] \quad (10)$$

where C is the data rate at which data packets are transmitted (6, 9, 12, 18, 24, 36, 48, and 54 Mbit/s) and C_{con} is the control rate at which the RTS, CTS and ACK control packets are transmitted (6, 12 or 24 Mbit/s). Note that the data and control rate may not be the same. In order to ensure that the vital information contained in the RTS and CTS packets will be received by all stations in range and to cope with potential hidden stations, control packets are transmitted at a lower data rate which increases reception distance.

B. Average packet delay

Our analysis also calculates the average delay $E[D]$ for a successfully transmitted packet. In fact, packet delay is defined to be the time interval from the time a packet is at the head of its MAC queue ready for transmission, until its successful reception in the destination. $E[D]$ is given by:

$$E[D] = E[X] \cdot E[slot] \quad (11)$$

where $E[X]$ is the average number of slot times for a successful packet transmission and $E[slot]$ is the average length of a slot time. The values of $E[X]$ are

³ Note that the derived mathematical analysis for the case of BER=0 has been validated by comparison with simulation results utilizing the OPNETTM simulation package in [6].

independent of the employed access mechanism (basic access or RTS/CTS) and finally can be found as [6]:

$$E[X] = \frac{(1-2p) \cdot (W+1) + pW \cdot (1-(2p)^m)}{2 \cdot (1-2p) \cdot (1-p)} \quad (12)$$

IV. Performance evaluation

Unless otherwise specified, the values reported in the following figures have been obtained using the system parameters summarized in table I for the Orthogonal Frequency Division Multiplexing (OFDM) physical layer used in the 802.11a [3].

Parameter	Value
Packet payload size, l	8184 bits
Slot time, σ	9 μ s
DIFS	34 μ s
SIFS	16 μ s
Channel data rate	54 Mbit/s
Control bit rate	24 Mbit/s
Minimum CW, CW_{min}	16
Number of backoff stages, m	6

Table I System parameters in IEEE 802.11a

Fig. 1 and 2 study the effect of transmission errors and network size by plotting throughput efficiency and packet delay versus n , for the basic access and the RTS/CTS schemes, respectively, for three BER values ($BER=10^{-4}$, 10^{-5} and 10^{-6}). Both figures 1 and 2 illustrate that as expected when the number of contending stations increases, throughput drops off and the packet delay increases in both basic access and RTS/CTS schemes as a result of more packet collisions. However, it appears that the throughput performance of RTS/CTS scheme is less sensitive on the network size than the basic access scheme. An interesting observation is that the performance achievable by the basic access is very close (for $BER=10^{-5}$) or higher (for $BER=10^{-4}$) to that achievable by the RTS/CTS scheme. The explanation is twofold; firstly, because a transmission error penalizes performance when the RTS/CTS is utilized compared to the basic access (note T_{er} values in equation (7)) and secondly due to the fact that the overall WLAN performance suffers significantly when the lower rate RTS/CTS exchange reservation scheme is combined with higher transmission data rates. In fact, performance results show that only very large network size values render the RTS/CTS beneficial, for high data rates (54 Mbit/s), unlike common expectation. This result holds true even when the highest control rate (24 Mbit/s) is utilized and is explained due to the exchange of the RTS and CTS reservation packets at a much lower control rate, which results in a significant delay in communication.

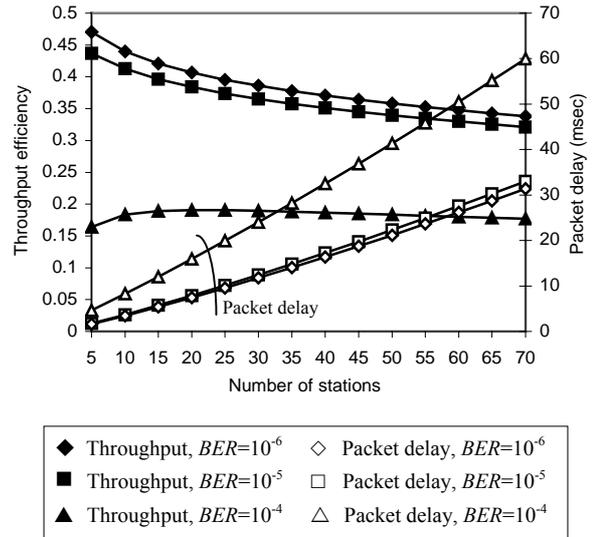


Fig. 1 Effect of BER on the basic access scheme

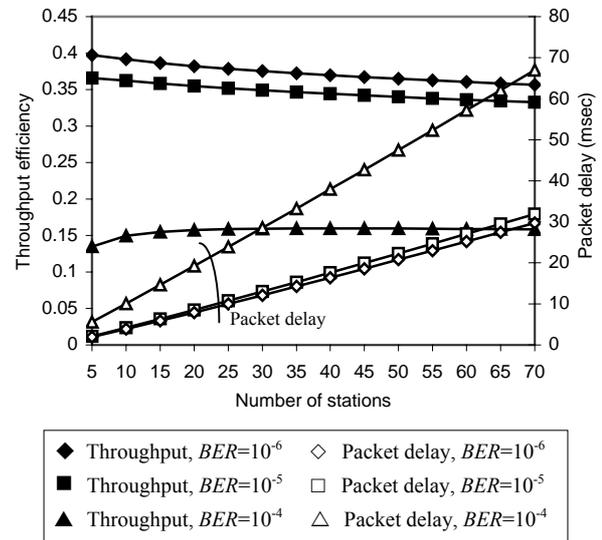


Fig. 2 Effect of BER on the RTS/CTS scheme

In order to better understand the impact of transmission errors on performance, we study what occurs in a randomly selected time slot. Dividing numerator and denominator of equation (5) by $P_{tr} \cdot P_s$, we obtain:

$$S = \frac{l}{\frac{1-P_{tr}}{P_{tr} \cdot P_s} \cdot \sigma + T_s + \frac{P_c}{P_s} \cdot T_c + \frac{P_{er}}{P_s} \cdot T_{er}} \quad (13)$$

The denominator of equation (6) expresses the average time spent on the channel for a successful transmission. This time is further decomposed into four components. It is important to study the third and fourth terms at the denominator of equation (6). The third term represents the time W_{col} wasted due to collisions per successful packet transmission. In fact, P_c/P_s is the average number of collided transmissions per successful transmission, which is multiplied by the

average duration T_c that the medium is sensed busy due a collision. Following the same approach, the fourth term at the denominator of equation (6) denotes the time W_{er} wasted due to transmission errors per successful packet transmission.

Figure 3 plots the average amount of time spent in collisions W_{col} and transmission errors W_{er} per successful packet transmission, normalized with respect to the slot time σ . The figure shows that the time wasted due transmission errors is not affected by the network size. This is justified by noting that in equation (13) the term P_{er}/P_s results to be independent of n . When the BER increases, the time wasted due transmission errors increases in both basic access and RTS/CTS schemes. In fact, transmission errors slightly affect W_{er} when $BER=10^{-6}$ but significantly increase W_{er} for higher BER values ($BER=10^{-4}$). We also observe that basic access achieves a significantly lower W_{er} value under high BER values in respect to RTS/CTS; the average duration T_{er} that the medium is sensed busy due a transmission error is considerably larger when RTS/CTS scheme is utilized as it is shown in equation (7). Furthermore, the figure shows the significant dependence of the time spent in collisions both from the number of contenting stations and transmission errors. In fact, the introduction of the RTS/CTS mechanism, the collision duration is reduced drastically since collisions only occur to the RTS packets that are much shorter than the data packets. As a result large network sizes do not have a significant impact on the performance regardless the increased number of collisions. For the same reason, basic access proves to be more sensitive on high values of n that penalize overall performance.

V. Conclusions

In this paper, we have extended an analytical model that calculates throughput and delay performance for IEEE 802.11a WLAN protocol in the presence of transmission errors. In order to better understand the impact of transmission errors on performance, we have studied what occurs in a randomly selected time slot. For this reason, we have derived simple expressions for the time utilized during collisions or transmission errors per successful packet transmission. Analytical results illustrate that transmission errors considerably affect protocol performance. When BER increases, throughput degrades and packet delay increases. Results also indicate that the performance of RTS/CTS scheme is less sensitive on the network size than the basic access scheme but is highly affected by transmission errors. Furthermore, we have found that there is a significant dependence of the time spent in collisions or errors from the number of contenting stations and transmission errors in both the basic access and the RTS/CTS schemes.

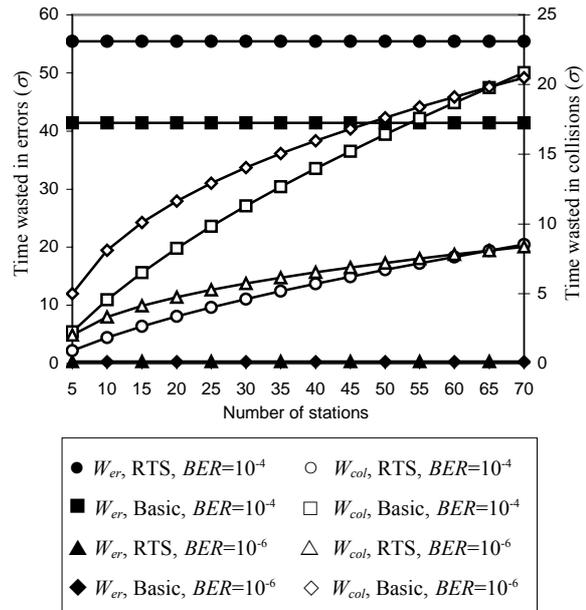


Fig. 3 Average number of slot time units wasted due to errors and packet collisions, per successful transmission

References

- [1] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ISO/IEC 8802-11:1999(E), 1999.
- [2] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High-Speed Physical Layer Extension in the 2.4 GHz Band, IEEE 802.11b WG, 1999.
- [3] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High-Speed Physical Layer Extension in the 5 GHz Band, IEEE 802.11a WG, 1999.
- [4] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function", IEEE Journal on Selc. Areas in Com., vol.18, no. 3, pp. 535-547, 2000.
- [5] P. Chatzimisios, A. C. Boucouvalas and V. Vitsas, "IEEE 802.11: Packet Delay - A Finite Retry Limit Analysis", Proc. of IEEE Global Telecommunications Conference (Globecom), vol. 2, pp. 950-954, 2003.
- [6] P. Chatzimisios, A. C. Boucouvalas and V. Vitsas, "Packet delay analysis of the IEEE 802.11 MAC protocol", IEE Electronic Letters, vol.39, no. 18, pp.1358-1359, 2003.
- [7] B.P. Crow, I. Widjaja, J.G. Kim and P.T Sakai, "IEEE 802.11 wireless local area networks", IEEE Communications Magazine, vol. 35, no. 9, pp. 116-126, 1997.
- [8] B.P. Crow, I. Widjaja, J.G. Kim and P.T Sakai, "Investigation of the IEEE 802.11 medium access control (MAC) sublayer functions", Proc. of IEEE INFOCOM, vol. 1, pp. 126-133, 1997.
- [9] Jianhua He, Zuoyin Tang, Zongkai Yang, Wenqing Cheng and Chun Chou "Performance evaluation of distributed access scheme in error-prone channel", Proc. of IEEE Conference on Computers, Communications, Control and Power Engineering (TENCON), vol. 2, pp. 1142-1145, 2002
- [10] Zuoyin Tang, Zongkai Yang, Jianhua He and Yanwei Liu, "Impact of bit errors on the performance of DCF for wireless LAN", Proc. of IEEE Conference on Communications, Circuits and Systems and West Sino Expositions, vol. 1, pp. 529-533, 2002.

Improving a Simulated Annealing Approach for Network Optimisation using Inexact Forecasts.

GILBERTO FLORES LUCIO, MARTIN J. REED, IAN D. HENNING

Electronic Systems Engineering Department, University of Essex
Wivenhoe Park, Colchester, UK

gflore@essex.ac.uk, <http://privatewww.essex.ac.uk/~gflore/>

Abstract—Simulated Annealing with a modified perturbation cooling schedule is applied to the design of communication networks where traffic forecasts are uncertain and the problem is *NP*-complete. The SA approach is compared with an alternate strategy, an optimised GA. It is found that in terms of implementation, the parameterisation of the GA technique is significantly more complex than that for a SA. Extensive experimentation was carried out to determine the near optimal values for the parameters used in SA for this problem. To compare numerical performance, both techniques (SA and GA) were tested over a number of network sets varying in size, number of commodities (network bandwidth demands) and levels of uncertainties on the commodities. The results show that the SA approach outperforms a GA in terms of both solution quality and optimisation speed.

Keywords—Network optimisation, Simulated Annealing, Forecasting.

I. INTRODUCTION.

Network traffic in practical networks has a tendency to grow over time, and this growth is uncertain [1]. Within a business, forecasting is used to predict demands and this may include migration to different services with different needs for QoS in the network [2]. This paper starts with the premise that a network operator defines a set of future network demands as a set of traffic matrices each specifying predicted demand at a specific point (or epoch) in the future. In reality, predicting such demands is not an exact science and there is likely to be some uncertainty as to whether the demand will be actually required. Determining the demand and likelihood (stated here as a probability of demand) is a business level activity; this paper assumes the operator specifies them. At each point (or epoch) in the future the routing of each traffic demand (a commodity) has to be determined such that the operator achieves some business goal and this is the focus of this work. Specifically, this paper considers the goal of maximising the use of the network at lowest cost whilst balancing load across the network; however, the approach is highly applicable to other goals.

The planning of routes in a QoS enabled network with the associated constraints of integral

fixed bandwidth path allocation and capacity constraints is classed as an *NP*-complete problem [3] (specifically an *integer multicommodity flow optimisation* [4]). Consequently, finding an optimum solution is not practicable and even finding “good” solutions in an efficient manner is a challenge. This type of non-linear constrained optimisation is often solved using some form of heuristic approaches, primarily nature-based [5]. Nature-based stochastic approaches have been widely applied to design and planning problems for communication networks [6][7]. One widely used is the genetic algorithm (GA) as it often provides good solution quality with reasonable optimisation speed. However, a GA has some difficulties in implementation, as it requires a number of internal variables to be tuned. Furthermore, these variables are not directly related to the problem specification.

The motivation for this work is to demonstrate improvements in optimisation quality, speed and implementation for network planning problems when using Simulated Annealing (SA) with a modified cooling [7]. The application studied here uses a tuned version of SA to compare it with an optimised hybrid GA for the design of communication networks where the traffic forecasts are uncertain.

The paper is organized as follows: Section II provides the problem description and application of SA; Section III provides a general explanation of the network scenarios and experimental procedures used to test the efficiency of the techniques; Section IV presents results and draws comparisons between SA and hybrid GA approaches.

II. PROBLEM FORMULATION AND OUR SIMULATED ANNEALING APPROACH.

A. Multicommodity Flow Model with uncertainties.

Consider a network with a set of nodes $n \in N$ and a set of links $l \in L$, where link l has capacity m_l at time t . The traffic flow for a commodity $m \in M$ at time t , is defined as \mathbf{X}_m from vectors \mathbf{x}_{mt} each representing the flow of m on links $(1 \dots L)$, x_{mtl}

defines the flow of m in link l at time t . \mathbf{X}_m is subject to the constraint:

$$\mathbf{g}_l = \sum_{m \in M} x_{ml} \leq \mathbf{m}_l \quad \forall l, \forall t \quad (1)$$

Where (1) defines the constraint that for all routes ($m \in M$) passing through a link l the assigned traffic (vector \mathbf{x}) is less than the total capacity μ_l on that link l . Note that all of these variables are dependent on time t .

The objective is to maximize the network usage (by deploying the maximum number of commodities M) at the lowest cost. Therefore:

$$\text{Maximize } \mathbf{b} = \sum_{m \in M} d_{mt} p_{mt} \quad (2)$$

Where: p_{mt} is the certainty of commodity m at time t of being deployed, and:

$$d_{mt} = \begin{cases} 1 & \text{if the commodity } m \text{ is accepted} \\ 0 & \text{if the commodity } m \text{ is not accepted} \end{cases}$$

And, for the maximum value of \mathbf{b} minimize:

$$C = \mathbf{x} \sum_{m \in M} K(s_m, v_m) + \mathbf{y} \sqrt{\sum_{l \in L} \left(\bar{l} - (\mathbf{m}_l - \mathbf{g}_l) \right)^2} / L \quad (3)$$

Where \mathbf{x} and \mathbf{y} represent weight factors to suit specific application requirements, $K(s_m, v_m)$ represent the number of hops for each commodity accepted, \bar{l} is the mean of the remaining resources in the whole network, $(\mathbf{m}_l - \mathbf{g}_l)$ is the remaining resource of edge l . (3) aims to reduce the hop count and provide homogeneous flow throughout the network.

B. Simulated Annealing approach.

Simulated Annealing (SA) is a well known nature-based search heuristic technique, for solving combinatorial optimisation problems [8], and has been used to solve problems in many applications [9]. We generate an initial solution \mathbf{X}_c using the modified custom Dijkstra Algorithm [10]. This generates a solution following shortest paths in terms of hop count but which obeys capacity constraints. The *Perturbation* function of our SA generates random 2-Opt [11] moves according to the size of the network and the current temperature of the system. If the moves improve the total cost of the current solution then it is accepted and saved as the best solution so far. If no improvement is found, the cost of the current solution is accepted with a certain probability. The stop criteria for the algorithm are: no improvements for the best solution found so far within a certain number of iterations, and reaching the final temperature T_f . The SA implementation used here is:

Algorithm SA ($T_o, T_f, \text{maxiter}, \text{total}, C, G, r$)

```

1 Create initial feasible solution  $\mathbf{X}_c$ 
2  $T = T_o, \mathbf{X}_p = \mathbf{X}_B = \mathbf{X}_c, \tau = 1.0;$ 
3 repeat until  $T \leq T_f$  or  $\text{iter} == \text{total}$ 
3.1 for  $i = 1$  to  $\text{maxiter}$ , step by 1
3.1.1  $\mathbf{X}'_c = \text{Perturbation}(\mathbf{X}_c, G, \tau)$ 
3.1.2 if  $C(\mathbf{X}'_c) < C(\mathbf{X}_B)$ 
3.1.2.1  $\mathbf{X}_B = \mathbf{X}'_c, \mathbf{X}_p = \mathbf{X}'_c, \mathbf{X}_c = \mathbf{X}'_c$ 
3.1.3 else if  $C(\mathbf{X}'_c) < C(\mathbf{X}_p)$ 
3.1.3.1  $\mathbf{X}_p = \mathbf{X}'_c, \mathbf{X}_c = \mathbf{X}'_c$ 
3.1.4 else
3.1.4.1  $P = \min \left\{ 1, \exp \left( \frac{-(C(\mathbf{X}'_c) - C(\mathbf{X}_c))}{T} \right) \right\}$ 
3.1.4.2 if  $P > \text{random}[0,1)$ 
3.1.4.2.1 Set  $\mathbf{X}_p = \mathbf{X}'_c, \mathbf{X}_c = \mathbf{X}'_c;$ 
3.1.5 end if
3.2 end for
3.3 Set  $T = T * r$ 
3.4 Set  $t = \frac{t}{t+1}$ 
3.5 if  $\mathbf{X}_p == \mathbf{X}_B$ 
3.5.1  $\text{iter} = \text{iter} + 1$ 
4 end repeat
5 return  $\mathbf{X}_B$  as the best solution;

```

Where T_o is the initial temperature, T_f is the final temperature, C is the problem cost function, G represents the problem specification (e.g. network topology, demand matrix in our case) r is the cooling rate value and $\exp(x) \equiv e^x$. \mathbf{X}_B , \mathbf{X}_c represents the best and current solution respectively. \mathbf{X}_p is a solution to allow comparison in a local minima and is allowed (controlled by randomness and temperature) to accept poorer solutions. Furthermore, if \mathbf{X}_p shows no improvement in total iterations the algorithm is terminated. This significantly reduces the running time compared to using final temperature alone as the finishing condition. One of the significant features of our SA implementation is to introduce an alternate cooling variable t in the perturbation function. This allows the 2-Opt to have a large number of k -steps initially but after a small number of SA iterations the value of t forces this to 1-step. This is important as a larger number of 2-Opt k -steps are useful to explore a wider search space at the start of the algorithm, but larger k -steps later reduce convergence.

The inner *Perturbation* function is defined as:

function *Perturbation* (\mathbf{X}_c, G, t)

```

1 Create random number  $\text{num} = \text{random}[1, N)$ 

```

2 \mathbf{X}'_c is a new solution obtained after running 2-

Opt algorithm on \mathbf{X}_c with k -steps (where

$$k = \max\{1, \lfloor (N-1) * t \rfloor\}$$

3 **return** \mathbf{X}'_c

Where N is the number of nodes of the current network. Using the definition of k above allows the algorithm to have some independence from problem size.

III. EXPERIMENTAL PROCEDURES

To evaluate the effectiveness of these methods, 3 different networks with different levels and inexact traffic forecasts were used. The first network scenario is formed by 10 nodes and 15 edges, the second 19 and 30 edges and the last 35 nodes and 50 edges. The values of temperatures, iterations and cooling rate r were varied over first a wide range and then an increasingly small range of so optimal values could be determined. The experiments were conducted in an Intel® Pentium® 4 PC 1.7Ghz with 512 MByteRAM in a Linux O.S using C++.

Two performance measures are used to compare the SA approach and the hybrid GA: the overall cost, and the computation time. The computation time comparison was based on the time taken for each of them to obtain an equivalent solution in terms of number of commodities "routed". The stop criterion for the SA has already been defined in section II.B. The stop criteria for the GA was that it should "route" at least as many commodities as the equivalent SA. The Genetic Operators used for the GA optimisation are Uniform Crossover with a probability rate of 100% and Simple Mutation with a probability rate of 7.5%. To obtain these optimum values of Genetic Operators required a significant amount of work, as described in [12].

IV. COMPUTATIONAL RESULTS.

A. Comparison between tuned SA and the Hybrid GA.

For the 3 network scenarios with 10, 19 and 35 nodes, the best values for overall cost, computation time and number of accepted requirements were obtained for both the tuned SA and the Hybrid GA techniques. These scenarios were used with 20, 30 and 40 different commodities and the certainty of these planned commodities expressed as a probability p . For example a commodity m with $p_m=1$ represents a planned commodity that is certain to be required whereas a value of $p_m=0$ represents a commodity that will never be required.

Table I shows the 3 different sets of commodities used along with the number of certainty to happen, amount accepted for both techniques depending on their probability, quality of the solution (overall cost) in "economic units"

and computation time in seconds. The overall cost of each solution is calculated by combining two values. The number of requirements supported by the established network and the cost of that network in "economic units". In order to calculate the amount of economic units that a solution can have, two factors are considered. The first factor is the total number of hops that each path of each requirement has (every hop is considered an economic unit). The second factor is obtained by calculating the standard deviation of the resources that the solution leaves after deploying the requirements (the total value is passed as economic units), this factor is to define how well the load in the network is distributed (wider distribution is given lower cost). This set of commodities was used for the three network scenarios. Commodities with a higher value of certainty will have more priority to be deployed in the network.

The first comment on the results is that regardless of the technique used, both showed good performance in terms of number of commodities routed. The SA gave slightly better results with larger networks. The SA gave solutions with a better traffic load distribution (fewer highly loaded links). The priority given by both approaches to fulfil the commodities with a higher probability is shown to be independent of network size. In this work we are most interested in ranking this priority through the likelihood of the commodities being needed, however, it is may be applicable to other applications for example prioritised QoS routes.

The second comment is on the quality of the solution as well as the computation time. Again, the SA approach shows improvement over the GA with lower cost solutions (regardless of the size of the network and number of commodities). But the most pronounced difference is in terms of computation time for each technique, demonstrating that SA outperforms the Hybrid GA being up to 6 times faster than the GA approach in the smallest scenario (10 nodes and 20 commodities) and 4.3 times faster on average.

V. CONCLUSIONS.

An evolutionary computation approach to optimise the design of communication networks where traffic forecasts is uncertain was presented. A nature-based heuristic method called Simulated Annealing (SA) was used. To fine-tune the parameters involved in SA, a series of evaluations were performed in several network scenarios to make the search more efficient for a specific problem class. The optimisation was required to route different sets of commodities with diverse levels of certainties and with minimum network cost. The results showed the effectiveness of the proposed methodology in this Multicommodity Flow problem (MCF) in 3 different network scenarios when the parameters are fine-tuned. In terms of computation time, the SA approach

outperformed the hybrid GA by a typical factor of four times. In terms of solution quality given as cost, the difference was not that significant but the

SA approach was always found to perform better than the GA.

TABLE I. COMPARISON OF SA AND GA FOR 3 NETWORK SIZES AND THREE DIFFERENT SETS OF COMMODITIES. RESULTS SHOW THE NUMBER OF COMMODITIES PLACED, QUALITY OF THE SOLUTION (COST ACCORDING TO FUNCTION C) AND COMPUTATION TIME (IN SECONDS).

Commodities	Certainty of commodities being required	For 10 nodes (Commodities accepted, cost, computation time)		For 19 nodes (Commodities accepted, cost, computation time)		For 35 nodes (Commodities accepted, cost, computation time)	
		GA	SA	GA	SA	GA	SA
20	16 with $p=1.0$, 4 with $p=0.75$	20 (all) 4775, 1135.45 s	20(all) 4673, 192.78 s	20 (all) 8907, 3159.54 s	20 (all) 8801, 639.54 s	20 (all) 16945, 4983.2 s	20 (all) 15918, 2065.6 s
30	21 with $p=1.0$, 6 with $p=0.75$, 3 with $p=0.5$,	27(21,5,1) 5873, 1253.89 s	27(20,5,2) 5672, 372.67 s	27(20,5,2) 11985, 4543.9 s	28(20,6,2) 10786, 887.65 s	28(20,6,2) 20546, 10965.7 s	29(21,6,2) 19345, 2567.98 s
40	24 with $p=1.0$ 10 with $p=0.75$ 4 with $p=0.5$ 2 with $p=0.25$	32(23,7,2,0) 7643, 1435.6 s	32(23,7,2,0) 7015, 503.45 s	34(22,8,3,1) 14289, 6763.4 s	34(22,8,3,1) 12810, 1352.67 s	35(24,8,2,1) 24074, 14078.5 s	36(24,8,3,1) 22109, 2984.72 s

References:

- [1] Marbukh V., "Network Provisioning as a Game Against Nature" in the IEEE International Conference on Communications ICC 2003, Anchorage Alaska USA 11-15 May 2003
- [2] Tohru Ueda; "Demand Forecasting and network Planning Methods under Competitive Environment"; in the IEICE Transactions in Communications, Vol. E80-B, No. 2, February 1997, pp. 214-218.
- [3] Dengiz Berma, Altıparmak Fulya, Smith E. Alice; "Local Search Genetic Algorithm for Optimal Design of Reliable Networks"; in the IEEE Transactions on Evolutionary Computation, Vol. 1, No. 3; September 1997, pp. 179-188.
- [4] Awerbuch Baruch, Leighton Tom; "A Simple Local-Control Approximation Algorithm For Multicommodity Flow"; in the Proceedings of the IEEE 34th Conference on Fundamentals of Computer Science, Oct. 1993.
- [5] Chakraborty Mandira, Chakraborty K. Uday; "Applying Genetic Algorithm and Simulated Annealing to a Combinatorial Optimization Problem"; in the Proceedings of 1997 International Conference of Information Communications and Signal Processing (ICICS 1997), Vol. 2, September 1997, pp. 929-933.
- [6] KF Poon, A. Conway, Gilbert Wardrop, John Mellis; "Successful application of genetic algorithms to network design and planning"; in the BT Technological Journal, Vol. 18, No. 4, October 2000, pp. 32-41.
- [7] Costamagna Eugenio, Fanni Alessandra, Giacinto Giorgio; "A Simulated Annealing for the Optimization of Communications Networks" in the IEEE International Symposium on Signal Systems and Electronics 1995 (URSI/ISSSE 95), October 1995, pp. 405-408.
- [8] Kirkpatrick S., Gelatt C.D., Vecchi M.P.; "Optimization by Simulated Annealing"; in the Science Journal 13 May 1983; Vol. 220, No. 4598, 1983, pp. 671-680..
- [9] Diekmann R., Lüling R., Simon, J.; "Problem Independent Distributed Simulated Annealing and its Applications", in the R.V.V.Vidal (ed.): Applied Simulated Annealing, Lecture Notes in Economics and Mathematical Systems, Springer LNEMS 396, 1993, pp. 17-44.
- [10] Kershenbaum Aaron; "Telecommunications Network Design Algorithms"; Computer Science Series, McGraw-Hill Ed., International Editions 1993. Chapter 4, pp. 157-159.
- [11] Sengoku Hiroaki, Yoshihara; "A Fast TSP Solver Using GA on JAVA" in the Proceedings of the 3rd. International Symposium on Artificial Life and Robotics Vol. 1, March, 1998.
- [12] Flores Lucio Gilberto, Paredes Farrera Marcos, Reed Martin J., Henning Ian D.; "A Performance Study for Basic Genetic Operators in GA's using Binary Encoding" Presented in the 3rd. International Congress of Electro and Systems Engineering 2002", ICESE 2002, Mexico City.

Improving performance through optimization of the RTS/CTS mechanism in IEEE 802.11 Wireless LANs

P. Chatzimisios and A. C. Boucouvalas

Multimedia Communications Research Group, School of Design, Engineering and Computing
Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, UK
{ pchatzimisios, tboucouv }@bournemouth.ac.uk

ABSTRACT - IEEE 802.11 Medium Access Control (MAC) protocol employs two techniques for packet transmission; the basic access scheme and the RTS/CTS reservation scheme. In this paper, we carry out an analysis in order to derive an all-purpose expression for the threshold value, which determines when the RTS/CTS scheme should be employed, under ideal channel conditions without the presence of hidden stations or transmission errors. The main advantage of our proposed approach is that it is simple and gives insights of the RTS/CTS mechanism. Results based on the presented analysis for the IEEE 802.11b transmission rates and delays study the effect of the different protocol parameters on the RTS/CTS threshold. Moreover, results indicate that proper selection of protocol parameters such as retry limit and physical packet overhead for the specific data rate is of great importance in minimizing packet delay and improving overall performance.

I. Introduction

Wireless Local Area Networks (WLANs) are becoming more and more popular since they provide high data rates while maintaining a relative low price. The IEEE 802.11 protocol [1] is the dominant standard for WLANs and employs Distributed Coordination Function (DCF) as the essential MAC method. DCF defines two access mechanisms to employ packet transmission; the default, two-way handshaking technique called basic access and the optional four-way handshaking RTS/CTS reservation scheme.

The RTS/CTS scheme involves the transmission of the short *request-to-send* (RTS) and *clear-to-send* (CTS) control packets prior to the transmission of the actual data packet. Since collisions may occur only on the RTS packets and are detected by the lack of CTS response, the RTS/CTS scheme results in an increase of the system performance by reducing the duration of a collision, especially when long data packets are transmitted. The RTS/CTS scheme is also employed to obtain a better performance in the presence of hidden stations. However, authors in [2] and [3] have reported several potential difficulties in the ability of the RTS/CTS scheme to eliminate the hidden stations problem and to reduce interference. On the other hand, RTS/CTS decreases efficiency since it transmits two additional packets without any payload. Hence, the 802.11 standard specifies the *RTS_Threshold* (RT), a manageable parameter that indicates the data length under which the data packets should be sent without RTS/CTS. The value of the *RT* parameter is not specified in the standard and has to be set separately by each station. The data packet size is the only

parameter used for deciding whether the RTS/CTS reservation scheme should be employed or not.

There are a number of studies in the literature on performance of wireless data protocols as well as the RTS/CTS mechanism in IEEE DCF [2]-[7]. The authors in [6] first studied the RTS/CTS mechanism in the IEEE 802.11 through simulations. Work in [2] and [3] has pointed out that the RTS/CTS handshake does not work as well as expected in theory. The authors in [7] have performed a simulation study and suggested that the RTS/CTS mechanism must be employed at all times. Bianchi in [8] calculated the RTS/CTS threshold for throughput maximization but without taking into account packet retry limits¹. In [9] we have presented a method capable of calculating packet delay by taking into consideration retransmission delays with or without packet retry limits. Moreover, in [10], we have evaluated the dependency of the RTS/CTS scheme on network size, however, without providing any general expression for the RTS/CTS threshold.

In this paper, we extend Bianchi's approach in [8], as well the analysis in [10], in order to derive an all-purpose expression for the RTS/CTS threshold. The proposed analysis takes into account packet retry limits and aims at minimizing the delay for data packets in 802.11 DCF by optimally employing the RTS/CTS scheme. Our work is carried out under the hypothesis of ideal channel conditions without the presence of hidden stations or transmission errors. The main advantage of our proposed approach is that is simple and gives insights of the RTS/CTS mechanism. We investigate the dependency of protocol performance on packet retry limit, data rates as well as physical packet overhead and network size. The derived framework in our paper can be useful and valuable for simple but effective performance improvements in WLANs, through the optimal use of the RTS/CTS reservation scheme.

The remainder of the paper is organized as follows. In section II, we briefly review the DCF mechanism used in IEEE 802.11 MAC. In section III, we carry out an analysis that gives a simple but general expression for the RTS threshold. Section IV presents performance results that examining the relationship between the RTS threshold and protocol parameters. Finally, section V concludes our paper.

¹ Every station maintains a retry count that indicates the maximum number of retransmission attempts of a RTS packet or of a data packet when RTS/CTS is not used. When the retry count reaches the specified limit, retry attempts cease and the data packet is discarded.

II. DCF of IEEE 802.11 MAC

DCF is based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) technique and adopts a slotted Binary Exponential Backoff (BEB) scheme to reduce collisions due to stations transmitting simultaneously.

Each node with a packet to transmit first senses the medium to ascertain whether it is in use. If the medium is sensed to be idle for a time interval greater than the Distributed Inter-Frame Space (DIFS), the station proceeds with the packet transmission. If the medium is sensed busy, the station defers transmission and initialises its random backoff timer. The value of the backoff timer value for each station is uniformly chosen in the interval $[0, W_i - 1]$, where W_i is the current contention window size, i is the backoff stage, $i \in [0, m]$ and m represents the station's retry limit. The backoff timer is decremented by one when the medium is idle, is frozen when the medium is sensed busy and resumes again only after the medium has been idle for longer than DIFS.

A station initiates a packet transmission (transmits the data packet in basic access or a short RTS packet first in the RTS/CTS scheme) when its backoff timer reaches zero. The value of W_i depends on the number of unsuccessful transmissions of a packet; at the first transmission attempt, $W_0 = CW_{min} = W$. After each retransmission due to a packet collision, W_i is doubled up to a maximum value, $W_m = CW_{max} = W \cdot 2^m$ where m' is the number of backoff stages. Once W_i reaches CW_{max} , it will remain at this value until it is reset to CW_{min} after the successful data packet transmission or when the retry limit for this packet is reached. After the successful reception of a data packet, the receiver sends back an acknowledgment (ACK) packet.

III. Mathematical analysis

We employ the same discrete-time Markov chain model as in [9]. Using the same assumptions with [7] and [9], we can calculate the probability p that a transmitted packet collides (independent of the number of collisions occurred in the past) as:

$$p = 1 - (1 - \tau)^{n-1} \quad (1)$$

where n is the number of contenting stations and τ is the transmission probability of a packet. When retry limits are taken into account [9], τ is given by equation (2), where W is the minimum contention window size. Equations (1) and (2) form a non-linear system with two unknowns p and τ , which can be easily solved using numerical methods.

Our analysis in [9] has calculated the average delay $E[D]$ for a successfully transmitted packet. Packet delay is defined to be the time interval from the time a packet is at the head of its MAC queue ready for transmission, until its successful reception in the destination. $E[D]$ is given by:

$$E[D] = E[X] \cdot E[slot] \quad (3)$$

where $E[X]$ is the average number of slot times for a successful packet transmission and $E[slot]$ is the average length of a slot time. The values of $E[X]$ are independent of the employed access mechanism (basic access or RTS/CTS) and can be found in [9]. The average length of a slot time is:

$$E[slot] = (1 - P_{tr}) \cdot \sigma + P_{tr} \cdot P_s \cdot T_s + P_{tr} \cdot (1 - P_s) \cdot T_c \quad (4)$$

where $P_{tr} = 1 - (1 - \tau)^n$ is the probability that there is at least one packet transmission in the considered slot time, $P_s = n\tau(1 - \tau)^{n-1} / P_{tr}$ is the probability that an occurring packet transmission is successful, T_c , T_s and σ are the time durations the medium is sensed busy due to a collision and a successful transmission and of an empty slot time, respectively.

The values of T_s and T_c depend on the medium access mechanism and are defined for the basic access and the RTS/CTS access mechanisms as follows:

$$\begin{cases} T_s^{bas} = DIFS + T_{header} + \frac{l}{C} + SIFS + T_{ACK} \\ T_c^{bas} = DIFS + T_{header} + \frac{l}{C} + SIFS + T_{ACK} \end{cases} \quad (5)$$

$$\begin{cases} T_s^{RTS} = DIFS + T_{RTS} + SIFS + T_{CTS} + SIFS + T_{header} + \frac{l}{C} + SIFS + T_{ACK} \\ T_c^{RTS} = DIFS + T_{RTS} + SIFS + T_{CTS} \end{cases} \quad (6)$$

where l is the length of the transmitted packet, C is the data rate, $C_{control}$ is the rate that the control packets (ACK, RTS, CTS) are transmitted (1 Mbit/s), T_{header} , T_{ACK} , T_{RTS} and T_{CTS} are the time intervals required to transmit the packet payload header, the ACK, RTS and CTS control packets, respectively. The above time intervals are given by:

$$T_{header} = \frac{MAC_{hdr} + PHY_{hdr}}{C} \quad , \quad T_{ACK} = \frac{l_{ACK}}{C_{control}} \quad (7)$$

$$T_{RTS} = \frac{l_{RTS}}{C_{control}} \quad , \quad T_{CTS} = \frac{l_{CTS}}{C_{control}} \quad (8)$$

where l_{ACK} , l_{RTS} and l_{CTS} is the length of ACK, RTS and CTS control packets respectively, MAC_{hdr} is the MAC header and PHY_{hdr} is the physical header. In fact, a physical layer preamble (PLCP preamble) and a physical layer header (PLCP header) exist in both data and control frames. Hereafter, we will refer to the sum of PLCP preamble and PLCP header as PHY_{hdr} .

$$\tau = \begin{cases} \frac{2 \cdot (1-2p) \cdot (1-p^{m+1})}{W \cdot (1-(2p)^{m+1}) \cdot (1-p) + (1-2p) \cdot (1-p^{m+1})} \quad , \quad m \leq m' \\ \frac{2 \cdot (1-2p) \cdot (1-p^{m+1})}{W \cdot (1-(2p)^{m'+1}) \cdot (1-p) + (1-2p) \cdot (1-p^{m+1}) + W \cdot 2^{m'} \cdot p^{m'+1} \cdot (1-2p) \cdot (1-p^{m-m'})} \quad , \quad m > m' \end{cases} \quad (2)$$

The IEEE 802.11b protocol supports data rates of 1, 2, 5.5 and 11 Mbit/s. The standard defines two different formats for the preamble and header (PHY_{hdr}): the mandatory supported Long PLCP PHY_{hdr} which interoperates with the 1 Mbit/s and 2 Mbit/s data rates and an optional Short PLCP PHY_{hdr} . The Short PLCP PHY_{hdr} allows performance at the high rates (2, 5.5 and 11 Mbit/s) to be significantly increased. In fact, the Short PLCP PHY_{hdr} is intended for applications where maximum performance is desired and interoperability with legacy is not a consideration. Figure 1 shows the format of the Long and Short PLCP PHY_{hdr} of a data packet.

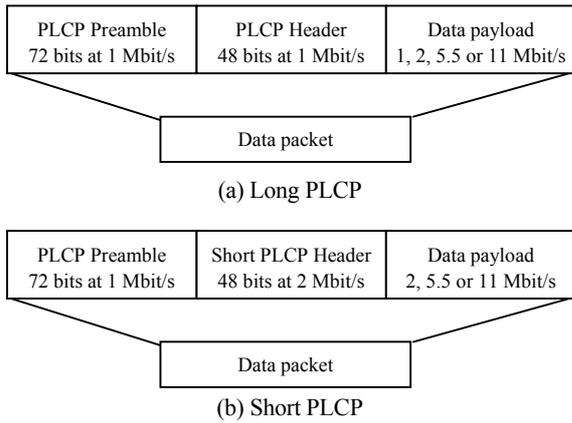


Fig. 1 Long and short PLCP data packet formats

In order to quantify the threshold value for the packet size over which it is best to switch to the RTS/CTS mechanism, we indicate with D^{RTS} and D^{BAS} the average delay of a packet transmitted by the basic access and RTS/CTS mechanism, respectively.

$$D^{RTS} < D^{BAS} \Leftrightarrow E[X] \cdot E[slot]^{RTS} < E[X] \cdot E[slot]^{BAS}$$

$$P_s \cdot T_s^{RTS} + (1 - P_s) \cdot T_C^{RTS} < P_s \cdot T_s^{BAS} + (1 - P_s) \cdot T_C^{BAS}$$

$$P_s \cdot (T_s^{RTS} - T_s^{BAS}) < (1 - P_s) \cdot (T_C^{BAS} - T_C^{RTS}) \quad (9)$$

Let $O_{RTS} = T_s^{RTS} - T_s^{BAS} = \frac{l_{RTS}}{C_{control}} + 2SIFS + \frac{l_{CTS}}{C_{control}}$ be the overhead introduced by the RTS/CTS scheme, T_{data} be the time required to transmit the packet payload and let $O_h = T_{header} - T_{RTS} = (\frac{MAC_{hdr}}{C} + \frac{PHY_{hdr}}{C_{control}}) - \frac{l_{RTS}}{C_{control}}$ be the extra length of the data packet header with respect to the RTS packet size. After some rearrangements, equation (10) finally gives the threshold value $l_{threshold}$ over which it is convenient to switch to the RTS/CTS mechanism. The value of the threshold size depends on the probability of a successful transmission P_s , the control and the data rate as well as the packet overhead.

$$\frac{P_s}{1 - P_s} \cdot O_{RTS} < O_h + T_{data}$$

$$l_{threshold} > \left(\frac{P_s}{1 - P_s} \cdot O_{RTS} - O_h \right) \cdot C \quad (10)$$

IV. Performance evaluation

Fig. 2 and 3 study the effect of packet retry limit and the network size by plotting the probabilities P_s and p versus m , and the packet size threshold versus m respectively, for four representative network sizes ($n = 5, 25, 50$ and 70), $W=32$ and data rate of $C=1$ Mbps. Figure 2 shows that both the probability that an occurring packet transmission is successful P_s and the packet collision probability p are highly dependent on the network size; more contenting stations cause the increase on packet collisions and the decrease of successful packet transmissions. The figure also illustrates that the retry limit significantly affects the probabilities P_s and p . An important observation is that large network sizes appear to be more sensitive on retry limit. A small increase of m results in a greater increase in the successful transmission probability for large networks (equivalent decrease in the collision probability) than for small networks. In fact, for small network sizes ($n = 5$), both the probabilities P_s and p are independent of the retry limit.

Fig. 3 provides the threshold value above which the performance of the RTS/CTS mechanism is considerably enhanced. When the number of the contenting stations is relatively small ($n = 5$), it appears that it is not necessary to employ the RTS/CTS reservation scheme due to the low collision probability (fig. 2). On the contrary, when the network size increases, the RTS/CTS threshold decreases to lower values. This can be justified since large network sizes and a low retry limit cause more packet collisions and a much lower successful transmission probability (fig. 2). The figure illustrates that the packet retry limit has a significant effect; when retry limit increases, the RTS/CTS threshold values also increase due to the improved successful transmission probability and the reduced number of collisions (fig. 2). An interesting outcome in fig. 2 and 3 is that for $m > 6$, the probabilities P_s and p as well as the RTS/CTS threshold are only marginally affected, indicating the proper choice of the retry limit value in the standard.

Fig. 4 plots packet size threshold versus network size for three data rates ($C = 1, 5.5$, and 11) as well as for a short and long PHY packet overhead. According to fig. 4, the packet size threshold is highly dependent on the data rate. When the data rate increases, the threshold values significantly increase. The reason is that although high data rates reduce the transmission time for data packets, the RTS and CTS control packets are still being transmitted by the low control rate, resulting in delay in communication. Moreover, the use of a short PHY header, which results in a lower transmission time comparing to the long PHY header's transmission time, considerably decreases the packet size threshold. This can easily be explained by considering that smaller packet overhead mainly reduces the overhead of RTS and CTS control packets. Thus, the main drawback the RTS/CTS scheme is minimized and it can be employed for even smaller data packets.

V. Conclusions

In this paper, we have presented a simple analysis to derive an all-purpose expression for the threshold value, which determines when the RTS/CTS reservation scheme should be employed, under ideal channel conditions without the presence of hidden stations or transmission errors. Based on our analysis, we have studied and concluded that the RTS/CTS threshold significantly depends on both protocol parameters and network size. Performance results show that high data rates and a high packet retry limit, bring about the considerable increase of RTS/CTS threshold values. Conversely, for large network sizes the RTS/CTS scheme appears to be beneficial due to the increased collision probability. The use of a short physical packet overhead minimizes the main drawback of the extra overhead for the RTS/CTS scheme and makes beneficial its employment for even smaller data packets. The derived analysis could be useful for simple performance improvements, through the optimal use of the RTS/CTS reservation scheme, however, it brings about the question of effectiveness and necessity of the RTS/CTS reservation scheme in high-speed IEEE 802.11 WLANs and in the absence of hidden stations.

References

- [1] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: High-Speed Physical Layer Extension in the 2.4 GHz Band, IEEE 802.11b WG, 1999.
- [2] S. Xu and T. Saadawi. "Revealing the problems with 802.11 medium access control protocol in multi-hop wireless ad hoc networks", *Computer Networks*, vol 38, no. 4, pp. 531-548, 2002.
- [3] K. Xu, M. Gerla and S. Bae, "Effectiveness of RTS/CTS Handshake in IEEE 802.11 Based Ad Hoc Networks", *Ad Hoc Networks Journal*, vol. 1, issue 1, pp. 107-123, 2003.
- [4] V. Vitsas, "Throughput analysis of linear backoff scheme in wireless LANs", *IEE Electronic Letters*, vol.39, no. 1, pp. 99 – 100, 2003.
- [5] V. Vitsas, "Link layer Protocol performance of indoor infrared wireless communications", PhD thesis, Bournemouth University, UK, 2002.
- [6] J. Weinmiller, H. Woesner, JP Ebert and A. Wolisz, "Analyzing the RTS/CTS mechanism in the MAC protocol for wireless LANs", *Proc. Of Personal Wireless Communications Workshop*, 1995.
- [7] S.-T. Sheu, T. Chen, J.-H. Chen, F. Ye: "The Impact of RTS Threshold on IEEE 802.11 MAC Protocol", *Proc. of IEEE International Conference on Parallel Distributed Systems (ICPADS)*, pp. 267-272, 2002
- [8] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function", *IEEE Journal on Selec. Areas in Com.*, vol.18, no. 3, pp. 535-547, 2000.
- [9] P. Chatzimisios, A. C. Boucouvalas and V. Vitsas, "IEEE 802.11: Packet Delay - A Finite Retry Limit Analysis", *Proc. of IEEE Global Telecommunications Conference (Globecom)*, 2003.
- [10] P. Chatzimisios, A. C. Boucouvalas and V. Vitsas, "Packet delay analysis of the IEEE 802.11 MAC protocol", *IEE Electronic Letters*, vol.39, no. 18, pp.1358-1359, 2003.

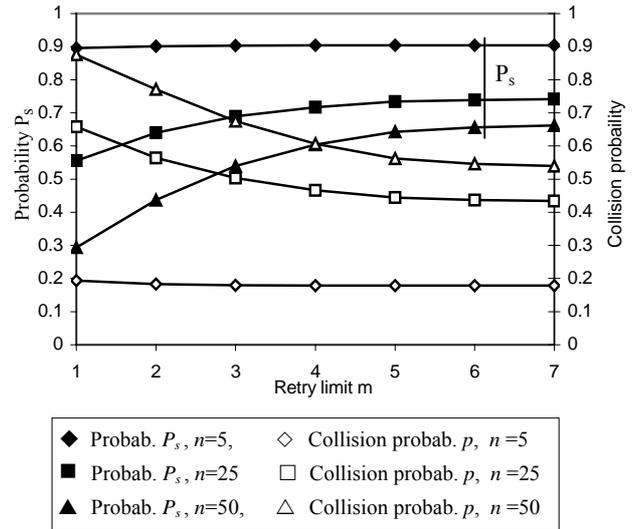


Fig. 2 Effect of retry limit on probabilities P_s and p

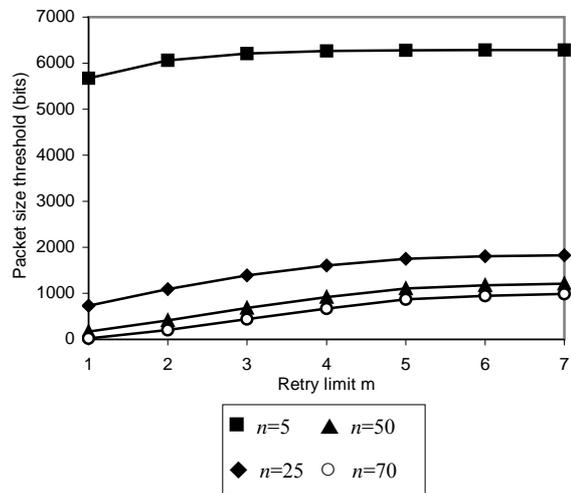


Fig. 3 Packet size threshold versus retry limit

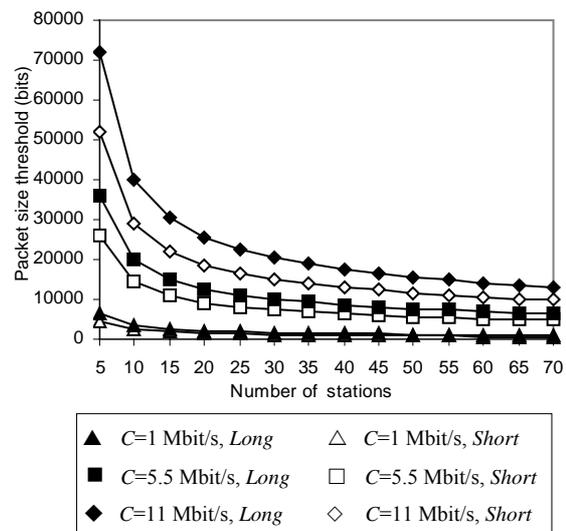


Fig. 4 Effect of data rate and physical packet overhead

MOBILITY OF WIRELESS NETWORKS FOR M-BUSINESS.

Julius .N. Aguni
PowerBank Systems Limited, Warri Nigeria
<aguni@ieee.org >

Abstract:

The emerging mobile device as an alternative service point will have profound impact on many industries and their participants. M-business is a result of the convergence of two core technologies: mobile wireless communication and internet clearly, 2.5 and 3G technologies based on packet switching will offer full time services and instant connectivity in the changing business environment of today. Although the third generation (3G) of mobile communication systems will soon be implemented in full, it will be digital mobile multimedia offering broadband mobile communications with voice, video, graphics, audio and other applications significant to m-business.

In this paper, it is expressed that the presence of 3G mobile and related technologies will take the concepts of m-business further to realization. This paper studies the basic concepts of mobility of wireless networks for m-business, while attempting to answer the fundamental questions on PAN to support individuals participating in m-worlds.

1.0 Introduction

Nowadays, rapid developments in wireless networks and mobile information systems are constantly emerging and have become a major stake the m-business game. There are several ways to come to an assessment of the mobile technology, and identify research issues in m-business. The assessment opportunities include research into the areas of design, use and adoption, and impacts. Within the context of this conference paper, the relevant design approaches will be discussed to cover mobility of wireless networks, heterogeneous networks, ubiquitous connectivity of users and probably m-business concepts and distinction.

With the increasing variety of mobile devices such as personal Digital Assistants (PDAs), a Laptop computer, a digital cellular phone to wearable computers, the new scenarios of the mobile application has gradually developed. As a user moves between different environments, these Mobile Devices co-operate as a coordinated Local area network i.e. Personal Area Network (PAN) interacting with and adapting to the current environment. [9] The challenge lies in supporting such co-operation and co-ordination through intelligent communication management while emphasizing the efficient use of available resources in

ubiquitous manner. To support the needs of the user in the m-business, the PANs may interact with networks and services in the current environment surrounding the user in order to determine local connectivity and service availability. Finally, the design of the network architecture must be enabled by mobility management of the mobile networks to guarantee ubiquitous connectivity as it delivers the desired QoS to the m-business player.

2.0 Concepts

In mobile and ubiquitous computing, the notion of context is often equated simply with location, but is actually more complex. Mobile applications use can vary continuously because of changing circumstances and differing user needs. A context model in [10] can be created using three broad categories of context – environment, participants and activities. The “environment” category is concerned with the location, orientation availability, quality and physical properties. “Participants” includes the status of the user(s) and other participants in the mobile environment. “Activities” covers users, participants, and environmental activities. Additionally, the conceptual model includes any interactions or relationships that may exist between these contexts.

In the m-business environment, users and application have to deal with a large diversity of devices (phones, handheld, telematics and intelligent) that continues to shrink in size and weight. While this achieves high device portability and mobility, networking of the devices can pose potential challenges. Since devices and networks are completely mobile in m-business, the outside environment can change rapidly from moment to moment and network to network. To satisfy these changes require mobile and heterogeneous networks enabling seamless connectivity for the users. [8][10]

3.0 M-Business Distinction

Many companies are just getting to grip with e-business and its impact on their the business model, strategies and systems. Significantly there is a shift to m-business. M-business differs from e-business, and from the current mobile telecommunications. Its distinct characteristics include: -

Ubiquitous- It's not constrained by wired infrastructures although constrained by network coverage limits.

Always on (24/7)- People can choose to be in touch.

Instantaneous – Provides rapid access to Internet and data sites.

Convenient – Resulting from above it's convenient to use.

Personal- It recognizes the individual. The mobile device can be used as an identification tool, and for wireless electronic billing payments.

Location sensitive- Mobile technologies such as GPS (Global Positioning System) and EOTD (Enhanced Observed Time Difference) will provide the ability to locate a mobile device. [5]

M-business is thus best suited to transactions where speed and convenience are important, and where the service required may change with the users' location. M-commerce refers to business transaction while on the move.

4.0 Network Mobility

This section aims to analyse network mobility issues and requirements in the context of m-business and explore various approaches to the solutions. Network Mobility is concerned with a mobile IP subnet changing its point of attachment to an IP backbone, which is illustrated in fig 1. [13]

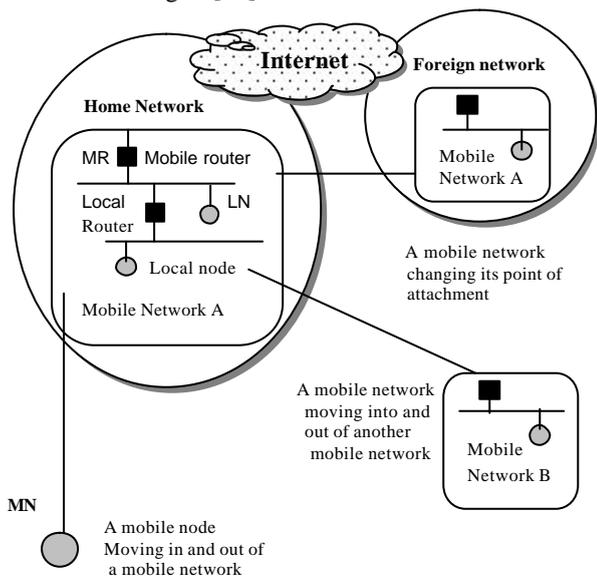


Figure 1. Mobile network scenarios.

The formation of a mobile network can exist at various levels of complexity. Global roaming among different types of wireless technologies could be provided by mobile IP, since it scales reasonably well with the number of users into a network. However, relevant drawback of mobile IP is the so-called triangular routing problem; all packets sent to the mobile network must transmit through its Home Agent (HA) as demonstrated in [1]. The main challenge is keeping the connection with the Internet as the m-business user roams and to provide multiple real-time services while achieving high QoS support.

In complex scenarios, mobile network is a set of IP subnets interconnected by Local routers forming an aggregate able to move as a unit, and interconnect to the

IP backbone through one or more mobile routers. [13] A general mobility model is shown in fig 2. Mobility of wireless networks is the fundamental technology used to automatically support the m-business user enjoying services while simultaneously roaming freely without disruption of communications. [4]

As to mobile communications, mobile objects are mostly physical components and can span all the path of m-business service delivery. Typical mobility scenarios include:

1. **Service Mobility**, means that a personalized service available to the user with one mobile device in one network can still be accessible by another mobile device and/or in another network of different region or operator.
2. **Network Mobility**, refers to the wireless networks that support the connections of mobile devices. Some wireless connections may be a collection of wireless nodes that dynamically form a network with any pre-existing fixed network infrastructure, also known as mobile ad hoc networks.
3. **Terminal mobility**, is the ability of a user device that can roam within a network or between networks with on-going or following communications still reachable.
4. **User Mobility**, means that end-users can access personal services regardless of moving to any network or using any terminal, through unique user identification [14].

5.0 Mobile IP

Because IP addresses by design refer to particular subnet, IP traffic will not be able to find its way to the new location. Given the context of m-Business, mobile applications must be carefully designed to account for this limitation. This is where mobile IP enters the environment to establish mobile Internet connection. [3]

Operational principles:

As shown in figure 2, the basic Mobile IP components include, aside from the Mobile Node (MN), the Home Agent (HA) in the home network, the Foreign Agent (FA) in the foreign network, the Correspondent Node (CN) and the Access Router (AR). The MN has a home address (H@), typically registered in the domain name sever (DNS). The upper 8 bytes of the H@ match the home subnet prefix of the MN's home link. The home link has at least one AR that can offer HA services to the MN. When the MN moves out of the home network into any other foreign network, it can be reached through obtaining a Care of Address (CoA). The general working process of Mobile IP is as follows:

- a) The mobility agent (HA and any FA) continuously sent agent advertisement message to the local link, to announce their existences.
- b) The MN receives the advertisement messages, according to which it determine whether its current attachment point is in the home network or any foreign network.

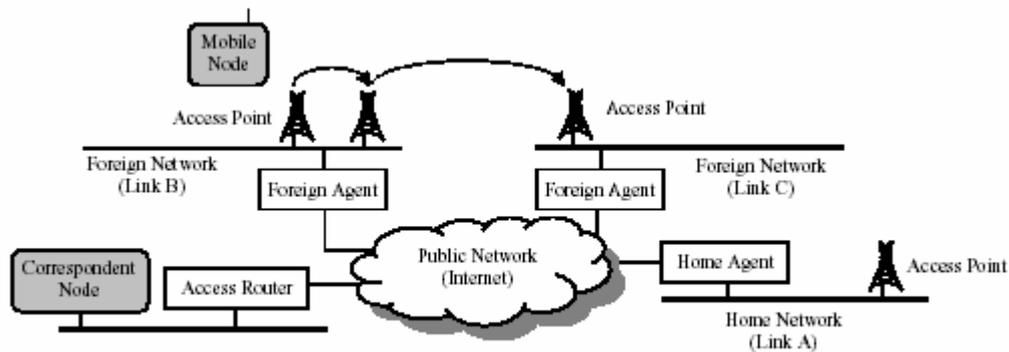


Figure 2: General mobility model.

- c) If the MN finds itself still in its home network (the advertisement message comes from the HA), then it will not start the mobility management functions and the process is terminated. If it comes back from a foreign network, it should cancel the registration on the HA by sending a corresponding message to the HA.
- d) If the MN finds itself in a foreign network, then it will obtain a COA that is either the IP address of the corresponding FA or a new temporary address.
- e) The MN should then register to its HA to inform it of the new CoA. [1] [4]
- f) After successful registration, data packets from any CN originating to the MN's H@ are intercepted and sent by the HA to the MN through "IP tunneling".
- g) The packets sent by the MN to any CN will be routed as normal, and do not need the assistance from HA.

5.1 Mobility with Mobile IP

The purpose of mobile IP is to enable a node to move all around the Internet while preserving the continuity of its communications [13]. This section discusses the mobile IP model for seamless connectivity for m-business. As presented here, mobility is divided into two complementary tasks: macro mobility and micro mobility. Mbusiness in this context requires mostly macro mobility solution along micro mobility. Macro mobility, i.e. mobile nodes movement between different domains, to which inter-domain mobility management schemes can be employed, acting as a global mobility solution with the advantages of flexibility, robustness and scalability. [4]

Macro Mobility: In the context of this paper, fig 3 shows the hierarchical mobility management model as the proposed architecture for mbusiness connectivity. The hierarchical mobility model is enabled by mobile IP. In the basic idea of hierarchical structure, the MN's HA needs not to be informed of every movement that the MN performs inside the foreign network domain. Recent proposals from Ericsson and Nokia employ a hierarchy of FA's to handle the local registration. [4] Limitations posed by FA's are overcome by developing mobile Access Points. What we seek is a scaling or hierarchical approach of FA's and access terminals that can offer reasonable coverage over a range of fixed and mobile wireless networks. [2]

6.0 Personal Area Networks (PANs)

As users of m-business collect small computing devices the amount of communication resources available to the user increase and so the demand for coordination of resources between these devices increase. The collection of these devices to support a mobile individual demands the extension of the mobility paradigm from an individual device to a network of devices [9]. In order to better support the needs of m-business user, the PANs may interact with networks and services in the current environment surrounding the user in order to determine Local Connectivity and service availability.

6.1 Mobile PANs

Providing wireless services at any time and anywhere requires terminal and personal mobility in the m-business infrastructure.

Terminal mobility in [6] allows mobile clients in the m-business to roam across geographic boundaries of wireless networks. There are two main issues in terminal mobility: Location Management and Hand off management. The detailed study of these issues is not considered in the context of this paper. However, with location management, the system tracks and locates a mobile terminal in the PAN for possible connection. Location management involves handling all the information about the roaming terminals such as original and current located cells, and authentication information. On the other hand, hand off management maintains ongoing (seamless) communications when the terminal roams.

Personal mobility is a principal concern in m-business. Personal mobility concentrates on the movement of users instead of users' terminals, and involves the provision of personal communications and personalized operating environments. In personal mobility, when there is a message addressed to the mobile user for a transaction, no matter where the user is located or what kind of terminal is being used, the message will be sent to the mobile user correctly. As demonstrated in [6] when somebody makes a call request to a user in m business, the caller's agent first locates the user's agent by making a location request to her home networks. By looking up the user's profile, the home network sends back the location of the user's agent to the caller's agent. Once the caller's agent identifies user's location,

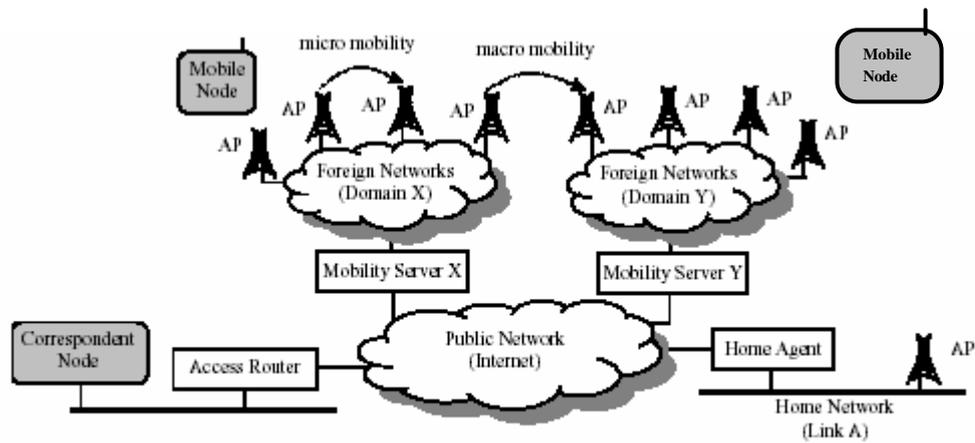


Figure 3: Hierarchical mobility management model.

the caller's agent can directly communicate with the user's agent. Thus, a mobile user reads a message using desktop PC in the office at a time, then reads the message using PDA when driving and finally reads the message using Laptop PC at home. [6] [12]

7.0 Conclusion

True mobility presents many challenges for mobile communications and networks. This paper has provided in brief mobility techniques for providing Ubiquitous Connectivity as a new methodology for m-business. There are no effective conclusions on the approaches for providing and improving ubiquitous connectivity for m-Business. In all proposals advanced the hierarchical framework model with FA's scalability stands considerable for enabling ubiquitous connectivity in mobile wireless network especially when heterogeneity is involved. And this framework can direct systematic research on seamless connectivity for future mobile networks.

Mobile IP, an extension to IP allows mobile nodes to roam transparently from place to place within the Internet usually with no disruptions of service. Standardization process is here recommended for the cooperation of various aspects of mobile networks especially FA's scalability in a hierarchical network.

8.0 References

1. C. E. Perkins, "Mobile IP", *IEEE Communications magazine*, May 2002, pp. 66-82
2. M. Welzl and M. Mühlhäuser, "Scalability and Quality of Service: A trade-off ?", *IEEE Comm. magazine*, Vol. 41 No. 6. June 2003, pp. 32-45.
3. <http://www.ietf.org/html.charters/mobileip-chater.html>
4. J. Sun, D. Howie and J. Sauvola, "Mobility management techniques for the next generation wireless networks", *Proc. SPIE Vol. 4586, Wireless and Mobile Communications, Beijing, China 2001*, pp. 155-166.
5. Official website of Arthur D. Little, www.adl.com

6. S. Y Hui and K. H. Yeung, "Challenges in the Migration to 4G Mobile Systems", *IEEE Comm. Magazine*, Vol. 41, No. 12, Dec. 2003, pp. 72-81.
7. W. Qadeer, T. Simunic, J. Ankcorn, V. Krishnan and G. De Micheli, "Heterogeneous Wireless Network management" 2003. pp. 2-4
8. R. Kalakota, "*m-business – The Race of Mobilty*". McGraw Hill (200), pp. 302-310.
9. R. Kravets, "Moving from Mobile devices to Mobile Networks", 2002.
10. P. Tarasewich, "Designing Mobile Commerce Applications", *Communications of the ACM*, 46(12), 2002, pp. 57-60.
11. H. Lach, C. Janneteau and A. Petrescu, "Networks Mobilty in Beyond- 3G Systems", *IEEE Communication Magazine*, Vol. 4, No. 7, July 2003, pp. 52-57.
12. J. Sun & J. Sauvola, "On fundamental concept of mobility for mobile communications" *Proc. 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Lisbon, Portugal, 2002, Vol. 2*, pp: 799-803

Biography

Julius.N.Aguni holds a B.Eng in Electrical/Electronic Engineering. At present, he is a Project Engineer/Consultant with PowerBank Systems Limited, Warri Nigeria. His research interests include Integration of Wireless Communications and Power Electronics technologies with concentration on Wireless Integrated Networks (WIN's) and Electronics Signal Conversion. He is a Member (M) of the IEEE & IEE. Contact him at <aguni@ieee.org>



Presenter: Julius Aguni

Modelling IrDA Transport Layer TinyTP

Pi Huang and A. C. Boucouvalas

Multimedia Communications Research Group,
School of Design, Engineering and Computing,
Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, UK
{phuang, tboucouv}@bournemouth.ac.uk

ABSTRACT. TinyTP, developed by Infrared Data Association (IrDA), is a transport protocol dedicated for the indoor point to point infrared applications. This article presents a comprehensive mathematical model for TinyTP over the IrDA protocol stacks accounting for the presence of bit errors. By implementing different TinyTP buffer sizes and various numbers of TinyTP connections, the performance of TinyTP is examined. The results show that TinyTP buffer size has a significant effect on the overall throughput.

1. Introduction

Two major transmission technologies have been used to achieve indoor wireless communication: radio frequency (RF) and infrared. Having numbers of advantages in terms of high available bandwidth, radio interference immunity and low cost component, infrared is preferred in certain cases. A large number of portable devices on the market today have been equipped with infrared ports ranging from mobile phones and digital cameras to portable computers and printers [1]. IrDA developed a set of protocols for the indoor infrared communications. The low level IrDA protocols including physical (IrPhy) [2][3], link access (IrLAP) [4] and link management (IrLMP) [5] are adopted as the industry standard and implemented on the infrared devices. Tiny transport protocol (TinyTP) is an optional IrDA layer, although it is so important and widely implemented that it is generally considered a required layer [6].

In [7], an inclusive IrLAP model is presented as the first significant work on the IrDA link layer. Subsequently many link layer performance evaluations and improvements have also been undertaken recently to address different infrared link issues including the impact on link throughput of device processing speed [8] and future increase in data rates [9]. All the previous publications focus on link performance by assuming always an infinite size packet of a single application ready to transmit. However, upper layers (e.g. TinyTP) practically offer finite size packet to link layer in a certain period of time due to protocol behaviour and limited buffer size. TinyTP also allows multiple applications to operate the IrDA link concurrently. It is therefore of interest to examine the system throughput at the TinyTP level.

The rest of this paper is organised as follows: first, we briefly describe the IrDA protocol stacks. Then, we derive a mathematical model for TinyTP which allows derivation of throughput taking into account the lower IrDA protocol stacks. Furthermore, the analysis and results are presented which compare the TinyTP throughput efficiencies by implementing different TinyTP receiver window sizes.

2. IrDA Protocol Stacks

The IrDA protocol stack illustrated in Fig.1 is the layered set of protocols particularly aimed at point-to-point infrared communications and the applications needed in that environment. A brief description of the IrDA protocol stack is as follows:

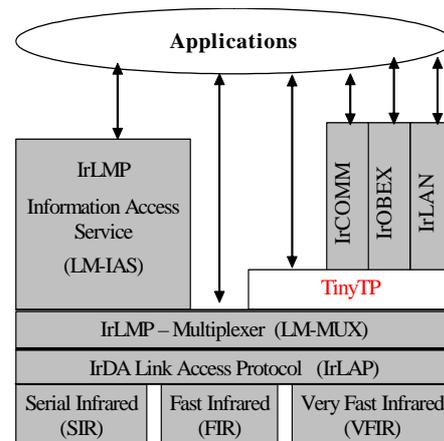


Figure 1: IrDA protocol stacks

I. IrPHY (IrDA Physical layer):

The IrDA Physical layer defines a directed half duplex serial infrared communications link established through free space to facilitate point-to-point communication. The maximum data rate currently defined by IrDA is 16Mbit/s [2] and [3].

II. IrLAP (IrDA Link Access Protocol):

IrLAP is the link access layer. By using mechanisms including retransmission, low-level flow control and error detection, IrLAP provides reliable data transfer. IrLAP transmits data in the form of frames and organizes the transmission in the manner of go-back-N (GBN) error recovery. As the physical layer defines a half duplex link, IrLAP manages the

transmission by assigning primary and secondary stations. The primary station initiates transfers to the secondary station and manages the link. When the primary completes the transmission of a window N —number of information (I) frames that can be sent before link turnaround, it then sets the Pull (P) bit in the last I-frame to signal link turnaround and request the acknowledgement from the secondary. Once P bit is set, the secondary can start sending data. It changes P bit to 0 to turnaround the link when it finishes transmission. The window size and frame size range from 1-127 and 128-16384bit respectively [3] and [4].

III. IrLMP (IrDA Link Management Protocol):

IrLMP provides Multiplexer (LM-MUX) for the support of multiple software applications or entities to operate independently and concurrently, sharing the single link provided by IrLAP between the transceivers [5]. IrLMP delivers upper layer data segments based on the ‘First In First Out’ (FIFO) queuing. We assume the multiple channels equally share the link in this paper.

IV. TinyTP

TinyTP (TTP) is a light transport protocol serving as a flow control mechanism to work with IrLMP [6]. Due to the possible deadlock problem of multiplexed channels introduced by IrLMP Multiplexer, TinyTP is very important to ensure the end to end data delivery for the application. TinyTP provides two functions: Segmentation and reassembly, and Flow control.

For TinyTP, the entire data packet from upper layers can be segmented and reassembled in Service Data Units (SDU). The maximum SDU size is negotiated at the TinyTP/IrLMP connection establishment. One SDU has to fit within one IrLAP frame. In this paper, we consider the challenge of having large application files to transmit. To make TinyTP efficient, we assume TTP is set at its maximum value $l_{TTP} = l_{LAP} - l'_{LMP} - l'_{TTP}$.

To perform flow control, TinyTP maintains a value of receiver window w for each TinyTP channel. The sender will send SDU if $w > 0$ and subtract w by 1. Therefore, each TinyTP application can send maximum w SDU without receiving acknowledgement but it has to stop when $w = 0$. w is updated by the TinyTP acknowledgement from its peer. We assume every TinyTP application has the same value of w in this paper. A flow chart of the data transmission with multiple TinyTP connections is provided in Fig.2.

3. Mathematical Modelling

For the purpose of deriving the mathematical model,

$$\begin{aligned}
 T_{send}(A) &= t_A + p_1 \left(\frac{1}{2} At_I + p(t_{Fout} + t_s) + t_{ack} + 2t_{ta} \right) + \dots + p_X \left(\frac{1}{2^X} At_I + p(t_{Fout} + t_s) + t_{ack} + 2t_{ta} \right) \\
 &= (1 + \sum_{i=1}^A \left(\left(\frac{1}{2} \right)^{\frac{1}{2}i(i+1)} (Ap)^i \right)) At_I + (1 + Ap + \sum_{i=1}^A \left(\left(\frac{1}{2} \right)^{\frac{1}{2}i(i-1)} (Ap)^i \right)) (p(t_{Fout} + t_s) + t_{ack} + 2t_{ta})
 \end{aligned} \tag{3}$$

we assume large application file (e.g. mp3) to be sent from the primary. The TinyTP segment therefore is always at the maximum size to accommodate the application data. We make use of Table I for symbol details.

Table I: Parameters used in the modelling

sym	Parameter Description	Unit
C	Link data rate	bit/s
B	Number of TinyTP connections	-
p	Frame error rate	-
N	Maximum IrLAP window size	-
w	TinyTP receiver window	-
l_{LAP}	Maximum IrLAP frame data length	bit
l_{TTP}	Maximum TinyTP segment size,	bit
l'_{phy}	Physical layer header: BOF+EOF+CRC	48bit
l'_{LAP}	IrLAP header	24bit
l'_{LM}	IrLMP header	16bit
l'_{TTP}	TinyTP header	8bit
t_I	Transmission time of an Information (I)-frame	sec
t_s	Transmission time of a Supervision (S)-frame	sec
t_{ack}	Time to transmit an IrLAP ack	sec
t_{ta}	IrLAP minimum turnaround time	sec
t_{Fout}	IrLAP F-timer time-out period	sec

I. IrLAP modelling

In this section we will derive the average time to successfully transmit one IrLAP window at a given bit error rate. As elaborated in [7], the average time to transmit one IrLAP window with length of A frames is given as follows,

$$t_A = At_I + p(t_{Fout} + t_s) + t_{ack} + 2t_{ta} \tag{1}$$

The probability of having error/errors in an IrLAP window with A frames is,

$$p_1 = 1 - (1 - p)^A \approx Ap \tag{2}$$

While error occurs in the IrLAP window with probability p_1 , due to the randomness of error occurrence, we assume that on average the error occurs in the middle of the window, and a retransmission will trigger to recover the error with window length of $0.5A$. If further error occurs in the retransmission with probability of $p_2 = p_1(1 - (1 - p)^{0.5A}) \approx 0.5A^2 p^2$, another retransmission window is needed with window length of $0.25A$, and so on. When the retransmission window is less than 1, we consider the whole window has been successfully transmitted. The average time to successfully transmit the IrLAP window is given in (3), Where X is an integer with the value of $\lceil \log_2 A \rceil$, which satisfied the retransmission window is no less than 1.

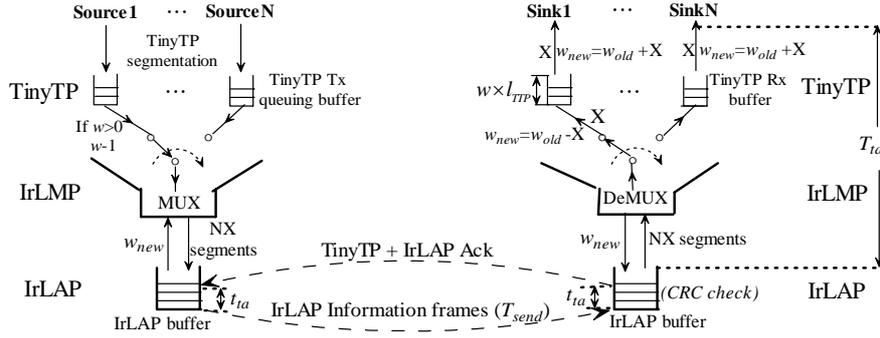


Figure 2: TinyTP data transmission

II. Derivation of TinyTP throughput

According to the standard [6], the transmission time of the TinyTP acknowledgement is given by

$$T_{ack} = (l'_{phy} + l'_{LAP} + l'_{LMP} + l'_{TTP}) / C \quad (4)$$

As shown in Fig.2, the time to consume the TTP segments in the buffer (T_{ta}) is the time accounting from passing the IrLAP frames to IrLMP to get the TinyTP ack ready at the secondary. In this paper, we assume the IrDA device uses 8bit processor and each 8bit data takes average 2 CPU cycles. As processing time T_p is the major factor of T_{ta} , we assume $T_{ta} \approx T_p$.

$$T_{ta} \approx T_p = 2Al_{TTP} / 8v = Al_{TTP} / 4v \quad (5)$$

Where A is the incoming IrLAP window size and v is the processor speed in Hz.

When a value of w is allocated for each of B TinyTP connections, the IrDA receiver has to assign a buffer with size of $B \times w \times l_{TTP}$. Given the fact that memory is highly constrained for resource-limited wireless device, such devices often can not afford large memory size for TinyTP. For a given N , three possible scenarios by implementing different w are investigated as follows. By mapping TinyTP segments to IrLAP frames, the TinyTP model is illustrated in Fig.4.

A. $Bw \leq N$

In Fig.4(A), parameters $w=2$, $B=2$ and $N \geq 4$ are employed which satisfies $Bw \leq N$. The IrLAP window will be always less than 4 due to the w constraint. As T_{ta} depends on v of the receiver and is normally much longer than t_{ta} and t_{ack} , we assume $T_{ta} > t_{ta} + t_{ack}$. After IrLAP delivers the IrLAP frames, the secondary has to wait T_{ta} before the TTP acks get ready. Since two TTP connections are considered, the secondary needs to send 2 TTP acks. Then, following the same routine another window will be sent from the primary. Therefore, we only need to consider one window transmission for the TinyTP throughput derivation.

By using (3), the average time for one TinyTP window transmission T_1 is given in (6),

$$T_1 = T_{send}(Bw) + T_{ta} + BT_{ack} + t_{ta} \quad (6)$$

The TinyTP throughput which is defined as information bits per second is,

$$D = Bwl_{TTP} / T_2 \quad (7)$$

B. $N < Bw < 2N$

In Fig.4(B), $w=3$, $N=4$ and $B=2$ are used. The 1st window has 4 segments and makes use of maximum IrLAP window. Since the secondary is fed by 4 TTP segments and has no time to consume, the secondary sends 2 TTP acks to inform the available buffer size ($w_1=w_2=1$). Then, the primary sends 2 segments in the 2nd window. Assuming the 4 TTP segments of the last window have been processed, the secondary acknowledges with $w_1=w_2=2$, and so on.

By using (3), the average time for one TinyTP window transmission T_2 is given in (8),

$$T_2 = T_{send}(N) + T_{send}(Bw - N) + 2BT_{ack} + 2t_{ta} \quad (8)$$

TinyTP throughput is,

$$D = Bwl_{TTP} / T_2 \quad (9)$$

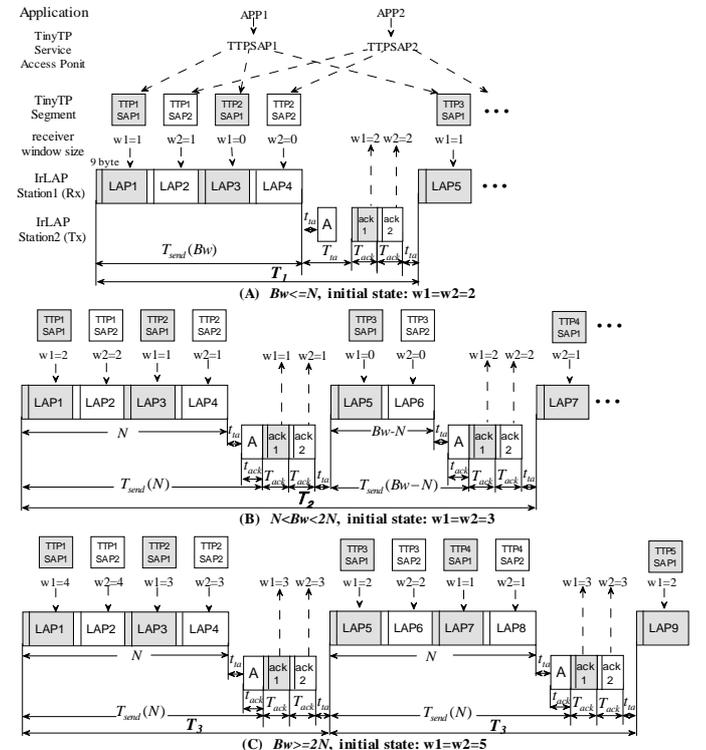


Figure 4: TinyTP transmission model.

C. $Bw \geq 2N$

In Fig.4(C), $w=5$, $N=4$ and $B=2$ are used which satisfies $Bw \geq 2N$. The 1st TinyTP window has 4 segments. The secondary acknowledges with $w_1=w_2=3$. Then, the primary sends another 4 segments in the 2nd window. Assuming the TTP segments of last window have been consumed, the secondary then acknowledges with $w_1=w_2=3$. The same procedure repeats.

As each IrLAP window has a length of N . The average transmission time for each IrLAP window is,

$$T_3 = T_{send}(N) + BT_{ack} + t_{ta} \quad (10)$$

The TinyTP throughput is,

$$D = Nl_{TTP} / T_3 \quad (11)$$

To normalise D , throughput efficiency is defined as,

$$TPE = D / C \quad (12)$$

4. TinyTP Throughput Analysis

I. Effect of TinyTP receiver window size (w)

In Fig.5, TinyTP throughput efficiencies are compared by implementing different receiver window sizes. The following parameters are used for the figure: $C=16\text{Mbit/s}$, $N=20$, $l=16\text{kbit}$, $v=10\text{MHz}$, $t_{ta}=10^{-3}\text{s}$ and $B=2$. The throughput efficiencies are plotted against the bit error rate (BER) in the range of 10^{-4} to 10^{-8} . All of the 3 throughput efficiencies (TPE) deteriorate with the increase in the BER. In the case of $w=15$ and 30, the system obtains much better TPEs than when $w=5$ especially in low BER. The TPE of $w=30$ is slightly better than $w=15$. The graph shows that the system achieves the best throughput for any BER when $Bw \geq 2N$. However, a good throughput level is also reached by using $w=15$. Therefore, a w in the range of $N < Bw < 2N$ results in good system performance as well as requiring relatively smaller buffer size.

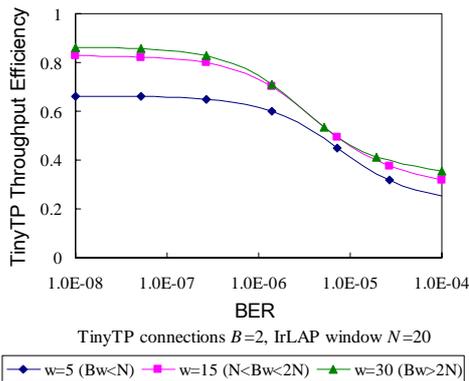


Figure 5, TinyTP TPE comparison for different w

II. Effect of IrLAP turnaround time (t_{ta})

In Fig.6, the TinyTP throughput efficiencies are plotted against t_{ta} in the range of 10^{-5} to 10^{-2}s . The same parameters as Fig.5 are used. The TPEs are compared by implementing a different value of w . The link using TTP window size of 15 and 30 has larger TPE than in

the case of $w=5$. The TPEs only drop slightly when $t_{ta} < 10^{-4}\text{s} = 0.1\text{ms}$, however, the TPEs decrease significantly with t_{ta} afterwards. By considering the system performance and hardware requirement trade off, t_{ta} of 10^{-4}s is a suitable parameter for the 16Mbit IrDA links.

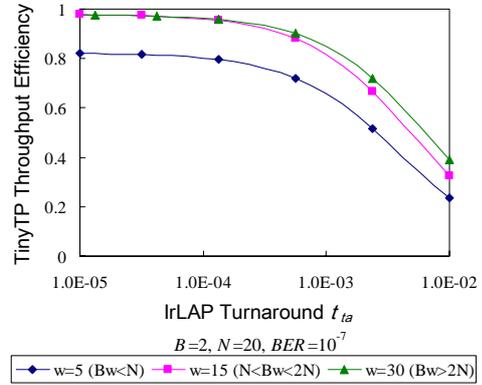


Figure 6, Effect of t_{ta} on TinyTP throughput efficiency

6. Conclusion

In this paper, we derive a comprehensive model for TinyTP in the presence of BER by considering multiple IrLMP connections and taking the underlying IrDA protocol stacks into account. Then, we compare the TinyTP throughput by implementing different TinyTP buffer sizes for various BER. The result shows that the system always achieves its best performance when $Bw \geq 2N$. However, for the memory scarce devices, TinyTP can use a receiver window size in the range of $N < Bw < 2N$ as this range also achieves good throughput as well as requiring relatively small buffer size. Finally, the effect of IrLAP turnaround time on the throughput is examined. t_{ta} of the order of 10^{-4}s is a suitable parameter for 16Mbit/s links.

References

- [1]. Williams S., "IrDA: Past, Present and Future", IEEE Personal Communications, 7(1), 11-19 (2000)
- [2]. IrDA, Serial Infrared Physical Layer Specification Version 1.1.
- [3]. IrDA, Serial Infrared Physical Layer Specification for 16Mb/s Addition (VFIR)-Errata to version 1.3, 1999
- [4]. IrDA, Serial Infrared Link Access Protocol, Version 1.1.
- [5]. IrDA, Link Management Protocol, Version 1.1.
- [6]. IrDA, Tiny Transport Protocol, Version 1.1.
- [7]. P. Parker, A.C. Boucouvalas & V. Vitsas, "Performance Modelling of the IrDA Infrared Wireless Communications Protocol," International Journal of Communication Systems, vol.13, pp589-604, 2000.
- [8]. A. C. Boucouvalas, V.Vitsas, "100 Mb/s IrDA Protocol Performance Evaluation", International Conference on Wireless and Optical Communications, pp.49-57.
- [9]. P. Chatzimisios and A. C. Boucouvalas "IrLAP IrDA Protocol Throughput Dependence on Processor Speed", International Symposium on CSNDSP 2002, pp. 272-275.

Node Discovery and Self Initialization Protocol for Distributed Underwater Acoustic Networks

A.K Othman, A.E. Adams, C.C. Tsimenidis

Underwater Acoustic Group, School of Electrical, Electronic and Computer Engineering,
University of Newcastle upon Tyne, NE1 7RU, United Kingdom.

Tel. +44 (0)191 2227273, Fax. + +44 (0)191 2228180, email: al-khalid.othman@ncl.ac.uk

ABSTRACT

There has been a large volume of research invested over the last decade into overcoming the difficulties inherent with propagation of information bearing signals through the underwater acoustic communications. This has been driven by an increasing demand for reliable and high capacity Underwater Acoustic Networks (UANs). Applications of interest in this area include oceanographic information gathering, environmental monitoring and coastal defence. One of the major challenges in the UANs area of research is the development of a networking protocol that can cope with the adverse underwater environment. In this paper we suggest a protocol that is capable of node discovery and self initialization in distributed UANs. The methods and strategies of nodes discovery will involve a primary seed node which is capable of determining neighbouring nodes and eventually all nodes in the network. By the end of node discovery, each node will have the knowledge of locations all other nodes. The self initialization involves assigning secondary seed nodes for further discovery.

I. INTRODUCTION

Underwater acoustic network may be a static network whose nodes are dropped by a plane into a sea or dynamic such as autonomous underwater vehicle (AUV). The primary seed node and nodes are equipped with an acoustic modem where they can interface with each other across the underwater acoustic network. The primary seed node could be equipped with the high speed interface such as high frequency (HF) transceiver, very high frequency (VHF) or ultra high frequency (UHF) so it can pass the data or information to the onshore station. The primary seed node is the node that responsible to start the discovery in the network with known coordinate. In underwater

acoustic network, one considers the problem of organizing nodes whose connectivity is unpredictable. This connectivity is depending on a few factors. At low frequency acoustic signal range is not constant but varies in time and frequency due to variety of factors such as

1. relative node orientation (location of the remote nodes)
2. noise level
3. propagation losses (sea state affects [1])
4. fading (bottom and upper sea wave interference)

As a result of these variations, the connectivity in the network is subject to change. Furthermore, this connectivity is affected by relative movement of the node in such environment, node and link failures and the addition of the new nodes. Consequently a very important characteristic of a mobile and widely dispersed underwater communication network is a changing of topology and long propagation delay (approximately 1550 m/s). Another characteristic should be considered in such environment is the half duplex communication where the modem cannot transmit and receive at the same time. The near far problem may occur in such underwater environment. To prevent the near far effect, scheduled transmission is required. Therefore, any organization of the network's architecture must take account of these characteristics. For these reasons, we do not used transmit power adjustment for data transmission. We assume there are only two states in the node operation either transmitting or listening. The technique to gain distance is by using the time of arrival. Recently a few approached in RF have been proposed to gain location that require few anchors (node with known coordinate) [2, 3, 4, and 5]. We propose of using one anchor (primary seed node) while tackle the limitations of underwater environment. As GPS is not used in our algorithm, we provide relative positions of the nodes with respect to the network topology. We use a term of broadcast and transmit signal to differentiate the transmission. Broadcast packet is a packet for every node that can be reached. Broadcast packets are flooded through the network, with nodes storing a history of broadcast packets so that flooding is controlled. Transmit packet is to the intended node only with identification number in its packet.

The rest of the paper is organized as follows. In section II we review the network topology suggested for underwater communication networks and the models of the system. In section III we describe the operation of

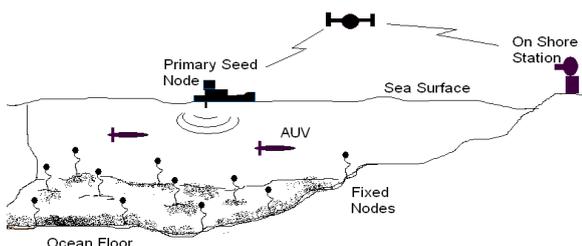


Figure 1: Typical Underwater Acoustic Network

the proposed algorithm for the primary seed node, hereafter called node discovery algorithm. In this section we also describe the Algorithm for More than Three Packet Receives at the same time and the Algorithm of Secondary Seed Nodes. Section IV shows the simulation result of the relative location of primary seed node in its local network. Conclusion and future works are concluded and discussed in section V.

II. NETWORK TOPOLOGY

What is distributed network in underwater acoustic communication?

- Nodes able to communicate each other.
- Each node knows the knowledge of distance, propagation delay, neighbours, routes information and etc.

Why distributed network?

- Better chance of providing surviving connection of communication if one node fails.
- Cover large area in the network since there is limitation of single modem transmission.
- Topology information of each node in the network can be used to:
 - Avoid congestion
 - Compute routing tables
 - Discover resources
 - Gather data
 - Scheduling
 - Traffic Monitoring

Network Models :

- Upon deployment, primary seed node knows its own coordinate
- Nodes have no knowledge of location with respect to the other nodes, number of remote node and remote node ID in the network.
- Nodes are distinguishable by an ID number
- S_0 and N_s have the synchronize clock.
- The nodes are placed randomly in some area in underwater.
- Each node has some internal memory to record local topology information.
- Nodes in the network have a multiple reception with 3 as maximum.
- The node in the network is fixed (static).

III. INITIALIZATION AND NODE DISCOVERY ALGORITHMS

A. Primary Seed Node (S_0)

Upon deployment the primary seed node, S_0 broadcast first discovery packet (DISC_COMM) to its local network. It will wait a reply from remote nodes that in its range. If there is no reply after predetermine time that means there is no node available in that area. The predetermine time is depending on maximum propagation delay. Assume that with transmission power of 180dB can reached up to 3km . Therefore

predetermine time = $\frac{2 \times 3 \text{ km}}{1550 \text{ m/s}}$. When the predetermine

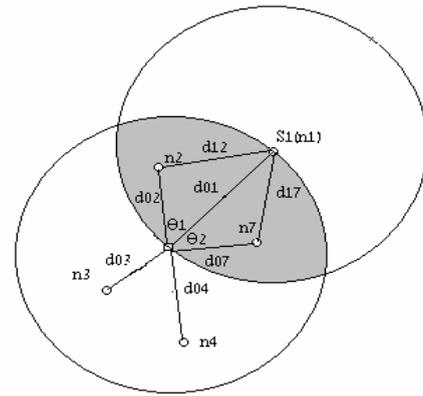


Figure 2: Selecting the Secondary Seed Node

time expired and still there is no reply from the remote node then primary seed node is still far away from the remote node and need to change its position in the network. It repeats the same procedure until it finds at least one or more nodes in the network. When the remote node fully receives DISC_COMM packet, it immediately replies by transmitting REP_PACK packet back to the sender. REP_PACK may contain node identification number. When the primary seed node receives a reply packet (REP_PACK) from the remote node, the primary seed node creates a new table and stores the information received in its memory. When the primary seed node successfully receives all information from its neighbours within a predetermined time, it immediately broadcasts this information back to its neighbours. With the distance information gained previously, S_0 is able to decide which node is the first secondary seed node. Assume S_1 is the node that has the furthest distance from S_0 . S_0 transmits a command (MORE_DISC) to S_1 for further discovery. When S_1 receives this command, it does the same procedure as S_0 and broadcasts the information to its neighbours when it receives all replies from its neighbours. When S_0 receives this information, it will check whether there are non-overlapping nodes available. If there are non-overlapping nodes in the S_0 region, it will set the next secondary seed node with the furthest node from its location. If there are no non-overlapping nodes found, then S_0 will request the next new node found in S_1 to become the next secondary seed node. The process continues until S_0 gains all information to build up a map for the relative location of the nodes in its local network. The algorithm for finding the next secondary seed node will be discussed in detail below. The map of its local network is created when S_0 finds all nodes in its network are connected at least once to each other. With this local map, S_0 requests further discovery to a new node found via a relaying node (secondary seed node). Distance (m) = $(\text{Time receive REP_PACK} - \text{Time broadcast DISC_COMM}) / 2 \times \text{Speed of sound (1550 m/s)}$

B. Algorithm for More than Three Packets Received at the Same Time

When the nodes receive more than three packets from other nodes in the network, it will choose any three

packets according to the highest power as priority and discard other packets. The nodes with high power of transmission will be ask do not reply to avoid producing interference while transmission. To do this the primary seed node will transmit a multicast packet to these nodes requesting them to not reply the second discovery packet. The primary seed node then set the power level to maximum for broadcasting a second discovery packet (DISC_COMM). In this mode the previous three nodes will not reply to any broadcast discovery packet until a START_RESPOND command is send by primary seed node or sender. The process will continue until there is no reception of more than three packets from nodes in the network.

C. Algorithm of Secondary Seed Nodes

When S_0 receives all REP_PACK from its local remote nodes, first secondary seed node (S_1) for the next discovery always furthers node from S_0 . This information can be found in the table at primary seed node by comparing the distance among the remote nodes. Below, we discussed three types of algorithm in selecting the next secondary seed node.

Algorithm with Non-Overlap Node

This algorithm uses distance in the non overlap node to find the next secondary seed node. The next secondary seed node (S_2) is the node that has furthers distance in non overlap nodes from S_0 and S_1 . Similar process in S_0 is taken by this second secondary seed node. When it successfully receives REP_PACK from its neighbours then it broadcast the information back to its neighbours including S_0 . S_0 again repeat it process to find furthers non-overlap node from S_0 and S_2 and request this node, third secondary seed node (S_3) for further discovery. If there is non-overlap node found then the next secondary seed node (S_3) is node with furthers distance from S_0 . Figure 2 shows the example of selecting the secondary seed node. Node n_1 is furthers node from S_0 and become first secondary seed node (S_1). Nodes n_3 and n_4 is the non-overlap node. If $d_{03} > d_{04}$ then the next secondary seed node is n_3 . If $d_{04} > d_{03}$ then the next secondary seed node is n_4 . If $d_{03} = d_{04}$ then S_0 will choose the node with strongest signal to become the next secondary seed node. If there is a new node found at the secondary seed node, S_0 first will request furthers new node from secondary seed node to become the next secondary seed node. The information from this secondary seed node will be passed to S_0 via it relaying node. The next secondary seed node will be the node that not overlaps with the previous secondary seed node.

IV. RESULTS

Figure 3 shows the example of underwater network topology. Let consider the nodes are separated by 3 km as maximum. The propagation delay of a signal from one node to other node requires = $\frac{3 \text{ km}}{1550 \text{ m/s}} = 2 \text{ s}$. If a

packet is 10,000 bits in length and the transmission rate is 10,000 bps, the transmission duration is 1.0 s. Then the total time required for one node to receive a packet

from the other is 3.0 s and the round trip time is 6.0 s. Processing time is assumed negligible. Figure 4 shows the first discovery by the primary seed node. When receives reply from remote nodes in its local network, its only gain the distance of the nodes replied. It does not know where exactly the location of the remote nodes. Nodes from First Discovery, $S_0 \{1, 13, 5, 10\}$
First Secondary Seed Node, $S_1, \{10\}$

Figure 5 shows the discovery of the first secondary seed node. When receives a reply from it neighbours, it will broadcast the information back to the neighbours including the primary seed node. Nodes from Second Discovery, $S_1, \{10, 12, 15, 13, 14\}$
Overlap Nodes from First and Second Discovery, $\{10, 13\}$

Non Overlap in First Discovery, $\{1, 5\}$
Second Secondary Seed Node, $S_2, \{5\}$

Figure 6 shows the discovery from the second secondary seed node (S_2) where it gain distances and node identification from its local network. Nodes from Third Discovery, $S_2 \{5, 16, 8, 7, 1\}$.

No Non Overlap Nodes Found from Third Discovery
Third Secondary Seed Node, $S_3 \{13\}$.

Figure 7 shows the third secondary seed node discovery. Nodes from Fourth Discovery, $S_3 \{13, 1, 10\}$. Figure 8 shows the final connectivity gain from the 3 secondary seed node (S_1, S_2 and S_3). The primary seed node and the remote nodes in its local network now have the knowledge of relative location to each others.

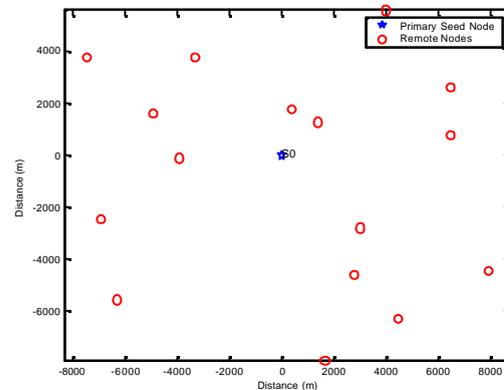


Figure 3: Nodes Deployment in Underwater Acoustic Network

V. CONCLUSION AND FUTURE WORKS

We have presented the node discovery and self initialization protocol for UANs. Their simplicity and flexibility make them attractive in static wireless ad hoc network where delay is critical. The protocols provide a concrete way to success in discovering neighbours and delay. We can use the link as other alternative to communicate the other nodes if one of the nodes failed. At the end of the process the primary seed node and its local network:

- have knowledge of their location with respect to the other nodes
- have knowledge of number of node
- have knowledge of each node's identification
- have knowledge of the distances between other nodes
- know the broken link through its status in table list
- know the fail node in the network

Future work will consist of developing the protocol and algorithm that able for the network to do further discovery and gain relative location beyond the first stage of secondary seed node.

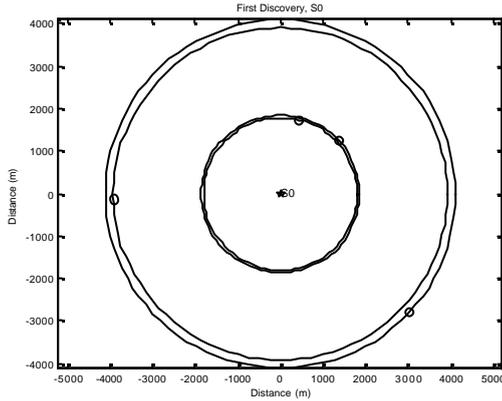


Figure 4: Discovery from Primary Seed Node, S0

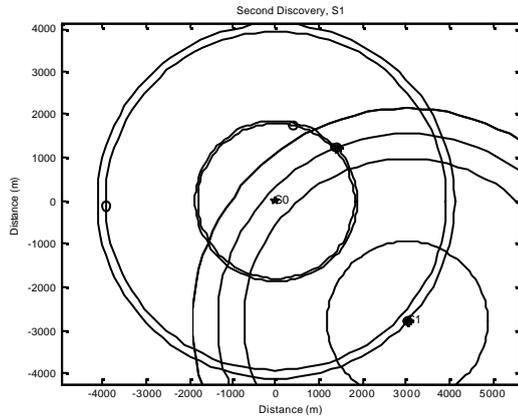


Figure 5: Discovery from First Secondary Seed Node, S1

REFERENCES

[1] D. Barrick "Theory of HF and VHF propagation across the rough sea: Application to VH and VHF propagation above the sea", Radio Sci. vol 6 pp 527-533 May 1971.

[2] Srdjan Capkun, Maher Hamdi, Jean-Pierre Hubaux, "GPS-free positioning in mobile ad hoc networks", In Hawaii Int. Conf. on System Sciences (HICSS-34) pp 3481-3490, Maui, Hawaii, January 2001.

[3] Niculescu D., Nath, B. "Ad-hoc Positioning system (APS)" Global Telecommunication Conference 2001, GLOBECOM'01 IEEE, VOL 5, pp 2926-2931, 25-29 Nov. 2001.

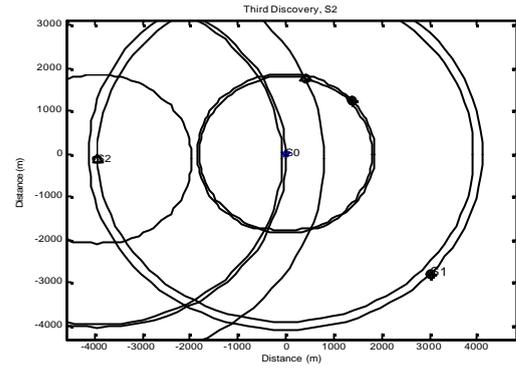


Figure 6: Discovery from Second Secondary Seed Node, S2

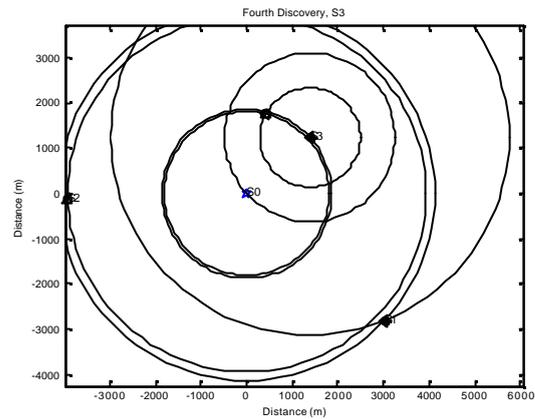


Figure 7: Discovery from Third Secondary Seed Node, S3

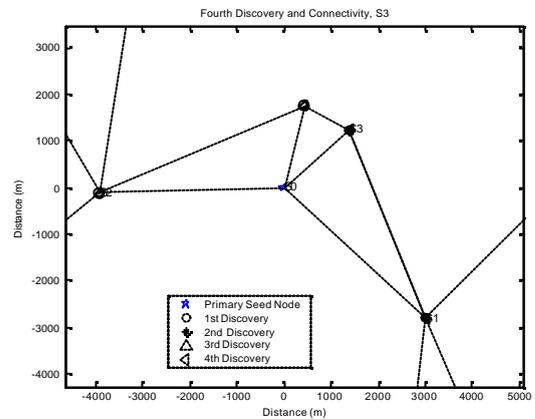


Figure 8: Connectivity from S0, S1, S2 and S3

[4] Niculescu D., Nath, B. " Ad-hoc Positioning system (APS) using AOA" INFOCOM 2003, Twenty-Second Annual Joint Conference of the IEEE Computer and Communication Societies, IEEE, Vol. 3, pp 1734-1743, March – April 2003.

[5] Chris Savarese, Jan M. "Locating in Distributed ad hoc wireless sensor network" In IEEE Int. Conf. on Acoustic, Speech and Signal Processing (ICASSP), pp 2037-2040, Salt Lake City, UT, May 2001.

Novel Rotated Channel Assignment Schemes for Non Line-of-Sight Fixed Wireless Networks

Konstantinos E. Ntagkounakis^{1,2}, Panos I. Dallas¹, Bayan S. Sharif²

¹International Co-operations R&D, INTRACOM S.A.

19, 5 Km Markopoulou Ave, Peania 19002, Greece, E-mail: {knta, pdal}@intracom.gr

²School of Electrical, Electronic & Computer Eng, University of Newcastle

Mertz Court, Newcastle upon Tyne NE1 7RU, UK, E-mail: bayan.sharif@ncl.ac.uk

Abstract

In this paper, we consider frequency Channel Assignment (CA) in the scope of non line-of-sight fixed wireless networks. First we introduce a new, highly accurate, method to characterize CA schemes based on the changes of the sector antenna orientation in co-channel cells along cell rows and columns in the network. With the new characterization scheme as a basis, we propose new CA schemes, such as Rotated-Interleaved CA (RICA) and Rotated-Rotated CA (RRCA) and compare them with traditional schemes, such as ICA. Results show that RICA outperforms the other CA schemes, for re-use factors $F_R=1, 2$ both without considering polarization assignment (PA) and when considering cell-by-cell alternative PA.

I. Introduction

Broadband Fixed Wireless Access (BFWA) systems are considered as a promising solution for the last mile access because they can achieve fast deployment, high scalability and provision of high data-rate services in a cost efficient manner. The main challenge in BFWA systems is to overcome the need for Line-Of-Sight (LOS). Great technological effort has been devoted to BFWA systems, below 11 GHz, that are able to operate in non LOS (NLOS) conditions by exploiting the presence of strong multi-path components to reconstruct the signal in the absence of a LOS path [1].

The significance of NLOS operation can be found in the fact that service can be also provided to terminals that do not have a direct signal path to the base station and as a result, the service coverage for a given area is significantly increased. The drawback is that the terminal antenna should have a wide beam to be able to receive enough indirect paths inside the main lobe to successfully reconstruct the signal. Compared to LOS systems where typical terminal antennas have a beam-width of 10° , in NLOS systems the optimum beam-width is 60° according to [1], which makes the system sensitive to co-channel interference (CCI).

The CCI effect is addressed in wireless networks through frequency planning. In mobile networks, the concept of directional antennas is considered only in the base station, while polarization assignment (PA) is not applied. In this case, CCI is reduced through advanced Channel Assignment (CA) schemes as in [2],

[3]. The interleaved CA (ICA), where channels are assigned in an interleaved orientation manner to sectors, is proposed in [2]. Compared to fixed CA (FCA) where all co-channel sectors have the same orientation, ICA experiences lower interference outage because some interferers are received through side-back lobes. In [3], the ICA and rotated CA (RCA) are considered for different re-use factor scenarios with the first scheme to achieve the best performance in most cases.

In LOS fixed wireless systems the case is different as terminals are stationary and use highly directional antennas with strong side-back lobe rejection capabilities. Considering also high sectorization schemes with narrow base antennas, the CCI is eliminated mainly through the antenna directivity while the Polarization Assignment (PA) further reduces the correlation between a link and interferers. An LMDS case is presented in [4], where RCA and strip-by-strip PA are mainly considered, however no simulation evaluation is performed.

The motivation behind our work is an NLOS-BFWA study case. Due to the use of wide base and terminal antennas that tolerate interference, we focus mainly in the CCI rejection performance of CA schemes. We introduce a new, more accurate method of characterizing CA schemes, which is used as a basis for determining the differences between existing schemes, as well as for developing criteria and methodologies to design new ones. On top of this new characterization, we introduce new CA schemes that combine the properties of both interleaved and rotated methodologies and compare their performance to traditional schemes.

The remaining parts of this paper are organized as follows: In Section 2, the new characterization scheme is described and new CA schemes are proposed. In Section 3, the simulation model and results are presented and in section 4, conclusions are drawn.

II. New CA Schemes

A. New CA Characterization Scheme

CA schemes, as described in the literature, are characterized by the way the co-channel sector orientation changes in co-channel cells (e.g. interleaved in ICA). This characterization is accurate if the co-channel sector orientation changes only in one dimension of the network layout (e.g. along cell rows) and

preserves orientation in the other dimension. In the case that co-channel sector orientation changes in both dimensions (e.g. interleaved along rows and rotated along columns), the characterization scheme is incomplete.

We propose the characterization of CA schemes by the way the co-channel sector orientation changes in both dimensions in an X-Y CA manner. X and Y stand for the way the orientation changes along the rows and columns of the network respectively, and can be *Monotonous* if orientation is preserved, *Interleaved* when assignment occurs between a pair of orientations, or *Rotated* when co-channel sectors alternate among all possible orientations in a rotated manner.

Consider a square network layout, as it appears in figure 1, where cells are perfectly organized in rows and columns. Each cell has four sectors with 45° , 135° , 225° , 315° orientations respectively. Assuming a re-use factor $F_R=1$ and four frequency channels available, each channel is assigned to one sector inside each cell. In figure 1, the co-channel sector orientation (identified by grey color for channel 1), is monotonous along rows and interleaved along columns resulting in a Monotonous - Interleaved CA (MICA) scheme. In the same sense, PA schemes can be also characterized by the way antenna polarization changes along cell rows and columns in an X-Y PA manner.

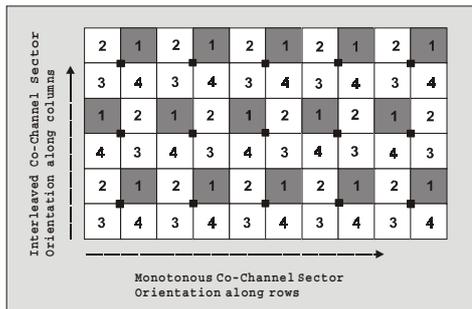


Fig. 1. Monotonous - Interleaved CA (MICA)

B. CA Methodology Optimized for NLOS-BFWA

Compared to frequency planning in cellular networks, in fixed wireless networks, CA methodology is optimized to exploit terminal directivity, while PA can be used to further reduce CCI from selected sources that the CA cannot isolate through antenna rejection.

Directional antennas are capable of receiving some of the interfering signals from co-channel cells in the side or back lobes with high attenuation and thus reducing overall CCI. In BFWA systems, where terminal antennas are directional, a higher re-use factor can be achieved compared to cellular networks that have omni-directional antennas. In this case, CA methodology should be based on the interference rejection capability of directional antennas rather than large re-use distance, applied in cellular networks. Considering that if the direction of interference and signal is the same, both experience maximum gain in the transmitting and receiving antenna, channels should simply be assigned

in a way that all co-channel sectors in close distance from a link should have a different orientation if possible. If the co-channel sector set size is greater than the number of possible orientations, there will be a subset of co-channel sectors with the same orientation. For these sectors, a large re-use distance should be kept among them.

CA scheme performance is closely related to the number of orientations the scheme exploits during assignment. By definition, FCA uses one, ICA uses two, while RCA uses all possible orientations. In the latter case, the number of co-channel sectors with the same orientation is half compared to ICA and a quarter compared to FCA. Performance of ICA and RCA is expected to improve further for a higher sectorization scheme due to the narrower sector antennas and the higher number of orientations.

On top of CA, PA can be applied to further reduce signal and interference correlation. A desired property of PA over CA is to use the antenna cross-polarization discrimination (xPD) between horizontal and vertical polarizations to further reduce interference contribution from dominant interferers (i.e. with the same orientation or small re-use distance). Since for different CA schemes, the dominant interferers are in different sectors in respect to the sector of interest, different PA schemes would interact differently with each CA scheme.

C. CA Schemes

In the view of the new characterization scheme, existing CA schemes are reconsidered and new schemes are proposed. FCA, where all co-channel sectors have the same orientation, is characterized as Monotonous-Monotonous CA (MMCA).

In ICA, as described in [2], the co-channel sector orientation can interleave along one or both dimensions, thus exploiting two or four orientations respectively. According to the new characterization, these two cases are considered as different schemes. In the first case, where orientation is preserved along rows and is interleaved along columns, the scheme is called Monotonous-Interleaved CA (MICA), as it can be seen in figure 1. When orientations are interleaved in both dimensions, the scheme is called Interleaved-Interleaved CA (IICA) and is presented in figure 2a. MICA and IICA are expected to perform differently as IICA exploits more sector orientations and thus distributes interference in more directions. Note that in MICA a dominant interferer appears in the adjacent cell along rows, while in IICA appears every two cells in both dimensions.

RCA, where the sector orientation rotates in one direction and is preserved in the other, corresponds to Rotated-Monotonous CA (RMCA). In RMCA, as it can be seen in figure 2b, the sector orientation is clockwise rotated along rows. By default, the rotated methodology exploits all possible orientations; however, along the columns a dominant interferer occurs in the adjacent cell.

Based on the new characterization scheme and the analysis of existing CA schemes, two new CA schemes for BFWA systems are proposed in this study. The Rotated-Interleaved CA (RICA) is a combination between interleaved and rotated methodology. In RICA the sector orientation is clockwise rotated along rows and interleaved by 180° in columns as it is shown in figure 2c. Compared to RMCA, a dominant interferer occurs every two cells along columns. The second CA scheme is Rotated-Rotated CA (RRCA), where the co-channel sector orientation is clockwise rotated along both rows and columns as shown in figure 2d. IICA and RICA schemes are also presented for re-use factor $F_R=2$ in figures 3a, 3b respectively. It should be noted that independently of the re-use factor, each CA scheme, as defined above, is applied by considering only the co-channel cells in the network.

The new characterization scheme can be applied in PA schemes too, as mentioned above. In this study we consider a cell-by-cell alternative polarization assignment, which according to the new characterization scheme can be identified as interleaved along rows and interleaved along columns (IIPA). IIPA pattern is presented in figures 2, 3 on top of each CA scheme, where dark colored sectors denote horizontal antenna polarization, while light colored sectors denote vertical polarization.

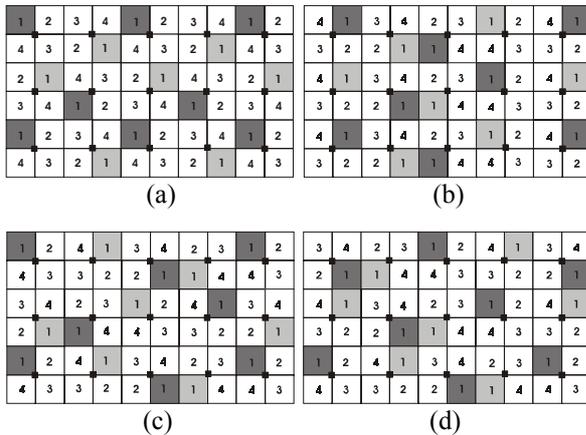


Fig. 2. CA for $F_R=1$. a) IICA, b) RMCA, c) RICA, d) RRCA.

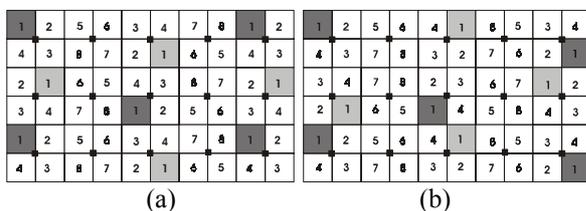


Fig. 3. CA for $F_R=2$. a) IICA, b) RICA

III. Performance Results

A. System Model

The BFWA network layout has a format of 8×8 quad-sector square cells. Each sector has a diagonal length of 5.7 Km with 300 terminals uniformly dis-

tributed inside its area. Both sector and terminal antenna gains are modelled as a function of the direction offset from the orientation-angle according to a realistic radiation pattern, as in [2]. The sector antenna has 80° beam-width, 10dB gain and 30dB Front-To-Back (FTB) ratio. The optimum terminal antenna beam-width, for NLOS operation, is 60° according to [1] with a Gain Reduction Factor (GRF) of 3dB, 11.5dB gain and 30 dB FTB. The antenna cross-polarization factor is chosen to be 7dB according to [3], [4]. The signal path loss is calculated by the “Log-distance” model considering a propagation factor of $\gamma=4$. Shadow fading is also considered, which follows a normal distribution with 0 mean and 8 dB of standard deviation.

The BFWA system, in this study, is based on Orthogonal Frequency Division Multiplexing (OFDM) technology and a TDMA / FDD access method. It should be mentioned that the proposed CA schemes can be applied in a TDMA / TDD access method also, if the downlink-uplink separation is fixed and synchronized among base station frames. CA schemes are evaluated on the downlink over a period of 10000 timeslots. Downlink SIR estimations are mapped, considering a perfect link adaptation function, to a constellation scheme and coding rate pair to produce an average throughput in Mbps per channel, according to table 1, [5]. The performance of CA is also evaluated on the outage probability of an interference-corrupted link that cannot support the lowest constellation scheme and coding rate pair (i.e. $SIR < 6\text{dB}$) at a given timeslot. The model assumes system operation at the 10.5 GHz band with a channel bandwidth of 7 MHz.

TABLE I
OFDM Modulation & Coding Spectral Efficiency

Modulation & Coding	SIR Threshold (BER 10^{-6})	Spectral Efficiency (Bits/s/Hz)
QPSK - $1/2$	6	0.72
QPSK - $3/4$	9	1.09
16QAM - $1/2$	12	1.44
16QAM - $3/4$	16	2.17
64QAM - $2/3$	20	2.89
64QAM - $3/4$	23	3.25

B. Performance Results

The first part of the simulations was devoted to an evaluation of the new characterization scheme and an investigation of interference rejection capabilities between existing and new CA schemes without considering PA. As shown in tables 2, 3 the new characterization scheme is proved to be more accurate as MICA and IICA that were previously considered both as ICA, achieve different performance. Normalized to MMCA, as shown in “no PA” column, “throughput” section in table 2, MICA achieves an improvement of 20%, while IICA offers 47.5%. Note that IICA maintains improved performance when IIPA is considered.

Among other CA schemes, the schemes that exploit more orientations, such as IICA, RMCA, RICA and RRCA achieve higher performance compared to MMCA, MICA. However, among schemes that exploit all orientations, there are still differences in performance. In this case an important factor is how close a co-channel sector with the same orientation occurs along rows and columns. For example, the same co-channel sector orientation occurs every four cells along rows in RICA, in figure 2c, while in IICA, in figure 2a, every two cells. This suggestion is justified by the results as the newly proposed RICA improves throughput by 52%, compared to MMCA, while IICA by 47%, as shown in table 2. For $F_R=2$, as shown in table 3, the performance superiority of RICA over IICA (both compared to MMCA) is 15%.

In the second part of the simulations the performance of CA schemes is investigated when IIPA is applied. Results in tables 2, 3 show that RICA outperforms the other CA schemes in all cases. In the $F_R=1$ case, RICA improves throughput by 37% compared to MMCA and by 8% to IICA. Notice that among CA schemes, the improvement due to IIPA is different. This can be explained by the fact that the position of interferers is in different sectors among CA schemes and the PA pattern affects them in a different way.

The importance of the re-use factor on the performance of CA schemes is also identified. Notice, in table 3, that opposite to $F_R=1$, in the $F_R=2$ case the RRCA performs better than IICA both with and without PA. This can be explained by the fact that the re-use factor and method (i.e. column-by-column, refer to fig. 3) changes the co-channel set properties as along rows, re-use distance is now increased.

The CA performance, as analyzed above in terms of throughput, is also verified in the outage probability results in tables 2, 3. It should be mentioned that the above performance comparison results are verified for "10dB" cross-polarization factor and a 40° beam-width terminal antenna directivity also.

IV. Conclusions

In this paper we propose a new, more accurate, characterization methodology for CA, PA schemes. On top of this characterization, we optimize CA methodology for the distinct features of NLOS-BFWA and we propose new CA schemes, such as RICA. A simulation comparison of the new versus traditional CA schemes is performed, in terms of throughput and interference outage, to determine the best scheme under different frequency re-use factors. Results show that RICA achieves the best performance compared to traditional and new CA schemes both without PA and when IIPA is considered.

Future work can be focused on the performance of CA schemes for higher sectorization and re-use factors. Another interesting topic would be an investigation on a joint CA-PA evaluation methodology.

Acknowledgments

This work is partly supported by the BroadWay project (IST-2001-32686). All partners are acknowledged. These are Motorola Labs, France, INTRACOM and University of Athens, Greece, TNO-FEL, Netherlands, IMST, IRK and Dresden University of Technology, Germany and FARRAN, Ireland. The BROADWAY project is partly funded by the Commission of the European Union.

References

- [1] D. Gesbert, L. Haumont, H. Bolcskei, R. Krishnamoorthy and A. Paulraj, "Technologies & Performance for Non-Line-of-Sight Broadband Wireless Access Networks", *IEEE Communications Magazine*, April 2002.
- [2] L. Wang and K. Leung, "A high-capacity wireless network by quad-sector cell and interleaved channel assignment", *IEEE JSAC*, no. 3, March 2000 pp. 472-480.
- [3] O. Ata, H. Seki, A. Paulraj, "Capacity enhancement in quad-sector cell architecture with interleaved channel and polarization assignments", *IEEE International Conference on Communications*, no. 1, June 2001 pp. 2317-2321
- [4] V. I. Roman, "Frequency reuse and System Deployment in Local Multipoint Distribution Service", *IEEE Personal Communications*, Dec 1999.
- [5] IST-1999-10731 - ADAMAS, 4D3 "Adaptive Algorithmic Design", Consortium Internal Report.

TABLE II
Performance of CA Schemes for $F_R=1$

CA/PA	Pr (SIR<6dB)		Throughput (Mbps)	
	no PA	IIPA	no PA	IIPA
MMCA	0.243	0.146	10.805	13.381
MICA	0.168	0.078	13.014	15.952
IICA	0.082	0.050	15.936	17.321
RMCA	0.110	0.048	15.059	17.745
RICA	0.071	0.036	16.434	18.381
RRCA	0.098	0.060	15.432	17.169

TABLE III
Performance of CA Schemes for $F_R=2$

CA/PA	Pr (SIR<6dB)		Throughput (Mbps)	
	no PA	IIPA	no PA	IIPA
MMCA	0.133	0.058	14.000	16.645
MICA	0.071	0.023	16.369	18.909
IICA	0.053	0.016	17.467	19.098
RMCA	0.096	0.038	15.907	18.507
RICA	0.017	0.010	19.594	20.436
RRCA	0.045	0.014	18.139	19.810

Offering Broadband services via the Telephone system in a Volatile Economic Environment

Alberto Fernandes
first.name.last.name@cl.cam.ac.uk
University of Cambridge

Abstract

Spectrum management in DSL protects lines of potential customers from spectrum pollution. A spectrum management decision is equivalent to estimating a price for the future revenue from these potential customers. Making explicit this pricing decision would allow mixing economic and technical knowledge into a single estimator. Keeping them separate may be leading to double conservatism where the engineers produce a conservative estimate and the economists start from there to add their own pessimism to arrive at an investment decision. We propose a model where varying economic parameters would be included in the estimator used to compute the data rate offered to each customer. This would allow to maximise expected revenue and implement a version of a financial “hedge”, where the risk is minimised by mixing optimistic and pessimistic assumptions. ¹

1 Introduction

Digital Subscriber Line (DSL) technologies owe their existence to an economic motivation: the exploitation of the existing telephony system as a medium for “broadband”, or high speed data connections. Telephone wires are far from ideal for the job, but this technical hurdle is compensated by its cost advantage, as the lines are already installed.

Thus, DSL is about exploring a finite, non-renewable resource to the best possible eco-

nomical result. This exploitation, though, occurs in a deflationary and volatile environment, with the cost of bandwidth in constant fall and competing technologies in development.

The physical make up of the system determines how it can be exploited. Telephone lines serving different customers are collected together in binders, and can be modelled as an interference channel, ie, a Multiple Input Multiple Output (MIMO) system where users are not centrally coordinated. ².

Although the number of lines in a binder is fixed, the number of active lines running a DSL service varies with time. As users gradually adopt DSL, is not known how many users will end up sharing a binder.

Every time a new customer decides to buy a DSL service, the telco is faced with the problem of offering the highest bit rate possible while avoiding damage to older and to future systems. There is an inherent trade-off between present and future users data rates. The data rate made available to the new customer is limited by the necessity of protecting the other lines in the same binder from interference.

This protection is called spectrum management. It can be done by setting fixed limitations on power and spectrum of new entrants, or by directly aiming for some feasible set of data rates to be distributed between users. The former method is the current practice today, and is detailed by standards agreed by the committee T1.E1.4 in the United States and by the European Telecommunications Standard Institute in Europe ([1], [2]). The latter is the state of current research in the area and is actively being

¹The author thanks the helpful comments and support of Dr David Greaves, Dr Miguel Rodrigues, Prof Jon Crowcroft and Prof. John M. Cioffi. Thanks to Dr Ken Kerpez of Telcordia for providing measurements and New Visual Inc for funding this research.

²As of today there are few or no exceptions to this, but there will probably be some centrally coordinated systems in the future.

discussed in T1.E1.4.([3],[7])

Any spectrum management policy has very direct economic implications. Using the basic axiom that in general telephony companies (telcos) will be able to charge more for higher data rate, a spectrum management problem is one of trade-off between revenue from current customers and revenue from potential future customers. This can be expressed as a pricing exercise.

The price in question here is the expected revenue to be achieved from future customers. For example, suppose that telco X receives an order from a new DSL customer whose line has potential to support 2 Mbps³, but only offers 1 Mbps in order to cut on “spectrum pollution”.

This decision is equivalent of that of a well-informed telco that computed the expected extra revenue from that binder if the “pollution” is kept low, and decided that the price that the current customer is willing to pay for data rate above 1Mbit/s is not enough to compensate for the potential loss of revenue. From the point of view of an external observer, the behaviour of telco X and the well-informed telco is indistinguishable; in this sense we view the decisions motivated by the spectrum management policies as equivalent to pricing decisions.

The analysis in this paper is focused on the case where a single operator controls the physical layer of a telephone binder. This applies for instance in VDSL scenarios in the US. VDSL is designed to support high data rates over short distances, as a “last-mile” solution, in conjunction with fibre. For those “hybrid” networks, that mix fibre and copper, a recent FCC ruling decided that the incumbent telephony companies are not forced to unbundle, i.e., to allow other companies to provide DSL using the incumbent’s telephone lines [6]. This decision allows telcos to adopt new strategies such as those explained here.

This paper is organised as follows. In section 2, we examine why careful pricing is necessary; in section 3, we show how it is possible to change practices by incorporating the most recent research results from DSL; and in section 4 we propose new practices and discuss benefits.

³While not interfering with current customers

2 Motivation

The “co-existence” in DSL of current and potential future customers invalidates the paradigm used in most of MIMO literature, that the goal is to maximise the bit rate sum. Bit rate sum implies that data rate is worth the same, whatever line it is in. We argue that bit rate in current customers lines is better than bit rate for potential customers, for the following reasons:

1. Bit rate in a new customer line can certainly be exchanged by revenue. The expected revenue from future customers has to be weighted by the estimated probability that these customers will take up the service eventually.
2. The estimate in item(1) has to take into account the fierce competition in provision of broadband services by other means rather than the telephone line. In particular, cable has a bigger share of the market in many OECD countries [5].
3. Future deflation of prices charged per data rate unit is generally expected.
4. Revenue comes in the form of a stream of monthly payments by the customer. Future customers will initiate their stream of payments later on. If everything else is equal, their streams of payments will last for a shorter period than the stream that is starting now.
5. Interest rates or capital return rates mean that money received now is more valuable in actuarial terms (has a higher present value) than the same amount in the future.

A second source of uncertainty comes from technical limitations. A practical assumption is that the MIMO system itself is not completely known, but is gradually “discovered” as new users take up the service. In the typical case, a telephone exchange or cabinet will serve several customers – for instance, different homes. To measure the direct channel would involve a visit to each user’s home, including to those who haven’t express interest in DSL yet. Some technology for measuring a DSL channel from a single end has recently been presented [4], but is still challenging from a theoretical point of view, so it is not reasonable to expect it to be in widespread use soon with sufficient accuracy

for this task.

Both economic and technical sources of uncertainty must be taken into account for an unbiased estimate of the value of data rate for a future customer. There is nothing new about mixing different components such as these into price. Computing price estimates that incorporate all available knowledge about a certain asset, while recognising the implicit underlying uncertainty, is the basis on which financial markets operate, for instance.

As mentioned before, spectrum management decisions being made today can be seen as implicit pricing decisions. Because they ignore most of the factors listed above that should have been in the model, there is a high risk that these are bad decisions. For instance, companies may be “paying” too much for the potential of data rate in future DSL customer lines if the rate of deflation is not being taken into account.

The economic uncertainty factors are likely not to be ignored by companies. The problem is that they are kept in a separate forum. This may lead to duplication of the conservatism, where different people in the same company build their own margins of safety. For instance, data rate capacities may be conservatively estimated by the engineers and passed on to the economists, who include their own conservatism to decide on investment policies.

The result is underinvestment in DSL, and eventually the industry missing some of the window of opportunity represented by the existing telephone infrastructure. The evidence is that, despite their initial advantage, telephone companies lost the first move advantage in providing broadband Internet services in practically all of the richest countries ([5]).

3 Potential for Implementation

A price estimator depends on knowledge about the terms of trade, i.e., how much extra data rate is gained in one line when a unit of data rate is lost in another.

Fortunately, recent advances in DSL technology much increased our knowledge about this and

allow better terms of trade to be achieved. Dynamic Spectrum Management techniques (see for example [7], [3]) moved the spectrum management discussion from simply considering whether technologies are “compatible” into identifying the “rate-region”, i.e., the set of all feasible sets of data rates for the MIMO system users. To be efficient, any MIMO system would have to operate in the convex hull of the rate region.

Even though it is not always possible to know the complete MIMO system, the downstream (i.e., from the telephone company to the user) Near End (NEXT)⁴ crosstalk functions may all be measured on the telephone companies’ side. These happens to be the most important crosstalk functions, as the downstream rates are often higher, as in ADSL or in applications like video on demand. In most cases, NEXT is far higher than FEXT.

This measurement is assumed in many new technologies and there are now less technical impediments for it to be performed. This would allow a much greater insight into which rates are achievable. Spectrum management techniques in current use a very pessimistic scenario, designed to be worse than 99% of practical crosstalk functions. The amount of conservatism used is illustrated by figure 1.

4 Recommendations

The main conclusion is that new engineering models are needed that take economic parameters into account. The practice today, when a new customer requests DSL, is to make a decision on the data rate offered based on loop measurements. We propose a new practice where loop measurements and economic parameters serve as input on which individual data rate decisions should be based. Thus these new models would provide a function that takes this data and returns a recommended data rate for each individual case.

Such a function would provide the means for a company to react to changes in the economic

⁴Near End crosstalk means a transmitter on the telephone company exchange interfering with a receiver in the same exchange; Far End crosstalk (FEXT) would be a transmitter in a customer’s home interfering with a receiver on the telco exchange.

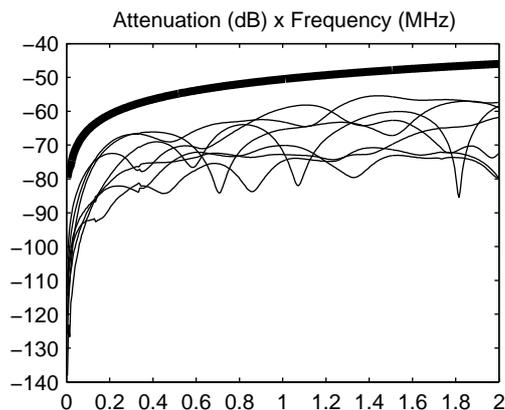


Figure 1: The thick line is the conservative estimate of a crosstalk NEXT transfer function currently in use (Unger model). The other lines are real measurements by Telcordia, for pairs close together.

and technological environment by simply adjusting parameters. This capacity of rapid reaction is very important in a sector where changes are occurring fast.

The model can be expressed as finding the rate r for the incremental new customer that maximises total revenue : $R(r, \theta, \phi) = R_c(r, \theta) + R_f(r, \phi)$, where R_c is the revenue expected from the new customer, $R_f(r)$ from future customers and θ and ϕ are vectors including all the other parameters mentioned. As a function of random variables, R is itself a random variable and the expectation of R is computed via computing the integral over its probability distribution function (pdf). Protection of older systems is considered a hard constraint and thus not included in this trade-off.

This pdf would incorporate data and beliefs about all the economic and technical parameters, including whatever inferences are possible about the shape of the rate region based on the measurements available. Even a simple model here would be better than the current situation where no explicit model is used.

A second possible application of the model is to maximise revenue subject to keeping the variance of revenue small; i.e., reduce the risk for the company that the expected revenue won't be achieved. This maximisation would be done over many rate decisions and would be similar to the concept of "hedge" in finance.

In short, the data rate offered in some lines would be based a pessimistic view of the future, following a greedy strategy, taking close to the maximum revenue possible from the current customer. In some lines, the data rate offered would reflect an optimistic view of the future, following a parsimonious strategy, keeping spectrum pollution to a minimum in order to achieve future profits from the other lines in the binder. Therefore, whatever scenario materialises would not drive the actual revenue too far from the initial expectation.

Investors in telecommunications are notoriously more risk-averse following the crisis in the sector in 2000. Such a tool for reducing the implicit risk in their returns could be essential for bringing back more investment in DSL.

References

- [1] Spectral management: Definitions and signal library tr 101 830-1. Technical report, ETSI.
- [2] Spectral management t1.417-2001. Technical report, T1.E1.4.
- [3] R. Cendrillon, M.Moonen, W.Yu, J.Verlinden, and T. Bostoen. On the optimality of iterative waterfilling. *T1.E1.4 contribution 2003-325*, December 2003.
- [4] Stefano Galli and David Waring. Loop makeup identification via single ended testing: Beyond mere loop qualification. *IEEE Journal on Selected Areas in Communications*, June 2002.
- [5] Irene Wu Sherille Ismail. Broadband internet access in oecd countries: A comparative analysis. Technical report, Federal Communications Commission (USA), October 2003.
- [6] Federal Communications Commission (USA). Fcc adopts new rules for network unbundling obligations of incumbent local phone exchange carriers. Press Release, 2003.
- [7] W. Yu, W. Rhee, S. Boyd, and J. Cioffi. Iterative water-filling for gaussian vector multiple access channels. In *2001 International Symposium on Information Theory, Washington, DC*, June 2001.

Optimization of OFDM modulation for Broadband Power Line Communication Implemented Through Wavelet Transform

Cheng Xiao-rong Hou Si-zu Yuan Jin-sha Zhao Hui-lan

North China Electric Power University, Baoding, Hebei Province, 071003 cheng3100@sohu.com

Abstract In this paper, we analyze the characteristic and performance of all kinds of modulation techniques involved in the broadband power line communication, and contrast them under the condition that adapts to power line communication. Based on wavelet transform technique and OFDM modulation technique, we propose an idea of substituting multi-scale orthogonal wavelet basis for the cosine function in OFDM for the purpose of the optimization design of OFDM system. This method presented in this paper can obviously improve the performance of OFDM, overcome the deficiency of OFDM and enforce the synthetic anti-jamming capability of OFDM.

Keywords: Broadband Power Line Communication, Modulation Technique, Orthogonal Wavelet, OFDM

1 Introduction

PLC (Power Line Communication) based on the distribution network has implemented high-speed broadband communication network through the medium of low-voltage power lines. Contrasted with telecommunication networks and cable television networks, broadband PLC has several advantages, such as resourcefulness, multi-use of a cable, all kinds of electric components can directly be used for network terminals so that it becomes the focus of the research of communication and network nowadays. The development of transmission technology and the application and improvement of modulation methods enable the development of PLC. Now the application of PLC is in operation, but there are some problems such as the transmission distance limitations, the noise effects, the nonstandard protocols, which result in the degradation of communication quality and the frequent phenomena of dropping packets. The employment of modulation technique in PLC system has directly affected data transmission quality and noise restraint in channels. The performance of several typical modulation techniques are presented and contrasted in this paper. And then we will also discuss optimizing the OFDM modulation using the wavelet transform theory, which can improve the transmission performance in PLC.¹

2 Performance Analysis and

¹ Support by the National Natural Science Foundation of China under Grant No. 60372035.

Comparison of PLC modulation

2.1 PLC modulation and its performance analysis

Conventional PLC employs FSK (frequency shift keying) modulation technique. The principle of FSK is that the carrier changes with the modulating signal, i.e. 0 or 1. It is easy to implement and widely applied to low-speed data transmission. FSK modulation has low usage of spectrum and bad anti-noise ability so that it cannot adapt to high-speed PLC. There is no channel equalizer in FSK system, because of channel distortion and unevenness, it is normal to communicate when the transmission rate is below 0.5Mbps, but if the transmission rate is higher than 1Mbps, almost the entire frequency band cannot be used. FSK cannot implement high-speed PLC when the transmission rate is between 5Mbps and 10Mbps.

QAM (Quadrature Amplitude Modulation) is a modulation mode of multi-scale combined amplitude with phase. The content of QAM has overlaid that of PSK and it has high usage of frequency band and good anti-noise ability so that it adapts to high-speed signal transmission. If applied it to PLC to obtain better transmission performance, adapt a specific design to the characteristics of PLC channel.

Spread spectrum communication employs direct sequence spread spectrum and its operating frequency varies from 4 to 20MHz. PAKE (the receiver) which is used in radio communication cannot directly be applied to power line communication because there is

severe inter-symbol interference and cannot classify several main ways with more energy in impulse response, but forward error correction method and nonlinear quantizer can be used, such as convolutional codes with constraint length 4 and code efficiency 1/2. After quantization process, interleaved convolutional encoding is followed and the output of quantizer is delivered to convolutional encoders for decoding through Viterbi algorithm. Accordingly, spread spectrum communication transmits signal during the entire available frequency band. The receiver receives signal energy from the spectrum with positive signal-to-noise ratio and constricts the short peak interference through error correction coding.

The basic idea of OFDM is to divide the available spectrum into several narrow channels, low-speed carriers or subcarriers. OFDM is a system of parallel data transmission composed of N subcarriers with equal interval at frequency. The subcarriers modulate a kind of independent data. After modulation, N subcarrier signals are sent at the same time and they maintain the orthogonal characteristic of spectrum so that the usage of the system spectrum is increased. The technique has several characteristics, as follows:

(1) it can reduce the effects of time-dispersive channels caused by multipath reflects.

(2) it can effectively overcome the effects of frequency selectivity upon the system performance and achieve channel equalization.

(3) it has high usage of bandwidth and can upgrade to high speed.

(4) OFDM signal possesses high peak-to-average power ratio and the nonlinear distortion of power amplifier will deteriorate the performance of OFDM transmission. Therefore there are high demands of linearity from power amplifier.

(5) OFDM employs the method of adding symbol period and guard interval in order to overcome the multipath fading of low-voltage power line. And the addition of guard interval makes the spectrum of OFDM signal unevenness.

OFDM possesses good performance and it is the most effective modulation mode applied to PLC. The

effect is obvious especially in frequency selective fading channels. OFDM can obtain high-speed data transmission through predicting the channel and adaptive channel process. However OFDM also has several disadvantages, such as complexity of technique, performance deterioration when there are severe attenuation and interference in a broad range of channels. Thus it is necessary to improve OFDM performance using new technology.

2.2 Comparison of modulation techniques

After testing and analyzing the performance and characteristic of PLC modulation, we draw several conclusions:

(1) FSK can only apply to low-speed data transmission, it is impossible to obtain a reliable power line communication at high speed transmission rate.

(2) Spread spectrum is applicable to PLC with the transmission rate equal or below 10Mbps. Spread spectrum communication is easy to implement and has a lower cost compared with OFDM at this transmission rate. Assuming that OFDM employs simple communication protocol similarly to spread spectrum, the performance of spread spectrum system is better than that of OFDM.

(3) If OFDM employs the channel adaptive technique and complex channel coding, OFDM possesses exceptional performance, especially adapts to frequency selective channels with channel distortion such as power line.

(4) Spread spectrum technique cannot meet the needs of higher speed power line communication during the available frequency because it needs higher spread-spectrum gain. Spread spectrum communication is generally used when data transmission is below 10Mbps.

(5) OFDM does not spread spectrum and can use modulation techniques with high availability of frequency, so that it is applicable to higher speed power line communication, especially the case that data transmission is higher than 10Mbps.

(6) OFDM can not adapt to power line communication with data rate below 10Mbps because of its complexity and high cost.

(7) Making use of spread spectrum and OFDM can resolve the problems of channel attenuation and

unevenness, big load changes, and severe noise interference.

3 Optimization design of OFDM

modulation

Optimization design of OFDM modulation using wavelet transform is presented in this section in order to overcome the multipath effect in the channels and enhance the ability to constrict severe attenuation and interference during greater ranger of frequency. Conventional OFDM system uses complex exponents sub carriers based on Fourier transform. In fact it is composed of a series of orthogonal cosine functions. If substituting new orthogonal basis for these subcarriers, the inter-carrier interference and inter-symbol interference can be decreased and thus the OFDM performance can be increased without the addition of guard interval. Signal analysis based on wavelet belongs to time domain analysis. Substituting orthogonal wavelet basis for traditional cosine

function basis, we can analyze the signal by synthesis both in time domain and frequency domain, while implementing the optimization design of OFDM modulation in broadband PLC. According to the wavelet transform theory, select scaling function and wavelet as the filter function and using equivalent filters of each subchannel in orthogonal multiplexing system based on orthogonal wavelet basis, the corresponding scaling function of low pass filters is obtained through the computation of orthogonal mirror filters. High pass filters correspond to the selected wavelet, so the sub carrier signal synthesis is corresponding to the wavelet reconstruction at the sender. At the receiver, the subcarrier signal abstraction corresponds to the wavelet decomposition. The relation of reconstruction and decomposition can be depicted as binary wavelet with transformation and reverse transformation. Making use of the reconstruction of binary wavelets we will obtain $N-1$ -order reconstruction circuit as follows:

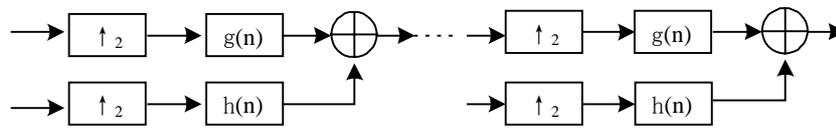


Figure 1 A circuit diagram of $N-1$ -order wavelet reconstruction (algorithm)

$$h(n) = -(-1)^{n-1} g(n-l) = (-1)^{n-1} g(l-n)$$

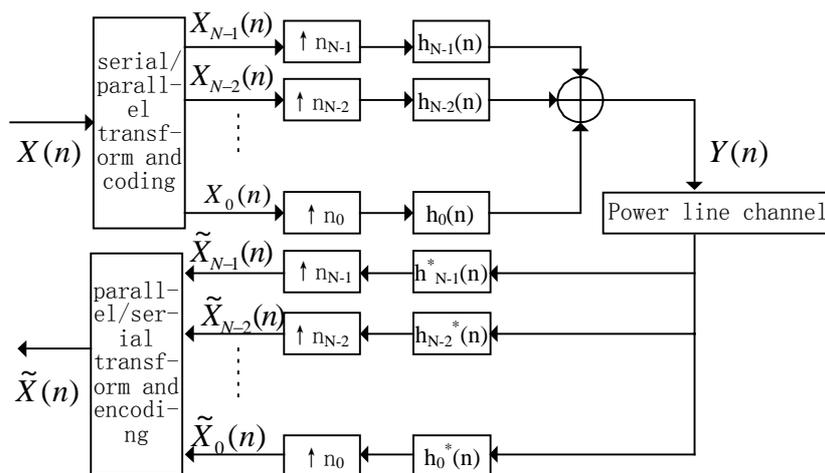


Figure2 An OFDM system with a N -channel filter bank

h possesses low pass characteristic and g has high pass characteristic. l is a positive integer and it is the length of filters. After multi-rate transform and the the sample coefficients and equivalent filters in each channel is indicated as follows:

cascade operation of filter, N -channel can substitute for conventional OFDM system, as is illustrated in figure 2.

$$n_i = \begin{cases} 2^{i+1} & i = 0, 1, \dots, N-2 \\ 2^{N+1} & i = N-1 \end{cases} \quad (1)$$

$$h_i(n) = \begin{cases} h(n) \Big|_{\uparrow 2^i} * \prod_{k=0}^{i-1} g(n) \Big|_{\uparrow 2^k} & i = 0, 1, \dots, N-2 \\ \prod_{k=0}^{N-2} g(n) \Big|_{\uparrow 2^k} & i = N-1 \end{cases} \quad (2)$$

II means the cascaded convolution of several filters, the length of each channel filter is:

$$l_i = \begin{cases} (2^{i+1} - 1)(l - 1) + 1 & i = 0, 1, \dots, N-2 \\ l_{N-2} & i = N-1 \end{cases} \quad (3)$$

where the orthogonality can be given as follows:

$$\{h_i(n) * h_k^*(-n)\} \Big|_{\downarrow n_k} = \sum_{m=0}^{\lambda_i} h_i(m) h_k^*(m - n_k n) = \delta[i - k] \delta[n] \quad (4)$$

The orthogonal wavelet filter bank composed of the parameters in equations (1),(2),(3) and (4) represents an OFDM transmission system, in which the subchannels also maintain good orthogonal characteristic. The filters of OFDM system based on orthogonal wavelet basis are generated by the equation (2) and the length of filters in each subchannel is obtained through the equation (3). The length of filters determines the summation intervals of k and m . Therefore, after optimization of OFDM modulation using wavelet method, the multipath effect on each orthogonal subchannel is reduced and the interference decreases with the increase of the length of filters, so the OFDM performance has been greatly improved.

4 Conclusion

In this paper, we analyze and discuss different modulation techniques based on broadband power line communication. Then we draw a performance comparison of FSK,PSK,QAM,OFDM and spread spectrum communication under the conditions of coding, quantization, synchronization, and power line adaptation. We also obtain the speed and condition to which each technique adaptive. Finally, we propose an optimization design method for OFDM system based on orthogonal wavelet basis. This will improve the OFDM performance, overcome the effect of multipath in channels and increase anti-jamming of OFDM during larger range of frequency. It provides

favorable techniques for broadband power line communication.

REFERENCES

- [1] Yuhua Pengy Wavelet transform and engineer application, Beijing: Science press,2000
- [2] Wei Shen,Qimei Chen,Qiang Ma, power line communication coming into the future (Basic PLC modulation and its improvement) Automation of electric power systems.Vol.27 NO.5 P71~75
- [3] Meng H,Chen S,Guan Y L,Law C L,So P L,Gunawan E,Lie T T.A transmission Line model for high frequency power line communication channel. Power System Technology.2002,Proceedings,PowerCon2002.International Conference on.Volume 2,2002.P1290-1295
- [4] N M Moyo, N.M Ljumba A C Britten. High levels of Noise Generation in the PLC system of a Ionline.Africon Conference in Africa, 2002.IEEE, Africon.6th,volume:2, 2002.P859-864
- [5] F Issa,T Sartenaer,E Marthe,F Rachidi,N korovkin,S Duteau,A pacaud.Analysis of Power Line Communication Networks using a new Approach Based on Scattering ParametersMatrix.ElectromagneticCompatibility,2002IEEE International Symposiu on, volume:2,2002. P1043-1048.
- [6]Paul,H,Moose,Member,IEEEA technique for Orthogonal Frequency Division Multiplexing Frequency offset Correction. IEEE Transactions on Communications.. VOL.42.NO. 10,October,1994,P29 08-2914

The introduction of author:

Cheng Xiao-rong: female, born in April, 1963,North China

Electric Power University, Major of electric power system and automation ,doctor, associate professor, Research field: Power line communication and network communication.

Hou Si-zu: male, North China Electric Power University, Professor, Department of communication, Research field: Electric power communication network.

Yuan Jin-sha: male, North China Electric Power University, the tutor of doctor, Major of electric power system and automation Research field: Electric power communication network.:

Zhao Hui-lan: North China Electric Power University, master of science, Research field: network communication.

OverTester: A Peer-to-Peer, Modular Toolkit for Network Performance Evaluation.

Theodore K. Apostolopoulos
Professor

George C. Oikonomou
Ph.D. Candidate

Department of Informatics
Computer and Communication Systems Laboratory
Athens University of Economics & Business
76, Patission str., 104 34 Athens,
Greece

Tel.: +30 210 8203234

Tel.: +30 210 8203155

Fax: +30 210 8203159

E-mail: tca@aueb.gr

E-mail: g.oikonomou@aueb.gr

Abstract

The field of network performance evaluation is of great significance for researchers and network administrators alike. This has resulted in a great number of applications designed to aid the process. These applications vary quite a bit, ranging from simple tools, such as the famous ping program all the way to complex and sophisticated ones. OverTester resides on top of other performance measurement programs, like a management console, automating and co-coordinating the process. It is Modular, while its most significant features are its Peer-to-Peer and Cross-Platform nature.

1. Introduction

The field of network performance evaluation is of great significance for researchers and network administrators alike. This has resulted in a great number of applications designed to aid the process. These applications vary quite a bit, ranging from simple tools, such as the famous ping program all the way to complex and sophisticated ones.

Active measurements and passive traffic monitoring [1] are two distinct approaches trying to resolve the issue. All tools fall within one, or both in some cases, of those categories. In the former case of active measurement, the common denominator of most available tools is their client – server nature. Thus, starting a test often requires the physical presence of a person, in both client and server side. This can be frustrating or outright impossible in some cases.

We envisioned a Peer -to- Peer application that would solve this problem by automating the process, combining, at the same time, all advantages of the most popular performance evaluation applications.

OverTester is the result of our vision. Just as an Overseer looks from above, supervising and managing, OverTester resides on top of other performance measurement programs, like a management console, automating and co-coordinating the process. The primary duty of the toolkit is to make a networks tester's life easier. OverTester is Modular, while its most significant

features are its Peer-to-Peer and Cross-Platform nature.

OverTester itself is not a performance evaluation tool. Re-inventing the wheel was out of scope for this project right from the start. There is a variety of great performance analysis tools, which have been around in the market for many years. They have been tested in the field and have more than proven their value. They are available either as commercial products or as free, downloadable software.

The famous Ping program [2], TTCP (stands for "Test TCP") [3], Netperf [4], the Iperf and Jperf pair of tools [5][6], Netlogger [7] and IPTraf [8] are some well-known, widely-used tools for performance measurement and/or monitoring.

In the following pages we present the reader with "A Brief Overview" of the OverTester toolkit. Its main features and functionality are documented in this section. Some details of a more technical nature can be found in the "Design Issues" section. The "Security Considerations" section, tries to illustrate the most important security issues. A brief "Conclusions – Further Work" section follows. "Acknowledgements" conclude this document.

2. OverTester: A Brief Overview

As we previously mentioned, in the brief introduction, the main reason for developing this software was to automate and co-ordinate the testing process. One can imagine OverTester as a management console. It may be used to launch a variety of other performance analysis applications.

When the testing process is complete, it gathers the results and presents them to the person in charge.

One of the main features of the program is its modular nature. OverTester does not explicitly “support” specific testing programs. The user, through the use of configuration files [Figure 1], which we call OTCFs (stands for OverTester Configuration Files), has a way of describing the tool he wants to use and how the tool is used. Thus, OverTester may be used in conjunction with virtually any testing program. All that is needed is the performance analysis software and the relative configuration. The configuration process itself is quite simple and straightforward.

```
[SomeTestingTool]
ver:1.7
path:/usr/bin/sometestingtool
PacketSize;-z;int;textbox
Send;-r;bool;checkbox
Target;-t;string;textbox
```

Figure 1 - Portion of a Configuration File

When a user launches OverTester, the configuration files are parsed. Any supported testing tool is made available for use, through the OverTester’s GUI.

For instance, in Figure 1 we describe a testing tool named “SomeTestingTool”. When OverTester launches, three options will be available to the user.

- “Packet Size”, which will be set via a text box. It will be represented as an integer and corresponds to parameter `-z` of the testing tool.
- “Send”, which corresponds to parameter `-r`. It is a Boolean value controlled by the use of a check box.
- “Target”. A string value corresponding to the command-line parameter `-t`.

Through the above example it becomes clear that the user does not need to memorize application specific, command line arguments. He simply sets the parameters of the test, in a text-box, check-box manner. The appropriate command line is generated by OverTester and is passed to the relative program.

One may also notice that the configuration file mentions the version number of the testing software. The need for this will become clear later on in this text, where we will discuss a simple scenario.

Thus said, we proceed with the description of a typical scenario. Imagine a network administrator who wishes to measure the performance of the link between two PCs (see Figure 2). We will, henceforth, call the user’s PC the “Test Initiator” and the other PCs will be called “Participants”. All three PCs have the OverTester software installed. In the case of the Participants, OverTester runs in the background, listening for requests. The user, sitting at the initiator’s console, decides to start a

test. He specifies the participants of the test and sets the parameters to be used. He then issues a command to start the test.

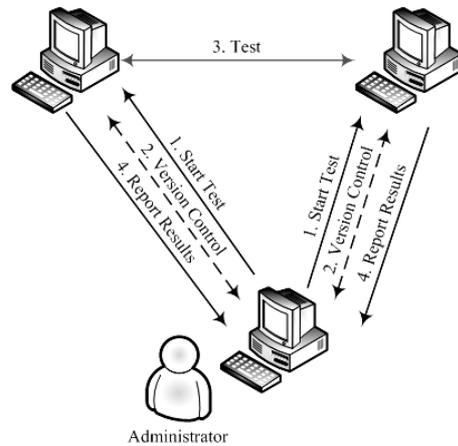


Figure 2 - A Typical Scenario

OverTester generates the commands that need to be sent to the participants. Afterwards, the following actions take place.

1. A “Start Test” message is sent to the two PCs. OverTester is running as a service there. The appropriate commands are also dispatched at this stage.
2. A version control phase follows. During this phase, all three PCs will tune to the latest version of the OverTester software as well as the software that will be used for the test (for instance “SomeTestingTool”). This is done transparently and needs no user intervention of any kind.
3. The two PCs perform any test specified according to the parameters obtained by the initiator in the initial message exchange.
4. The test’s output is redirected to files, which are sent back to the initiator at the end of the test.

The entire process is asynchronous, from the initiator’s point of view. That is, the initiator does not become “blocked waiting” for the results. Instead, one thread is listening for the results, while the administrator may order new tests between other network nodes.

On the other hand, a participant node is a state machine. When the state is “testing”, it will not accept any commands to initiate a new test and will respond with an error message stating the reason. This is done in order to avoid heavy network traffic initiating from or directed to a single node.

3. Design Issues

When designing a network application, the protocol specification is one of the most crucial steps. We ended up with an adequate protocol after quite a few brainstorming sessions. The protocol runs on top of UDP for all message exchanges. Messages share the same structure (Figure 3).

A detailed explanation of the message format is out of scope for this paper. However, we will briefly describe the overall logic.

Category	Type	Data
4 bytes	4 bytes	Variable

Figure 3 - Message Format

Messages may belong to one of four categories. “Version” messages (the value for “Category” is “vers”) have to do with the version control subsystem. “Test” messages (value “test”) have to do with test parameters, commands and results. “Management” messages (value “mgmt”) are used for various management issues. Finally, “Authentication” messages (value “auth”) are reserved for future use. They will be used for user authentication and permission control, features described in section 4, further on in this document.

For each category, there are various message types, which are recorded in the “Type” field of the Protocol Data Units (PDUs) exchanged.

Finally, the data field is optional and carries various information, depending on the PDU “Category” and “Type” fields.

Right from the start, we wanted the OverTester software to be cross-platform. For this reason we started developing two separate versions.

One version is purely for the Microsoft Windows Operating System and is coded using Visual Basic. This version is almost a hundred percent complete. We are, currently, in the final testing stages.

The second version is being written in Tcl, a well-known scripting language, much like Perl and Python. This version will run in any UNIX based system, as well as Windows-based systems. However, development and testing, up to now, has taken place in a PC running Debian/GNU Linux. The software had not been tested on other UNIX platforms as of the time of writing this paper.

No specific libraries are required in order to use the software, apart from the Visual Basic and Tcl runtime environments. The two versions will be fully inter-operable.

OverTester has a restriction, which is worth pointing out. It may only be used in conjunction with applications that are capable of launching a test from a single command line, without requiring any further user intervention. This fact does not imply that GUI-based testers are excluded. It merely means that the testing application must be capable of reading all the parameters of the test (IP addresses, protocols and so on) from a single shell command. The same applies for both windows and Linux versions.

As a final note, we would like to point out that “Test Initiators” as well as “Participants” use the same binary file. Thus, a user may initiate a test from any node. This, of course, implies that the toolkit is designed according to the peer-to-peer

philosophy. All nodes may act as “Initiators” and “Participants” at any time without need for modifications or fine-tuning. One installation is enough in order to have all functionality available.

4. Security Considerations

Security issues are worth discussing. The current design of the toolkit makes it vulnerable to various security hazards.

Denial of Service (DoS) attacks may be initiated based on the nature of the toolkit. Imagine a malicious user, having access to a host with OverTester installed. He could initiate a large number of tests, which might involve transmitting an immense number of over-sized packets for long periods of time. This would result in very high network resources utilization. In the worst-case scenario the infrastructure could even collapse altogether.

OverTester involves a protocol that results in remote execution of commands. Care should be taken in order to avoid execution of malicious programs, instead of the supposed testing tools. The version control system relies on the exchange of binary executables. Imagine a compromised host, initiating tests and claiming that it is equipped with the new version of “SomeTestingTool”. All other hosts will eventually get this version and use it for tests. If the executable is not really a newer version but some kind of Trojan, some hosts will eventually end up executing unauthorized, probably even harmful, code.

For the solution of the above problems, we consider implementing a user scheme, with each user having specific permissions. For instance, some user A may initiate tests, while user B may only review the results of previous tests. There may be yet another user C, who may be able to initiate a test but if a version inconsistency is detected, the system will not proceed to binary updates.

All information related to users, passwords and permissions will be stored in Lightweight Directory Access Protocol (LDAP [9][10]) servers. All hosts will check if an authorised user has initiated any request, before taking any action.

In order to protect from hazards related to packet interception techniques, such as packet sniffing and eavesdropping, we plan protecting all message exchanges using encryption techniques, thus securing the communication protocol itself. This could be done with the Secure Sockets Layer (SSL [11][12]) protocol or the Transport Layer Security (TLS [13][14]) protocol.

Using X.509 certificates [15] to verify the identity of the hosts is another security-related feature of the toolkit, which will be implemented in the future. This is how we intend to prevent unauthorized hosts from issuing commands to the OverTester’s network.

5. Conclusions – Further Work

A few paragraphs above we mentioned using LDAP servers for storing the user scheme. However, LDAP will play another important part in the toolkit. Apart from storing user information, the server will be a repository for binaries. The version control will be based on information stored in this repository. Thus, all hosts that have out-of-date versions of software will update from a central location. Notice the differences between Figure 4 (below) and Figure 2 in page 2. In order to avoid having a single point of failure, there is the capability of adding one or more backup LDAP servers replicating the master.

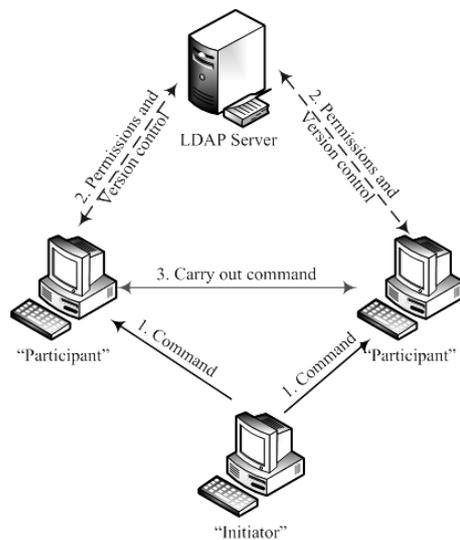


Figure 4 - Using LDAP Servers

In the future, a Java-based version of the toolkit will be developed and made available for use. It will incorporate all features of the other releases and will use the same style of configuration files. Needless to say, there will be full interoperability among all versions.

Along with the Java version, we are planning a major re-engineering of the configuration files (OTCFs). Instead of using the current format, mentioned in paragraph 2 above, the files will be written in XML (XOTCFs). We have found out that this will have many advantages. Describing them is out of scope at this time.

We consider implementing a module designer and adding it to the toolkit. This little utility will provide a friendly user interface for creating and editing OTCFs and XOTCFs. This will make the task of building a configuration file an even more trivial process. The user will not be forced to dive into the specifics of the language used. The development of the module designer will take place at a later stage.

OverTester will soon have its home in Cyberspace. As soon as the project is finished, overtester.ccslab.aueb.gr will be the URL from which, interested parties will have the chance to

download the software (binaries and source) along with all relative documentation.

6. Acknowledgement

The authors wish to send an ACK frame to Dionysis Bastas and Stratos Tsarouxs. At the time of writing this paper they were both in the final year of their graduate studies in the Department of Informatics of the Athens University of Economics and Business. Their contribution to the project was of utmost significance.

References

- [1] N. Brownlee, C. Loosley, "Fundamentals of Internet Measurement - A Tutorial", *CMG Journal of Computer Resource Management*, Issue 102, April 2001
- [2] "The Story of the PING Program"
<ftp.arl.army.mil/~mike/ping.html>
- [3] "Test TCP (TTCP) Benchmarking Tool for Measuring TCP and UDP Performance"
www.pcausa.com/Utilities/pcattcp.htm
- [4] "The Public Netperf Homepage"
www.netperf.org/netperf/NetperfPage.html
- [5] "Iperf 1.7.0 - The TCP/UDP Bandwidth Measurement Tool"
dast.nlanr.net/Projects/iperf/
- [6] "Jperf version 1.0 - The Iperf Front-End"
dast.nlanr.net/Projects/jperf/
- [7] B. Tierney, W. Johnston, B. Crowley, G. Hoo, C. Brooks, D. Gunter, "The NetLogger Methodology for High Performance Distributed Systems Performance Analysis", *Proceeding of IEEE High Performance Distributed Computing conference (HPDC-7)*, July 1998, LBNL-42611
- [8] "IPTraf - An IP Network Monitor"
cebu.mozcom.com/riker/iptraf/
- [9] M. Wahl, T. Howes, S. Kille, "Lightweight Directory Access Protocol (v3)", *Request for Comments 2251*, December 1997
- [10] J. Hodges, R. Morgan, "Lightweight Directory Access Protocol (v3): Technical Specification" *Request for Comments 3377*, September 2002
- [11] Hickman, Kipp, "The SSL Protocol", *Netscape Communications Corp.*, February 1995.
- [12] A. Frier, P. Karlton, P. Kocher, "The SSL 3.0 Protocol", *Netscape Communications Corp.*, November 1996
- [13] T. Dierks, C. Allen, "The TLS Protocol Version 1.0", *Request for Comments 2246*, January 1999
- [14] S. Blake-Wilson, M. Nystrom, D. Hopwood J., "Transport Layer Security (TLS) Extensions", *Request for Comments 3546*, June 2003
- [15] R. Housley, W. Polk, W. Ford, D. Solo, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", *Request for Comments 3280*, April 2002

Packet Scheduling Configuration in Proteo Network-on-Chip

David A. Sigüenza-Tortosa and Jari Nurmi
Tampere University of Technology, Finland
david.sigüenza-tortosa@tut.fi

June 1, 2004

Abstract

A mathematical analysis of packet scheduling oriented to design-time configuration of an asynchronous on-chip network is presented. This analysis is motivated by the need to incorporate information about the physical layout of the system. For a restricted class of problems, the analysis serves the purpose of determining the optimal parameters for the schedulers in the network, involving only basic arithmetic operations. The analysis technique is demonstrated through a design example.

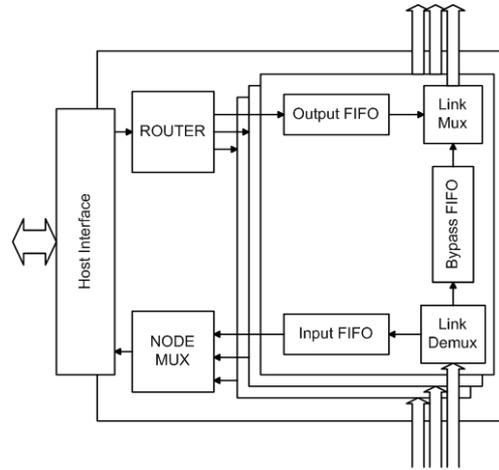


Figure 1: Extended node architecture.

1 Introduction

Internal communication structures in future Systems-on-Chip (SoCs) will face important challenges [1]. It seems clear that part of the solution includes the adoption of asynchronous interconnection structures [2] and Networks-on-Chip (NoCs) [3].

At Tampere University of Technology (TUT), we are developing a NoC architecture that we call *Proteo*. Based on a small library of parameterized components, we are able to implement a range of different network configurations. A comprehensive description was published in [4].

A fundamental part of this project is the definition of a design methodology and the prototyping of software tools to help choosing and configuring a particular instance of the network, according to a set of requirements. The use of asynchronous communication protocols forces us to consider explicitly the physical geometry of the system in the process. In this paper we present a simplified mathematical analysis that allows easy incorporation of parameters, such as link length, to the packet scheduler configuration. A design example is included.

2 Proteo Node Structure

We have chosen a highly modular structure that makes easy its configuration and tuning (figure 1). The *Output*, *Input* and *Bypass FIFOs*, plus the *link-mux* and *link-demux* blocks, form what we call a *port-module* block. The node architecture is built using at least one port-module and an *interface block*. This interface block translates the functional IP's (host) interface signals, subpacketizes the bursts coming from the host, and place the resulting subpackets in the output buffers. Conversely, it translates the format of the packets placed in the input buffers to the correct interface signals. In order to provide support for multi-dimensional networks we connect several port-module blocks, together with a *router* and a *node-mux*. The router block contains a configurable routing table used to decide to which port-module to send each packet presented by the interface block. The node-mux block selects the next packet from the ones offered by the pool of port-modules, based on some local priority scheme. A network node works either as a *requester* (being able to initiate transactions) or as a *responder* (only able to serve other's requests), but not both.

3 Packet Scheduling in Proteo

Packet scheduling in Proteo NoC takes place at the link-mux block, using one of three different work-conserving schedules we defined in [5]. When the next packet to be transmitted over the output link is selected from the Output FIFO, it is said that link-mux has *injected* a packet. The minimum guaranteed average rate of injection at each link-mux, which we call ρ_i , depends on the scheduling algorithm parameters. We would like ρ_i to use values as big as possible. However, if the values are too high, the network becomes unstable. An informal definition of *network stability* can be found in [6]: a network is said to be stable if the number of packets in the system remains bounded as the system runs for an arbitrarily long period of time. Since we can not implement infinite buffers on-chip, instability manifests itself in two ways: *overflow*, when any further increase of ρ does not produce any change, and *deadlock*. The state of the system just before entering either overflow or deadlock is called *saturation*. Saturation is attractive because network utilization is highest and scheduling algorithms are more predictable. An analysis of the problem of conveniently choosing the different ρ_i requires not only to consider communication requirements and network topology, but also the physical properties of the links. For example, differences in link length translate into different link capacities.

3.1 Definitions

In the rest of the article, we will use the following naming convention: nodes in the network are labelled with an index i . When we connect two nodes together with an unidirectional link, the upstream node takes the smallest index, and the downstream node takes the index $i+1$. The link connecting them receives the label of the node driving it, i. e. the upstream node. All operations on these indices are understood to be modulo N (the number of nodes in the network).

Paths and induced traffic

A path P_j in the network is an ordered set of network links that represents a session between two nodes in the network (a requester and a responder). The *injection (absorption)* rate of a path P_j node is the proportion of packets belonging to that path that are present in the requester's (responder's) outgoing (incoming) flow. If the network is in a steady state, in order to keep the network stable, the injection rate for each path must be equal to its absorption rate.

We define the total *induced traffic*, ϕ_i , in link i as:

$$\phi_i = \sum_j^M R_j I_{ij} \quad (1)$$

where R_j represents the injection rate of path or session P_j , and I_{ij} is a truth function that takes the value 1 if link i is in path P_j and 0 otherwise.

Node injection and absorption rates

The *injection rate* of node i , r_i , is defined as the proportion of the outgoing traffic that is taken from the Output FIFO. Note that r_i is a variable and depends on dynamic aspects of the network, while ρ_i is a constant defined at configuration time. The *absorption rate* of node i , represented by μ_i , is defined as the proportion of incoming packets that are addressed to node i , and thus removed from the flow. The value of μ_i depends on the traffic distribution pattern in the network and can be calculated as:

$$\mu_i = \frac{\phi_{i-1} - \phi_i}{\phi_{i-1}} \quad (2)$$

Activity rate

We call *activity rate*, χ_i , to the number of packet transmissions across link i per unit time. The activity rate of a link is constrained in the following manner:

$$\chi_i = \min\left\{\chi_{i-1} \frac{1 - \mu_i}{1 - r_i}, \chi_{i,\max}, \chi_{i+1} \frac{1 - r_{i+1}}{1 - \mu_{i+1}}\right\} \quad (3)$$

where the min operator evaluates its argument expressions and returns the smallest.

We call *hop time*, h_i , the time it takes to transmit one unit of data across link i . As a first approximation, we could consider that transmission delay is directly proportional to the link length, l_i . If h is the hop time for a link of unit length, and t is the unit time, the maximum activity rate of link i can be expressed as:

$$\chi_{i,\max} = \frac{t}{h_i} = \frac{t}{h \cdot l_i} \quad (4)$$

Stability condition

In order to present our own stability definition, we will have to define two concepts first. The *total available work* is defined as: $X = \sum_i^N \chi_i$. The *total demanded work* is defined as: $\Phi = \sum_i^N \phi_i$. We postulate (without proof) that the stability condition for the network is:

$$\Phi \leq X \quad (5)$$

Saturation state

If a link is able to work at an activity rate $\chi_{i,\max}$, we will say that node (or link) i is *unconstrained*. Saturation will be reached when the activity rate of the unconstrained links is maximum. Under these conditions, the r_i parameter of an unconstrained node i is equal to ρ_i .

3.2 Design example

In order to keep the analysis simple, in this example we will consider the system depicted in figure 2. First, we will consider all links to have the same length, equal to the unit length. If we choose the time unit such that $t = h$, from equation 4, we deduce that the maximum activity rate for every node is 1.

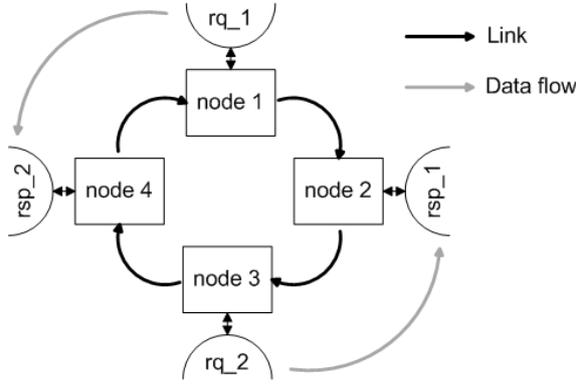


Figure 2: Example system.

For requester nodes, the constraint equation 3 is simplified:

$$\chi_i = \min\left\{\frac{\chi_{i-1}}{1-r_i}, \chi_{i,\max}, \frac{\chi_{i+1}}{1-\mu_{i+1}}\right\} = \chi_{i,\max}$$

Since we will assume that $r_i > 0, \forall r_i$ (the nodes are actually injecting traffic), and looking at the traffic pattern, it is obvious that $\mu_i > 0, \forall \mu_i$. It follows that the requester nodes (indices 1 and 3) are unconstrained. For responder nodes, both neighbouring nodes are requesters:

$$\chi_i = \min\{\chi_{i-1}(1-\mu_i), \chi_{i,\max}, \chi_{i+1}(1-r_{i+1})\}$$

In saturation state, in order to keep the network stable, the injection rate for each path must be equal to its absorption rate. Since there is only one path coming out of each requester and each path has a different target responder, for path P_1 we have:

$$\chi_3\mu_4 = \chi_1r_1$$

From which we derive $r_1 = \mu_4$, and similarly $r_3 = \mu_2$. Thus, the first and the third bounds for the activity

rate of a responder are equal, and smaller than the second bound (maximum rate). We select the right hand expression. Applying saturation conditions, the total available work is:

$$X = \chi_1(2-r_1) + \chi_3(2-r_3) = 4 - \sum_i^N \rho_i$$

Since there is only one path per requester, $R_j = r_i$, being node i the node where P_j originates. Taking into account equation 1, and applying saturation conditions, the total work demand:

$$\Phi = 3r_1 + 3r_3 = 3 \sum_i^N \rho_i$$

Finally, applying the stability condition expressed in 5:

$$\sum_i^N \rho_i \leq 1$$

Since the traffic distribution is symmetric, it makes sense to choose ρ_i such that $\rho_i = \rho, \forall \rho_i$. We deduce that the network will be stable for $\rho \leq \frac{1}{2}$.

Now, let's make link 2 three times longer than the rest of the links ($l_2 = 3$). Applying equation 4, we deduce that $\chi_{2,\max} = 1/3$. Since link 2 is much slower than the rest, we will assume that it is unconstrained and thus $\chi_2 = 1/3$. We consider the following equations:

$$\chi_1 = \frac{1}{3(1-\mu_2)} \quad (6)$$

$$\chi_3 = \frac{1}{3(1-r_3)} \quad (7)$$

$$\chi_4 = \min\{\chi_3(1-\mu_4), 1, \chi_1(1-r_1)\}$$

The injection rate for each path must be equal to its absorption rate:

$$R_1 = \chi_3\mu_4 \rightarrow r_1 = \frac{\chi_3}{\chi_1}\mu_4$$

$$R_2 = \chi_1\mu_2 \rightarrow r_3 = \frac{\chi_1}{\chi_3}\mu_2$$

Applying the constraint equations 6 and 7, we find that:

$$r_3 = \frac{1-r_3}{1-\mu_2}\mu_2 = \mu_2$$

$$r_1 = \frac{1-\mu_2}{1-r_3}\mu_4 = \mu_4$$

Then, the total available work:

$$X = \frac{1}{3} \left(\frac{4 - \sum_i^N r_i}{1-r_3} \right)$$

And the total work demand:

$$\Phi = 3(R_1 + R_2) = 3 \frac{1}{3(1-r_3)}(r_1 + r_3) = \frac{\sum_i^N r_i}{1-r_3}$$

And applying the stability condition 5, we reach the result:

$$\sum_i^N r_i \leq 1$$

Notice the similarity with the case in which we considered all link lengths equal. The difference is that in this example the requester nodes are all constrained and we can not directly derive any specific value for the parameters ρ_i . In order to do so, we assume their outgoing links are working at maximum activity rate, so we can equate r_i and ρ_i . From constraint equations 6 and 7, making $\chi_1 = \chi_3 = 1$ we conclude $\mu_2 = \rho_3 = 2/3$. Using expression 2:

$$\begin{aligned} \mu_2 &= \frac{R_2}{R_1 + R_2} = \frac{2}{3} \rightarrow R_2 = 2R_1 \\ \mu_4 &= \frac{R_1}{R_1 + R_2} = \frac{1}{3} \end{aligned}$$

Finally, we deduce $\rho_1 = 1/3$. All the values obtained are valid, and they comply with the stability condition previously found, so we conclude that this configuration is correct.

4 Conclusion

In the context of NoC, conventional network design practices and results must be adapted to the specific constraints of implementation in the submicron domain. We have introduced an analysis technique with the aim of determining the optimal values for the scheduler parameters in an asynchronous network, incorporating physical layout properties. We are working in adapting our analysis to make it suitable for automatization.

References

- [1] J. Cong, Lei He, Kei-Yong Khoo, Cheng-Kok Koh, and Zhigang Pan. Interconnect design for deep submicron ICs. In *Proceedings of ICCAD*, San Jose, USA, November 1997.
- [2] J. Muttersbach, T. Villiger, H. Kaeslin, N. Felber, and W. Fichtner. Globally-asynchronous locally-synchronous architectures to simplify the design of on-chip systems. In *Proceedings of the 12th Annual IEEE International ASIC/SOC Conference*, Washington DC, USA, September 1999.
- [3] W. J. Dally and B. Towles. Route packets, not wires: On-chip interconnection networks. In *Proceedings of DAC*, Las Vegas, USA, June 2001.
- [4] I. Saastamoinen, D. Sigüenza-Tortosa, and J. Nurmi. An IP-based on-chip packet-switched network. In A. Jantsch and H. Tenhunen, editors, *Networks on chip*, chapter 10, pages 193–213. Kluwer Academic, 2003.
- [5] D. A. Sigüenza-Tortosa and J. Nurmi. Packet scheduling in Proteo NoC. In *Proceedings of IASTED PDCN*, Innsbruck, Austria, February 2004.
- [6] M. Andrews, B. Awerbuch, A. Fernández, T. Leighton, Zhiyong Liu, and J. Kleinberg. Universal-stability results and performance bounds for greedy contention-resolution protocols. *J. ACM*, 48(1):39–69, 2001.

Performance Investigation of a Broadband OFDM System Incorporating Selection Diversity and Turbo Coding

EMMANOUIL I. KALANTZIS^{1,2}, BAYAN S. SHARIF², PANAGIOTIS I. DALLAS¹

¹INTRACOM S.A. Telecom Industry, 19.5 Km Markopoulo Avenue, 19002 Paiania, GREECE.

Phone: (+30) 210 6679042, Fax: (+30) 210 6671312

email: ekal.pdal@intracom.gr

²University of Newcastle upon Tyne, EECE Department, Newcastle upon Tyne, NE1 7RU, UK.

Phone: (+440) 191 222 6539, Fax: (+440) 191 222 5685

email: Bayan.Sharif@newcastle.ac.uk

Abstract: -Wireless communication systems and terminals operate in an environment in which multipath effects and interference limit performance. These effects are most severe when a wireless device engaged in a communication is moving and there is no direct line of sight between itself and the receiving station. To help offset these limitations, diversity techniques (i.e., Spatial, Temporal, Polarization) are used, many of which are very practical and provide significant link improvement with little added cost. In this paper a well-known diversity technique, known as selection diversity (SD), will be investigated. This scheme will be applied in systems utilising several well known (convolutional coding-CC) and recently presented (turbo coding-TC) techniques used to combat the impact of selective fading on the OFDM system performance. Also the adaptivity technique through bit loading has been considered for a further improvement in system performance.

It will be shown that the turbo coding scheme, especially when combined with SD, provides sufficiently robust performance, rendering it a potential candidate for next generation broadband wireless communication systems, with the aim of providing extremely high data rate, bandwidth efficient signalling and improved quality of service (QoS).

Key -Words: - Adaptive Modulation, Turbo Coding, BFWA System, Diversity, Selection Diversity.

1 Introduction

Orthogonal Frequency Division Multiplexing (OFDM) [1] has become an important method of digital transmission in several wire line and wireless applications such as xDSL, digital audio broadcasting (DAB), digital terrestrial television and wireless LANs. The reasons for such wide use are: high spectral efficiency, large flexibility of matching to the transmission channel properties and relatively easy high-speed modem implementation as compared with sophisticated high bit-rate serial transmission supporting the same data rate in the same bandwidth.

Forward error correction (FEC) can be incorporated with OFDM to improve the bit error rate (BER) performance. In the IEEE 802.11a standard, convolutional coding and Viterbi decoding have shown to yield low BER in wireless data applications.

In order to further improve the performance of convolutional coding, several adaptive modulation schemes have been recently suggested [2].

Turbo codes [3] have also drawn much interest particularly since they provide greater BER performance and flexibility in controlling bandwidth utilisation and QoS.

In our simulation studies we have compared the OFDM BFWA system performance when the following coding schemes were applied: CC, AMCC, TC, AMTC, and SDTC.

2 System Model

An OFDM signal [4], as depicted in Fig. 1, consists of a sum of subcarriers, which can be generated by applying the inverse Fourier transform. Modulation can be based on either phase shift keying (PSK) or quadrature amplitude modulation (QAM).

The subcarriers spacing is a multiple of $1/\tau = 1/NT$ ($N =$ number of subcarriers and T is the symbol time), thus they will be linearly independent or orthogonal where their sidebands overlap, while still allowing reception without interference from adjacent carriers.

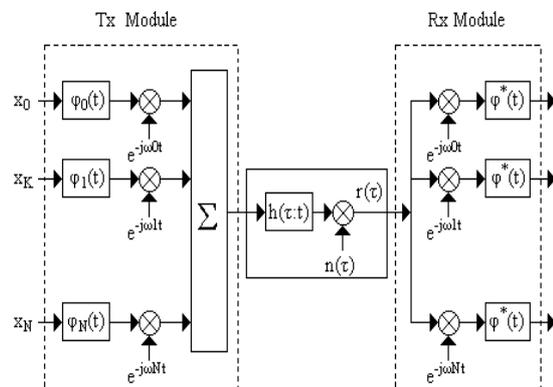


Fig. 1: Model of an OFDM system.

The OFDM signal's orthogonality can be easily maintained by applying the inverse fast Fourier transform (IFFT, 128 size for our simulations) to

the vector m_j to transform from frequency-domain to time-domain. The only obstacle to using the variant of Fourier transform for generating the OFDM symbol is the non-periodical nature of the time domain signal.

Therefore, inserting cyclic prefix (CP), Fig. 2, that comprises the last L length of an active OFDM symbol makes the symbol appear periodically at the receiver. The addition of a CP also creates a guard band around individual OFDM symbols to accommodate for the delay spread from the previous symbol which greatly reduce inter symbol interference (ISI) and inter channel interference (ICI) [1].

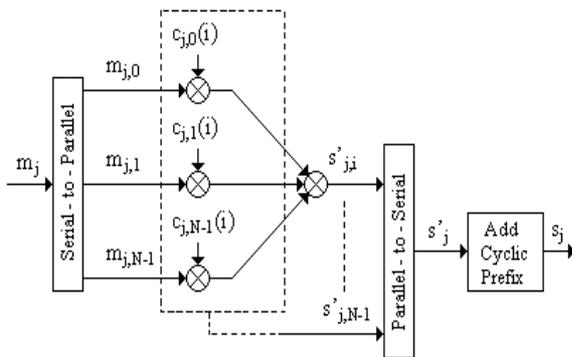


Fig. 2: OFDM modulator.

At the receiver, the reverse process is performed. The OFDM signal is sampled into discrete-time sequence and the CP is removed from the OFDM symbol, Fig. 3. Assuming that the CP completely eliminates ISI and ICI, the time-domain sequence is converted to frequency-domain by the OFDM demodulator (FFT) where the in-phase and quadrature component are generated.

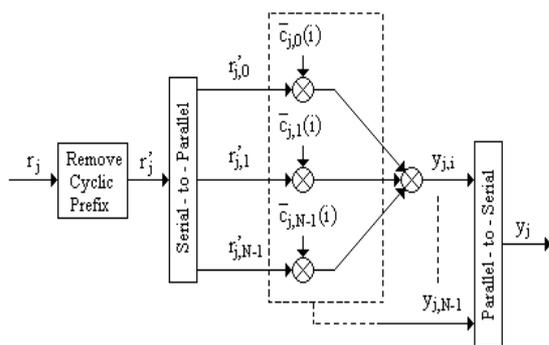


Fig. 3: OFDM demodulator.

3 Channel Model

In this study an indoor line-of-sight (LOS) model was used in the 17GHz band with a coherence time of 1 ms and a Doppler spectrum of 96 Hz (speed of 6 Km/h). Other channel main parameters based on experimental measurements are 20.72MHz for coherence bandwidth, 101ns maximum delay (T_{MAX}) and 17.5ns for the RMS delay spread. The

impulse response was generated assuming Wide-Sense Stationary and Uncorrelated Scattering (WSSUS). In addition to that, the channels path loss exponent is 1.68 and its K-factor equal to 9. Finally during our simulations, perfect timing and frequency synchronization were assumed.

4 Bit Loading Algorithms

In an OFDM system, it is possible to adjust the modulation level for each subcarrier individually. This technique, known as bit loading, is also referred to as *adaptive modulation*. In this approach several bits are allocated to subcarriers with a high signal-to-noise ratio (SNR), whereas on subcarriers with low SNR only few or no bits at all are transmitted. Loading algorithms, which mainly differ in their optimisation criteria and computational load perform bit allocation. In [10] we examined three well-known adaptive algorithms - such as Cioffi, Fischer and Piazzo - and saw that the Piazzo algorithm offered the best performance. An algorithm, illustrated in [2], based on knowledge of channel's frequency-domain attenuation vector. Consequently, all the bits are initially assigned to the highest modulation type (64QAM). Then a series of bit reallocations is performed, reducing the transmit power upon each reallocation, thus leading to an optimum bit allocation in each subcarrier. Thus, in the simulations where adaptivity is applied, this is done using the Piazzo algorithm.

5 Turbo Codes

Since its introduction in 1993 [3], turbo codes (TC) have attracted tremendous interest due to its remarkable performance. TC is also known as parallel-concatenated convolutional code (PCCC). The code [5] employs two identical recursive systematic convolutional (RSC) encoders connected in parallel, with an interleaver preceding the second RSC, as depicted in Fig. 4. Depending on the code rate desired, the parity bits from the two constituent encoders are punctured before transmission.

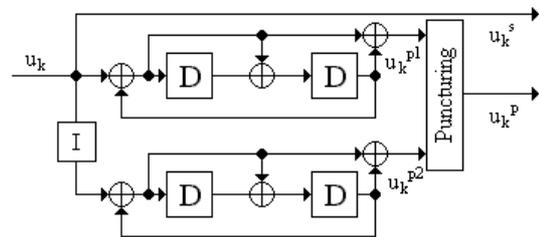


Fig. 4: Turbo Encoder with $g = (1, 5/7)_8$.

The TC data can be decoded by using either Log-MAP [6] or SOVA [7] algorithms. The two decoders are arranged serially in a simplified turbo decoder structure. The OFDM demodulator's outputs are de-multiplexed into three streams of M data, while the punctured bits are replaced with zeros, and are then fed to the constituent decoder.

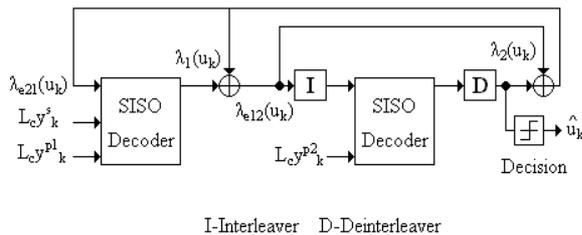


Fig. 5: Turbo decoder.

As can be seen from Fig. 5, the purpose of the first SISO decoder is to yield the extrinsic information, $\lambda_{e12}(u_k)$, which would be interleaved and applied to a second SISO decoder as the *a priori* information. In this simplified iterative decoder, the extrinsic information, $\lambda_{e12}(u_k)$, and parity information, L_{cy}^{p2} , are passed together to the second SISO decoder. The second SISO decoder computes the extrinsic information, $\lambda_{e21}(u_k)$, as the *a priori* information to the first SISO decoder, who will use it - along with systematic L_{cy}^s and parity information L_{cy}^{p1} - to produce a new $\lambda_{e12}(u_k)$. After the iterative process is completed, the decoder outputs the LLR (\bar{u}_k) values.

6 Selection Diversity

Selection diversity is one of the simplest diversity techniques. A block diagram of this method is similar to that shown in Fig. 6, where m demodulators are used to provide m diversity branches whose gains are adjusted to provide the same average SNR for each branch. The receiver branch having the highest instantaneous SNR is connected to the demodulator. The antenna signals themselves could be sampled and the best one sent to a single demomulator.

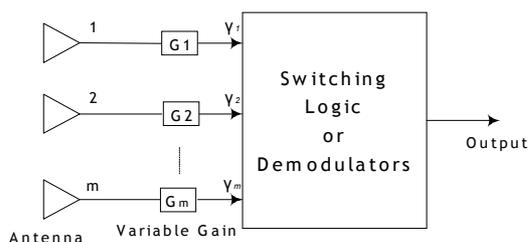


Fig. 6: Block diagram for selection diversity.

In practice, the branch with the largest $(S+N)/N$ is used, since it is difficult to measure SNR alone. A practical selection diversity system cannot function on a truly instantaneous basis, but must be designed so that the internal time constants of the selection circuitry are shorter than the reciprocal of the signal-fading rate.

Selection diversity offers an average improvement in the link margin without requiring additional transmitter power or sophisticated receiver circuitry. The diversity improvement can be directly related to the average bit error rate for

various modulations by using the equivalent principles (reference). It is also easy to implement because all that is needed is a side monitoring station and an antenna switch at the receiver.

7 Simulation Results

Simulation studies have been performed using the MATLAB simulation package for all the OFDM systems considered above. Below we will present selected BER vs. SNR curves for all the implemented scenarios.

Fig. 7 presents the BER vs. SNR curves using CC and AMCC, which is implemented using the Piazza algorithm. Clearly the system incorporating AMCC has better performance, than CC. This improvement is interpreted, in terms of dB, in a gain of about 2.1 and 3.6db for 13,34 and 53,34Mbps data rates.

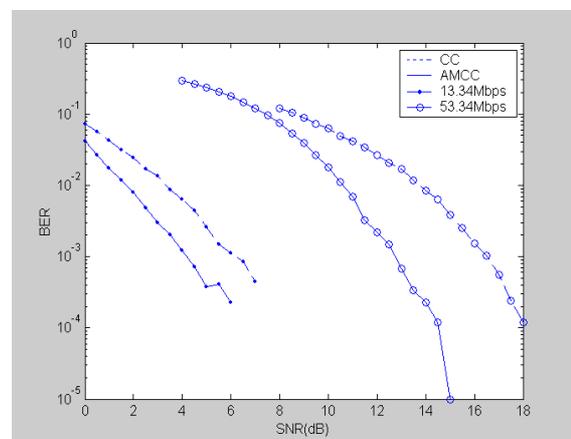


Fig. 7: BER vs. SNR for CC and AMCC.

The next step involved the implementation of a PCCC turbo coder [7], using a random interleaver. For the decoding process two decoding algorithms were used, Log-Map and SOVA, whose performance is depicted in Fig. 8. It can be observed that the Log-Map outperforms the SOVA one in all modulation schemes offering a step up of 0.7dB in average, which is in accordance with theory [9].

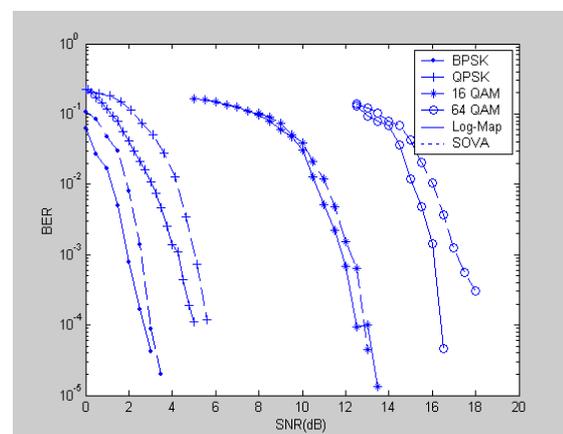


Fig. 8: BER vs. SNR for the TC system.

Fig. 9 illustrates the comparison in performance of a TC system with and without the SD technique being utilized. From this figure, we can observe that the combination of SD and TC offers an improvement in performance by almost 1.4dB and 1.15dB using the Log Map and SOVA decoding algorithms respectively.

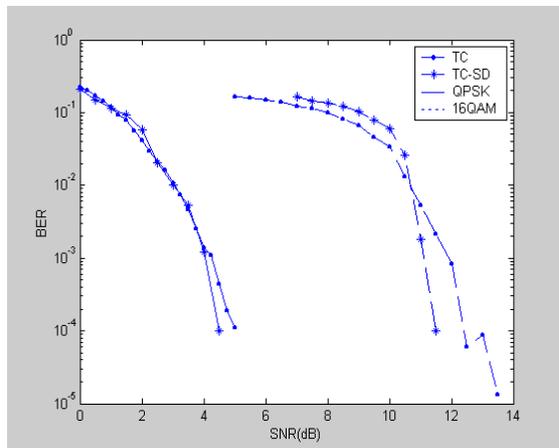


Fig. 9: BER vs. SRN for the TC system using Log Map decoding algorithm with and without applying SD for QPSK and 16QAM.

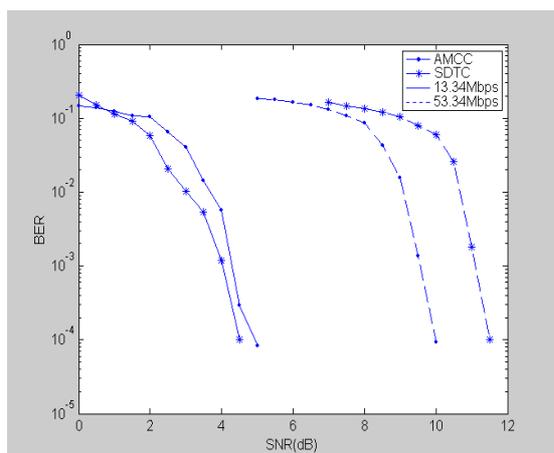


Fig. 10: BER vs. SNR using AMTC and SDTC at 13,34Mbps and 53,34Mbps.

Concluding with our simulations we compared the improvement that is offered in a TC system by two completely different techniques, such as AM and SD are. The results from this comparison are given in Fig. 10 from which we can see something very interesting. From 13,34Mbps data rate the SD technique offers an extra improvement, compared to AM, of about 0.4dB. For the case of 53,34Mbps the AM technique outperforms SD by nearly 1.3dB. This shows us that the AM algorithm performs better when it has to allocate a large number of bits.

8 Conclusion

The paper demonstrates that combining turbo coding with selection diversity provides sufficiently robust performance. The results are a representative

part of the simulation results achieved for the system with different PSK and QAM modulations and coding schemes.

Generally, from Fig. 7 one concludes that the bit loading scheme results in better performance compared to CC. The gain was as much as 3.6dB at the error rate of 10^{-3} .

Similar improvement, compared to TC, was offered by the AMTC and SDTC systems that utilized an adaptive and a diversity technique respectively.

Therefore, it can be asserted that TC and especially AMCC or SDTC have a strong potential for use in OFDM packet transmission. Further simulations will follow in order to examine other diversity techniques and channel scenarios.

Reference:

- [1] VAN NEE, R.: "OFDM for wireless multi-media comms", Artech House, 1999.
- [2] HANZO, L., WEBB, W., KELLER, T.: "Single and multi-carrier Quadrature Amplitude Modulation", John Wiley and Sons LTD, 2001.
- [3] BERROU, C., GLAVIEUX, A., and THITIMAJSHIMA, P.: "Near Shannon limit error - correcting coding and decoding: Turbo codes", in ICC'96, 1993, pp. 1064-1070.
- [4] SALIGA, S.V.: "An introduction to IEEE 802.11 wireless LAN's", Radio Frequency Integrated Circuits (RFIC) Symposium, 2000, Digest of Papers. 2000 IEEE, 2000.
- [5] BATTAIL, G.: "A conceptual framework for understanding turbo codes", IEEE Journal on selected areas in comms. Vol. 16, February 1998.
- [6] ROBERTSON, P., VILLEBRUN, E., and HOEHER, P.: "A comparison of MAP decoding algorithms operating in the log domain", IEEE Int. Conf. on comms, pp. 1009-1013, June 1995.
- [7] TAN, T., STUBER, L.G.: "Soft Output Viterbi Algorithm (SOVA) for Turbo Codes", ISIT 2000, Sorrento, Italy, June, 2000.
- [8] Kahl, L.: "Ratio Squarer" Proceedings of fire, Vol. 42, pp 1074, November 1954.
- [9] Jakes, W. C.: "A Comparison of Specific Space Diversity Techniques for Reduction of Fast Fading in UHF Mobile Radio Systems" IEEE TVC, Vol. VT-20, pp 81-93, November 1971.
- [10] E.I.Kalantzis, P.I.Dallas and B.S.Sharif: "Performance Evaluation of Adaptive Modulation and Turbo Coding on an OFDM BFWA System", 3rd International Symposium on Turbo Codes, 1-5 September, Brest, France, 2003

Drop Precedence Mapping on Dynamic Class Assignment Method: On Simplified One Link Model

Kenta Yasukawa[†], Ken-ichi Baba^{††}, and Katsunori Yamaoka^{†,†††}
[†]Tokyo Institute of Technology, ^{††}Osaka University

[†]knt@net.ss.titech.ac.jp, ^{††}baba@cmc.osaka-u.ac.jp, ^{†††}yamaoka@gsic.titech.ac.jp

1 Introduction

In this paper, we term multimedia streaming application traffic “stream flows”, and the other usual application traffic “non-stream flows”. In many cases, most non-stream flows are TCP flows, while most stream flows are UDP flows, because of their needs. Where stream and non-stream flows are aggregated on a shared link, there are many problems when the link is congested because different TCP and UDP behaviors cause negative interactions. We have assumed a situation where stream flows account for most of traffic in our research. Under this situation, the QoS of non-stream flows are especially degraded when congestion occurs because responsive TCP flows close their window size, while non-responsive UDP flows occupy the link bandwidth [1].

For solving these problems, we proposed a dynamic class assignment method in our previous work [5]. It is workable on the Differentiated Services (Diffserv) architecture which has been standardized to improve IP network QoS [2]. Our method is especially designed to be suitable to AF (Assured Forwarding) class [3] which has four subclasses and three levels of drop precedence. The behavior of our method is as follow. Non-stream flows are classified into four classes, and statically assigned to four AF subclasses. Stream flows are dynamically assigned to any class, taking conditions and interactions of non-stream flow classes into consideration. We term this behavior “dynamic class assignment”. The two purposes of our method, i.e., classification and dynamic class assignment, can be summarized as follows.

- To improve the QoS of stream and non-stream flows where these are evaluated in a comprehensive manner.
- To adapt to changing traffic without reconfiguring bandwidth allocation ratio to the AF subclasses.

The system also does admission control behavior, i.e., if no class has sufficient bandwidth for arriving flows, it assigns no class to them.

In our previous works, we classified non-stream flows into four classes (Table 1) and investigated their characteristics (Table 2) in [4]. Also in [5], we proposed an algorithm for our dynamic class assignment method and proved its effectiveness through simulation experiments.

However, only using classification and dynamic class assignment, our method could not handle a problem of unpredicted increase of non-stream flow traffic which will be mentioned in the following section. We therefore integrate a new function into our method using the drop precedences of the AF PHB. In this paper, we will describe the new function and discuss the appropriate way for implementing it.

Table 1: Classification of Non-stream Flows

	High Arrival Rate Short Duration	Low Arrival Rate Long Duration
Greedy for Bandwidth	HTTP, SMTP, POP3 –Class a	FTP data, NNTP, (P2P) –Class b
Non-greedy for Bandwidth	DNS, ICMP –Class c	FTP control, TELNET –Class d

Table 2: Characteristics of Non-stream Flow Classes

	Class a	Class b	Class c	Class d
Friendliness with Stream Flows	Low	Low	High	High
Tolerance against Stream Flows	Low	High	-	Low

2 Unpredicted Increase of Non-stream Flow Traffic

Let us imagine a situation where the amount of traffic of non-stream flows in a class increases after the system assigned the all remaining bandwidth of the class to stream flows. In this situation, non-stream flows in the class are prevented from gaining sufficient bandwidth after the amount of traffic of themselves increased. This situation is a case of class assignment failure, and the system should avoid it by predicting the peak amount of traffic based on the historical information when making decisions of assigning classes. Because the policy of our method is assigning stream flows without degrading the QoS of non-stream flows, the system has to recover such a failure if it happens.

We therefore propose to use the drop precedences of the AF class for saving non-stream flows from bandwidth suffering when the non-stream flows in the class increase more than predicted. On assigning a class, we map stream flows to a higher drop precedence and non-stream flows to a lower one.

This mapping will cause more packet loss in stream flows and degrade their QoS, however, it can be a notification of congestion for the stream senders. If they adaptively control their bit rate when they notice the congestion, it can be avoided. Where stream flows account for most of the traffic as we have assumed in this paper, their bit rate adaptation is the most effective way of avoiding congestion.

Senders of stream flows using RTP over UDP can be informed of packet losses by receiving RTCP packets, so they can potentially be adapted to congestion through such information. Moreover, scalable video coding has attracted a great deal of interest for video streaming over the Internet. In scalable video coding, media data is encoded into a Base Layer (BL) and one or several Enhancement Layers (EL), and its total bit rate is controllable by changing the number of ELs to send, without

losing the base quality. Bit rate adaptation will be thus achieved through these techniques. ISO/IEC MPEG-4 fine grain scalability (FGS) video was recently proposed as a standard coding tool for video streaming applications [6], so bit rate adaptation will become more popular in the near future.

This is the reason we propose this drop precedence mapping even though it causes more packet drops in stream flows when congestion occurs. We will discuss the way to achieve the drop precedence in the following section.

3 Implementation of Drop Precedence

We first discuss the drop precedences further. Three drop precedences are independently available on each subclass in the AF class. These can be described as $DP0$, $DP1$, and $DP2$, and their relationship is $DP0 < DP1 < DP2$, in this paper. Here, we consider to map non-stream flows to $DP0$, and stream flows to $DP1$.

To implement the three drop precedences in the AF class, multi-level RED (MRED) schemes are generally utilized (Fig. 2). The MRED schemes are extended RED (Random Early Detection) [7] schemes that have multi priority levels. RED drops incoming packets with a certain probability depending on the average queue size q_k . q_k is calculated recursively based on the previous average q_{k-1} and the instantaneous queue size q (moving average):

$$q_k = q_w q + (1 - q_w) q_{k-1},$$

where q_w is the averaging weight for calculating q_k . It is set to 0.002 in most cases because this is the recommended value in [7]. The drop probability increases from 0 to maximum drop probability (max_p) as q_k increases from minimum threshold (min_{th}) to maximum threshold (max_{th}). If q_k goes beyond max_{th} , all incoming packets are dropped. The MRED generally has max_p , min_{th} , and max_{th} independent of each drop precedence, and there are a couple of MRED schemes depending on the way q_k is handled. Goyal et al. classified these [8].

3.1 Difficulties and Problems in MRED

It is generally difficult to determine effective RED or MRED parameters, and much research has been done for this [9, 10]. However, RED or RED-like schemes are originally proposed to avoid global synchronization problems of responsive TCP flows, but for non-responsive UDP flows. In other words, the characteristics of RED, such as probabilistic dropping or the average queue size calculations based on the moving average, are not needed for UDP flows that we will map to a higher drop precedence. Moreover, we obtained some results that indicated the average queue size calculation using the moving average causes needless packet drops in flows mapped to $DP1$ to provide sufficient protection with flows mapped to $DP0$. Bonald et al. also mentioned issues of the same kind in [10]. We here discuss why such needless packet drops happen in MRED schemes.

We show an example of queue size fluctuation in a MRED scheme in Fig. 1. Fig. 1 shows two reasons of needless packet drops in the following.

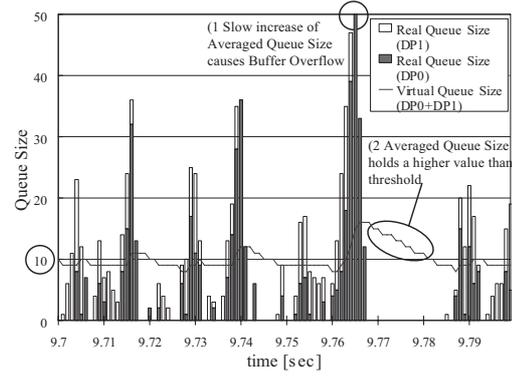


Figure 1: Example of Queue Size Fluctuation

- (1) Even when the queue size increases, the moving average calculation makes the drop decision be delayed and this delay may cause buffer overflow. It means that we must set the threshold of $DP1$ lower taking this delay into consideration, to provide a sufficient protection with flows mapped to $DP0$.
- (2) Even when the queue size decreases, the moving averaged queue size holds the higher value than the actual for a while. If it exceeds the threshold, it may cause needless packet drops.

These two reasons are both based on the fact that the moving averaged queue size varies slower than the actual queue size.

Considering these issues, we introduce a more simpler scheme for implementing the drop precedence and call this the Multi-level Drop Tail (MDT) scheme (Fig. 3). MDT has multi thresholds, th_0 , th_1 , and th_2 for each drop precedence, but does not use probabilistic dropping and average queue size calculations. It observes the queue size of each drop precedence and drops packets that are marked DPi if the queue size of DPi exceeds th_i (Fig. 3). In the queue size of each drop precedence, MDT does RIO-like coupled queue size calculations [11]:

$$\begin{aligned} q_{DP0} &= p_{DP0} \\ q_{DP1} &= p_{DP0} + p_{DP1} \\ q_{DP2} &= p_{DP0} + p_{DP1} + p_{DP2}, \end{aligned}$$

where q_{DPi} is the queue size of DPi , and p_{DPi} is the number of packets in the queue that are marked DPi . Makkar et al. verified that this calculation archives effective protection for lower precedence packets by comparing RIO-C (RED with In/Out profile and Coupled average queues) with WRED (Weighted RED) [11].

4 Comparison of MRED and MDT

It is assumed that MDT can provide sufficient protection with flows mapped to $DP0$ by fewer packet dropping in flows mapped to $DP1$ because of the discussion in the previous section. Hence, we will now compare MDT with MRED (RIO-C) through NS-2 simulations to confirm this assumption [13]. We will also provide some guidelines for parameter setting from the obtained results.

The simulation scenario is as follows.

1. Generate non-stream flows according to Table 3.

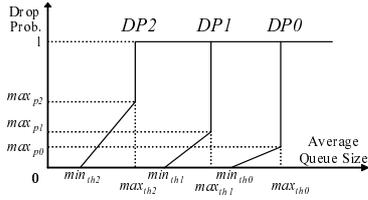


Figure 2: Drop Prob. in MRED

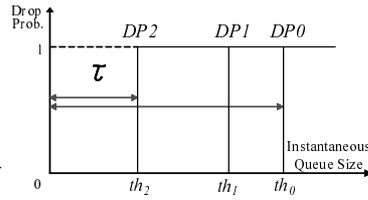


Figure 3: Drop Prob. in MDT

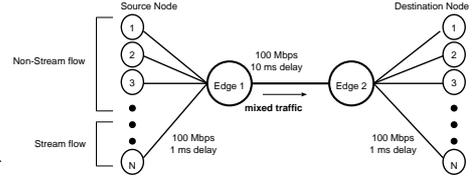


Figure 4: One Link Model

2. Add stream flows one after the other until the total amount of traffic reaches 100% of the link bandwidth.
3. Make the arrival rate of non-stream flows increase until its peak.

The parameters in Table 3 are based on statistical information by Ata et al. [12]. The other details in the simulation are explained in the following. Stream flows were UDP CBR (Constant Bit Rate) flows, with an average rate of 500 Kbps, and a packet size of 1 KB. The network topology was the simplest one, “one link model” in Fig. 4. The bottle-necked link bandwidth was 100 Mbps. The total queue size of nodes was normally set to 50 (the default value in NS-2).

The degree of drop precedence in MRED or MDT is almost determined by the ratio of the threshold (max_{th_i} or th_i) to the total queue size. We term this ratio τ in the following. If $\tau = 1.0$, drop precedence is lowest, and the lower the τ , the higher the drop precedence.

We mapped non-stream flows to $DP0$, and stream flows to $DP1$ and set the τ of $DP0$ to 1.0, and observed variations in the QoS of stream and non-stream flows, changing the τ of $DP1$ from 0.1 to 1.0. Here, we will introduce two different QoS measures in this paper that we used for the TCP and UDP flows. The QoS measure for the TCP flows is the “duration of service”, which is the time needed to finish transferring data received from application layers, and the QoS measure for the UDP flows is the packet loss probability.

As space is limited, we show the results for the case of *class a* in the following. Our drop precedence mapping is most important in *class a* because non-stream flows in this class has a low degree of the tolerance against stream flows [4] and its amount of traffic widely fluctuates in general cases.

Fig. 5 shows the results where the total amount of traffic was set to three levels, 100%, 110%, and 120%. The x-axis is the packet loss probability of stream flows and the y-axis shows the duration of service of non-stream flows in *class a*. We can say from these results that MDT can maintain the QoS of non-stream flows by dropping fewer packets than MRED in any levels of traffic.

Fig. 6 shows the results where the total queue size of the nodes was set to three levels, 20, 50, and 100, but the

Table 3: Non-stream Flow Setup

	Transport Layer Protocol	Average Data Size [bytes]	Average Packet Size [bytes]
<i>class a</i>	TCP	5,000	500
<i>class b</i>	TCP	150,000	1,000
<i>class c</i>	UDP	200	100
<i>class d</i>	TCP	60	60

amount of traffic was fixed to the 120% of the link bandwidth. The x-axis and the y-axis are the same as Fig. 5. This figure shows that the smaller the total queue size was, the wider the difference between MDT and MRED became. In other words, MDT can effectively work even where the total queue size is small, while MRED cannot. This is an important benefit of MDT because the queue size in nodes tends to be set smaller based on the demand of suspending the queuing delay.

Let us now show the results where the ratio of stream and non-stream flows in the traffic was changed as shown in Table 4, but the amount of traffic and the total queue size were fixed to the 120% of the link bandwidth and 50, respectively. Fig. 7 shows the packet loss probability of stream flows needed to suspect the duration of service of non-stream flows below a level (Fig. 7 shows the case of below 0.6 sec., for example). It clearly shows that MDT can maintain the QoS of non-stream flows by dropping fewer packets than MRED in any traffic conditions.

Let us now discuss the influence of τ to clarify a guideline for configuring it. Fig. 8 shows the duration of service of non-stream flows under conditions in Table 4. The x-axis is the τ of $DP1$ and the y-axis is the average duration of service of non-stream flows in this figure.

Fig. 8 shows that the influence of changing τ is smaller and less sensitive in MDT, compared to MRED. In MRED, it is needed to set τ to less than 0.5 to protect the non-stream flows but the influence of τ is still sensitive in this range. This is one of the reason of difficulties in configuring the parameters in MRED. We can say from these results that MDT is easier to configure the threshold than MRED.

Fig. 8 shows that non-stream flows were sufficiently protected where the τ of $DP1$ is set to a little less than 1.0, i.e., $0.7 < \tau < 0.9$, and the effects can be obtained in any conditions, in MDT. We note this as the guideline in configuring τ in MDT. Of course, there is

Table 4: Experimental Conditions

	Non-stream flow Peak Arrival Rate [flows/sec]	Number of Stream Flows
Condition 1	1800	60
Condition 2	1600	80
Condition 3	1400	100
Condition 4	1200	120
Condition 5	1000	140
Condition 6	800	160
Condition 7	600	180
Condition 8	400	200

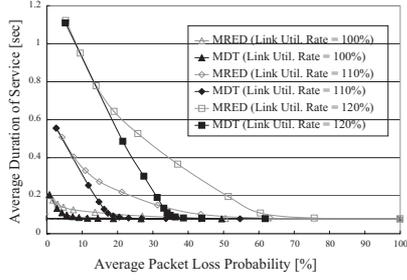


Figure 5: Changing the Total Amount of Traffic

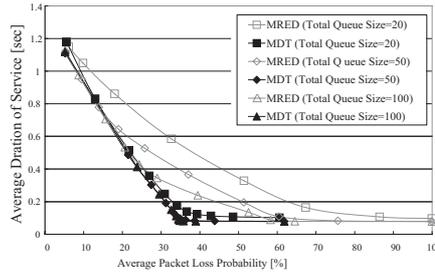


Figure 6: Changing the Total Queue Size

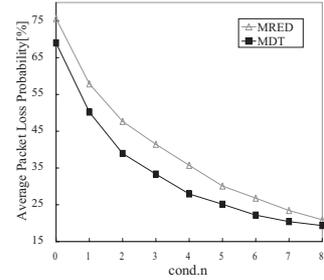


Figure 7: Changing the Traffic Condition (Table 4)

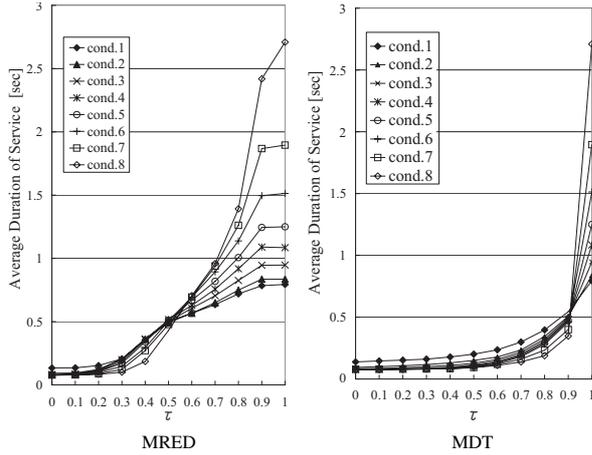


Figure 8: Parameter Sensitivity of τ

a tradeoff in configuring it. Therefore, it should be configured depending on the policy of the domain indeed. We intend to prove the validity of this guideline through experiments in more varied situations in other work.

We also obtained good effects for non-stream flows in *classes b* and *d*, with this guideline, however, in *class b*, this setting made non-stream flows gain bandwidth unfairly, while stream flows were restricted. This problem was caused by the long duration characteristics of *class b*. Flows in this class are generally in the congestion avoidance phase of the TCP, probing for bandwidth, because of the long duration characteristics. Therefore, the more stream flow packets are dropped, the wider bandwidth they will gain. The same problem seldom happens in *classes a* and *d*, because flows in these classes finish their transmission in the slow-start phase as we mentioned in Sect. 2. Considering these things, setting the τ of *DP1* in *class b* should be done carefully.

5 Conclusion

We addressed a problem in Dynamic Class Assignment Method that was proposed in [5] and integrated a new function into our method, using the drop precedences. In this paper, we also addressed difficulties and problems of general MRED schemes for applying it to our method, and introduce MDT that is simpler and more suitable for our purpose. We proved following things through simulation experiments.

- MDT can maintain the QoS of flows mapped to *DP0* by dropping fewer *DP1* packets than MRED, even though the amount or condition of

traffic is changed.

- MDT can effectively work where the total queue size is set to small, while MRED cannot.
- MDT is easier to configure its parameter than MRED.

To integrate the above mentioned drop precedence mapping with MDT, our method can handle the unpredicted increase of non-stream flows. We intend to prove the validity of the guideline which we mentioned in future work.

Acknowledgements

Some parts of this research were supported by International Communications Foundation ('ICF') in Japan.

References

- [1] N. Seddigh, B. Nandy, and P. Piedad, "Bandwidth Assurance Issues for TCP Flows in a Differentiated Services Network," *Proc. of GLOBECOM '99*, pp.1792–1798, Dec. 1999.
- [2] S. Blake, D.Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, "An Architecture for Differentiated Services," RFC2475, IETF, Dec. 1998.
- [3] J. Heinanen, F. Baker, W. Weiss and J. Wroclawski, "Assured Forwarding PHB Group," RFC2597, IETF, Jun. 1999.
- [4] K. Yasukawa, K. Baba and K. Yamaoka, "Classification of Non-stream Flows to Reduce Negative Interactions between Stream and Non-stream Flows," *Proc. of PACRIM 2003*, pp. 772–775 Aug. 2003.
- [5] K. Yasukawa, K. Baba and K. Yamaoka, "Class Assignment Management for Stream Flows Considering Negative Interactions with Non-stream Flows," to appear in *Proc. of Networks 2004*, Jun. 2004.
- [6] W. Li, "An overview of fine granular scalability in mpeg-4 video standard," *IEEE Trans. on Circuits and Systems in Video Technology*, Vol. 11, No. 3, pp. 385–398, Mar. 2001.
- [7] S. Floyd, and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Trans. on Networking*, Vol. 1, No. 4, pp. 397–413, Aug. 1993.
- [8] M. Goyal, A. Durresi, R. Jain, C. Liu, and P. Misra, "Effect of number of Drop Precedences in Assured Forwarding," *Proc. of GLOBECOM 99*, pp. 188–193, Dec. 1999.
- [9] C. V. Hollot, V. Misra, D. Towsley and W. Gong, "A Control Theoretic Analysis of RED," *Proc. of INFOCOM 2001*, pp. 1510-1519, 2001.
- [10] T. Bonald, M. May and J. Bolot, "Analytic evaluation of RED Performance," *Proc. of INFOCOM 2000*, pp. 1415-1424, Mar. 2000.
- [11] R. Makkar, I. Lambadaris, J. Hadi-Salim, N. Seddigh, B Nandy and J. Babiarz, "Empirical Study of Buffer Management Schemes for Diffserv Assured Forwarding PHB," *Proc. of ICCN*, pp. 632–637, Oct. 2000.
- [12] S. Ata, M. Murata, and H. Miyahara, "Analysis of Network Traffic and Its Application to Design of High Speed Routers," *IEICE Trans. on Information and Systems*, Vol E83-D, No5, pp.988–995, May 2000.
- [13] UCB/LBNL/VINT, "Network Simulator NS," <http://www-mash.cs.berkeley.edu/ns/>

Provision of Virtual Line-switched Service over DWDM with PFTS

Dengyuan Xu, Huaxin Zeng, Haiying Wang
Sichuan Network Communication Key Laboratory,
Southwest Jiaotong University,
Chengdu, Sichuan, China, 610031

xudave@tom.com, huaxinzeng1@vahoo.com.cn

Abstract --This paper is devoted to switching techniques over DWDM aiming for future Integrated Networks for cable TV, telephone switching and computer network services.

In this paper, a novel technique called PFTS (Physical Frame Time-slot Switching) is discussed and its technical and application aspects are analyzed. The format of the Ethernet MAC frame is borrowed in defining the physical frame for PFTS and the transmission time for the maximum size of the MAC frame is defined as a PFT (Physical Frame Time-slot). Consequently, user data can be fed into PFTs and switched in a single layer sub-network in an asynchronous mode.

With out-band signaling concept, a PFTS sub-network in User plane (U-plane) together with existing Internet protocol stack in Signaling and Management planes (S & M planes) create a new architecture for Next Generation Internet, called the Single User Plane Architecture (SUPA).

The SUPA network with a Physical Frame Time-slot switching layer over DWDM can provide a virtual line switched service for time-critical traffic such as in cable TV and telephone switching service with guaranteed QoS.

Keywords--PFTS (Physical Frame Time-slot Switching), 3D-EMAN (Three Dimension Ethernet MAN), EPF (Ethernet-like Physical Frame), SUPA (Single User Plane Architecture)

1. Introduction

The advent of Dense Wavelength Division Multiplexing (DWDM) raises the raw bit rate in a single fiber from a few Gigabits per second (Gbps) to tens of Terabits per second (Tbps), and enables the total bit rate in a single fiber reaching the order of magnitude of Tera bits. This achievement opens a new era for high-speed networks, and potentially, enables existing telephone, cable TV, and computer networks to be merged into a unified network.

However, in a merged network, service requirement for time critical traffic cannot be satisfied

through statistically sharing a high-speed channel as is widely adopted today in Internet. With DWDM techniques, even the transmission capacity of a single lambda is much too big a channel for individual user-data transfer (10/20/40 Gbps at present, and 80 Gbps soon), therefore, a lambda channel needs to be further decomposed into sub-channels. Moreover, Quality of Service (QoS) for time-critical data flows requires a sub-channel with guaranteed throughput and other time-relevant parameters. In other words, a virtual line switched service is most desirable.

This paper promotes a novel technique called the "Physical Frame Time-slot Switching (PFTS)" and a framework of PFTS is also presented. The format of the Ethernet MAC frame is borrowed in defining the physical frame for PFTS and the transmission time for the maximum size of the MAC frame is defined as a PFT (Physical Frame Time-slot). Consequently, user data can be fed into PFTs and switched in a single layer sub-network in an asynchronous mode.

This paper is organized in 4 parts. Section 2 provides with a comprehensive study of present switching techniques over DWDM, and their shortcomings are analyzed in some detail. Section 3 introduces a novel switching technique – PFTS and its technical details are discussed. Finally in section 4 PFTS technique applied to 3D-EMAN network create a new architecture called SUPA, which can provide a virtual line switched service for time-critical traffic such as in cable TV and telephone switching service with guaranteed QoS.

2. Existing Switching Techniques over DWDM

2.1 Burst Switching [1]

Burst Switching or Optical Burst Switching (OBS) does focus on relay of individual data, called "bursts" – a representation of data with variable lengths from a few kilobytes up to few megabytes, hence a burst switching network handles user data in terms of big data-block.

People are in favor of pure optic approach in realization of burst switching in order to avoid the electric/optic conversion. Consequently, they have to try to push most processing and buffering functions from pure optic core-nodes to the photoelectric edge-routers in a burst switching network because of

This research is sponsored in part by NSFC (National Science Foundation of China) under contract number 60372065

difficulty in optic buffering and lack of optic processing capability.

Burst switching adopts out-band signaling concept, where user data and control data are transmitted in different lambdas. This implies that burst switching has to rely on controllable lambda switching techniques within a node, otherwise it degenerated to pure fixed lambda switching.

Ilija Baldine, et al., gave up pure optic control approach in core nodes, and suggested a signaling approach called JET (Just Enough Time) [1], with which signaling “bursts” are done in an electric/optic combined manner.

The shortcomings of burst switching can be summarized as:

A. Lack of mechanism for guaranteed QoS –

There is no mentioning of mechanism for ensuring user required QoS, especially for time-critical traffic where guaranteed throughput, transit delay, and other time parameters.

B. Complexity in parameter selection and adjustment– Switching of optic lambdas has to be done through electronically controlled optic lambda switches. Because of the difference in speed between optic transmission and electronically controlled lambda switching (including head processing), there is a need for user data buffering. However, for lack of buffering and processing capability in optic domain, different transmission strategies have to be suggested:

- **TAG (Tell And Go)** – a data burst is sent on the data lambda right after the control packet is sent out. This demands buffering of the burst in the next switch, which has to rely on FDL (Fiber Delay Line), otherwise transmission contention and data loss may arise. Reference [3] provides a figure which says a 10^4 bits optic “buffer” needs a delay line of one kilometer fiber. Alternatively a fiber delay loop longer than the length holding a maximum burst is employed, where controlled times of circulations in the loop can multiply the delay time [4]. The difficulties with this approach lies in calculation and control of times of circulation.
- **JET (Just Enough Time)** - Instead of sending data burst immediately in the ingress router, JET postpones data sending time called the “offset time” which is just enough for the establishment of lambda path, so as to avoid buffering in intermediate optic switches and possible data losses. With this approach, it would be difficult to determine the required offset time.
- **Determining Length of Burst** – Burst length and offset time (or buffering time) have a complex interrelationship. The longer the maximum length of the burst is, the higher the channel utility can

be obtained; and consequently a longer offset time is needed. On the other hand, waiting for collection up to a maximum burst might result in longer waiting time at the ingress router. All these have made the tradeoff in determining parameters difficult.

In view that the optic computing is still in its infancy, there is a long way to go before an optic burst switching node capable of handling hundreds of wavelengths, with acceptable price on the market.

2.2 Optic Packet Switching [2]

Optic Packet Switching differs from burst switching chiefly by use a “packet” measured by a fixed duration. One influential OPS project - KEOPS (KEys to Optic Packet Switching) employs a $1.646\mu\text{s}$ time-slot including a 14-byte header and $1.35\ \mu$ for payload

Generally speaking, there is not much difference between OBS and OPS in KEPOS, except that a burst length is variable while the length of an optic packet is fixed. Variable length might incur inefficiency in channel utility caused by fixed path establishment time, while the inefficiency in OPS occurs when user data is shorter than the fixed optic packet length.

OPS inherits some of the shortcoming of OBS as discussed in section 2.1.

3. Physical Frame Time-slot Switching (PFTS)

This section is focused on a novel technique called Physical Frame Time-slot Switching (PFTS), where the format of the Ethernet MAC frame is adopted as the Physical Frame (called Ethernet-like Physical Frame, EPF) and the transmission time for EPF is defined as the basic time-slot. PFTS works in an asynchronous mode without the need for global clock synchronization and can provide a virtual line-switched service for time-critical data flows. For lack of granular multiplexing mechanism within a lambda of DWDM, the DWDM architecture should be augmented to support multi-class of QoS requirement.

3.1 Augmentation to existing DWDM layer

Existing DWDM devices can be conceptually viewed as having a two sub-layer structure: the MDTRS (Medium Dependent Transmission and Receiving Sub-layer) at the bottom and the LMDS (Lambda Multiplexing/De-multiplexing Sub-layer) on the top. The LMDS does multiplexing and de-multiplexing functions on the basis of lambdas, while MDTRS does transmission and receiving functions for multiplexed lambdas to or from a fiber.

Lambda switching cannot satisfy the general requirement in a multi-user network environment. To cater for this requirement, the architecture can be enhanced with a Physical Frame Sub-layer on top of LMDS.

In view that the Ethernet is the most popular technique in LAN, MAN, and soon in WAN as well, the format of the Ethernet MAC frame is deliberately

chosen for the physical frame, and consequently, the new sub-layer is referred to as EPFS (Ethernet-like Physical Frame Sub-layer). Figure 1 depicts an augmented architecture for DWDM physical layer

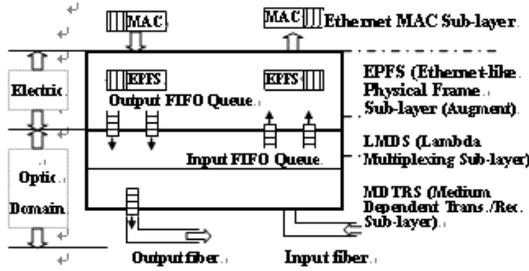


Figure 1: An Augmented DWDM Architecture

People in favor of pure optic switching have argued that data handling in electronic domain is much slower than that in optic domain, and E/O conversion also takes time; therefore, E/O combined approach towards physical frame switching introduces excessive time delay and might be a source of jitters. The opposite argument would be: such a delay indeed will postpone the time of the arrival of the first frame, but it will not change the time-relationship between frames within a data stream, nor it will generate jitters as long as such delay and conversion time remain comparatively constant. Indeed, such argument can be backed up by the success in 10 GE products[5] and 40Gbps Chip [6]. More importantly, it would be much more comfortable to deal with buffering and necessary processing electronically than to do it optically considering the state of art in two domains.

With respect to E/O conversion, an experiment with super-conducting techniques has demonstrated that a conversion rate of 100 Gbps is reachable [7]. To save the processing time for frame header, parallel techniques, optimal look-up algorithm for routing, and hardware based processing should be applied.

Figure 1 also illustrates the relationship between the EPFS and its upper Ethernet MAC Sub-layer. Mapping between MAC frame and the EPF is straightforward since they have the same format except for a minor modification to the destination address field as explained in the next section.

3.2 Operation of Physical Frame Time-slot Switching

The addition of EPFS on top of existing DWDM enables users to mingle their data into physical bit streams in the form of formatted frames. Here, a novel concept called the “Physical Frame Time-slot” (PFT) plays important roles in multiplexing within a lambda, in relaying frames within a DWDM node, and in synchronization between DWDM nodes.

A PFT is defined as the transmission time for an EPF with the maximum MAC frame length (1542 bytes) defined by IEEE 802.3 (2002 version), and used as the basis for frame-based asynchronous operation. An ingress edge router of PFTS domain

will simply to hand received MAC frames down to the physical layer, which will, in turn, treat them as EPFs and multiplex them into the bit stream of different lambdas towards PFTS nodes according to routing decision. DWDM switching nodes will then forward PFTs on the DWDM physical layer network along the predetermined path to egress edge router and synchronize their operation by detecting the frame preambles. This enables an asynchronous operation based on PFTs without a need for clock synchronization among switching nodes.

The format of an EPF resembles an Ethernet MAC frame except that the destination address field in the MAC frame is used in the EPF as the switching field, which consists of a PN (Port Number), an LN (Lambda Number), and a VC (Virtual Circuit) Id. Figures 2 and 3 illustrate the switching model for PFTS nodes and the format of switching field in an EPF frame respectively.

Figure 2 is essentially two sets of sub-layer stratum in figure 1 interconnected by a Physical Frame Time-slot (PFT) relay entity, which switches individual PFTs according to the fields of the Physical Frame shown in figure 3.

As shown in figure 3, the PN (Port Number) is used in DWDM nodes to identify a physical port (corresponding to a fiber), while the LN (Lambda Number) uniquely identifies a single lambda within a given fiber. VC Identifier is designed to distinguish different user data streams. A VC is also associated with certain transmission capability decided by the number of PFT(s) per second assigned to that VC. Finally, the Router Id is reserved for shared media topologies, such as a ring, where a simple number is used to replace complicated MAC address.

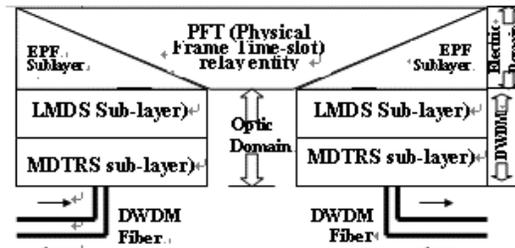


Figure 2: PFT Switching Node Model

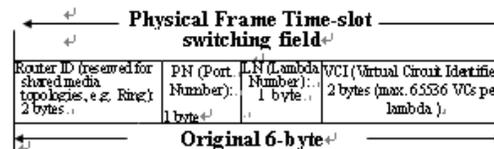


Figure 3: Reassignment of the Destination Address

3.3 Advantages of PFTS

PFTS differs from the existing switching techniques such as OBS, OPS in many ways, and its advantages can be summarized as:

A. High channel utility

First of all, its theoretical channel utility can approach 100% as long as all the time-slots are used in carrying user traffic, since PFTS is time-slotted. Furthermore, the theoretical effective data rate can be as high as 98.04% with the maximum extended frame length (1530 bytes) [IEEE 802.3 2002] against the maximum payload length (1500 bytes), which are much superior over other switching techniques discussed in section 2.

B. Capable of provision of virtual line-switched service with guaranteed QoS

PFTS service is connection-oriented and most suitable continuous traffic flows, which are dominated traffic in B/S application model in Internet today and will form an even greater percentage in the traffic of future merged networks (cable TV streams, switched telephone traffic, and other audio/ video streams). Communication path or virtual physical layer connection between an ingress router and egress router needs to be set up once for one communication session, which simplifies the routing process into "IP routing once, and fast switching for whole session". Compared with manner of "setting up a path for each burst or packet" in OBS and OPS, it is much more efficient.

QoS provided by PFTS can be guaranteed when enough time-slots are assigned to a virtual physical layer connection with fast transmission in the fiber and fixed processing time in each node. Techniques such as RSVP and CR-LDP (Constraint-based Routing Label Distribution Protocol) developed for Internet or MPLS, can all be utilized.

C. Merging the data link and physical layer into a single layer

The PFTS enables a user plane to be reduced to a single physical layer, in which some of the data link layer functions such as framing, error checking, and etc., can be embedded. As the Ethernet has been widely adopted not only in end-systems but also in intermediate systems (sub-networks), MAC frame used as physical layer frame in PFTS can save lots of troubles of segmentation and reassembly between various data link frames. It has been envisaged that the data link layer and physical layer will eventually be merge into a single layer, not only in the user plane but also in signaling and management planes.

4. Single-layer User Plane Architecture (SUPA) for Next Generation Internet

Introduction of PFTS technique is part of the research result in the 3D-EMAN project carried out at Sichuan Network Communication Key Laboratory. The 3D-EMAN architecture [8] utilizes current Internet protocol stack in C & M planes and the PFTS on top of DWDM in U-plane. Operation between systems supporting 3D-EMAN architecture will operate in 3D-EMAN mode, where Internet protocol

stack plus a moderate enhancement will play the C & M functions and transport user data on a PFTS network. For SVC service, a path between ingress and egress edge routers can be established by use of Internet routing protocols and QoS negotiation is done through 3D-MAN specific protocol. PVC service, such as cable TV broadcasting service, can be subscribed in advance. Detailed discussion concerning 3D-EMAN can be found in [12].

PFTS applied to 3D-EMAN can create a new Architecture called Single-layer User Plane Architecture (SUPA). The SUPA network with a PFTS sub-layer over DWDM can provide a virtual line switched service for time-critical traffic such as in cable TV and telephone switching service with guaranteed QoS.

Through the experience gained with 3D-EMAN at Sichuan Network Communication Key Laboratory, the authors are convinced that SUPA with the PFTS for next generation Internet is feasible.

REFERENCES

- [1] C.Qiao and M.Yoo, "Optical burst switching(OBS)-a new paradigm for an Optical Internet," *Journal of High Speed Network*, **vol. 8**, pp.69-84 (1999).
- [2] C. Guillemot, et al. "Transparent Optical Packet Switching: The European ACTS KEOPS Project Approach," *Journal of lightwave technology*, **vol. 16**, pp. 2117-2134, Dec (1998).
- [3] H. F. Jordan, *Pipelined Digital Optical Computing*. Technical Report OCS Technical Report 89-34, Optoelectronic Computing Center, University of Colorado, Boulder, 1989
- [4] Mustafa Cengizhan Cardakli, PhD dissertation: *Experimental Demonstrations of Optical techniques to Implement Networking Functions in Dynamically Reconfigurable Wavelength-Division-Multiplexed Fiber-Optic Communication Networks*, University of Southern California, USA, May 2001
- [5] IEEE 802.3ae - Local and metropolitan area networks—Specific requirements, Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications, Amendment: Media Access Control (MAC) Parameters, Physical Layers, and Management Parameters for 10 Gb/s Operation, June 13, 2002.
- [6] Faraydon Karim et al., "On-Chip Communication Architecture for OC-768 Network Processors," *Proceedings of the 38th Conference on Design Automation Conference*, PP.678-683, Oct, 2001
- [7] Roman Sobolewski, "Ultrafast optoelectronic interface for digital super conducting electronics," *Superconductor Science and technology*, vol. 14, pp.994-1000, Nov (2001).
- [8] Huaxin Zeng, et al., "On Three-Dimension Ethernet MAN (3D-EMAN) architecture," in: *Proceedings of PACAT03*, IEEE press, PP. 535-540, Aug (2003).

Sigmoid Function Based Dynamic Threshold Scheme for Shared-Buffer Switches

Boran Gazi[†], Zabih Ghassemlooy[‡]

[†]Electronics Research Group, Sheffield Hallam University, Sheffield, Phone: +44(0)114 225 3254,
Fax: +44(0)114 225 3433, Email: boran.gazi@student.shu.ac.uk

[‡]Optical Communications Research Group, University of Northumbria, Newcastle, Phone: +44(0)191 227 4902,
Email: fary@ieee.org

Abstract – Buffer space in packet switching nodes is an important network resource. Shared buffer switches are prone to high packet losses and unfair use of buffer space. The use of a buffer management scheme is necessary to overcome these problems. This paper investigates the performance of Sigmoid Function Threshold scheme by means of simulations. This scheme regulates the usage of shared buffer space by employing multiple thresholds for packet admission. The results show that it performs well under non-uniform input traffic; nevertheless buffer space reserved for lightly loaded output ports is wasted due to the strict packet admission control.

Keywords: Dynamic buffer management, shared buffer switch.

I. Introduction

An important issue of the overall control in communication networks is the management of buffers at the packet switching/routing nodes. Various studies have shown that shared buffer switch (SBS) is better than other buffering schemes in terms of packet loss performance [1]. It provides a notion of flexibility when allocating space for contending packets. Note that SBS greatly suffers from inefficient use of buffer space (e.g. packet loss rate, fairness) when network traffic is bursty and asymmetric [1].

A buffer management policy can be adapted for efficient and effective use of buffer space in SBSs. There have been many attempts in the literature to tackle this problem [1][2][3]. These policies can be mainly categorised into two main classes: Static and dynamic policies. The first one is based on static parameters set based on statistical information and the latter attempts to control the common buffer space based on the information from dynamic

environment parameters. As expected, dynamic policies outperform static policies as they are more aware of network dynamics (e.g. traffic load and pattern, link failure, priority traffic etc).

Push out (PO) or drop on demand (DoD) is known as the best policy as it is fair, efficient and naturally adaptive [2]. However practical implementation is almost impossible. Newly developed policies aim to achieve the same performance as DoD with low implementation overhead. Dynamic threshold (DT) is developed to achieve the dynamism of DoD together with simplicity of static threshold (ST) [3]. Another policy, maximum busy period (MBP), developed by Sharma and Viniotis, aims to keep output ports as busy as possible by simply pushing out a packet from the queue with the highest busy period [4]. Finally, adaptive fuzzy threshold scheme uses fuzzy rules and membership functions to set the threshold according to the overall occupancy of the shared buffer [5]. The main drawback is that the parameters of the membership functions have to be set and tuned through various simulations.

The aim of this paper is to adapt the Sigmoid function based fuzzy threshold policy [6] (developed for single server queue, M/D/1) to SBS; and contrast its performance to well-known DT. Section II provides a brief explanation of both DT and Sigmoid Function Threshold (SFT) schemes. Simulation model is presented in Section III and simulation results are discussed in Section IV.

II. Dynamic Buffer Management

A. Dynamic Threshold (Choudhury & Hahne)

DT scheme sets a single queue threshold for all of the dynamic length queues on the basis of the amount of empty buffer space. Purpose of this

scheme is to spare some space for unutilised output ports in order to prevent fully utilised output ports from dominating the usage of buffer. Simply, a packet is rejected if the queue length of that port exceeds the threshold value at time t , $T(t)$ [3]:

$$T(t) = \alpha.(M - Q(t)) \quad (1)$$

Where M is the total buffer capacity, $Q(t)$ the amount of buffer space occupied at time t and α a constant. The studies have shown that for different conditions, such as switch/buffer size and traffic phase, an α value between 2^{-1} and 2 is appropriate [1][3].

B. Sigmoid Function Threshold

The use of sigmoid shaped membership function was first studied in [6] on the basis of single server queuing model with the input traffic characterised by the Poisson arrivals. This scheme states that packets are admitted or blocked by using a notion of *fullness* (Fig. 1).

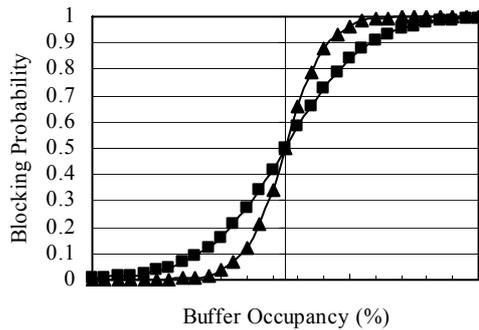


Fig. 1 – Sigmoid shaped packet-blocking scheme

This scheme can be easily utilised in SBSs. That is, probability of blocking a packet arriving at the queue of port i at time t is given by:

$$\mu_b^i(t) = \frac{1}{1 + e^{c.a}} \quad \text{and} \quad c = \frac{q^i(t)}{M.\beta} \quad (2)$$

Where $q^i(t)$ is the length of queue i at time t , M is the size of shared buffer and β is the maximum allocation ($0 < \beta \leq 1$). Parameter a defines the steepness of the curve. For very large a , the accept/reject policy is fixed rather than fuzzy. Regardless of the overall buffer occupancy, SFT takes into account the individual buffer usage within the permissible maximum allocation ($M.\beta$).

III. Simulation Model

In order to measure the performance and understand the characteristics of both of the schemes, a program is developed by using parallel virtual machine (PVM) platform. The aim of using this platform is to achieve a realistic environment with asynchronous and independent working processes (traffic generators, queue controller, and buffer management unit).

A. SBS Model

An $N \times N$ SBS consists of a buffer management unit that is responsible from storing buffer state, allocating space for newly arrived packets and calculating the threshold value. Threshold value is updated at every space request from one of the queue controllers. Queue controllers accept newly arrived packets and deliver packets at the head of the queue. Based on the threshold value, packet is granted a space or blocked. Each queue controller serves one packet per frame-time¹ (i.e. deterministic service time). A one-packet space is returned to the buffer management unit whenever a packet leaves the switch.

B. Traffic Model

Inputs of SBS are connected to N independent asynchronous traffic generators each of which generates fixed length packets. At most one packet is generated during a frame-time. Each traffic generator is a realisation of an Interrupted Poisson Process (IPP) [7]. The time spent in ON and OFF states is exponentially distributed with mean durations of T_{ON} and T_{OFF} respectively. Arrivals occur only at ON states with rate λ . Thus, the traffic load can be formally presented by:

$$p = \frac{\lambda.T_{ON}}{T_{ON} + T_{OFF}} \quad (3)$$

Traffic distributions considered in our simulations are symmetric and asymmetric scenarios.

IV. Results and Discussions

In our simulations we considered a 32×32 switch with 640-packet space shared among 32

¹ Frame-time is the duration of a packet.

output ports. Input load is 0.8 and the mean burst duration is 100 frame-times with symmetric and asymmetric distributions. Also, in all of the simulations DT α value is set to 1.0.

In Figure 2, we considered both symmetric and asymmetric loads. When the load is symmetric, where all of the output ports receive the same amount of traffic load (0.8), optimal β value is observed as 1.0 and SFT achieves the same packet loss rate (PLR) as DT. However, in case of asymmetric loads where hotspot ports receive 0.95 and 1.05 loads, β value has to be within the range of 0.20 and 0.25 in order to achieve the lowest possible PLR. Therefore, one could state that a full share is needed for SFT when the traffic is uniformly distributed and the parameter β defines the level of buffer space sharing.

The result of a similar test is depicted in Figure 3. In this case hotspot load is set to a fixed value (0.95) and the effect of different number of hotspots is observed. Regardless of the number of hotspot ports, optimal β value is between 0.20 and 0.25. Unlike DT, SFT employs a separate admission (threshold) for every queue and its performance is bounded with the β value. For this reason, share parameter (β) has to be at an optimal level in order to limit the queue lengths of heavily loaded output ports or a full share has to be employed in case when there is a uniformly distributed traffic.

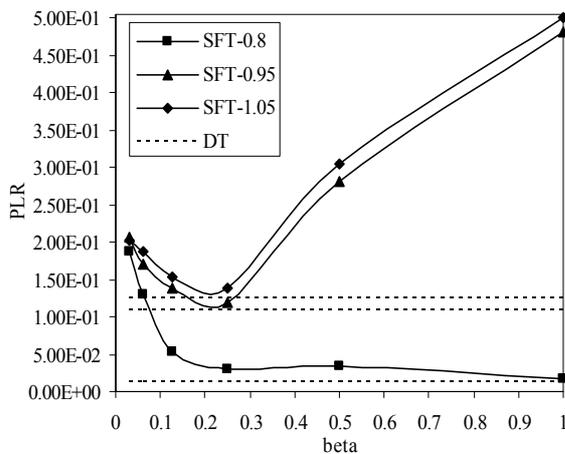


Fig. 2 – Packet loss rate against β value of SFT for various hotspot loads

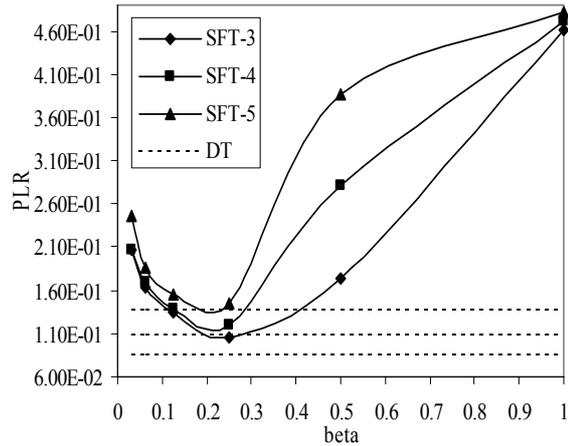


Fig. 3 – Packet loss rate against β value of SFT for different number of hotspot ports

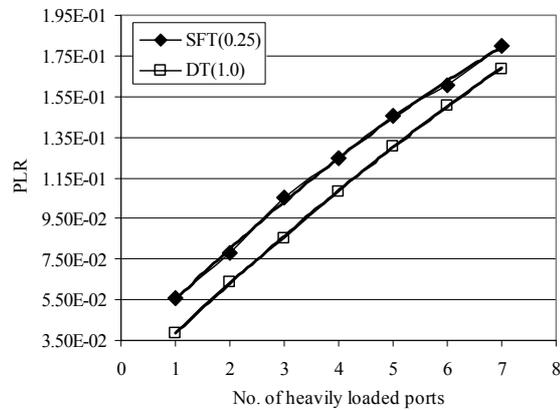


Fig. 4 – Packet loss rate against the number of heavily loaded hotspot ports

In Figure 4, we compared the performances of both schemes using a fixed hotspot load (0.95) while varying the number of heavily loaded ports. Also, β is selected as 0.25 for SFT. SFT scheme achieves a reasonably close performance to DT. As the number of hotspots increases, PLR performances of both of the schemes are expected to be approximately the same due to saturation of the buffer.

In Figures 5 and 6, we considered the buffer occupancies of hotspot ports and unused buffer space under asymmetric input traffic for a given duration of simulation time. As can be seen in Figure 5, the available buffer space in SFT scheme is affected by the ON and OFF periods of the input traffic. However, in DT scheme, the available buffer space is determined by the active ports (Figure 6). Besides, buffer occupancy of active ports in SFT is rather steady when compared to DT. The reason for this is

that the sigmoid shaped membership function is applied to all of the queues individually. As a result, the lengths of all of the queues, regardless of whether they are active or not, are controlled in the same way. On the other hand, DT scheme controls only those queues that are highly active. Hence, lightly loaded queues are allowed to grow and the maximum busy periods of each queue is increased [4].

The buffer space that SFT scheme spares for lightly loaded ports is underutilised due to the sigmoid membership control of the lightly loaded queues. However, DT scheme spares enough space for lightly and moderately loaded queues to grow freely. Nevertheless, this scheme still performs worse than the well-known DoD, because some space is still wasted.

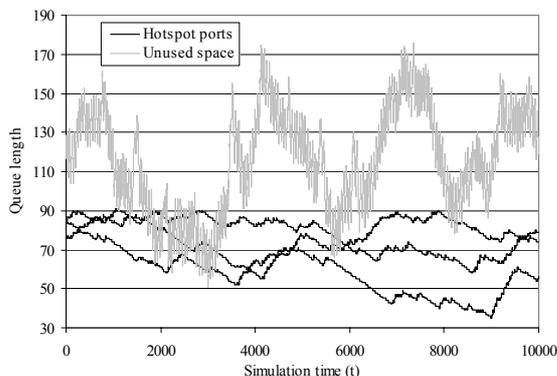


Fig. 5 – Transient buffer occupancy of SFT in time (3 hotspot ports with 0.95 load each)

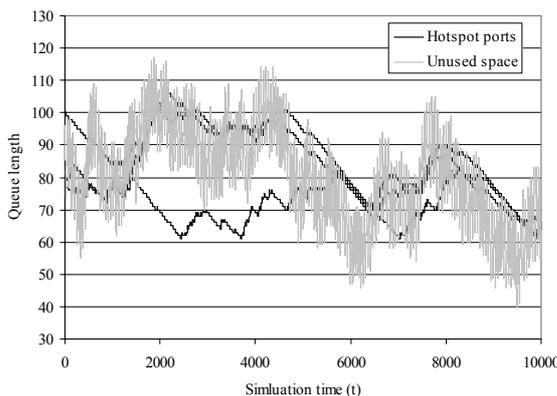


Fig. 6 – Transient buffer occupancy of DT in time (3 hotspot ports with 0.95 load each)

V. Summary

In this paper, we adapted the sigmoid membership function in order to be used in shared buffer switches. A simulation study has been carried

out to compare the performance of SFT with the well-known DT.

The results presented show that under symmetric ON-OFF input traffic the optimal β value for SFT is 1.0; whereas for asymmetric ON-OFF traffic β value is between 0.20 and 0.25. SFT achieves a reasonably good packet loss rate when compared to DT. It uses a multiple packet admission scheme rather than a single threshold. The main drawback is that even the lightly and moderately loaded queues undergo this control scheme and they are not allowed to grow freely. As a consequence, buffer space spared by very active queues is underutilised.

Although, SFT performs slightly worse than DT, its performance is promising and it can be tuned by applying the admission scheme only to very active queues.

VI. References

- [1] M. Arpaci and J. A. Copeland, "Buffer Management for Shared-Memory ATM Switches," *IEEE Communications Surveys*, First Quarter 2000.
- [2] S. X. Wei, E. J. Coyle, and M. T. Hsiao, "An optimal buffer management policy for high performance packet switching," *IEEE GLOBECOM'91*, Vol. 2, Dec. 1991, pp.924-928.
- [3] A. K. Choudhury, and E. L. Hahne, "Dynamic queue length thresholds for shared-memory packet switches," *IEEE/ACM Transactions on Networking (TON)*, Vol.6, No.2, Apr 1998, pp.130-140.
- [4] S. Sharma, and Y. Viniotis, "Optimal buffer management policies for shared-buffer ATM switches," *IEEE/ACM Transactions on Networking*, Vol.7, No.4, Aug. 1999, pp.575-587.
- [5] G. Ascia, V. Catania, and D. Panno, "An Efficient Buffer Management Policy Based On An Integrated Fuzzy-GA Approach", *IEEE/ACM INFOCOM*, June 2002.
- [6] A. Bonde, and S. Ghosh, "A Comparative Study of Fuzzy versus 'Fixed' Threshold for Robust Queue Management in Cell-Switching Networks," *IEEE/ACM Transactions on Networking*, Vol.2, No.4, August 1994, pp.337-344.
- [7] A. Adas, "Traffic Models in Broadband Networks," *IEEE Communications Magazine*, July 1997, pp. 82-89.

The Antnet Routing Algorithm – A Modified Version

F. Tekiner^a, F. Z. Ghassemlooy^a and S. Al-khayatt^b

a: Optical Communications Research Group, School of Engineering and Technology,
Northumbria University, Newcastle upon, Tyne, NE1 8ST, UK.

b: School of Computing and Management Sciences, Sheffield Hallam University,
Sheffield, S1 1WB, UK.

Emails: ftekiner@ieee.org, fary.ghassemlooy@unn.ac.uk

Abstract – Antnet is an agent based routing algorithm that is influenced from the unsophisticated and individual ant's emergent behaviour. Ants (software agents) are used in antnet to collect information and to update the probabilistic distance vector routing table entries. Modified antnet algorithm has been introduced, which improve the throughput and average delay. Results shows that by detecting and dropping 0.5% of packets routed through the non-optimal routes the average delay per packet decreased and network throughput can be increased. The effect of the traffic fluctuations has been limited with the boundaries introduced in this paper and the number of ants in the network has been limited with the current throughput of the network at any given time.

1. Introduction

In today's fast growing Internet traffic conditions changes and failures occurs at some parts of the network from time-to-time, in an unpredictable manner. Therefore, there is a need for an algorithm to manage traffic flows and deliver packets from the source to the destination in reasonable time.

In a network there are more than one optimal path exist. The challenge is to deliver packets to their destinations with minimum time delay, if necessary routing them via more nodes. However, it is neither realistic nor feasible to have a common control unit that controls all links and processing elements (nodes) in the network or part of the network. Moreover, such systems are not fault tolerant, since a node failure will result in unavoidable network delays and shut down of some parts of the network. Therefore, the solution would be to divide the task into a number of sub-tasks distributed across the entire network. Moreover, the network should have a capability of organising/reorganising itself whenever changes take place within the network.

Routing algorithm is the key element in networks performance and reliability, thus it can be seen as the "brain" of the network. An ideal routing algorithm should be node and link independent, and be able to deliver packets to their destination with the minimum amount of delay, regardless of their size and traffic load. Formally speaking, a routing algorithm should be intelligent, adaptive and fault tolerant. The only way to achieve this would be by employing an intelligent and distributed routing algorithm.

The routing algorithms currently in use [1] lack intelligence, and need human assistance and interpretation in order to adapt themselves to failures and changes. Moreover, they are mainly table-based and search process is highly undesirable. Routing is considered to be NP-Hard Optimization problem, therefore widely used optimization problems have been applied widely. To name a few, Genetic Algorithms, Neural Networks, Simulated Annealing, Software Agents and Reinforcement Learning [2].

In recent years, agent based systems and reinforcement learning have been widely applied to routing. This is because these methods do not need any supervision and are distributed in nature. Swarm intelligence particularly ant based systems [3] [4], Q-learning [5] methods and hybrid agent based Distance Vector algorithms [6] have also shown promising and encouraging results. In this paper, the focus will be on the antnet routing algorithm.

2. Software Agent and Ant Based Routing

Unlike other optimisation problems, the routing problem has to be solved in real time since it is a dynamic optimisation problem. The ant-based approach applied to routing problem was first reported in [7]. This work was influenced from the work done on the software agents used for control in telecommunication networks [8]. Improved

version of the algorithm in [7] was applied to the connection oriented systems [9]. For the first time ant based routing was applied to the packet based connection-less systems [10]. This was followed by a mobile agents approach to adaptive routing in [11].

2.1 Antnet Routing Algorithm

Individual ants are unsophisticated and simple insects, but collectively that are capable of performing a variety of complicated tasks [12]. Antnet [13] [14] is an agent based routing algorithm that is influenced from the real ants behaviour. In antnet ants (nothing but software agents) explores the network to find the optimal paths from the randomly selected source destination pairs. Moreover, while exploring the network ants updates the probabilistic routing tables and builds a statistical model of the nodes local traffic. Ants use these tables to communicate with each other. The algorithm uses two types of agents namely, forward agents and backward agents to collect network statistics and to update the routing table. In each node there are 2 types of queues, low priority and high priority. The data packets and the forward agents use low priority queues, whereas the backward agents use the high priority queues. Later forward agents do also use the high priority queues [14].

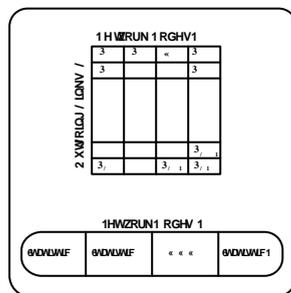


Figure 1: Probabilistic routing table and traffic statistic.

Agents communicate through the two data structures stored in every node (see figure 1) as outlined below [13,14]:

- i. A distance vector routing table T_k with distance metric defined with probabilistic entries where for each possible destination d and neighbour node n there is a probability value P_{nd} , which reflects the goodness of the link (path), given as:

$$\sum_{n \in N_k} P_{nd} = 1, d \in [1, N], \quad (1)$$

$$N_k = \{Neighbours(k)\}$$

- ii. An array $M_k(\mu_d, s_d^2, W_d)$ defines a simple statistical traffic model experienced by the node k over the network. Where W_d is the observation window used to compute the estimated mean μ_d and the variance s_d^2 parameters¹ given as, respectively:

$$\mu_d = \mu_d + \alpha(o_{k \rightarrow d} - \mu_d) \quad (2)$$

$$s_d^2 = s_d^2 + \alpha((o_{k \rightarrow d} - \mu_d)^2 - s_d^2) \quad (3)$$

Where $o_{k \rightarrow d}$ is the new observed trip time experienced by the agent while travelling from node k to destination d .

The antnet behaviour is summarised as [13]:

1. At regular intervals (defined by the user) from every node an agent A is sent to a randomly selected destination node d .
2. Each agent first defines the possible neighbour nodes (unvisited neighbour nodes) at the current node by using its routing table. Then, an agent chooses the next node n within the identified possible (unvisited) nodes by using the probabilistic values in the table by taking into account the state of the associated queue as follows²:

$$P'_{nd} = \frac{P_{nd} + \alpha l_n}{1 + \alpha(|N_k| - 1)} \quad (4)$$

Where l_n is the heuristic correction value with respect to the probability values stored in the routing table that gives a quantitative measure of the queue waiting time, which is defined as:

$$l_n = 1 - \frac{q_n}{|N_k|} \quad (5)$$

$$\sum_{n'=1} q_{n'}$$

Where, q_n is the bit length (or number of packets if packet size is fixed) waiting to be sent to the queue on the output port n of the node k .

3. If the selected port is full, then an agent is directed to the FIFO output buffer associated for that port and waits for its turn. It is assumed that the buffer size is infinitely large.
4. The identifiers of every visited node and time elapsed since the agent is despatched

¹ α weights the number of samples that affect the average and set to 0.05[13].

² α weights l_n with respect to routing table entries and set to 0.3[13].

from the source are stored in the stack $S(k)$ carried by the agent.

5. If an agent is forced to visit previously visited node, hence if cycle exists, then it deletes all the entries for the nodes associated with the cycle.
6. When an agent reaches the destination it changes its type as backward agent and by using the same path, to reach the destination, it returns back to the source node. On its way back, agent pops every visited node from its stack and updates the associated routing table entries (probabilities) for all the nodes along the path by using the following rules:

(i) $M_k(\mu_d, s_d^2, W_d)$ is updated with values stored in the stack memory $S(k)$. The time elapsed to arrive (for the forward ant) to the destination ($O_{k \rightarrow d}$) is used to update the estimated mean μ_d , variance s_d^2 and the best trip value over the observation window W_d .

(ii) The routing table T_k is updated by incrementing the probability of selecting neighbour f when the destination is d (P_{fd}) is given as:

$$P_{fd'} = P_{fd} + r(1 - P_{fd}) \quad (6)$$

Probabilities (P_{nd}) for rest of the neighbour nodes are updated as:

$$P_{nd'} = P_{nd} - rP_{nd}, n \in N_k, n \neq f \quad (7)$$

r defines the goodness of the path taken by the ant, which is used as the only feedback (reinforcement) information to the routing tables. r is defined as follows:

$$r = c_1 \left(\frac{W_{best}}{T} \right) + c_2 \left(\frac{I_{sup} - I_{inf}}{(I_{sup} - I_{inf}) + (T - I_{inf})} \right) \quad (8)$$

Where, T is the observed trip time, c_1 and c_2^3 are constants used to control the effect of the last T , W_{best} is the ant's best trip time for a given destination and last observation period. I_{inf} and I_{sup} are the lower and upper limits for the confidence interval μ , respectively. Where $I_{inf} = W_{best}$ and $I_{sup} = \mathbf{m}^* z^* (\mathbf{s} / \sqrt{|w|})^4$. More details and discussion can be found in [13].

³ c_1 and c_2 are chosen experimentally to be 0.3 and 0.7 respectively.

⁴ $z = 1/\sqrt{1-g}$, g lies is the confidence, ?? [0.75, 0.8].

3. Antnet Modifications and Results

Based on the original antnet routing algorithm three modifications has been proposed and tested.

i) Deleting aged packets

During the simulations run on the original antnet routing algorithm it was discovered that some packets travel within the network for a number of hops until they reach their destination. This problem occurs because the routing tables used in the antnet routing algorithm is probabilistic and therefore a few packets have chance to be directed to the non-optimal routes and cycle. A simple rule is defined to detect and drop these packets from the network as follows:

if PACKET AGE > 2 x NO_OF_NODES
then DROP PACKET

This rule is defined based on the information gained from experimental results. It was observed that only 0.5% of the packets experience this problem. Therefore, only 0.5% of the packets are dropped from the network. However, when the packet age condition is set to $1XNO_OF_NODES$, the packet loss increased to almost 7%. On the other hand, when the condition is set to $3XNO_OF_NODES$, the loss decreased to 0.3% with no further improvement in the performance.

ii) Limiting the effect of r due to traffic fluctuations.

The reinforcement r applied to the routing table entries is limited by the lower and upper bounds defined as follows:

if (NO_OF_NODES <= 5)
 $0.1 < r < (1 - 0.1 * NO_OF_NODES)$
else /* *if* (NO_OF_NODES > 5) /*
 $0.05 < r < (1 - 0.05 * NO_OF_NODES)$

The values used are based on experimental results and it is intended to limit the effect of the traffic fluctuations in the network at a given time. Similar method to control the effect of r has also been reported in [15]. However, if r does not satisfy these values for three consecutive times and is less than 0.95, then reinforcement is applied to the routing table entry.

It is believed that by limiting the impact of r on the routing table entries the algorithm

would not freeze as easily as it was on the originally proposed algorithm.

iii) Limiting the number of Ants in the system.

In the modified version of the algorithm, the number of ants is limited by the **ant/packet** ratio. Although, this has no major impact on the performance of the algorithm (when compared with results on $1.0 > \text{ant creation rate}^5 > 0.1$), it would increase the adaptability and suitability of the algorithm. In other words it makes algorithm more generic.

$$\frac{\text{NO OF ANTS CREATED}}{\text{NO OF PACKETS SEND}} = 0.001$$

4. Experimental Results

With the proposed modifications both the original and the modified antnet routing algorithms were implemented in the following environment.

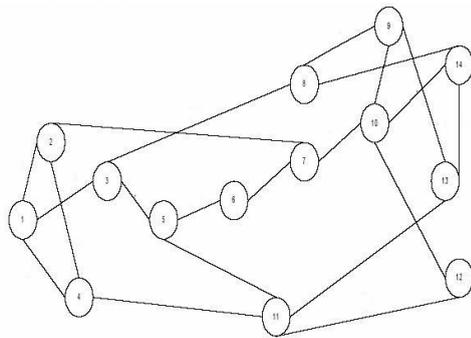


Figure 2: NFSNET topology.

- Algorithms are implemented in the C language in a parallel environment by using PVM.
- Parallel behaviour is simulated by assigning every node to a different process both on the same machine and different machines.
- Poisson traffic distribution is used with three different system loads low, medium and high.
- Algorithm is given 5 seconds to discover all the paths and initialise the probabilistic routing table entries.
- NFSNET (figure2) is used as the network topology with each link having equal cost.
- For each simulation, packet generation is stopped after creating 5000 packets per node and simulation is stopped after all

packets are arrived to their destinations or detected and deleted from the network.

- Every simulation is run 8 times and the average of the results is used for accuracy.
- It is assumed that there is no packet loss.
- Packet size is fixed.
- All experiments are implemented for varying ant creation rates, since it has a significant effect on the performance of the algorithm.

The performance parameters for the simulations are:

- Average packet delay:** is the average delay a packet experiences while being routed from source to destination.
- Average throughput per packet:** is the average number of packets being forwarded by node for the duration of the simulation.

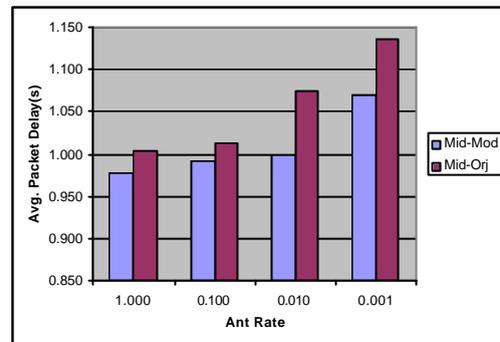


Figure 3: Average packet delay vs. ant creation rate.

From figure 3, it can be seen that as the number of ants increases in the network the average packet delay increases. However, with the modified algorithm packet delay is reduced compared with the original routing algorithm. This is because in the original routing algorithm some packets travel with very high no of hops within the network thus uses a lot of network resources and occupies a considerable amount of bandwidth.

Moreover, further simulations for ant creation rates greater than 1.0 showed that the average delayed increased and throughput of the algorithm decreased slightly in both algorithms. Therefore, the ant rate between 1 and 0.1 is optimal for our system. However, since node and link failures has not been implemented and investigated in this work it is not possible to comment on the effect of ant rate on the performance of the network in problematic conditions.

⁵ Ant creation rate defines the frequency (in time) of the forward ants to be created by the node. Therefore, a low ant rate, means a high number of ants in the system.

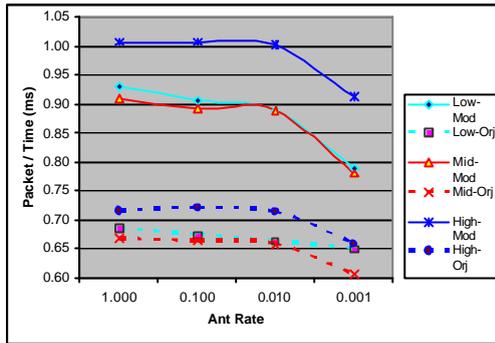


Figure 4: Throughput vs. ant rate for low, medium and high traffic loads.

Figure 4 shows the throughput against the ant rate for three different system loads for the original and the modified antnet routing algorithms. For all system loads there is an improvement of approx. 0.25 packet/time in the throughput when using the modified algorithm. As the ant rate increases the throughput decreases, but having more ants increases the accuracy of the probability entries of the routing tables. However, some of the traffic created is artificial due to the ants travelling in the system.

5. Conclusions and Further Work

In this paper, it has been shown that by detecting and dropping packets that travel continuously within the network can improve the antnet's performance in terms of network throughput and the average packet delay. The effect of traffic fluctuations on the network performance has been limited by the introduction of boundaries, by limiting the number of ants within the network at any given time. Although limiting the number of ants increases the network utilization, it reduces the chance of finding the best and new routes and detecting failures and problematic conditions. In future work the ratio proposed for the ant creation rate will be investigated for problematic conditions and hotspot traffic together with the other network topologies.

6. References

[1] M. E. Steenstrup, "Routing in communications networks", Prentice-Hall, 1st edition, ISBN: 0130107522, 1995.
 [2] R. S. Sutton, and A. G. Barto, "Reinforcement learning: an introduction", MIT Press, 1998, Chp. 1.
 [3] M. Dorigo, G. Di Caro, and M. Sampels, "Ant algorithms", Proc. of ANTS 2002, 3rd Internl. Workshop on Ant Algorithms, Brussels, Belgium, Sep. 12-14, 2002, Lecture Notes in Computer Science Vol. 2463, Springer-Verlag, 2002.

[4] K. M. Sim, and W. H. Sun, "Ant colony optimization for routing and load-balancing: survey and new directions", Systems, Part A, IEEE Transn. on , Man and Cybernetics, 33 (5), Sept. 2003, pp. 560-572.
 [5] J. A. Boyan, and M. L. Littman, "Packet routing in dynamically changing networks: a reinforcement learning approach", Advances in NIPS 6, Morgan Kauffman, SF, 1994, pp. 671-678.
 [6] K. Amin, J. Mayes, and A. Mikler, "Agent-based distance vector routing" Proceedings of the Third International Workshop, MATA 2001, Montreal, Canada, August 2001.
 [7] R. Schoonderwoerd, "Collective intelligence for network control", M.S.c Thesis, Delft University of Technology, May 1996.
 [8] S. Appleby, and S. Steward, "Mobile software agents for control in telecommunication networks", British Telecom Techn. J., 12, 1994, pp. 104-113.
 [9] E. Bonabeau, F. Hénaux, S. Guérin, D. Snyers, P. Kuntz, and G. Theraulaz, "Routing in telecommunications networks with "smart" ant-like agents", Proceedings of IATA '98, 2nd Int. Workshop on Intelligent Agents for Tele. Applns. Eds., S. Albayrak & F. J. Garijo, (pp. 60-71), Lecture Notes in AI, Vol. 1437, Springer-Verlag, 1998.
 [10] D. Subramanian, P. Druschel, and J. Chen, "Ants and reinforcement learning: A case study in routing in dynamic networks", Proc. of IJCAI-97, Palo Alto, Morgan Kauffman, 1997, pp. 832-838
 [11] G. Di Caro, and M. Dorigo, "AntNet: a mobile agents approach to adaptive routing", (Tech. Rep. 97-12) Université Libre de Bruxelles, IRIDIA, 1997.
 [12] R. Beckers, J.L. Deneubourg, and S. Goss, "Trails and u-turns in the selection of a path by the ant *Lasius Niger*", J. of Theoretical Bio., 159, 1992, pp. 397-415.
 [13] G. Di Caro, and M. Dorigo, "AntNet: distributed stigmergetic control for communications networks", J. of Artificial Int. Research (JAIR), 9, 1998, pp. 317-365.
 [14] G. Di Caro, and M. Dorigo, "Two ant colony algorithms for best-effort routing in datagram networks", Proc. 10th Intern. Conf. on Parallel and Distributed Computing and Systems, Las Vegas, Nevada, USA, October 28-31, 1998, pp. 541-546.
 [15] B. Baran "Improved AntNet routing", ACM SIGCOMM, Workshop on data com. in Latin America and the Caribbean, 31 (2), April 2001, pp. 42 - 48, ISSN:0146-4833.

The effect of using multiple ARs in MANETs for global connectivity

Vinh Dien HOANG, Zhenhai SHAO and Masayuki FUJISE

Wireless Communications Laboratory,

National Institute of Information and Communications Technology, Singapore 117674,

{hvdien, zhshao, fujise}@nict.com.sg

Abstract: Providing Internet access for nodes in MANETs (Mobile Ad hoc Networks) is a challenging task because of the dynamic nature of the ad hoc networks. This dynamic nature leads to the difficulty in addressing, routing and data delivering. In this paper we present the simulation results of using multiple Access Routers (ARs) in MANET for Internet connectivity. Using multiple ARs brings a lot of benefits such as soft-handoff, throughput improvement and load balancing between ARs. The simulation results using DSDV routing protocol show that using two ARs at the same time improves throughput especially in the small size network.

Keywords: Mobile Ad hoc network, Internet connectivity, Auto configuration, Hybrid networks.

I. Introduction

Internet access services for MANET nodes are provided by attachment points which are called Internet gateways or Access Routers (AR). Nodes in MANET requiring connection to the Internet have to register with ARs and obtain a global routable address. This is done through address autoconfiguration [1, 6], gateway discovery [2, 3] and registration processes. Depending on what type of ad hoc routing protocol is used in MANET, those processes could be manual or proactive or integrated with ad hoc routing protocol.

When a MANET node wants to send data, it first sends packets to the AR by using ad hoc routing protocol in MANET. The AR will forward those packets to the destination in the Internet using normal routing protocol in the wired network such as RIP or OSPF. A packet sent to a MANET node from the Internet will eventually reach AR and AR will forward the packet to this MANET node using ad hoc routing protocol.

However, it's not that simple in reality because of the dynamic nature of MANET where nodes could join and leave the network at any time. It becomes even worse when more than one ARs and/or large and dynamic nodes movement are involved. Some main challenges are:

How to discover and select suitable ARs?

How to create, to maintain connections efficiently when node is moving?

When should the node handover to new ARs and how?

Quite a number of solutions have been proposed so far to tackle those problems [2-5]. Among them, MIPMANET [4] is one of the first solutions to provide global access for nodes in MANET by

extending the Mobile IP to support nodes multiple hops away from access routers. Using MIPMANET, the MANET node will register with the Foreign Agent (FA) in the foreign network and use their home address in all connections. Packets from MANET nodes will be tunneled to the FA using MANET ad hoc routing protocol and forwarded to the CN (Correspondent Node) by the FA as in normal Mobile IP.

Another detail scheme for global connectivity is presented in [2]. This scheme can use proactive, reactive routing protocols or Neighbor Discovery Protocol (NDP) to discover the Internet gateway-Access Router. After a MANET node discovers and registers with an AR, it will use this AR as its default router for all connections outside MANET. Routing header option could be used to explicitly specify this AR.

To the best of our knowledge, there're no studies so far taking into account the fact that in a hybrid network, MANET nodes could have access to more than one ARs at the same time and could use these ARs simultaneously. This is the motivation for us to study the case when a MANET node uses more than one ARs simultaneously for data communication with nodes in the Internet. Whenever a MANET node wants to send a packet to the Internet, it will send the packet to one of its registered ARs. So the load on the connection will be divided between the two ARs which in turn will increase the performance of the connection. Additionally, when the MANET node hands over to a new AR, its data will not be lost because it still can use the remaining AR for the communication during the hand over process.

The rest of this paper is organized as follows. Detail information of using two ARs at the same time for global connectivity is presented in Session II. It includes addressing, registration, routing and data delivering. The performance analysis based on simulation using NS-2 Simulator is presented in Session III. The conclusion and suggestions for future works are shown in Session IV.

II. Detail operations

2.1 Address configuration & Gateway discovery

Initial address configuration is quite strait forward and could be followed [1]. The MANET node uses the 64 bits MANET_PREFIX (fec0::ffff/64) and its 64 bits EUI as its IPv6 address when it joins the MANET. After that, it has to perform the Duplicate Address

Detection- DAD operations to ensure that its above chosen address is unique in the whole MANET.

Then the MANET node has to discover the ARs (Access Router - Internet gateway) to obtain information, such as global network prefixes, address lifetime, etc., for registration & routing packet to the Internet. Gateway discovery process is described in [2]. Basically, there're two ways for MANET nodes to discover the ARs.

- In the first way, MANET node discovers the ARs by listening to the gateway advertisement messages (GWADV) periodically broadcasted by ARs as part of the ad hoc routing protocols or Neighbor Discovery Protocol (NDP). If a proactive ad hoc routing protocol is used, GWADV could be piggybacked to any broadcast routing message, such as HELLO message. If a reactive ad hoc routing protocol is used, GWADV could be stored in the sent route reply/notify messages. If NDP is used, GWADV message could be sent in router advertisement message of NDP. In all cases, the GWADV message will be broadcasted to all nodes in the MANET.

- In the second way, MANET node will explicitly ask for ARs to send it the GWADV information. Again the ad hoc routing protocol or NDP could be used. If a reactive routing protocol is used in MANET, MANET node could use the Route Request packet (RREQ) which is destined for all ARs as Gateway solicitation message (GWSOL). Source address used in this RREQ message could be the above chosen MANET address or home address. ARs will reply by route reply message (RREP) as a GWADV message. If NDP is used, MANET node could send the GWSOL message to Internet gateway multicast group (IGW_MCAST). ARs should belong to this group so they could be reached and send back GWADV information to the requesting MANET node.

Upon receiving those GWADV messages, MANET node will have enough information to choose the best ARs (which usually the shortest AR in term of hops) for connecting to the Internet.

2.2 Registration

To use two ARs, the MANET node has to register with two ARs. Based on the information advertised by ARs in GWADV message, MANET node will know the number of available ARs, their addresses, IP prefixes, and other metric such as number of hop to reach ARs, etc. The MANET node will choose the two best ARs (usually the two shortest ARs in term of hops) as its candidate ARs to register with. More than two ARs could also be used with the same principles presented here or the number of ARs node could be varied according to network conditions. However, it's out of scope of this paper.

At this time the MANET node already has the IPv6 prefixes of the two selected ARs and related information. Using those prefixed, two local on-link Care of Addresses (LCoA) are automatically created

by appending them to the MANET node's EUI. It then sends REG packets containing information about the newly constructed LCoAs to all the selected ARs. Upon receiving those REG packets, the selected ARs will update their list of registered MANET nodes. Only registered nodes could use ARs as gateways for communication with outside networks.

The MANET node now is ready to receive and send data from and to the Internet.

2.3. Routing & data delivering

Routing inside MANET is quite straightforward using ad hoc routing protocol. Routing outside MANET is done through two registered ARs. When MANET node wants to send a packet to the Internet, it will choose one of the two registered ARs to send this packet to, based on the MANET node load balancing and network conditions.

IPv6 header	Routing header	Data
Src addr: RCoA	Addr 1: LCoA	...
Dst Addr: AR	Addr 2: CN's addr	...

(a) Original packet headers

IPv6 header	Routing header	Data
Src addr: RCoA	Addr 1: LCoA	...
Dst Addr: CN's addr	Addr 2: AR	...

(b) After being processed by the selected AR

Figure 1: Packet headers before and after being processed by the selected AR

This packet's IPv6 headers are presented in Figure 1. The source address for this packet could be the LCoAs or other global reachable address (e.g using HMIP RCoA). The destination address for the packet is the selected AR address (Figure 1(a)). The final destination CN (Correspondent Node) address of this packet will be stored in the routing header extension [9]. Upon receiving the packet, the AR will exchange the AR address in the IPv6 header with the CN address in the routing header extension as specified in [9] (Figure 1(b)). This routing header extension told the selected AR that the packet was sent from the MANET node to the CN node through the registered LCoA node. The destination of the packet is now the CN address.

The AR will check this destination CN address. If this destination is inside the MANET, an ICMP error packet is sent back to the source node and forces it to use route within MANET for data communication.

If the destination is a host in the Internet, the AR will forward this packet normally using the normal Internet routing protocol. All nodes outside MANET will only see the global reachable address of the MANET node shown in the packet's header. Packets sent by CNs/HA to this global address will eventually reach registered ARs. ARs then route those packets to the destination MANET node by the ad hoc routing protocol.

III. Performance Analysis

Simulations using NS2 simulator [7] are used to evaluate the effect of using two ARs for the global connectivity. Direct Sequence Distance Vector-DSDV is chosen as the proactive ad hoc routing protocol used in the MANET. AR information is piggy-backed in routing packets which are regularly exchanged between MANET nodes.

The following metrics are used for performance comparisons:

Routing overhead is the total number of ad hoc routing messages transmitted during the simulation. If the routing packet is forwarded by a node, it will also be counted as one transmission. The higher number of routing overhead are, the less efficient the routing protocol is.

Throughput is the number of packet received at the destination during the simulation.

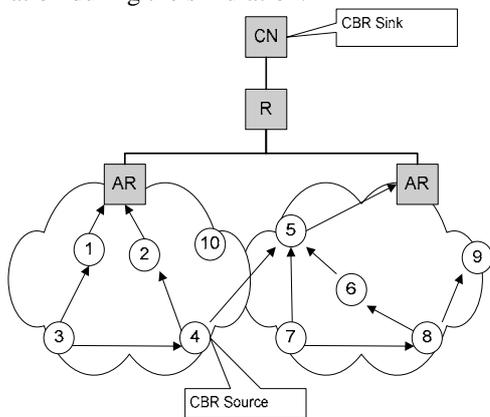


Figure 2: Network simulation scenario

Hybrid network consists of a fixed network (representing the Internet) and a wireless network (representing the MANET) as in the Figure 2 will be simulated in various simulation scenarios. The fixed network consists of 4 nodes (one CN, one Router node and two ARs). The wireless network is simulated in the small (10 nodes) and large (50 nodes) network scenarios. DSDV is used as the ad hoc routing protocol. Information about the ARs is piggybacked in the standard DSDV routing packet by adding a new flag to inform other nodes whether this routing packet is from a AR or not. This flag is one byte long. Since in DSDV, routing information is regularly exchanged between nodes, nodes will know which ones are the ARs and choose them for global connectivity.

3.1 Small network simulations

In the small network scenarios (10 MANET nodes) simulation, there're 3 CBR (Constant Bit Rate) traffic sources between MANET nodes, one more CBR source from MANET node to the router and one more CBR source from MANET node to the CN. So totally there're 5 CBR traffic sources. All MANET nodes, except one, are randomly moving according to the random waypoint mobility model [8]. The fixed node

is used to study the multi hops connection from this node to ARs. MANET nodes use two ARs only for packets originating from themselves. If they have to forward packets for other nodes, they will forward the packet according to their routing table. All the simulations are lasted for 900 seconds.

Parameters for random waypoint model are:

Max speed: 20m/s

Pause time: 10s

Topology: 450x450m

Parameters for the CBR source are:

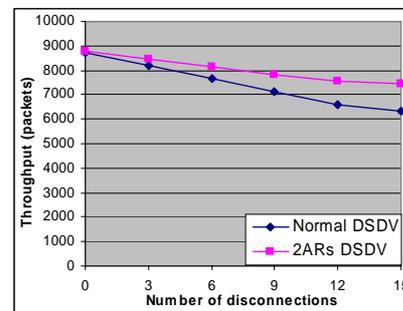
Traffic type: CBR

Packet size: 512 bytes

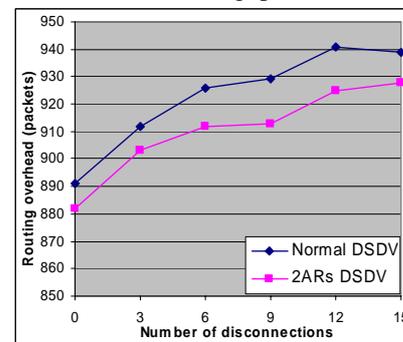
Packet rate: 10 packet/s

Links between CN, Router and ARs are 5Mbps with 2ms delay.

Simulation results



(a) Throughput



(b) Routing overhead

Figure 3: Small wireless network using DSDV

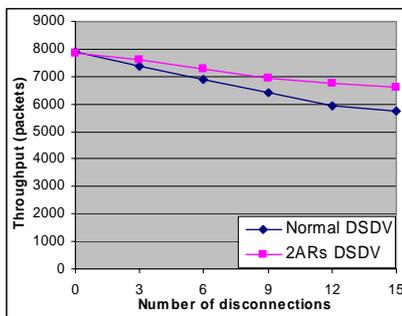
Small network simulation results are shown in Figure 3. The number of disconnection is the number of times that the MANET node is disconnected from its ARs. It's obvious that when the number of disconnection increases, the network performance decreases, because the network needs time to stabilize after each disconnection and packets have to be buffered or may have to take a longer route to the destination.

Figure 3(a) shows the throughput of node using normal DSDV and 2ARs DSDV (A node using two ARs simultaneously). When a connection with current AR is lost, it takes a longer time for a normal MANET node to react to the changes. A 2ARs node during this time still can use the other AR for its packets, so the

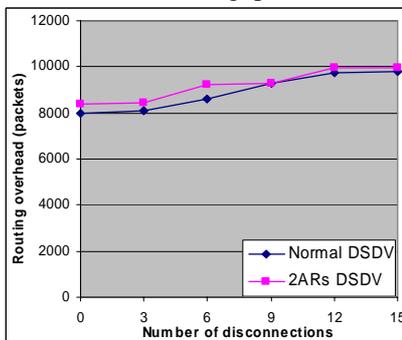
disconnection has less impact on the 2ARs nodes. That's reason when the number of disconnection increases, the throughput of both 2ARs node and normal node decreases but normal DSDV node decreases faster. Routing overhead (Figure 3(b)) of the 2ARs DSDV is slightly less than normal DSDV in the small network simulation due to the fact that the network is a little bit more stable using 2ARs, so there's less triggered update.

3.2 large network simulations

These second set of simulations are used to study the behavior of the MANET nodes in the larger scale network consists of 50 nodes. Other parameters are kept unchanged as in the first set of simulation except that there're four more CBR traffic connections between MANET nodes. And the topology is also increased to 1000x1000m.



(a) Throughput



(b) Routing overhead

Figure 4: Large wireless network using DSDV

Figure 4 shows the simulation results of the large scale network using DSDV routing protocol. It's obvious from Figure 3 and Figure 4 that when the number of MANET node increases, the network performance is decreased. However, simulation results in Figure 4 prove that a 2ARs DSDV node still outperforms normal DSDV nodes in large network environment in term of throughput (Figure 4(a)). Routing overhead is dominated by network size and mobility. Due to the mobility and the large size of the network, more changes are expected in the network topology resulting in more triggered updates. So routing overhead of DSDV nodes using or not using 2ARs is almost identical (Figure 4(b)).

IV. Conclusion

In this paper a new scheme of using more than one ARs for internetworking between MANET and the Internet has been studied. Using more than one ARs for Internet connection simultaneously, MANET node could maintain its connection even during handoff. Moreover, the connection reliability is increased because when one AR's down the other AR is still there.

Simulations using proactive DSDV routing protocol, in the small and large size network, show that using two ARs at the same time improves throughput especially in the small size network. However in larger networks, using two ARs produces more routing overhead. But overall performances of using two ARs are very promising.

In the future works, we would like to extend the simulation to use reactive routing protocols such as AODV as well as to study the load balancing algorithm at the MANET nodes. We will also want to study what's the optimum number of ARs that the MANET node should use, based on the network and node conditions. This will help to achieve better network performance and reduce routing overhead.

Acknowledgments

We would like to thank Dr. Nguyen Minh Hoang of Institute for Inforcomm Research, National University of Singapore for his valuable comments and suggestions on this paper.

References

- [1] C. E. Perkins, J. T. Malinen, R. Wakikawa, E.M. Belding-Royer, Y. Sun, "Ad hoc Address Autoconfiguration", Internet Draft, draft-ietf-manet-autoconf-01.txt, Nov. 2001 (Work in Progress).
- [2] R. Wakikawa, J. T. Malinen, C. E. Perkins, A. Nilsson, A. J. Tuominen, "Global connectivity for IPv6 mobile ad hoc networks", Internet Draft, draft-wakikawa-manet-globalv6-02.txt, Nov. 2002 (Work in progress).
- [3] Y. Sun, E. M. Belding-Royer, C. E. Perkins, "Internet Connectivity for Ad hoc Mobile Networks", Int'l Journal of Wireless Information Networks special issue on "MANETs: Standards, Research, Applications", 9(2), pp75-88, Apr. 2002.
- [4] U. Jönsson, F. Alriksson, T. Larsson, P. Johansson, G. Maguire, Jr., "MIPMANET - Mobile IP for Mobile Ad Hoc Networks", Proceedings of the MobiHOC'00, pp 75-85, Aug. 2000.
- [5] J. Broch, D. Maltz, D. Johnson, "Supporting Hierarchy and Heterogeneous Interfaces in Multi-Hop Wireless Ad Hoc Networks", Proceedings of the 4th International Symposium on Parallel Architectures, Algorithms, and Networks, pp 370-375, June 1999.
- [6] C. Jelger, T. Noel, "Gateway and address autoconfiguration for IPv6 adhoc networks", Internet Draft, draft-jelger-manet-gateway-autoconf-v6-00.txt, Aug. 2003 (Work in progress).
- [7] K. Fall and K. Varadhan. ns Manual. The VINT Project. <http://www.isi.edu/nsnam/ns/doc/>.
- [8] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, J. Jetcheva, "A Performance Comparison of Multihop Wireless Ad Hoc Network Routing Protocols", Proceedings of MobiCOM'98, pp 85-97, Oct. 1998.
- [9] S. Deering, R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, Dec 1998.

UNIVERSAL QUADRATURE BROADCAST MODULATOR WITH SEVERAL TRANSPORT STREAM SOURCES

Marcin Nowakowski

ADB Polska, Sp. z o.o.

Ul. Trasa Północna 16

65-119 Zielona Góra

POLAND

M.Nowakowski@adbglobal.com

Introduction

There are known different kinds of systems for digital broadcast signals managing and modulation. These are mainly complex systems based on PC architecture with PCI expanded cards (e.g. modulation or demodulation cards). They can be used for commercial or test purposes at broadcast companies. Because of their complexity, they are very expensive and their usage for tests or development is debatable.

In order to prepare simpler and less expensive solution ADB company prepared a device which joins functionality of complex systems (demodulator, modulator, HDD, Ethernet port, SPI ports, central unit) in one small chassis and at a very low price. The device, called Stream Wizard, is as simple as it is possible, basing on two central units (processor with MPEG decoder and one FPGA chip) and some extra parts.

The presented solution is patent pending.

Stream Wizard

Stream Wizard is a universal quadrature modulator for broadcast purposes with several

transport stream sources (MPEG-2 specification). In one chassis there are placed: modulator (cable and terrestrial modulation techniques), digital demodulator (satellite or cable or terrestrial), Ethernet port, SPI ports, HDD, two main processing units. The device has the possibility to modulate broadcast transmissions basing on three transport stream sources (demodulator, HDD and SPI input port). It also has the possibility to provide transport stream for four outputs: digital modulator, SPI output port, HDD and an MPEG decoder.

A general diagram of Stream Wizard is shown in the Figure 1. As the first main processing unit the STi5514 processor is used, designed for Set Top Boxes. It includes the MPEG decoder which, as the input, may have several transport stream sources. In the device there are two inputs used: one from the demodulator and one from the FPGA chip. Additionally the processor includes several interfaces (ATA, memory interfaces).

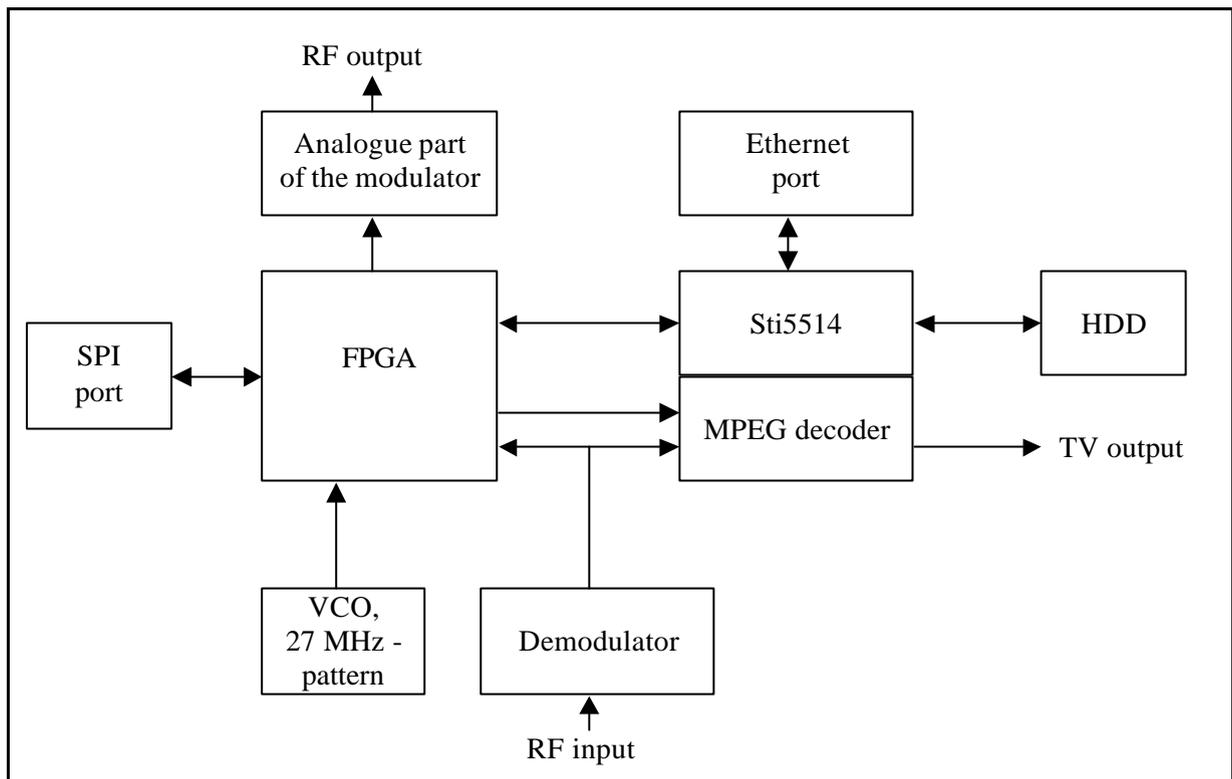


Figure 1. Stream Wizard – general device diagram.

FPGA unit

The second main processing unit is based on the FPGA chip. In case of ADB solution the XC2S600E chip was used. The FPGA chip provides two main tasks:

- realising digital part of requested modulation;
- managing transport streams provided from three sources and feeding a transport stream into four outputs.

So far ADB has designed modulations for Annex A, B, C of ITU J83 norm and COFDM standard terrestrial transmission. The digital part of the modulator, as the input, takes adequately prepared transport stream. Then the transport stream is modulated using standardised operations (for example the one described by DVB-C

specification) in digital modulator module and as the output Q and I signals are provided for the analogue part of the modulator.

Transport stream inputs and outputs are controlled by the manager of transport streams. The block diagram of the used FPGA chip functionality is shown in the Figure 2.

The transport stream manager consists of the following parts:

- Memory interface to transport stream that takes a transport stream from the memory interface and provides it in transport stream format from MPEG-2 specification.
- Input multiplexer that chooses the input transport stream to be used.

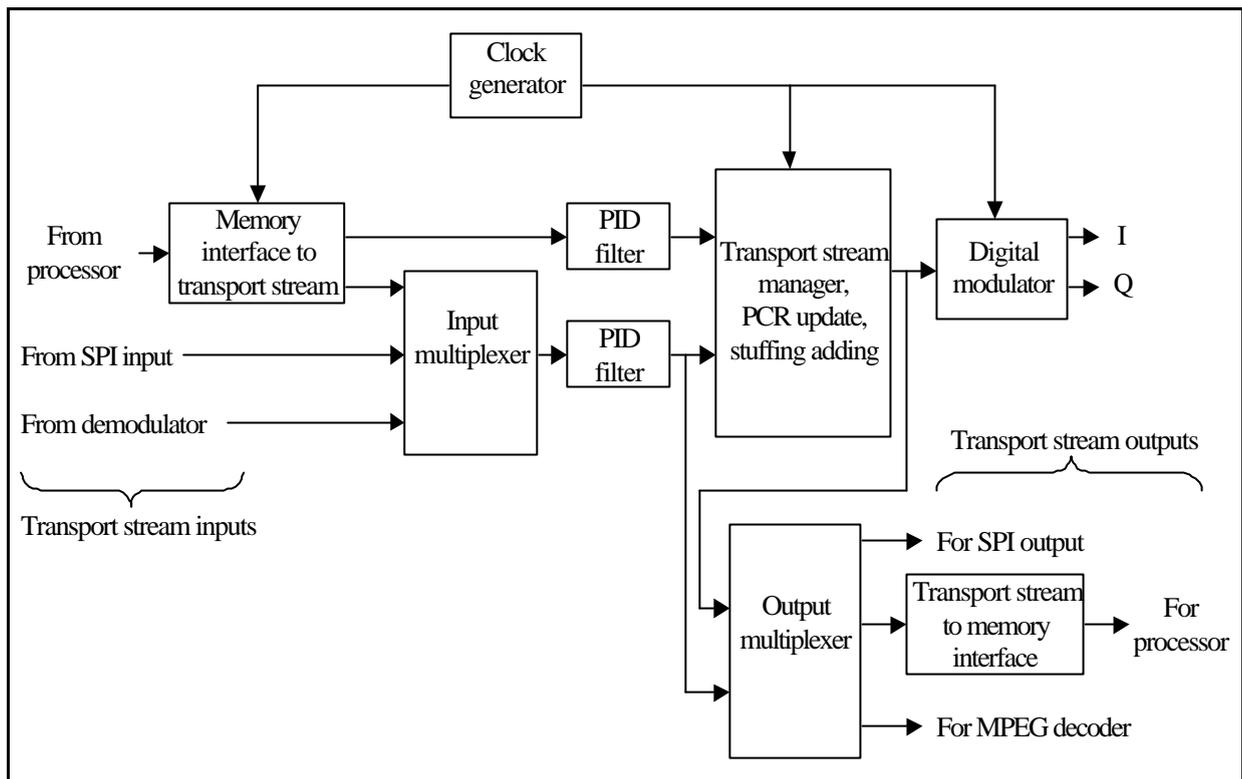


Figure 2. Stream Wizard – FPGA chip functionality.

- PID filters allow removing packets from the input transport streams, which is very useful to decrease bit-rate of used transport streams.
- Transport stream manager that may update PCR field, add stuffing packets to increase transport stream bit-rate or/and merge two input transport streams.
- Output multiplexer that chooses transport stream to be used on one of three external outputs.
- Transport stream to memory interface that provides transport stream to memory interface.

Clock generator

Clock generator is based on PLL solution and consists of two parts:

- VCO (Voltage Controlled Oscillator), which is placed on analogue part of the device (27 MHz crystal is used as the pattern);
- Digital part placed on FPGA chip (clock dividers and a phase comparator).

For proper modulation of higher QAM constellations it is required to provide very good parameters of generated clock source signal. This clock is next used as a frequency for I and Q symbols, which is of fourth factor of symbol rate frequency for the generated RF signals.

Analogue modulator

Analogue modulator consists of the following parts:

- Two DAC converters used to prepare I and Q analogue signals with anti-aliasing filters. The filters let to generate signals with 8 MHz bandwidth.
- Direct quadrature modulator which takes I and Q signals and modulates them with 1090 MHz IF frequency.
- Down-converter based on an analogue mixer and PLL generating signals in range of 1130 – 1940 MHz.

The analogue modulator output is 75 ohm RF signal in range of frequencies of 50 to 860 MHz and output power of 38 dBmV.

Additional features

The Stream Wizard has several additional features:

- Included HDD allowing storage of transport streams for two purposes: generating test patterns or recording transport streams received through the demodulator or SPI.
- Applied FTP server allowing copying of transport streams through the Ethernet port.
- HTTP server used for generating a web page for managing the device in a very user-friendly manner at any place.
- Programming FPGA chip through processor allowing the change of modulation techniques without hardware changes.

Usefulness

The usefulness of ADB Stream Wizard is very wide, especially for companies designing Set Top Boxes (as ADB).

Almost every team designing or working on their own digital modulation techniques finds very useful the presented solution, especially because of its modularity, flexibility and all these functionalities put together in one small chassis.

Using Mobile Agents for Data Collection, Data Correlation and Response in an Intrusion Detection and Response System

Nattapon Chatprechakul and Philip Nobles
Cranfield University

Royal Military College of Science
Shrivenham, Swindon, UK. SN6 8LA

{N.Chatprechakul, P.Nobles}@rmcs.cranfield.ac.uk

Abstract

Intrusion detection, the process of identifying and responding to malicious activity targeted at computing and network resources, despite being a subject for long research dated back to 1980, it is still not reliable as it should be. The main reasons are its failures to handle new unknown type of intrusions in case of misuse based and generating too many false positives in case of anomaly based. The lack of better understanding of overall network status has also contributed to its failure to handle intrusion detection effectively. In this paper, we propose a distributed mobile agent based intrusion detection and response system for limited bandwidth networks, such as wireless networks, that has the ability to detect and respond to intrusions by performing data mining and data correlation upon information gathered by mobile agents both in our system and from other proprietary security products. This approach is shown to provide to provide better coverage over the network and to provide a deeper understanding of network status.

1. Background

Security is always a major concern in the deployment of networking technology, especially a wireless network [1]. There are two approaches to improve network security; preventive and reactive. Preventive approach aims to prevent intrusions from occurring in the network. User authentication by means of logins and passwords, or the deployment of firewalls to filter network traffic are two examples of the preventive approach. Reactive approaches such as Intrusion Detection System (IDS), however, concentrate on how to detect intrusions and determine what should be an appropriate response to that intrusion.

1.1. Intrusion Detection System (IDS)

The intrusion detection concept for computer systems was first introduced in 1980 by Anderson [2] to complement available security mechanisms; however, this concept was not widely accepted until

Dorothy Denning published her paper on a model for intrusion detection [3] in 1987. Since that time, IDS has been a subject of much research. Various techniques have been applied to IDS in order to improve performance and reliability. IDS can be mainly categorised in two ways; by detection model or by operation.

There are two types of IDS if we categorise by how an IDS detects intrusions.

- Anomaly based IDS [5] – in this model, an IDS tries to define a normal operation profile. Action that differs from actions in this profile is considered suspicious. This model has the advantage that new unknown intrusions may be detected. Unfortunately, there is also the disadvantage that a high level of false-positives may be generated.
- Misuse based IDS – in this model, an IDS compares monitored information with predefined signatures. It detects known intrusions more accurately and has a lower level of false-positives. However, signatures need to be updated in order to be able to detect new types of intrusion.

There are three types of IDS if we categorise by how the IDS is operating.

- Host based IDS – an IDS analyses current system status and process running on that system and integrity of system files etc.
- Network based IDS – an IDS analyses every network packet transmitted over the network. The network adaptor must operate in “promiscuous” mode where all packets are captured and passed up from the network layer for analysis.
- Hybrid based IDS – this type of IDS combines advantages from both host based and network based approaches. For instance, the ability to work in an encrypted network or the ability to detect insider attacks from host based IDS whilst wider coverage from network based IDS.

Information processing methods used by the IDS are either centralised where all information gathered is sent to one location or distributed where information is processed at various locations in a network. The latter has been said to have better scalability [4].

There are many approaches to achieve distributed information processing. In traditional client/server

based architectures, information is passed to various locations for processing. While this approach distributes the computational load over various locations it does not, however, reduce network traffic. Another approach is to move information processing to the sources of information. This approach both distributes the computational load and reduces network traffic at the same time. Reducing network traffic is an important step towards improving IDS performance in an environment such as a wireless network. Research relating to a wireless network IDS has been published by Chirumanilla and Ramamurthy [14].

1.2. Mobile Agents

Mobile agent, one type of software agent, is a piece of software that can function continuously and autonomously in a particular environment [8]. It differs from mobile code in that it has extra state information, which enables it to suspend its operation at the current location, move to a new location and continue operation where it left off. Advantages of using the mobile agent approach over the traditional client/server approach are discussed in [9]. Applying mobile agents to IDS, we can overcome the limitations of the current generation of IDS. Hwang and Gangadharan had proposed an architecture [13] that utilises mobile agents in an intrusion detection and response system.

The security of the agent platform itself against attacks is beyond the scope of this paper. Fortunately, there is much research [10, 11] that proposes solutions for solving or working around these problems.

2. Our Contribution

In our work, we assume that however hard we try to secure our network there will always be some weak points waiting to be exploited. We therefore concentrate on the reactive approach by trying to detect an intrusion as soon as possible in order to limit damage to the integrity, confidentiality and availability of resources within the network, and to gain information about any new attack in order to prevent it from occurring again.

In this paper, we propose a distributed intrusion detection and response architecture that utilises the mobile agent paradigm. In a dynamic environment such as a wireless network where bandwidth is limited and connections are less reliable, compared to its wired counterpart, our prototype Distributed Intrusion Detection and Response System (DIDRS) can operate more efficient than a distributed IDS that use a client-server paradigm. By dividing intrusion detection, information gathering and intrusion response operations into smaller independent operations that can be put into different mobile agents and used when and where necessary, we can reduce the resources required by our DIDRS and thus reserve them for other tasks. Network bandwidth can

also be saved by moving information processing components to information sources and transmitting only processed information, rather than raw data, across the network.

In a diverse security network environment where different proprietary security products are being deployed that do not support information sharing standards, mobile agents in our DIDRS system also have advantages in information gathering and information correlation, providing better overall security for that network. The remainder of this paper is organised as follows. In section 3 we discuss the architecture of our mobile agent based DIDRS. In section 4 we discuss the details of using mobile agents for data mining and correlation on data generated by proprietary security products. In section 5 we discuss the implementation of our concept and discuss the prototype system we have developed, currently operational on a small test-bed network set up at our laboratory at Cranfield University (RMCS). In section 6 we discuss and conclude our work.

3. Mobile Agent Based Distributed Intrusion Detection and Response System

In a distributed system, information processing capabilities and functionalities are distributed into sub-regions or areas. To be able to coordinate between each sub-region a structure of processes must be established.

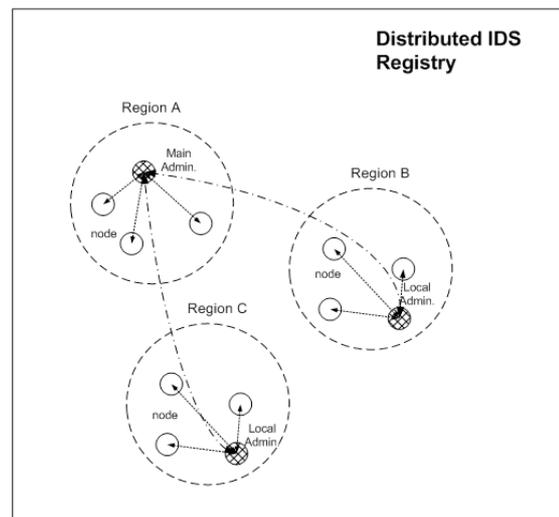


Figure 1 Hierarchical structures of distributed IDS

A hierarchical structure or a tree structure is well suited for distributed systems. This hierarchical structure reduces the amount of inter-region communication, which typically requires less bandwidth than local intra-region communication. Each region will have a centralised node via which other nodes in that region communicate. Inter-region communication takes place between the centralised nodes in each region. One region that has the

centralised root node that correlates and processes inter-region information. Figure 1 shows a distributed IDS organised in a hierarchical structure.

Functions of IDRS, such as information gathering, information processing and intrusion response, are implemented into mobile agents, which form the main building blocks of our DIDRS.

4. Data Mining and Data Correlation using Mobile Agents

Data mining is the process of identifying hidden patterns and relationships within data whilst data correlation is the process of evaluation and finding relationships between different sets of data. Different data mining methods have previously been applied to IDS [15, 16] but not for multiple heterogeneous data sources as described in this paper.

In our DIDRS, distributed decision making processes are arranged hierarchically from a root administrator node. Analysed information is communicated up the hierarchy using the message structure shown in Figure 2. Note that the raw data pertaining to a particular analysed result may be sent up since this may be necessary for higher level decision making processes to fully understand the analysed result in greater detail and/or to provide confidence in the analysis.

Action required	Analysis result	Raw data
-----------------	-----------------	----------

Figure 2 Hierarchical information structures

Another interesting aspect to be considered is the source of information. Our DIDRS system is capable of both gathering and processing raw information, however, in some situations, another network may be linked to the existing network. There may be a number of other proprietary security products, operating in that network, possibly in an attempt to provide security by diversity. Our system takes advantage of using information from those products as extra information sources for data mining and data correlation. Unfortunately, each individual product typically uses its own proprietary information structure. Even though there are efforts to develop standard information sharing processes for IDS data, namely the IDMEF [6] and CIDF [7] working groups, these efforts have gained little support, since each IDS vendor maintains a commercial advantage by preventing other products from using their information. Currently, correlating information from these products must be done manually by a human administrator.

Our proposed architecture solves this problem by automatically sending an appropriate correlation and data mining agent that is capable of handling each proprietary product to the corresponding node, that then extracts and reformats any useful information that it finds into a standard structure, such as IDMEF. Information in this standard format may then be used by our system.

5. Implementation

In this section, we discuss the implementation details of our system as it operates in a typical network and also in a situation where another network is subsequently linked to the existing network. We also discuss our working prototype system deployed in a test-bed network at Cranfield University (RMCS).

In our implementation, a mobile agent platform runs on every host to provide a suitable environment for mobile agents to perform their operations. Each mobile agent platform has a registration service that knows the location of other platforms and other agents within a Region from information stored within a Region Registry.

Figure 3 shows deployments of our DIDRS architecture for a typical network and an extended network. Referring to figure 3, network A and B form a typical network infrastructure layout that comprises both wired and wireless network nodes. Nodes, including gateway firewalls, are running the mobile agent system. There is one root administrator node (node A in Figure 3), which comprises a RM (registry monitoring) agent monitoring the system Registry, an ADM (administrator) agent performing centralised correlation and system administration, a RDB (rules database) agent providing local access to a database of misuse signatures, and a C-UI (control interface) agent that provides the human administrator with a GUI interface into the DIDRS.

When a new node connects to the network either wired or wirelessly (node G in Figure 3), the RM agent will notify the ADM agent that there is a new node connecting to the network. The ADM agent will dispatch an intrusion detection agent to that node. When an intrusion detection agent detects an intrusion (Node G in Figure 3), it will send analysed results back to the ADM agent for further analysis. The ADM will perform further analysis and respond appropriately by sending a ResA (response) agent to that node to enforce a response, for instance, updating a kernel level firewall to block any network connections from the intruder.

The system also correlates and analyses data gathered from external sources such as firewall audit logs and rules databases (node E and F in Figure 3) and any other proprietary IDS databases when there are other IDS running in the network. A C+D (correlation and data mining) agent retrieves and processes the relevant data on the relevant node.

To demonstrate how the DIDRS would operate in an extended network, in Figure 3, network C represents a network that is being deployed alongside existing networks. This network may already employ a number of different security products such as IDSes and firewalls (node K and L in Figure 3). Our C+D agent can travel across the network to these new nodes and perform data mining and correlation on these new information sources, thereby adding to the available information regarding potential new and previous intrusions.

Our prototype system test-bed consists of both wired and wireless node. There is one 3Com 8000 802.11b access point connected to a 10BaseT-wired network. A laptop computer wirelessly connects to the network and two desktop computers connect to the wired network. The computers run Linux and have the kernel integrated firewall – iptables - installed. Grasshopper [12] is used as the mobile agent platform. Grasshopper is a Java-based mobile agent platform, which requires a Java runtime environment (JRE). An SQL database, MySQL, runs on the root administrator node. The prototype is fully functional and is currently being used to validate the proposed DIDRS architecture and to measure its performance compared to existing architectures.

6. Conclusions

In this paper, we discuss different IDS approaches including their advantages and disadvantages. We have proposed a new distributed architecture using mobile agents for intrusion detection and response. Our proposed architecture overcomes the data mining and data correlating shortcomings that occur when different proprietary security products are deployed within the same network. The prototype system we have developed provides functionality beyond that of the current generation of IDS and is being used to develop more effective security methods for network environments.

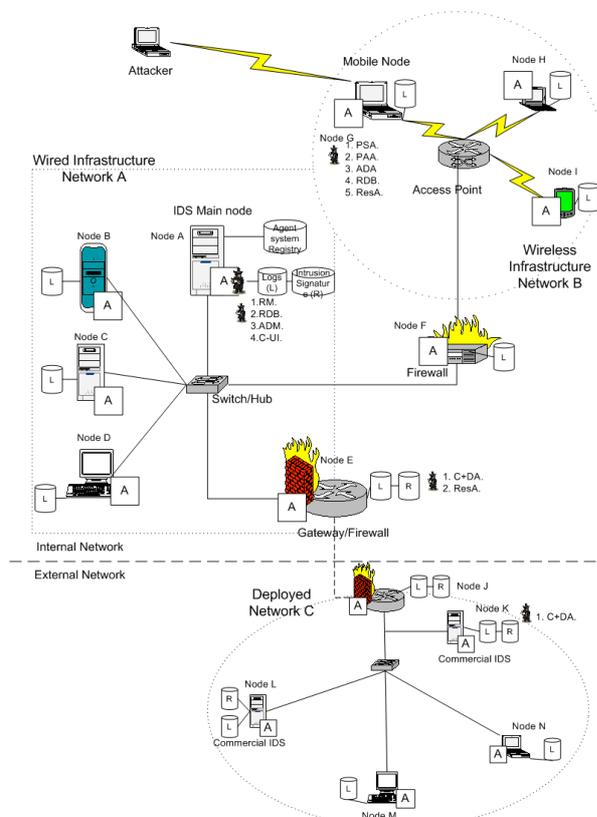


Figure 3 Deploying MADIDRS architecture in typical and tactical networks

7. References

- [1] B. Borisov, I. Goldberg, and D. Wagner, "Intercepting Mobile Communications: The Insecurity of 802.11", In Proceeding of the Seventh Annual International Conference on Mobile Computing and Networking, 2001, pp. 180-188.
- [2] J.P. Anderson, "Computer security threat monitoring and surveillance", Technical report, James P. Anderson Co., Box 42, Fort Washington, February 1980.
- [3] D.E. Denning, "An Intrusion-Detection Model", IEEE Transactions on Software Engineering, 13 (2), February 1987, pp. 222-232.
- [4] J.S. Balasubramanian, J.O. Garcia-Fernandez, D. Isacoff, E. Spafford, and D. Zamboni, "An architecture for intrusion detection using autonomous agents", Technical Report 98/05, Purdue University, 1998.
- [5] T.Y. Lin, "Anomaly Detection", In Proceeding of the New Security Paradigms Workshop, ACM SIGSAC, 1994, pp. 44-53.
- [6] Intrusion Detection Exchange Format (idwg) Charter website, <http://www.ietf.org/html.charters/idwg-charter.html> (4 October 2003).
- [7] S. Staniford-Chen, B. Tung, and D. Schnackenberg, "The Common Intrusion Detection Framework (CIDF)", Position paper accepted to the Information Survivability Workshop, Orlando FL, USA, October 1998.
- [8] D.B. Lange, "Mobile Objects and Mobile Agents: The Future of Distributed Computing?", In Proceeding of the European Conference on Object-Oriented Programming, 1998.
- [9] A. Bieszczad, B. Pagurek, and T. White, "Mobile Agents for Network Management", IEEE Communication Surveys, September 1998.
- [10] T. Sander and C. Tschudin, "Protecting mobile agents against malicious hosts" In G. Vinga (Ed.), Mobile Agents and Security, Springer-Verlag, Lecture Notes Computer Science No. 1419, 1998, pp. 44-60.
- [11] N. Karnik, "Security in Mobile Agent Systems", Ph.D. Dissertation, Department of Computer Science, University of Minnesota, October 1998.
- [12] Grasshopper 2 Agent Platform website, <http://www.grasshopper.de> (4 October 2003).
- [13] M. Gangadharan and K. Hwang, "Intranet Security with Micro-Firewalls and Mobile Agents for Proactive Intrusion Response", IEEE Int'l Conference on Computer Networks and Mobile Computing, Beijing, China, October 2001.
- [14] M.K. Chirumamilla and B. Ramamurthy, "Agent based intrusion detection and response system for wireless lans", IEEE Int'l Conference on Communications, 2003, Vol. 1, pp. 492-496.
- [15] W. Lee, S.J. Stolfo, and K.W. Mok, "A data mining framework for building intrusion detection models", In Proceedings of the 4th Int'l Conference on Knowledge Discovery and Data Mining, New York, August 1998.
- [16] K. Sequeira and M. Zaki, "ADMIT: Anomaly-based Data Mining for Intrusions", In Proceedings to the 8th ACM SIGKDD Int'l Conference on Knowledge Discovery and Data Mining, Cannada, July 2002.

Virtual Transmission Based MAC Protocol for Wireless Access

Tamás Radvánszki, Balázs Benkovics, Sándor Imre
Mobile Communications and Computing Laboratory
Budapest University of Technology and Economics
Email: sky@mcl.hu, {benkovics, imre}@hit.bme.hu

Abstract—The recently emerging possibility of ubiquitous wireless connectivity has increased the importance of the effective radio bandwidth usage. In the case of distributed multiple access MAC algorithms are needed that lower the probability of collisions. In our paper we focus on a new wireless medium access procedure that aims to increase the efficiency of the radio bandwidth usage. The presented VTBM (virtual transmission based MAC) procedure is based on the dynamic p-persistent algorithm and uses a virtual transmission mechanism. Simulation results presented in this paper demonstrate that the proposed algorithm outperforms the exponential backoff algorithm used in 802.11 MAC regarding throughput and collision rate.

I. INTRODUCTION

The evolution and wide-spreading of wireless communication devices during the past years made telephony available at anywhere regardless of being home or on the move. Recently the focus has moved towards data communications and wireless multimedia applications. One key issue in this field is the bandwidth of the wireless channel, which is a scarce resource thus it should be used as efficiently as possible. The broadcast nature of the wireless medium, however, poses a difficult problem for multiple access. To make a transmission successful a medium access control protocol is required to resolve access contentions among nodes and transform a broadcast wireless network into a logical point-to-point network. [1]

The common goal of the distinct MAC protocols is to let users transmit their packets successfully over the channel at the highest possible rate. That is, the protocols aim to avoid collisions that are regarded as inefficient usage. As in wireless data communications the collision detecting method would require a sophisticated and expensive radio instrument the collision avoiding method is used. [2]

The terminals having packets to transmit try to access the medium based on the rules defined by a MAC protocol. If a collision occurs the protocol attempts to resolve it according to some algorithms devised for that purpose. The most common methods in practice and in the literature are the exponential backoff, and the dynamic p-persistent algorithm [3], which is our algorithm built upon.

The rest of the paper is organized as follows. Section II gives a brief description of the generally used medium access schemes in MAC protocols. The effects of estimating transmission probabilities dynamically are also investigated. Section III introduces the novel VTBM algorithm and presents a detailed description of the protocol. The performance of the protocol is evaluated by simulations. The

simulation model, the test scenarios and the results can be found in Section IV. Finally, we conclude our paper in Section V.

II. THE IEEE 802.11 MAC

The 802.11 standard implements the CSMA/CA (Carrier Sense Multiple Access / Collision Avoidance) as medium access procedure. In this protocol, the nodes keep listening the wireless channel, and try to transmit if and only if it is found to be idle for a predefined time called DIFS (Distributed Inter Frame Space). The communication between the nodes is based on positive acknowledgement. That is, the receiver node must forward a receipt (acknowledgement packet, ACK) to the sender after each successfully received packet. This receipt indicates to the sender that there was no collision or data-loss in the radio channel and the packet was successfully transmitted.

If the acknowledgement is not sent or it is lost the sender node tries to resend its packet according to its scheduling algorithm. The 802.11 standard uses the wide-spread exponential backoff algorithm to schedule the resending process. [1]

The concept of the exponential backoff algorithm is to force the terminals to wait an exponentially increasing random time before trying to resend the collided packet.

When a collision occurs and the node has to repeat the last packet a random time is chosen from the interval $(0, CW_{min})$ where CW_{min} is the predefined initial size of the so-called contention window. This value is multiplicatively increased after each subsequent collision up to a maximum value (CW_{max}) . The exact time-slot in which the terminal will try to resend its packet is gained by choosing a random value according to a uniform distribution on the $[0, CW]$ interval.

Another commonly used algorithm in MAC protocols is the p-persistent or the more enhanced dynamic p-persistent algorithm. These algorithms use no contention windows but probabilities. Each time a terminal wants to send a packet a random variable is evaluated and the packet is transmitted only with probability p. This process is repeated in each time-slot as long as the terminal has packets to transmit. The dynamic p-persistent algorithm can change its p value anytime allowing the protocol to adapt to the changes in the network more efficient [4].

Terminals operating in a dynamic environment can only approximate the ideal parameters for communication irrespectively of the applied MAC algorithm. As adaptive techniques may have the opportunity to converge to the optimum it is not surprising why these algorithms outperform their static counterparts.

In what follows we apply the dynamic p-persistent model to investigate the slotted ALOHA system to point out the connection between transmission probability and the active population of the network and it is also shown how important to have a proper estimation for these parameters.

In the slotted ALOHA system the aggregated throughput can be calculated as follows

$$F(n, p) = np(1-p)^{(n-1)}, \quad (1)$$

where n is the number of stations ready to transmit and p is the probability at which the nodes access the medium. We remark that in this case all nodes use the same p value.

The optimal value at a given number of active terminals can be gained by calculating the maximum of $F(n,p)$ along the p variable

$$\frac{\partial F(n, p)}{\partial p} = 0 \Rightarrow p_{opt} = \frac{1}{n}. \quad (2)$$

Unfortunately, in case of collisions the affected terminals cannot get any feedback on the number of involved nodes. Therefore the collision resolving algorithms decrease the probability of transmission after each subsequent collision on a single packet. This process can be regarded as an estimation of the number of active stations that is a key point for performance optimization.

We investigated the impact of approximation error on the aggregated throughput of the system and the results for ten and one hundred nodes are shown in Figure 1.

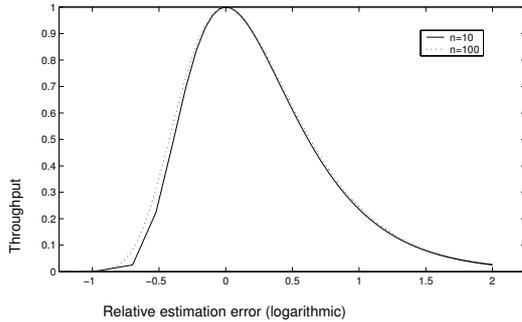


Figure 1 Effects of estimation error

The horizontal axis depicts the logarithm of the relative error ($\frac{n_{est}}{n}$) and the plotted throughput is normalized by the

theoretical optimum value that could be achieved if the error was zero. Note, that the two curves are close to each other leading us to the conclusion that the system reacts to estimation errors of the order of $\alpha(1)$ less sensitively. It can also be seen that under estimating the population has more serious effect on the throughput than over estimating it.

It can be seen that it is a difficult task to find the optimal values for the MAC algorithms. Moreover an improper estimation could result in either unutilized channel capacity or collisions during the medium access. In mobile networks the situation is even worse, as terminals keep changing their position and considering the relatively short range of the WLAN devices it can often happen that the population seen by a terminal is fluctuating. These factors render it more difficult to find an estimation algorithm that converges to the optimum.

Hereinafter, we propose an algorithm that aims to provide an eligible estimation for the transmission probabilities of the nodes that wish to communicate on the wireless channel at a given time.

III. THE VTBM ALGORITHM

Our VTBM (Virtual Transmission Based MAC) algorithm is based on the 802.11 MAC protocol but it uses the dynamic p-persistent method instead of the exponential backoff algorithm to resolve access contentions.

The concept of the algorithm builds upon a method called virtual transmission [5] or virtual thread that aims to provide more information for the dynamic p-persistent algorithm to obtain the optimal transmission probability.

A virtual transmission differs only from a "real" transmission of the dynamic p-persistent algorithm that the packet is not actually transmitted physically so it has no impact on throughput of the ongoing connections in the wireless channel.

The only purpose to perform such virtual transmissions is to gain additional information on the state of the medium: the throughput of these transmissions can be measured and based on this information the utilization of the channel can be estimated.

The estimation is carried out in each terminal by executing the following task:

1. virtual transmissions using different probabilities are performed simultaneously
2. the performance of these transmissions is evaluated periodically
3. the real transmission probability for the next period is computed
4. a new set of probabilities is determined for the virtual transmissions.

The performance of a virtual transmission is expressed by the utility factor that can be defined as the experienced throughput on the communication channel in a given T interval:

$$u_p^i = \frac{n^i}{T}, \quad (3)$$

where p is the transmission probability of the virtual transmission and n^i is the number of time-slots during the i^{th} T time-period in which packets have been sent over the channel successfully. It can be noticed that the virtual transmission probability is constant during the measurement period (T).

In order to calculate the utility factor the node using VTBM algorithm should know whether the last packet transfer was successful or not. In case of a real transmission the transfer is considered successful when the node gets the ACK receipt from the communicating party. In the virtual case, however, the situation is slightly complicated. The node wouldn't get any physical acknowledge as the packet is not actually transmitted therefore it can only draw conclusions from the state of the channel. After a virtual transmission is performed the node starts listening to the medium for SIFS time. If the channel became busy during this time it means that the virtual transmission would cause collision if it accessed the channel. Otherwise when the medium remains idle for SIFS long the

terminal considers the virtual transmission successful as if it were a real one the other nodes could sense it and would not interfere.

Since virtual transmissions require no packets to be sent over the wireless channel and the computation complexity of the utility factor is also very low the power consumption of the algorithm is neglectable. This may suggest using a fair amount of virtual threads to make the estimation as accurate as possible, however our simulations indicate that four threads are well enough and using any more of them has very little impact on the performance of the algorithm.

The transmission probability used for the real transmission is computed the following way

$$p_r^{i+1} = f(\underline{u}^i), \quad (4)$$

where p_r^{i+1} is the real transmission probability applied in the next period and \underline{u}^i is the vector of the utility factors measured in the i^{th} period. \underline{u} also includes the utility of the real transmission (u_0). The function $f()$ establishes the connection between the probability applied in future transmissions and the experienced utility factors during the last measurement. Equation 5 shows the form we used to gain the transmission probability

$$f(\underline{u}) = \frac{u_0 + \sum_{j=1}^{N_v} u_j^i}{1 + N_v}, \quad (5)$$

where N_v is the number of virtual threads, and u_0 is the utility factor of the real transmission as stated before.

This function is evaluated after each measurement period and the probability for real transmissions is determined.

Finally we have to provide a new set of probabilities (called virtual or pilot probabilities) for the virtual transmissions carried out in the next cycle of the operation. The elements of the set must be chosen carefully because these parameters can spoil the estimated transmission probability in the next iteration. It can be assumed that there are no rapid changes in the state of the network so the optimal transmission probability in the next iteration will be close to the previous one.

Figure 2 depicts the process of selecting probabilities. New pilot probabilities are picked according to a normally distributed random variable with mean value of p_r^{i+1} . Choosing this distribution ensures that the pilot probabilities will be in close proximity of p_r^{i+1} with a good chance. The deviance of the normal distribution can be expressed as

$$\sigma = \max\left\{u_0^i - u_k^i, \left|u_0^i - u_l^i\right|\right\}, \quad (6)$$

where u_k^i , u_l^i are the two highest utilities and u_0^i is the utility factor related to the transmission probability applied in the previous iteration.

Because of using normal distribution it can happen that the value of the random variable is outside the $[0, 1]$ range. To avoid this situation the density function of the applied distribution is truncated to the $[0, 1]$ interval.

As mentioned before the algorithm uses four virtual threads, thus four virtual probabilities. However, in order to eliminate

transient changes in the network state the algorithm replaces the two pilot probabilities only, which yielded the lowest utility and the others remain the same for the next measurement period.

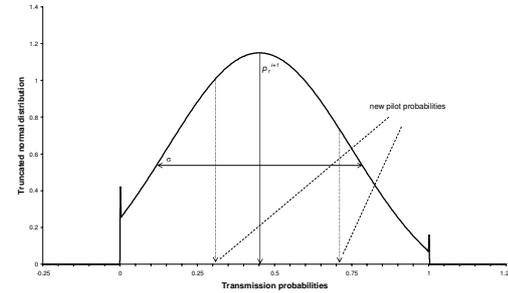


Figure 2 Picking candidate probabilities

The procedure described above produces the new set of pilot probabilities at the end of each measurement period. When a terminal is turned on, however, it has no information for the computation of the pilot probabilities and what is more important the transmission probability for real transmissions must be initialized, as well. As pointed out in Section II overestimating the network population has less impact on the performance the real transmission probability is close to zero. Theoretically, virtual probabilities can be any values between 0 and 1 but selecting values for typical network sizes make the convergence to the optimum faster.

IV. PERFORMANCE EVALUATION

The performance of the protocol is evaluated by running simulations in an event-driven simulation framework. The behavior of the protocol is investigated in a WLAN environment. The results of the simulation scenarios are compared to the ones obtained by the IEEE 802.11 MAC.

For our simulations we have taken the following considerations.

The communication medium is slotted, thus the nodes can access it at the beginning of the slots. The errors arising from signal-spreading are not taken into consideration, so there is no delay, fading and damping. The terminals use carrier sensing to decrease the chance of collisions. Both the DIFS and SIFS constants are one slot long. Our simulations scenarios aim to measure the MAC performance of the algorithms therefore all the terminals operate as an always on source. The packet size was 200 timeslot in all cases.

Table 1 summarizes the parameters used in the simulations for the exponential backoff and the VTBM algorithm, respectively (all units are measured in time-slots except for the probabilities).

Exponential backoff	Virtual probabilities (initial)	
CW_{min} 16 ts	p_r^0 0.05	p_3^0 0.1
CW_{max} 1024 ts	p_1^0 0.7	p_4^0 0.01
	p_2^0 0.2	
simulation time	200000 ts	
T (measurement period)	1500 ts	
L (packet size)	100 ts	

Table 1 Simulation parameters

On one hand our simulation scenarios focus on the performance of the protocol (throughput and collision rate) on the other hand we also investigate the fairness of the

algorithm, that is each station can send its packets and no starving is present in the system.

The throughput factor means the useful utilization of the wireless channel thus its investigation is very relevant question. The throughput is computed for each measurement period and it is interpreted as the rate of successfully utilized time-slots (in which packets have been sent successfully) in the given measurement window. Figure 3 shows the obtained throughput values as the function of the population size for the different MAC procedures.

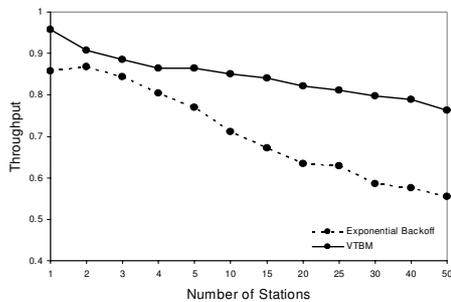


Figure 3 Throughput of the MAC algorithms

Note, that the VTBM protocol provides better efficiency than the exponential backoff in all cases that can be attributed to the estimation algorithm. As the exponential backoff algorithm uses a monotonically increasing contention window to delay transmissions when a collision occurs there can be numerous unutilized time-slots in the system even if only a few terminals access the channel. In contrast, the VTBM method does not change the transmission probability during the measurement period that can enable the collided packets to be resent in the consecutive time-slot.

The difference between the two collision resolution strategies can be studied in Figure 4 where the two algorithms are compared on the basis of generated collisions. Similarly to the throughput factor, the collision rate is the rate of the collided time-slots in a given measurement window.

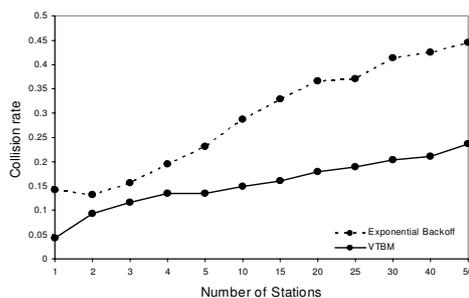


Figure 4 Generated collisions in the channel

It can be seen that the VTBM algorithm generates much less collision regardless of the number of active terminals in the system. The node, which executes the exponential backoff method can only adapt to the changes of the population by doubling of the contention window, while the VTBM algorithm uses eligible number of pilot probabilities enabling a much finer adaptation.

A medium access protocol is expected to ensure that each node can access the channel and not only a set of the active terminals have the opportunity to transmit via the medium. In this scenario we chose three of the nodes at random and investigated their throughput one by one at four different

network populations. The results are depicted in Figure 5. The average values in the figure are computed as the aggregate throughput achieved by a given population divided by the number of nodes. Apart from the fact that our algorithm provides high throughput it can be declared fair, as well.

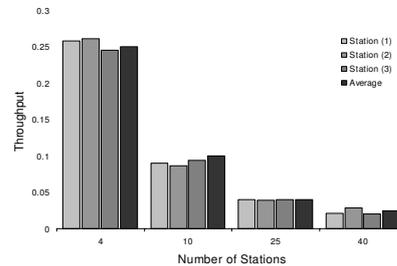


Figure 5 Fairness of the VTBM algorithm

V. CONCLUSION

In our paper we have focused on a novel approach to the problem of medium access control in wireless networks. It is shown that the transmission probability has a serious impact on the throughput and the optimal probability value depends on the number of the active stations. We pointed out the importance of accurate estimation for the number of active stations and found that the system tolerates the estimation error as long as it does not exceed the order of $O(1)$.

A novel algorithm is proposed that is based on the dynamic p-persistent method and it aims to estimate the transmission probabilities for the communicating nodes by using virtual transmission threads. We also presented our simulation results that indicate that the proposed algorithm can outperform the 802.11 regarding throughput. We also demonstrated the fairness of our algorithm.

In the near future we are planning to extend our analysis to the selection of pilot probabilities and we also want to evaluate scenarios with moving terminals to investigate the effects of mobility.

VI. ACKNOWLEDGEMENTS

This research is supported by OTKA F042590 and PHOENIX projects.

VII. REFERENCES

- [1] B. Bing: High-Speed Wireless ATM and LANs, Artech House, 2000.
- [2] A. C. V. Gummalla, and J. O. Limb: Wireless Medium Access Control Protocols, IEEE Communication Surveys, Second Quarter 2000.
- [3] D. P. Bertsekas and R. Gallager: Data Networks, 2nd Edition, Prentice Hall, 1992.
- [4] F. Cali, M. Conti, and E. Gregori: Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit, IEEE/ACM Trans. on Networking, 8(6):785-799, December 2000.
- [5] A. Veres, M. Barry, L-H. Sun, and A. T. Campbell: Supporting Service Differentiation in Wireless Packet Networks using Distributed Control, IEEE J. on Selected Areas in Comm., 19(10):2081-2093, Oct. 2001.

A design of softbots for intelligent, reactive intrusion detection systems

Islam M. Hegazy, Hossam M. Faheem, Taha Al-Arif
islheg@gawab.com
Computer Science Department,
Faculty of Computer and Information Sciences,
Ain Shams University

Abstract: Intelligent softbots are now used in several fields of computer science and artificial intelligence. Due to the arising of new attacks in the cyber world everyday new intrusion detection systems have to be built to cope with these new attacks. These intrusion detection systems should be reactive. Since intelligent agents are characterized by reactivity they can be used to build intrusion detection systems. In this paper, we will describe a model for intrusion detection system built with intelligent agents.

Keywords Agents, Intrusion detection, Security.

1. Introduction

Intelligent agents are a new paradigm for software development. An agent is a software application placed in a certain environment, and that is capable of autonomous action in this environment in order to meet its design objectives. Autonomy means that the agent should be able to act without the direct intervention of humans or other agents [1].

With the vast growth of information needs for organizations, a robust security system has to be applied to the organizations' computers to protect sensitive information. Threats to the information are not attributed to the outsiders only but to the insiders as well. Recent Computer Security Institute/Federal Bureau of Investigation Computer Crime and Security survey indicates that as much as 82% of losses were attributable to insider threats [2].

External network attacks can be categorized into IP spoofing attacks, packet sniffing, sequence number predication attacks and trust-access attacks. Categories of internal attacks include passwords attacks, session hijacking attacks, shared library attacks, and technological vulnerability attacks [3].

Computer network security can be categorized as follows [3]:

- *Security enhancement software:* thus replacing an operating system's built-in security software.

- *Authentication and encryption software:* encrypts and decrypts computer files.
- *Security monitoring software monitor:* monitors different operations of a computer network and outputs the results to system administrators.
- *Network monitoring software:* monitors user's behavior or monitors incoming and outgoing traffic.
- *Firewall software and hardware:* runs on the Internet/intranet entrance to a computer network, and checks all incoming network traffic for its contents at the network and transport layers of the OSI model.

Intrusion detection falls into the fourth category, Network monitoring software. Intrusion detection can be defined as the problem of identifying individuals who are using a computer system without authorization and those who have legitimate access to the system but are abusing their privileges [4]. Three techniques for illegal behavior detection are commonly used [5]: anomaly detection, rule based detection, and hybrid detection, a combination of anomaly detection and rule based detection.

Anomaly intrusion detection uses statistical methods to search for abnormal user behavior while the rule-based intrusion detection searches for attacks' signatures in the network traffic [6,7].

Intrusion detection systems (IDSs) can be classified into network-based IDS or host-based IDS. Network-based IDSs gather their information from the network traffic they are monitoring. Host-based IDSs gather their information from the log files of the hosts they reside on [8]. A special case of host-based IDSs is the application-based IDSs. Application-based IDSs analyze the events transpiring at a specific application [9].

During this paper, a modified Denial of service (DoS) agent will be introduced, a logging agent, an application based agent, and an update agent will be suggested. The following sections describe the function and the proposed architecture of each one of them.

This paper is organized as follows: section 2 explains the modified DoS agent, section 3 explains the logging agent, section 4 explains the application based agent, section 5 explains the update agent, and section 6 states the conclusion.

2. Denial of Service agent

The DoS attack is a type of attack where certain services are attacked to prevent them from legitimate users [6, 7]. DoS attacks constitute one of the major threats that make computer systems useless. Traditional techniques for implementing DoS modules are now obsolete since new attacks arise everyday. DoS modules should have some intelligence to cope with the upcoming attacks. We chose to implement DoS module using intelligent agents. Fig. 1 depicts the traditional DoS module and the DoS agent.

The DoS agent has to take into account the signatures of new attacks, detect automatically the signatures of repetitive attacks that are not listed in the signatures database and adds them. Table 1 shows the function of the DoS agent.

3. Logging agent

The logging agent writes out the log file of IDS. It receives two buffers, a buffer of packets and a buffer of suspected packets. It provides a code for every packet indicating whether or not it is suspected, the type of attack, and the attack severity. Table 2 explains the logging agent function.

4. Application log agent

Application log agent reads the log file written by the logging agent. This agent searches the log file for attacks codes, decodes them and writes the explanation of the attack in an attack description file. This agent acts as a simple application-based IDS such that it is specific to this system and do a simple mapping function. Table 3 explains application log agent function.

5. Update agent

The update agent is responsible for updating the signatures database. Every certain period of time the update agent connects to a certain server and downloads the new signatures database file. This database helps the analysis modules to recognize the new listed attacks.

6. Conclusion

IDSs are very complicated applications that require a lot of resources and much time. It should be as real time as possible. Using intelligent agents simplifies the building of the IDS because each

agent will be responsible for a specific part of the system. Using intelligent agents will give the ability to the system to learn new signatures. It is also a step toward a decentralized IDS where agents can be distributed over several machines.

Reference:

1. Amar Ramadan-Cherif, Samir Benarif, and Nicole Levy (2003). *An Adaptive Platform Based Multi-Agents for Architecting Dependability*, in proceedings of the third international conference in Intelligent Systems Design and Applications. Oklahoma, USA 10 – 13 August.
2. Paul E. Proctor. *The Practical Intrusion Detection Handbook*. Prentice Hall, New Jersey, 2001.
3. J. Pikoulas, W.J. Buchanan, M. Mannion, K. Triantafyllopoulos (2001). *An Agent-based Bayesian Forecasting Model for Enhanced Network Security*, in proceedings of eighth annual IEEE International Conference and Workshop On the Engineering of Computer Based Systems. Los Alamitos, CA, USA.
4. Jai Sunder Balasubramanian, Jose Omar Garcia-Fernandez, David Isacoff, Eugene Spafford, Diego Zamboni (1998) *An architecture for intrusion detection using autonomous agent*, in proceedings of the 14th IEEE Computer security applications conference. Scottsdale, Arizona.
5. J. Pikoulas, W. Buchanan, and M. Mannion (2002) *An Intelligent Agent Security Intrusion Systems*, in the proceedings of the ninth annual IEEE International Conference and Workshop On the Engineering of Computer Based Systems.
6. Stephen Northcutt, Judy Novak, Donald McLachlan (2001) *Network Intrusion Detection*. New Riders, USA.
7. Stephen Northcutt, Mark Cooper, Matt Fearnow, Karen Frederic (2001) *Intrusion Signatures and Analysis*. New Riders, USA.
8. Eugene Spafford and Diego Zamboni. *Data Collection Mechanisms for Intrusion Detection Systems*. CERIAS tech report 2000.
9. Rebecca Bace and Peter Mell. *Intrusion Detection Systems*. NIST special publication on intrusion detection system. <http://csrc.nist.gov/publications/nistpubs/800-31/sp800-31.pdf>. November 2001.
10. Islam M. Hegazy, Taha Al-Arif, Zaki. T. Fayed, and Hossam M. Faheem. *A Framework for Multiagent-based System for Intusion Detection*, in proceedings of the third international conference in Intelligent Systems Design and Applications. Oklahoma, USA 10 – 13 August, 2003.

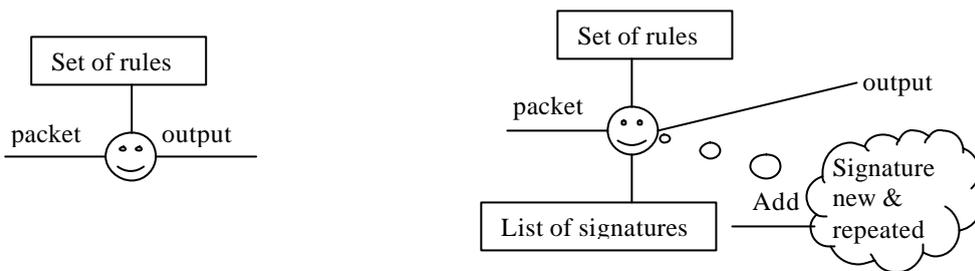


Fig. 1 Traditional DoS module and DoS agent

Table 1: DoS analysis agent algorithm

ALGORITHM DoS analysis agent

INPUT Buffer of packets //received from sniffing module

Buffer of signatures //read from the attacks signatures database

OUTPUT Buffer of packets and buffer of suspected packets

STEP 1 For i = 1 to number of packets in buffer

Check packet i for traditional attacks

If packet i is not suspected

Check packet i for listed signatures

If packet i is not suspected

Check packet i for misconfigurations

If packet i is misconfigured

Watch this misconfiguration signature

If misconfiguration is repeated from the same source

Add misconfiguration signature to the signatures buffer

End for

Table 2: Logging agent algorithm

ALGORITHM Logging agent

INPUT Buffer of packets

Buffers of suspected packets //received from the analysis or decision module

OUTPUT File of attacks codes

STEP 1 For i = 1 to number of packets in buffer

Code packet I

Write code of packet i in the file

If packet is suspected

Write its severity

End for

Table 3: Application log agent algorithm

ALGORITHM Logging agent

INPUT: Symbolic log file

OUTPUT: Attacks descriptive file.

STEP 1 While not end of log file

 Read symbol

 Decode symbol

 If it is an attack symbol

 Write attack description and its severity in the description file.

End for

A Low-Complexity Hybrid FIR/IIR Algorithm for Multi-Channel Line Echo Cancellation in VoIP Systems

Amjad Luna (luna@kics.edu.pk)
Al-Khwarizmi Institute of Computer Science

Imtinan Elahi
Texas Instruments

Muhammad Usman
Quartics

Abstract

This paper presents a low-complexity algorithm for multi-channel Line Echo Cancellation using an adaptive IIR filter. The scheme is based on a two-step procedure: first, an adaptive FIR filter is trained to model the echo impulse response, and second, the trained FIR filter is modeled by a stable IIR filter that replaces the FIR filter in the Echo Canceller. The order of an IIR filter needed to accurately model a typical hybrid impulse response is significantly lower than that of an FIR filter. Switching to the IIR model, therefore, yields substantial reductions in complexity. To ensure stability, adaptation in the IIR filter is restricted only to the zeros of the filter. The adaptation in both steps is based on the Least Mean Square (LMS) algorithm.

1 Introduction

The problem of Echo Cancellation is very well known in the Adaptive Filtering literature [1]. In the context of telephony applications, echoes arise out of an undesirable coupling between transmit and receive paths. The coupling is provided by the hybrid, which converts a two-wire signal into a four-wire signal and vice versa. The impedance mismatches in the hybrid cause a portion of the incoming signal to be reflected back thus causing an echo. The presence of echoes in the signal is annoying for the listener. A suitable echo-suppression mechanism is thus required for toll quality voice transmission. The most effective method of eliminating the echoes is to synthesize the echo from the signal that generates it, and subtract the synthesized echo from the signal that is corrupted by the echo. The most common method of synthesizing an echo is through the use of an adaptive Finite Impulse Response (FIR) filter that simulates the echo channel, thus generating a replica of the echo from the transmitted signal. Such a solution is shown in

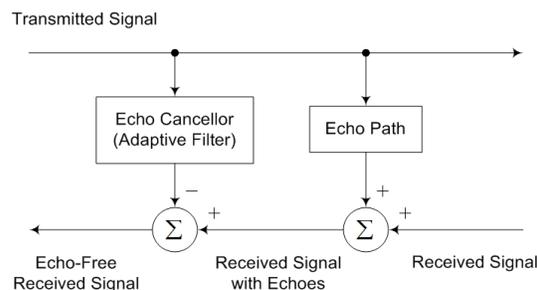


Figure 1: A Typical Echo Cancellation Setup

Figure 1.

In Figure 1, the block labeled “Echo Canceller” is an adaptive filter that simulates the behavior of the block labeled “Echo Path”, thus reproducing the echo. By subtracting this synthesized echo from the received signal, echo can be effectively eliminated. The LMS, or variants thereof [2], form the algorithms of choice for training the echo canceller filter. The LMS update forms the most computationally intensive part of a scheme such as the one shown in Figure 1. The computational complexity of the LMS algorithm is directly proportional to the order of the FIR filter. For Line Echo Cancellation, the Echo Path impulse response is typically specified to be 64 milliseconds in duration. For telephony applications the usual sampling rate of 8000 Hz would thus require a 512-tap FIR filter to cancel a 64-millisecond echo. The LMS algorithm would have to update all 512 taps of the adaptive filter every sample, i-e, every 125 microseconds. These 512 taps also need to be stored in the memory. Both these requirements—computations and the storage—impose serious limitations on the achievable channel density for a multi-channel Line Echo Canceller (LEC). Consider, for example, a 1024-channel LEC. Such a device would have to store $1024 \times 512 = 524288$ filter taps. A typical 16-bit resolution for the taps would thus require more

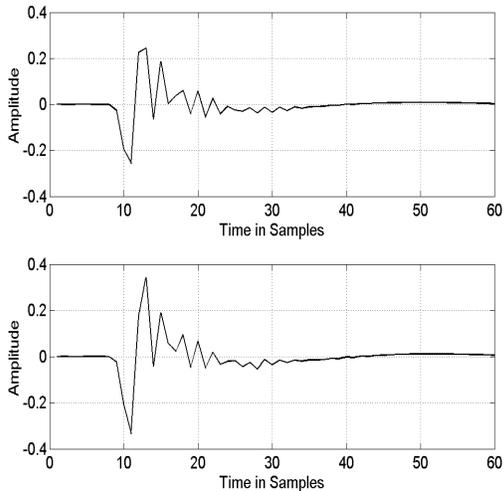


Figure 2: Two Snapshots of an Actual Hybrid Impulse Response Taken over a Period of 60-Minutes

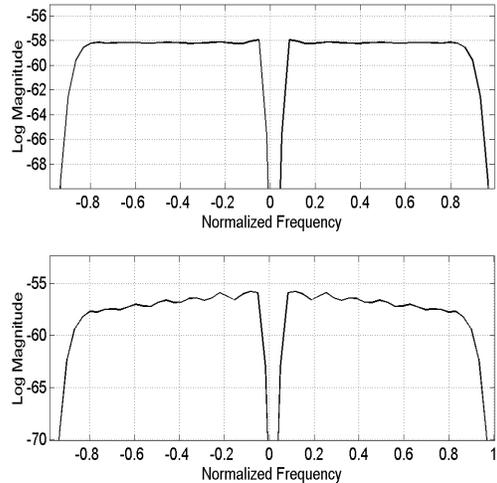


Figure 3: Time Variation in the Frequency Response of a Hybrid over a 60-Minute Interval 2

than eight Mbits of memory just to store the taps. To conserve on-chip storage, the filter taps for each channel can be loaded dynamically from an external memory. However, for the typical 10-msec processing-frame size, the required memory bandwidth of 800 Mbits/sec then becomes the bottleneck. A reduction in the number of filter taps would therefore reduce both the computational complexity and the memory requirements. Such a reduction translates directly into a proportional increase in the number of channels of echo cancellation that can be realized on a single-chip VoIP solution.

2 Adaptive IIR Filtering

The use of adaptive filters in practical applications has historically been restricted to the class of FIR filters. These filters have a decisive edge, in terms of stability, over their IIR counterparts for adaptive applications. However, given their superior modeling accuracy, IIR filters have the potential to significantly reduce the computational complexity as well as the memory bandwidth in an application such as multi-channel echo cancellation. However, the problem of ensuring the stability of the IIR filters during adaptation precludes their use in practical applications.

A typical hybrid response in the telephony network exhibits characteristically slow variations over time. This conjecture was validated by capturing several hybrid impulse responses and tracking their temporal variations. Figure 2 shows two snapshots of an ac-

tual hybrid impulse response, which were recorded in Irvine, California, over a duration of 60 minutes. Figure 3 shows the variation in the frequency response of this particular hybrid over the same interval of time. The key idea behind the proposed algorithm is the fact that the variations in the response, over a typical call-duration, are so small that they can be tracked effectively by adapting just the zeros of an IIR model.

The algorithm that we propose in the following uses a hybrid FIR/IIR approach to reduce the storage as well as the computational complexity required for Echo Cancellation.

2.1 Echo Cancellation Algorithm

1. Initialize the 512-tap FIR LEC filter $H(Z)$ with any suitable value.
2. Achieve convergence on $H(Z)$ using LMS algorithm. The converged FIR filter is denoted by $H_c(Z)$.
3. Identify the dispersive region in $H_c(Z)$.
4. Truncate $H_c(Z)$ obtained from Step 2, such that the truncated filter $H_T(Z)$ consists of only the dispersive region identified in Step 3.
5. Convert the truncated FIR filter $H_T(Z)$ obtained in Step 4 to a stable IIR model $N(Z)/D(Z)$ using the algorithm given in [4].

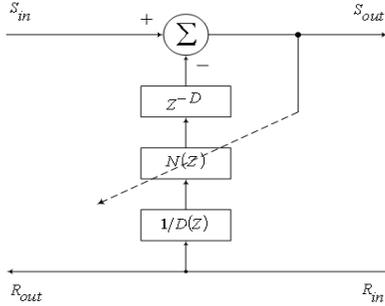


Figure 4: Adaptive IIR Filtering Setup

6. Filter the future voice samples using the IIR filter $N(Z)/D(Z)$ obtained in Step 5. The time variations in the system response are tracked by adapting the zeros of the IIR filter, while the poles are kept fixed.

Both Steps 2 and 6, use LMS algorithm for the update. However, the adaptation in Step 6 is restricted only to the zeros of the IIR filter. The typical hybrid response is slowly varying and the magnitude of these variations is extremely small. Keeping the poles fixed ensures the stability of the filter model, while adapting the zeros tracks the slow time variations in the hybrid response.

Figure 4 shows the Echo Cancellation setup in the steady state where D is a specific value of delay, in samples, that is needed to match the delay incurred due to conversion of the higher order FIR filter to a lower order IIR model.

3 Simulation Results

This section presents the results of applying the proposed algorithm to Line Echo Cancellation problem. We first demonstrate the validity of the basic conjecture behind the proposed technique; that the slow and small variations in a typical hybrid impulse response can be tracked by adapting only the zeros of the IIR model. We use the two impulse responses shown in Figure 2 to implement a step change in the hybrid model. Specifically, we train our FIR echo cancellor on the response shown in the upper part of Figure 2. Once the FIR echo cancellor is trained, we switch to an IIR model as outlined in Steps 3, 4, and 5 of Section 2.1. Shortly after switching to the IIR model we plug in the response shown in the lower part of Figure 2. This step change causes a sudden jump in the echo level as can be seen in the vicinity of sample number 100,000. However, as the IIR

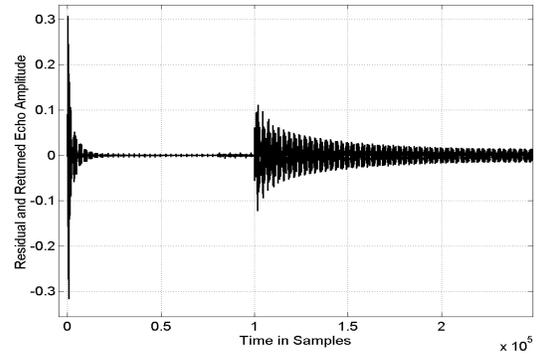


Figure 5: Adaptation Behavior of the IIR Model in Response to a Step Change in the Hybrid Model

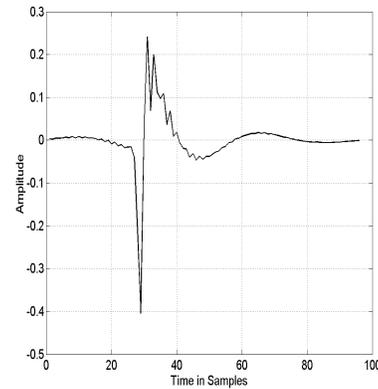


Figure 6: A Typical Hybrid Impulse Response as Specified in ITU-T Recommendation G.168

model starts to adapt its zeros, the echo level is seen to diminish quickly as shown in Figure 5.

The second part of the simulations tests the performance of the proposed IIR algorithm under G.168 Test 1, and again demonstrates the ability of the proposed algorithm to track temporal variations in the hybrid response. These variations are caused, primarily, by temperature changes and can thus be fairly accurately modeled as a Gaussian random process. We start the simulation with the response h_o , shown in Figure 6. This response is taken from the Annex C of ITU-T Recommendation G.168 [3] and represents a hybrid behavior that occurs most commonly on the North American telephony network. Shortly after switching to the IIR model, the response h_o is replaced by another response h_1 . The latter is obtained by adding white Gaussian noise to the former.

The response used in our simulation, as shown in Figure 6, has 96 taps. However, these 96 taps, which

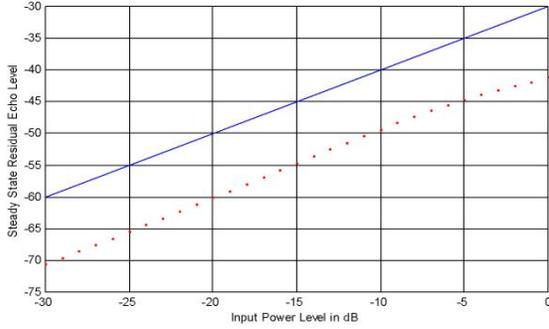


Figure 7: Steady-State Residual Echo Level

form the dispersive region of the echo tail, could occur anywhere in a 64-millisecond window. A typical Echo Canceller based on an adaptive FIR filter would thus take 512 taps to cancel the echo. Starting with a 512-tap FIR filter, $H(Z)$, in the proposed algorithm of Section II, we truncate the converged filter $H_c(Z)$ to a 96-tap filter $H_T(Z)$. The IIR model, $N(Z)/D(Z)$, obtained through the algorithm described in [4], uses 24 poles and 24 zeros.

The effectiveness of the proposed scheme is clear from Figure 7, which shows the results of the steady state residual and returned echo level test (Test No. 1, §3.4.2.1, ITU-T Recommendation G.168) for the hybrid response shown in Figure 6 to which a step change was applied as discussed earlier in this section. The solid line in Figure 7 represents the minimum acceptable echo levels in the steady state operation as a function of the input power level. The dotted curve shows the performance of the adaptive IIR solution. The use of IIR filter in this example case saves 48 taps in the filtering operation. However, the major part of the computational savings comes from the LMS update; instead of updating 96 taps of an FIR filter, the proposed scheme needs to update only 24 taps of the numerator polynomial in the IIR filter. There is a corresponding saving in the storage required as well; the proposed scheme needs storage for 48 IIR filter taps instead of 96 for the truncated FIR filter. The proposed Echo Cancellation algorithm thus results in substantial reductions, both in computational complexity and the storage required. Table 1 summarizes a comparison of the proposed scheme with a conventional FIR filtering approach. The proposed algorithm provides more than an order of magnitude reduction in the memory requirements and at least a factor of 14 reduction in the computational complexity.

Table 1: Storage (assuming 16-bit resolution) and Computational Complexity Analysis for the Proposed Algorithm.

		512-Tap FIR	96-Tap FIR	24th-Order IIR
Storage (in Bytes)		1024	192	96
Complexity	Filtering	⊗	512	96
		⊕	511	95
	LMS Update	⊗	512	96
		⊕	512	96
	Total	⊗	1024	192
		⊕	1023	191

⊗ ⇒ Multiplications; ⊕ ⇒ Additions

References

- [1] IEEE Signal Processing Magazine, “Adaptive Algorithms and Echo Cancellation,” Vol. 16, No. 4, July 1999.
- [2] S. Haykin, Adaptive Filter Theory, 2nd Edition, Prentice Hall, Englewood Cliffs, N. J.
- [3] ITU-T Recommendation G.168.
- [4] M. Usman, A. Luna, and I. Elahi, “A Stable and Adaptive Pole-Zero Filter,” *IEEE Transactions on Consumer Electronics*, Vol. 49, Issue 4, pp. 1279-1285, November 2003.
- [5] N. Al-Dhahir, A. H. Syed, and J. M. Cioffi, “Stable Pole-Zero Modeling of Long FIR Filters with Applications to MMSE-DFE,” *IEEE Transactions on Communications*, Vol. 45, No. 5, pp. 508-513, May 1997.

A minutiae extraction fingerprint algorithm

P. Chrisanthopoulos, Dr. S. S. Dlay, Dr. W. L. Woo
School of Electrical & Electronic & Computer Engineering.
University of Newcastle upon Tyne.
Merz Court, Newcastle upon Tyne, NE1 7RU
UNITED KINGDOM

Abstract

One of the most important tasks considering an automatic fingerprint recognition system is the minutiae biometric pattern extraction from the captured image of the fingerprint. Due to imperfections of the acquired image, in some cases certain minutiae can be missed by the extraction algorithm, and in others spurious minutiae can be inserted. We describe the design and implementation of an automatic fingerprint recognition system. The proposed method uses improving alternatives for the image enhancement process, leading consequently to an increase of the reliability in the minutiae extraction task.

1 Introduction.

Accurate automatic personal identification is critical in a wide range of application domains such as national ID card, electronic commerce, and automated banking.

IDs can be stolen, passwords can be forgotten or cracked. According to a UK poll one in three people write down their PIN number [1]. Security breaches resulting access to restricted areas of airports or power plants have caused terrorism. Although there are laws against false identification [2] incidents of intrusions and unauthorized modification to information/systems/ organizations occur daily with catastrophic effects. Credit card fraud is rapidly increasing causing bankruptcies [3], [4]. Biometrics incorporate authentication and identification technologies based upon unique biological characteristics and it refers to automated methods of identifying or authenticating the identity of a living person based on a physical or a behavioural characteristic. Inherent biological traits include voice, fingerprints, hand geometry, facial features, retinal patterns, etc.

Biometrics are inherently more reliable and more capable in differentiating between an authorised person and a fraudulent impostor than traditional methods such as passwords and PIN numbers.

Various forms of computer-based biometrics for personal authentication have been around for the past twenty years, but not until recently have some reached maturity and a quality/reliability that has enabled their widespread application. Among them, finger-scan is one of the most widely known and reliable techniques[5,6,7]. In finger scanning

the fingerprint is acquired, but instead of storing the full image, only data about specific points is stored for subsequent identification.

2 Algorithm implementation.

One of the most important tasks considering a fingerprint authentication/recognition system is the minutiae biometric pattern extraction from the captured image of the fingerprint. In some cases, the fingerprint image comes from an inked fingerprint, in other cases, the image is obtained directly scanning the fingerprint. Due to imperfections of the acquired image, in some cases certain minutiae can be missed by the extraction algorithm, and in other cases spurious minutiae can be inserted[8]. Image imperfections can also generate errors in determining the coordinates of each true minutiae and its relative orientation in the image. All these facts make remarkable decrease of the authentication/recognition system reliability.

The first step is to find the centre point of a fingerprint. Centre point location is done to find the point of most curvature by determining the normal of each fingerprint ridge, and then following them inwards towards the centre. The following is the procedure we used [8].

- Apply a pixel-wise adaptive 2-D Gaussian low pass filter to the fingerprint. The filter uses neighbourhoods of size 3 by 3 to estimate the local gradient mean and standard deviation. This will help in reducing any noise that may cause spurious results in the following gradient calculations.

- Divide the input fingerprint image into non-overlapping blocks of size 10x10.
- Determine the x and y magnitudes of the gradient at each pixel in each block, G_x and G_y . This is done by taking the average of the two neighbouring pixels.
- Apply the same 2-D Gaussian low pass filter on the x and y gradients as above to smooth out the gradients.
- With each block, compute the slope perpendicular to the local orientation of each block using the following formula

$$\Theta = \frac{1}{2} \tan^{-1} \left(\frac{\sum_{i=1}^{10} \sum_{j=1}^{10} 2G_x(i, j)G_y(i, j)}{\sum_{i=1}^{10} \sum_{j=1}^{10} (G_x^2(i, j) - G_y^2(i, j))} \right) + \frac{P}{2}$$

Formula 1: The mathematic formula used for the gradient.

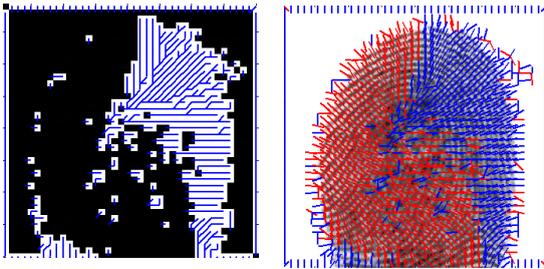


Figure 1: The normals of the gradients of the pixels of a fingerprint image.

The most salient property corresponding to ridges in a fingerprint image is that grey level values on ridges attain their local maxima along the normal directions of local ridges. In our algorithm two masks are convolved with the raw fingerprint image of size $W \times H$, which are capable of adaptively accentuating the local maximum grey level values along the normal direction of the local ridge directions. Finally this binary image is segmented and skeletonised in order to extract the minutia points. In the skeletonised images, each ridge line is of one pixel width.



Figure 2: The stages of the images in the pre-processing task.

3 Extraction of bifurcations and endings.

When we process the skeleton image and look for points that have 3 or more neighbouring pixels there tends to be a number of pixels clustered together near the desired intersection point that have more than 3 or more neighbouring pixels. Detecting these pixels is done by convolution. So that within this group the centre can be determined and the nearest pixel selected. This is to avoid multiple pixels selected at what is actually 1 intersection only. In our 8 connected neighbourhoods the horizontal and vertical connections come first and the diagonals come later. So by selecting an appropriate starting point like for example the top right corner of a border of a central pixel in a 3 x 3 neighbourhood, and by using the following sequence we can select the next pixel to be in a fixed order. So, for example, by starting at the top right corner, the next, in an anticlockwise order, pixel will be the left of it, bottom of it or the left-bottom pixel. Once a pixel is selected it is removed from the list of pixels that can be selected.

X-m	Y-m
X-m	Y
X-m	Y+m
X	Y+m
X	Y-m
X+m	Y-m
X+m	Y
X+m	Y+m

Table 1: The sequence used for the border coordinates

X-m, Y-m	X, Y-m	X+m, Y-m
X-m, Y	X, Y	X+m, Y
X-m, X+m	X, Y+m	X+m, Y+m

Table 2: a 3 x 3 neighbourhood and the pixel places

For a pixel not to be an endpoint it must have at least 2 connecting pixels and these pixels must be roughly opposite to each other. So it was obvious that we had to take a look at the continuity of the pixels in order to find the coordinates of the pixels that had to be looked at.

The, table 1, [2X8] matrix is a list of [X, Y] coordinates, and it denotes the coordinates of the border around a particular pixel to be tested. If for example, if the current pixel is at [X, Y] = [2, 2]

The border vectors are:

$$-1 -1 \quad \dots \gg 1, 1$$

```

0 -1  --->>2,1
1 -1  --->>3,1
1  0  --->>3,2
1  1  --->>3,3
0  1  --->>2,3
-1  1  --->>1,3
-1  0  --->>1,2

```

4 Results.

The algorithm takes total of 30 seconds to complete in an AMD 1.6GHz with 260MB RAM. It is absolutely successful in finding the ridge bifurcations and endings in a fingerprint image. Its outcome is 34 bifurcations and 47 ridge endings (this depends on the image).

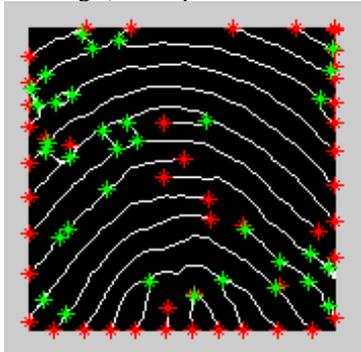


Figure 3: The centre of the fingerprint image, the red stars are the ridge endings & the green stars are the bifurcations.

Also it took 96 minutes to process 320 images and store the data.

The efficiency of the algorithm depends on the quality of the image, and the orientation of the fingerprint image in relation with the centre of the fingerprint acquisition mechanism. The algorithm's filtering process is based on extraction of the fingertip's background. Below two pictures of two different fingertips are shown and with that the background distribution of each picture. It can be seen that since the background for picture 1 is almost undistinguishable from the foreground the algorithm's reliability and accuracy falls drastically.

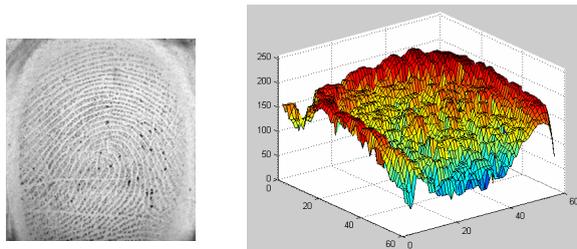


Figure 4: Bad fingerprint image & its background distribution

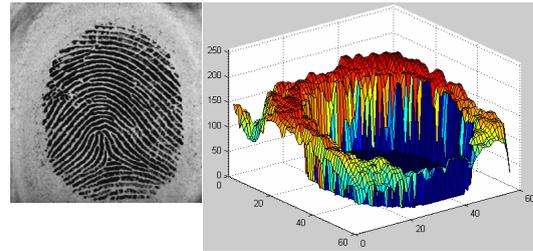


Figure 5: A good quality fingerprint picture & its background distribution

5 References.

- [1]. Newman, E. "The Biometric Report" SJB Services, UK, 1995.
- [2]. FACFI (1996) "The criminal of false identification" federal advisory committee on false identification, Washington, DC, 1976.
- [3]. Bright, R. "Smartcards: Principles, Practice, Applications" New York: Ellis Horwood, Ltd, 1988.
- [4]. Daugman, J. "High confidence visual recognition of persons by a test of statistical independence" IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 15, n. 11, Nov. 1993, pp.1148-1161
- [5]. Biometric Technology Today, "2001 Market Review Survey", volume 9, issue 1, January 2002.
- [6]. Federal Bureau of Investigation, "The science of fingerprints: classification and uses.", U.S. Government Printing Office, Washington, D.C., 1984
- [7]. H. C. Lee and R. E. Gaensslen, "Advances in Fingerprint Technology.", Elsevier, New York, 1991.
- [8]. Rao, A. Ravishankar, "A Taxonomy for Texture Description and Identification." Springer-Verlag, New York, 1990.

A New Image Compression Technique Based on Combining Feedforward Neural Network and Discrete Cosine Transform

Peter E. William^{*,#}, Tarik K. Abdel-Hamid^{*}, Magdy M. Doss^{*} and Hany Selim^{*}

^{*} Faculty of Engineering, Assiut University, Assiut, EGYPT.

[#] Student member, IEEE. (peter.ezzat@ieee.org)

Abstract: In this paper, an algorithm is proposed for image compression using One Hidden Layer Feedforward Neural Network (OHL-FNN). The algorithm combines OHL-FNN with Discrete Cosine Transform (DCT). The neural network performs the compression in a spectrum domain of DCT coefficients. Where by the OHL-FNN is only used to approximate the DCT coefficients representing the high detailed part of the image. Results, compared with baseline JPEG algorithm, demonstrate that the new algorithm dramatically increases compression for a given quality; conversely it increases image quality for a given compression ratio.

Index terms: Image compression, Discrete Cosine Transform (DCT), Feedforward Neural Network (FNN)

I. Introduction

DIGITAL image presentation requires a large amount of data and its transmission over communication channels is time consuming. To rectify and remedy this situation, a large number of techniques to compress the amount of data for representing a digital image have been developed to make its storage and transmission economical [1].

Neural networks are well suited to the problem of image compression due to their massively parallel and distributed architecture. Neural networks may break down some of the computational bottlenecks, which limit the performance of the serial machines [2].

Particularly, in the past decade numerous attempts have been made to pursue the possibility of using various neural networks for image compression. In [1] *Liying*, presented the application of OHL-FNNs to image compression. In [3] *Amerijckx*, presented a compression scheme based on vector quantization of the DCT coefficients by Kohonen map, differential coding by first order predictor and entropic coding of the differences.

The technique introduced in this paper, is based on dividing the image into blocks, obtaining DCT of each block. Then categorizing these blocks into low detail blocks (background) and high detail blocks (corners and edges), where each category is treated separately according to its importance in the representation of the image.

The main attraction of the proposed algorithm is that it can highly preserve the high frequency components for desired blocks while discarding the corresponding ones for background blocks, also categorization requires low computational effort.

This paper is organized as follows, section II gives an overview on OHL-FNN, section III discusses the basics of DCT, section IV introduces the proposed compression algorithm combining both OHL-FNN and DCT, section V demonstrates the experimental results performed on grayscale still images with comparison to baseline JPEG, and finally conclusion is given in section VI.

II. OHL-FNN

A neural network can be defined as “massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use” [4].

Since neural networks are trained using example data, they can be made to adapt to changes in the input data. Another advantage of training is that since data are presented individually, no overhead is required to store the entire training set. This is particularly important when processing very large data sets, of which images are an example [2].

When using the Multilayer Perceptron Feedforward Neural Network (MLP-FNN) in image compression, the hidden layer in the middle of the network has few numbers of nodes than the input layer. The output of this layer associated to all image blocks is considered as the compressed image or transmitted image to the receiver. In order to reconstruct the image at the receiver end, the weights for those connections on the right-hand-side (RHS) of the middle layer are also transmitted to the receiver [see Fig. 1 for details]. The fewer the number of units in the middle layer, the higher the degree of compression. The OHL-FNN considered here is symmetric in structure, with the network from input layer to the hidden layer acting as data compressor, while the network from the hidden layer to the output layer playing the role of a reconstructor. These two sub-networks have the same number of connections.

The idea of using only one hidden layer feedforward neural network leads to a faster training convergence at the expense of lower quality of the reconstructed

image for the same compression ratio as in the multi-hidden-layer FNN architecture [1].

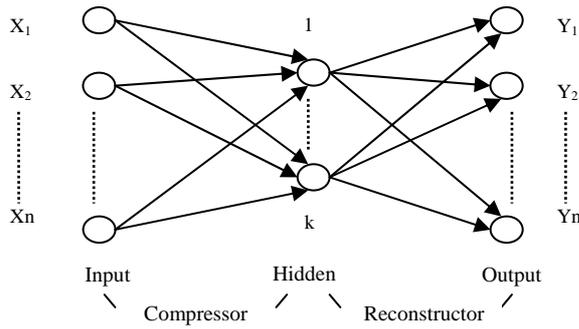


Fig.1 OHL-FNN for image compression, where ideally $n > k$

III. DCT

The DCT of an image is a transform that maps a block of pixel values in the spatial domain to values in the frequency domain. The DCT basis is image independent, this is an important issue in compression, since an image-dependent basis implies that additional computations need to be performed to determine the basis [5].

The DCT can operate mathematically in any dimension; however, an image is a two-dimensional (2-D) surface so the 2-D DCT transform is used.

The DCT is more efficient on smaller images. When the DCT is applied to large images, the rounding effects result in deteriorated image quality. In addition, as the size of the image is increased, the number of computations increases disproportionately. For these reasons, an image is subdivided into $N \times N$ blocks.

The 2-D DCT is applied to each block so that an $N \times N$ matrix of DCT coefficients is produced for each block. This is termed the DCT matrix. The top left component of the DCT matrix is termed the DC coefficient and can be interpreted as the component responsible for the average background colour of the block. The remaining components of the DCT matrix are termed the AC components. Usually the DC coefficient is much higher in magnitude than the AC components in the DCT matrix. Higher frequency components represented by higher order DCT coefficients are less visible to the human eye, and it is the property that is exploited in JPEG and in the proposed algorithm as these higher frequency components can be attenuated or removed with little noticeable effect on the quality of the image. The original image block is recovered from the DCT coefficients by applying the Inverse DCT (IDCT) [6].

In every block of an image, energy can be considered as being distributed over all pixels. The DCT is a sub-optimal orthogonal transform, which highly decorrelates the data in a block. That is, it packs most of the energy in few numbers of transformed coefficients [7].

One of the major drawbacks of the block based DCT compression methods is that they may result in visible artifacts at block boundaries due to coarse quantization of the coefficients [8].

IV. Image Compression Using OHL-FNN with DCT

The idea upon which this technique is based depends on classifying image blocks into two categories, one represents high detail blocks for edges and corners and the other represents low detail blocks for smooth regions.

All blocks that represent smooth parts are represented using only the DC coefficient and discarding all other AC coefficients. The remaining blocks that represent the high detailed parts of the image are treated in a different manner, in which the DC coefficients of these blocks are concatenated with those representing low detailed ones, and the remaining AC coefficients are compressed using OHL-FNN.

The problems that face this methodology can be stated as follows:

- How to find a suitable way to classify the image into these two categories.
- The size of overhead data needed to reclassify image blocks at the receiver should be held minimum.
- Finding the proper lossless coding techniques for each category.

The core computation pipeline employed in most of the lossy image compression standards based on DCT is discarding some of the DCT coefficients, referred to as the quantization process. The quantization process is irreversible and is the main source of loss in the baseline JPEG compression technique.

Illustration of the quantization process is as follows:

$$Y = \text{DCT} (X) \quad (1)$$

Where X is an $N \times N$ block represented in pixel values, Y is the $N \times N$ DCT coefficients of the X matrix. At this point, no compression has been achieved. The process of quantization of the elements of $Y (y_{kl})$ is expressed as

$$z_{kl} = \text{round} (y_{kl} / q_{kl}) \quad k, l = 0, 1, 2 \dots N \quad (2)$$

Where q_{kl} denotes kl^{th} element of an $N \times N$ quantization matrix Q . the process of quantization results in the zeroing out of many of the resulting DCT coefficients. The specific design of Q depends on psychovisual characteristics and compression ratio considerations. All compression standards provide default values for Q , Vasudev in [5] describes some of the techniques used to develop quantization tables.

An approximation of X can be obtained by performing the inverse quantization on Z then computing $N \times N$ IDCT.

Even if the quantized coefficients are not discarded, there will be an error between the original pixel values and the reconstructed ones because of the rounding action in (2). However, there is a great benefit obtained from the quantization process in the zeroing in a number of the DCT coefficients, where by the fewer the details in a block the larger the number of zeros we get.

In the proposed technique, quantization process is used only for the classification of blocks. Quantized blocks with higher number of zeros are considered low detail blocks. On the other hand, the coded coefficients needed for image reconstruction are the DCT coefficients before quantization.

The order of classified blocks needs to be sent for proper image reconstruction.

It was concluded in [9], that neural networks can operate more efficiently when both their inputs and outputs are limited to a range of [0,1]. Since usually the value of the DC coefficient is much larger than other AC coefficients and normalization of the whole coefficients will not improve the network performance, the AC coefficients of the high detail blocks are used only to train the neural network while the DC coefficients are sent separately.

All the compressed data from the transmitter is coded before transmission using a suitable form of encoding as Run Length Coding (RLC) and Variable Length Coding (VLC), both are widely used techniques for lossless data compression

V. Experimental Results

The image is first divided into blocks each is 8 x 8. The DCT transform is applied on each block, obtaining 64 coefficients. Arranging these coefficients in a zig-zag order starting from DC coefficient towards higher frequency coefficients, makes classification easy because higher frequency components are less visible to the human eye, and then lining up the coefficients in one column suitable for training the neural network.

For the OHL-FNN, the training data is applied to both the input layer and the output layer. Setting the transfer function of the hidden layer to log sigmoid and that of the output layer to linear. Each layer (hidden & output) has biases and weights connected between it and the preceding layer. The hidden layer contains only four neurons.

The OHL-FNN is trained in a supervised manner with a highly popular algorithm known as the error back propagation algorithm; this algorithm is based on the error correction-learning rule.

All the DC coefficients are differentially coded using Arithmetic Coding. In addition, the output of the neural network taken from the hidden layer representing the compressed AC coefficients of the high detail blocks is coded using Arithmetic Coding.

These coded data are associated with the biases and weights of the OHL-FNN that is required in the receiver for reconstruction.

Fig. 2 indicates the contribution of the OHL-FNN. The solid curve shows the Peak Signal to Noise Ratio (PSNR) versus the Compression Ratio (CR) for Lena image for the proposed technique, where the 63 AC coefficients are reduced to the four values taken from the OHL-FNN and arithmetically coded. On the other hand, the dotted curve depicts the same relation when the AC coefficients of the high detail blocks are only compressed by arithmetic coding.

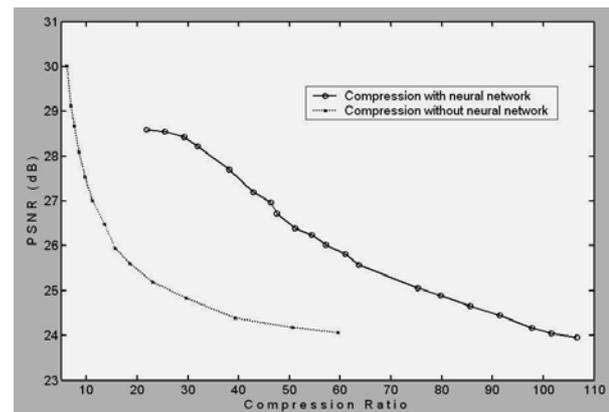


Fig. 2 Comparison between the proposed scheme for Lena image in case of usage and absence of OHL-FNN.

The proposed technique results are compared with those obtained with the JPEG standard; which represents the most lossy image compression method used on the internet where JPEG's good compression performance, low computational and memory complexity make it an attractive method for natural image compression [10].

Experiments are performed on different grayscale reference images like Lena, Plane, Mamo, and Tank. Similar results are obtained as with Lena image of size 512 x 512 with 8 bits/pixel. For relative performance of the proposed compression technique to JPEG, the PSNR versus compression ratio for Lena image is shown in Fig. 3.

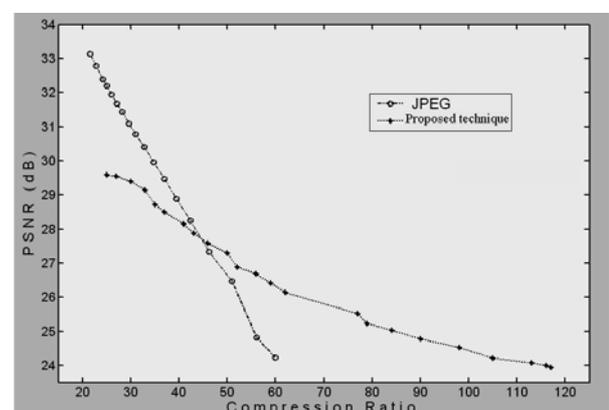


Fig. 3 Comparison between the proposed scheme and the JPEG algorithm for Lena image.

It is to be noted that it is easy and efficient to trade off between the compression ratio and the quality of reconstructed image using the OHL-FNN. This is because the PSNR of the constructive OHL-FNN depends closely on the number of hidden units used in the active network, which allows the user to decide on a trade-off as the network training evolves.

The reconstructed and the generalized images are shown in Fig.4.



Fig. 4 Lena image (a) original (b) Proposed technique, CR = 90. (c) JPEG, CR = 60. (d) Proposed technique, CR = 60. (e) JPEG, CR = 40. (f) Proposed technique, CR = 40.

Numerical results concerning Lena image are shown in table 1. They demonstrate both compression ratio and PSNR for each of the proposed technique and baseline JPEG.

In comparison with baseline JPEG, this technique gives results better than JPEG compression algorithm for compression ratios exceeding 40:1, these ratios may be needed in a number of applications whenever small image size is preferred while still concerning image quality. From the sample images shown above, it can be deduced that, although PSNR at compression ratio 40:1 is slightly better in case of baseline JPEG than the proposed scheme, it is obvious that the proposed scheme image has less visual error than the corresponding JPEG.

It is also shown that large compression ratio 117:1 is still possible with the new method while, retaining

reasonable image quality better than baseline JPEG at lower compression.

Table 1 Results obtained for Lena image

Compression ratio	PSNR (dB) <i>JPEG</i>	PSNR (dB) <i>Proposed Technique</i>
117	Not reached	23.93
90	Not reached	24.77
60	24.24	26.41
46	27.31	27.59
40	28.67	28.25
33	30.36	29.14
25	32.19	29.59

VI. Conclusion

An image compression algorithm is presented that takes advantage of OHL-FNN. The algorithm exploits the trend of the DCT coefficients after the image has been transformed from the spatial to the frequency domain. Combining DCT with OHL-FNN in image compression produces challenging results, which can be used in some applications concerning high compression ratios.

Further work could include applying 2-D OHL-FNN to the DCT coefficient matrix. This would bypass the need for the Zig-Zag transformation to produce 1-D set of DCT coefficients. Further work is also required to explore the results of applying adaptive OHL-FNN to the DCT coefficients, i.e., when the topology of the neural network is not fixed before training.

References

- [1] Liying Ma, "Application of adaptive constructive neural network to image compression", IEEE Trans. Neural Network, vol. 13, pp. 1112-1126 Sept. 2002.
- [2] Robert D. Dony, "Neural network approaches to image compression", proceedings of the IEEE, Vol. 83, No. 2, February 1995.
- [3] Christophe Amerijckx, "Image compression by Self Organized Kohonen Map", IEEE Trans. Neural Networks, vol. 9, pp. 503-507 May 1998.
- [4] S.Haykin, Neural Networks: A Comprehensive Foundation. New York: Macmillan, 1999.
- [5] Vasudev Bhaskaran, Image and Video compression standards, USA: Kluwer Academic publishers, 1997.
- [6] Jonathan Robinson, "Combining support vector machine learning with the discrete cosine transform in image compression", IEEE Trans. Neural Networks, vol. 14, pp. 950-958 July 2003.
- [7] Graham Wade, Coding Techniques, Great Britain: Creative Print and Design (Wales), 2000.
- [8] Ying Luo, "Removing the blocking artifacts of block based DCT compression Images", IEEE Trans. Image processing, vol.12, pp.838-842 July 2003.
- [9] J.Jiang, "Image compression with neural networks – A survey", Signal Processing: Image communication 14 (1999) 737-76.
- [10] Alfred M. Bruckstein, "Down-scaling for better transform coefficient", IEEE Trans. Image processing, vol. 12, pp. 1132-1144 September 2003.

A New Method for Linear-Phase FIR Filter Designing with Desired Shape

K. Mohamed-pour, M. Saadatmand-T,
kmpour@kntu.ac.ir m_saadatmand_tarzian@yahoo.com,

K.N.Toosi University of Technology,

P.O.Box: 16315-1355, Tehran, Iran

Abstract

This paper introduces a new method for designing linear-phase FIR filters with a desired shape. The proposed algorithm adjusts desired filter frequency response based on least squared (LS) method. It can design full-shaped filters with specified deviation in the entire the frequency band. Also, the transition bands can be considered as short as possible without any difficulty although it may not be necessary. The simulation results show that the proposed algorithm can successfully adjust and regulate the deviations of LS design methods for filters.

Key Words

Linear-Phase, FIR, Desired-Shape Filter, Recursive method

1. Introduction

This paper aims to introduce a simple recursive method for designing desired shape FIR filters. The proposed algorithm which is based on LS method, updates the coefficients of predefined filter FIR filter to close its specification as much as possible to desired one. The main criteria in this approach are the frequency response which has to be the same for design and desired filters. Therefore, it is clear that the transition bands between the frequencies are not considered. This is because that the transition bands are not really a demand of designing. Furthermore, the consideration of the transition bands [1,2] are usually introduced to reduce the Gibbs effect for the least square approximation or to use Chebyshev approximation. For example, a low pass filter normally may design in such a way to convey the pass, transition and stop bands respectively. In most practical cases there is no separation between the passband and stopband for the transition band. In other words, the spectrum of the desired and undesired signals often overlap and it is hard to specify a point that separates the pass and stop bands and certainly it is impossible to state a band for separating them. In this way, the transition band is introduced to reduce or remove the oscillations in the frequency response near the band edges caused by Gibbs effect. It should be noted that when there are large peaks in the transition band of filters such as Chebyshev filters, we must care and alter the specifications so that the peaks are eliminated [3].

In addition to transition band problem, to design a FIR filter with a special shape, some methods are introduced [4]. Some of these filter design algorithms are a kind of optimization, in which the deviation between the desired and designed filters are minimized [5,6,7]. Some other design algorithms are based on genetic algorithms [8,9]. Both these methods are computationally expensive. In this paper a simple effective recursive algorithm is proposed for desired shape FIR filters and predefined deviation in frequency band. The simulation results show that the algorithm is very flexible for obtaining the main characteristics of the desired filter such as transition bands, and less complexity at the same time.

The paper is organized as follows: first the designing of the linear-phase FIR filters using LS method is introduced. The proposed algorithm is described in section 3. The section 4 explains some considerations to reduce the computational complexity of proposed algorithm, and finally, the simulation results are presented and concluded in section 5.

2. Linear-phase Filter design with LS method

A linear-phase filter with impulse response $h[n]$, may have even symmetric as

$$h[n] = h[-n] \quad (1)$$

Where, its corresponding frequency response can be interpreted by

$$A(e^{j\omega}) = h[0] + \sum_{n=1}^L 2h[n]\cos(\omega n) \quad (2)$$

It is assumed that the values of the $A(e^{j\omega})$ at $0 \leq \omega_i \leq 2\pi$, which are called *designing frequencies*, are known, and denoted by $H(\omega_i)$. The problem is to compute $h[n]$ coefficients optimally. To minimize the error of designing according to LS method, this can be done by setting to zero the derivatives of the following objective function with respect to $h[n]$ [11].

$$E = \frac{1}{2} \sum_{i=1}^N [H(\omega_i) - A(e^{j\omega_i})]^2 = \frac{1}{2} \sum_{i=1}^N \left[h[0] + \sum_{n=1}^L 2h[n] \cos(\omega_i n) - H_i \right]^2 \quad (3)$$

$H_i = H(\omega_i)$

Hereafter, we call $H(\omega)$ "predefined filter". In continuous frequency domain the above function is reduced to

$$E = \frac{1}{2} \int_0^{2\pi} [H(\omega) - A(e^{j\omega})]^2 d\omega \quad (4)$$

Although the LS method is very simple and powerful, and the least square is the main criteria, other designing requirements can not be taken into account. One of the most important requirements is deviation limitation. It can be seen in the literature that many methods are proposed to overcome this drawback [5-7], which in turn the flexibility of the methods may be bst. The fundamental factors which determine the complexity and affect the results of the LS method are objective function, designing frequencies, and predefined filter. The method is usually built on changes of the objective function although the adjustment of the predefined filter may also be considered.

3. New proposed method

The new proposed algorithm for designing any FIR filters with desired shape is an iterative algorithm based on adjustment of predefined filter. At each step of the algorithm, with characteristics of the current predefined filter, the design is accomplished. Then, the error between the desired and design filters is computed as an error signal. Based on this error signal, the predefined filter specifications for the next step are updated. This process is continued; until the error signal tends to zero. Therefore, the filter coefficients will be optimized.

Let us define the frequency response of the desired filter as $d_0(f)$, and $b_n(f)$ and $d_n(f)$ denote the designed and predefined filters at n-th step, which are shown in Fig.1. At 0-th step the desired filter $d_0(f)$ is the same as predefined filter. With LS method $b_0(f)$ is obtained which the error can be defined as

$$e_n(f) = \text{sgn}(d_n(f) - b_n(f)) \times \max(|d_n(f) - b_n(f)| - \mathbf{d}(f), 0) \quad (5)$$

The acceptable deviation, $\mathbf{d}(f)$, is defined the difference between the obtained designed and desired filters,

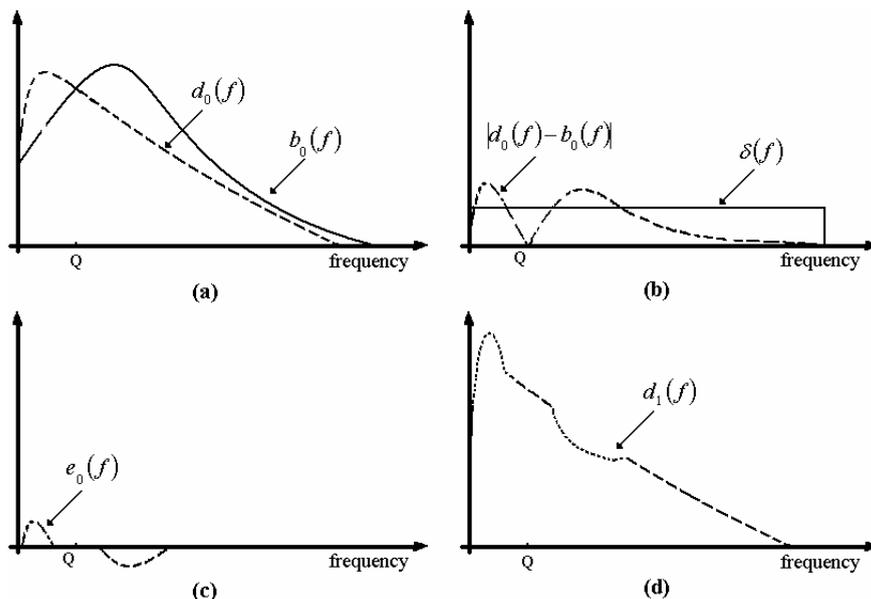


Fig.1

and sgn indicates the sign function. Fig.1-b and c show the computation process of the error signal. The error signal is nonzero at each frequency in which the designed filter deviation is larger than the acceptable deviation. In addition, we have,

1. if $d_0(f) \geq b_n(f)$ then, $e_n(f) \geq 0$.
2. if $d_0(f) \leq b_n(f)$ then $e_n(f) \leq 0$.

Now, the main problem is to minimize the error signal. It is obvious that to decrease the error, and make the designed filter shape as much as desired one, $b_n(f)$ has to be increased where $e_n(f) \geq 0$, and decreased where $e_n(f) \leq 0$. This can be done by increasing $d_n(f)$ any where $e_n(f) \geq 0$ and decreasing of $d_n(f)$ any where $e_n(f) \leq 0$. Therefore, to realize this mechanism, $d_{n+1}(f)$ is computed as follows

$$d_{n+1}(f) = d_n(f) \times (1 + \alpha \times e_n(f)) \quad (6)$$

which $\alpha > 0$ shows the convergence speed, and is selected appropriately. The small value of the convergence speed makes the computation time very long, and selecting it large, the algorithm may become unstable. During of the computing the next step predefined filter, it may be negative in some frequencies. To avoid this, we add a sufficient large number to desired filter source before starting the algorithm, and then subtract it from $h[0]$ after the algorithm convergence.

To consider the transition bands in the designing, It is simply realized by choosing a large acceptable deviation, $d(f)$, in the transition bands. Designing the filter according to LS and with no deviation consideration, it is sufficient to choose a large $d(f)$ in entire the frequency band. Therefore, from this point of view, the proposed algorithm is an extension of conventional LS method.

4. Computational complexity

Computational complexity of the LS filter design is related to the following equation set

$$P = \begin{bmatrix} N & 2\sum_{i=1}^N \cos(w_i) & \cdots & 2\sum_{i=1}^N \cos(Lw_i) \\ \sum_{i=1}^N \cos(w_i) & 2\sum_{i=1}^N \cos(w_i)\cos(w_i) & \cdots & 2\sum_{i=1}^N \cos(w_i)\cos(Lw_i) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^N \cos(Lw_i) & 2\sum_{i=1}^N \cos(Lw_i)\cos(w_i) & \cdots & 2\sum_{i=1}^N \cos(Lw_i)\cos(Lw_i) \end{bmatrix}, \quad P \times \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_L \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N H_i \\ \sum_{i=1}^N H_i \cos(w_i) \\ \vdots \\ \sum_{i=1}^N H_i \cos(Lw_i) \end{bmatrix} \quad (7)$$

Computation of P is the most expensive part of the LS computation. But, as eq.(7) shows, it is obvious that predefined filter (H_i) contributes only at right side of the equation. Thus, changing the predefined filter at each step, it is necessary to recompute this part only. To reduce the complexity, P can be simplified by choosing w_i as

$$w_i = (i-1) \times \frac{P}{N-1} \quad (8)$$

$$P = \begin{bmatrix} N & 0 & 2 & \cdots & 0 \\ 0 & N & 0 & \cdots & 2 \\ 1 & 0 & N & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 2 & 0 & \cdots & N \end{bmatrix} \quad (9)$$

Further computation decrease can be achieved when we update only summations parts of the eq.(9) in which H_i have been changed. For example, suppose only H_k is changed, it is sufficient to update $H_k \cos(jw_k)$ in summation parts of eq.(9). The proposed algorithm can be called Extended LS(ELS), or Recursive LS(RLS).

5. Simulation results

The proposed algorithm has been simulated for designing different filters and compared the results with LS method. As an example, the simulation results are shown in Fig.2-3 for three different samples. 0.01, 1000 and 1000 which are selected for α , N and R , respectively.. Each figure has 7 parts included,

- a. *Desired filter* characteristics such as transition bands, maximum acceptable deviation.
- b. *Designed filter by LS*.
- c. *Designed filter by ELS*.
- d. *Error signal of LS designed filter* computed by eq.(5).
- e. *Error signal of RLS designed filter* computed by eq.(5).
- f. *Changes of error* during convergence process.
- g. *The final filter* obtained by ELS.

Examples are designing of low-pass and band-stop filters. The comparison of Fig.2-d, and Fig.3-d with Fig.2-e and Fig.3-e, show that ELS designed filters have very smaller error than LS designed filters. By comparison of these figures, we can realize the performance of ELS.

Furthermore, the comparison of Fig.2-b and Fig.3-b with Fig.2-c and Fig.3-c is shown that ELS increases the oscillations in the entire of the band to reduce unacceptable large deviations.

The Fig.2-g and Fig.3-g show the final predefined filters. It can be seen that ELS increases predefined filter value at near band edges to enforce the deviations below the acceptable range. In Fig.(2-f), (3-f), & (4-f) the error changes of filter during convergence is shown.

In spite of the RLS flexibility, its main drawback is computational complexity of its computations compared to LS, which is because of its iterative structure. Thus, ELS may be not good for real-time processing, but very suitable approach to design any desired-shape FIR filters.

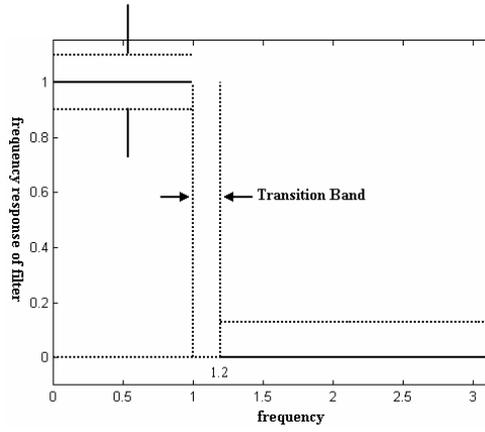
5. Conclusions

In this paper a new iterative algorithm for designing FIR filters with any desired shape is introduced, which is called ELS method. The simulation results show a good performance of ELS with computationally expensive with respect to other methods such as LS method. This new method may be not good for real-time processing, but very suitable approach to design any linear-phase desired-shape FIR filter.

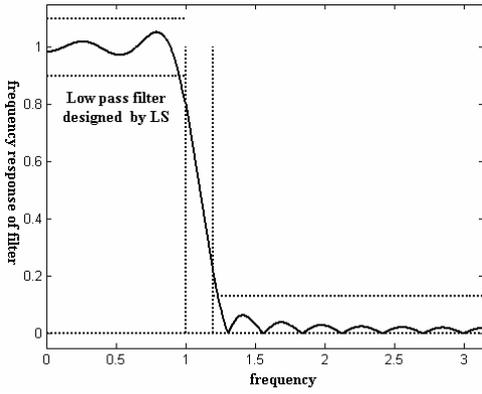
References

1. S.-P. Wu, S. Boyd, L. Vandenberghe, "FIR filter design via semidefinite programming and spectral factorization," *Proceedings of the 35th IEEE Conf. Decision and Control*, 1996.
2. A. V. Oppenheim, R. W. Schaffer, *Discrete-Time Signal Processing*, Prentice-Hall, 1989.
3. W. Selenick, M. Lang, C. S. Burrus, "Constrained least square design of FIR filters without specified transition bands," *IEEE Proc. Int. Conf. Acoustic, Speech, Signal Processing*, 1995.
4. H. H. Dam, K. L. Teo, "The dual parameterization approach to optimal least square FIR filter design subject to maximum error constraints," *IEEE Trans. Signal Processing*, 48(8): 2314-2320, 2000.
5. C. S. Burrus, "Multiband least squares FIR filter design," *IEEE Trans. Signal Processing*, 43(2): 412-421, 1995.
6. S.-C. Pei, C.-C. Tseng, "FIR filter design based on total least squares error criterion," *Proceedings of IEEE Int. Sym. Circuits and Systems (ISCAS)*, 3:283-286, 1999.
7. M. Z. Komodromos, S. F. Russell, P. Tak, P. Tang, "Design of FIR filters with complex desired frequency response using a generalized Remez algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing*, 43(2): 412-421, 1995.
8. M. Oner, M. Askar, "Incremental design of high complexity FIR filters by genetic algorithms," *Proceedings of the Fifth Int. Sym. Signal Processing and Its Applications (ISSPA)*, 2: 1005 -1008, 1999.
9. M. Oner, "A genetic algorithm for optimization of linear phase FIR filter coefficients," *Conf. Record of the Thirty-Second Asilomar Conf. Signals, Systems & Computers*, 2: 1397 -1400, 1998.
10. S. Haykin, *Adaptive filter theory*. Prentice-Hall, 3rd ed., 1996.

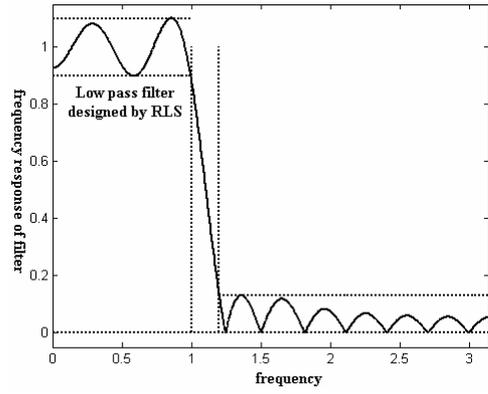
Acceptable Deviation



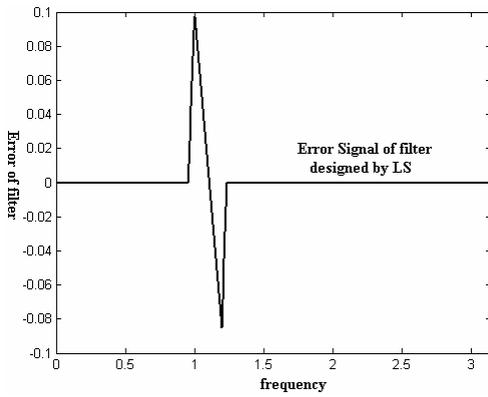
(a)



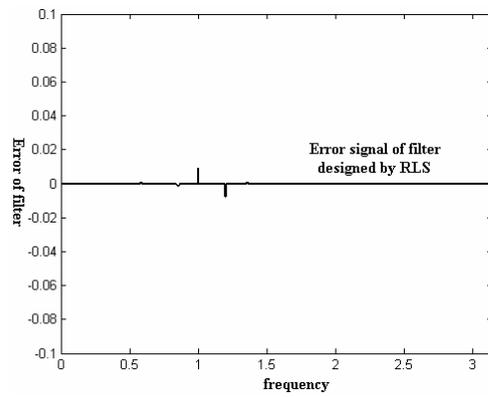
(b)



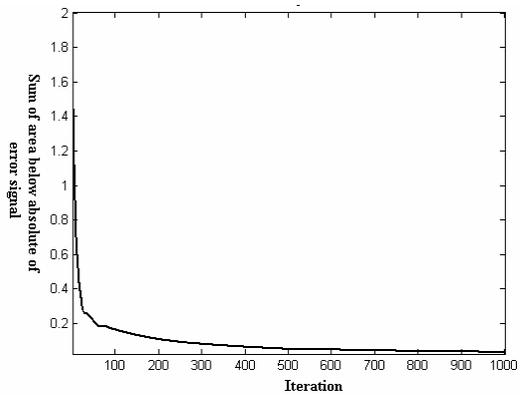
(c)



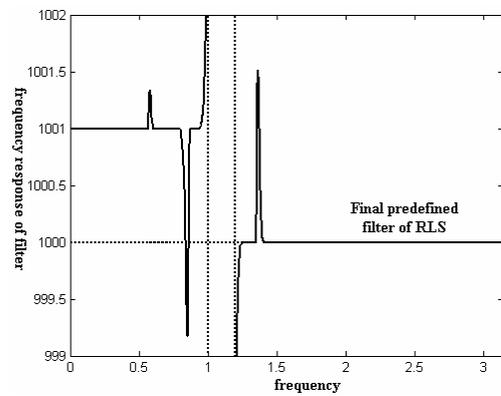
(d)



(e)



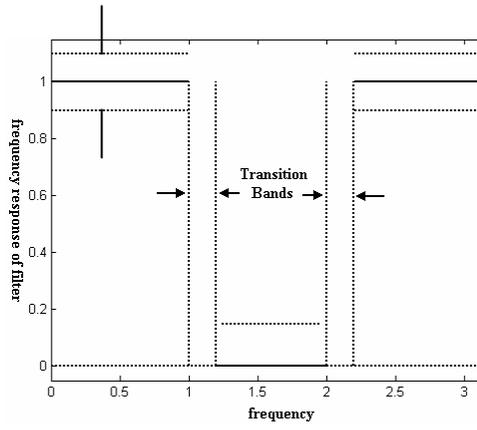
(f)



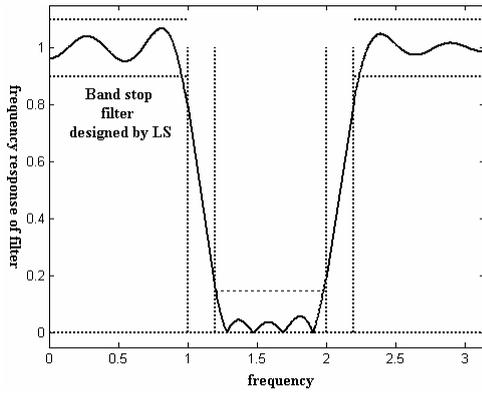
(g)

Fig. 2. A low-pass filter designing, with degree of 21

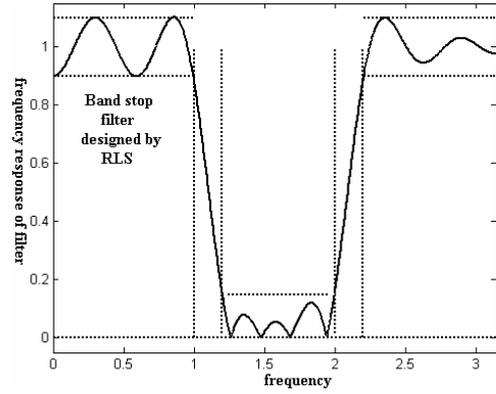
Acceptable Deviation



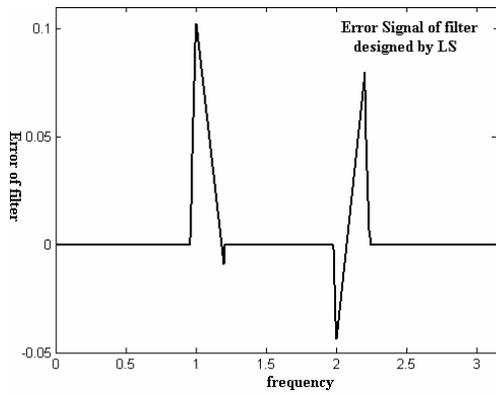
(a)



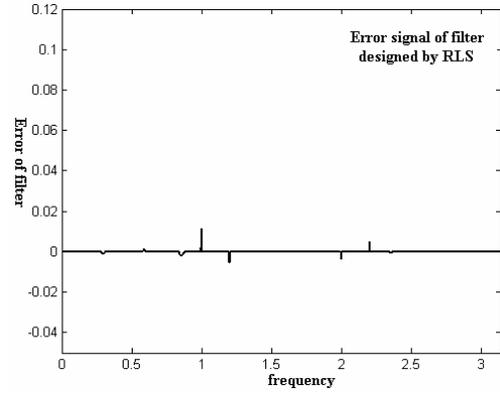
(b)



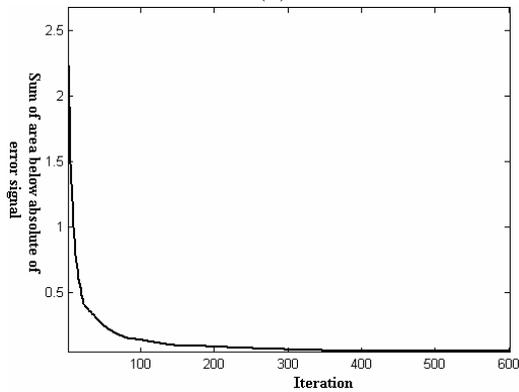
(c)



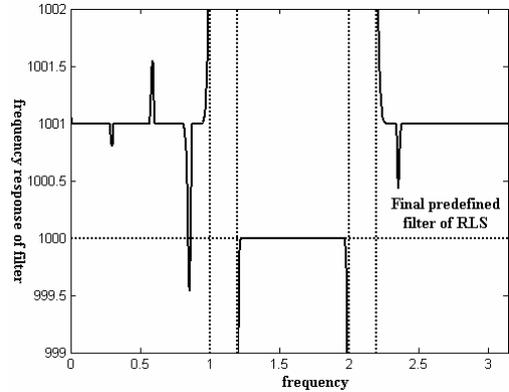
(d)



(e)



(f)



(g)

Fig.3. A band-stop filter designing, with degree of 21

A New Search Algorithm for Vector Quantization

M. M. AlDavood and S. J. Tabatabaian

Sistan & Baluchistan University
Zahedan, IRAN, 98164

and

Panasonic Mobile Communications Development of Europe Ltd.,
2 Gable Way, Colthrop, Thatcham, Berkshire, RG19 4ZB, UK

aldavood@hamoon.usb.ac.ir & seyed.tabatabaian@panasonicmobile.co.uk

Abstract:

Quantization is a source encoding method that adapts a continuous source for transmission over a digital channel in order to reduce the rate of transmission. In 1959 Shannon determined the minimum rate needed for encoding the source with a given fidelity criterion, this has been the basis for Quantization. One of the best methods to achieve the minimum rate is Vector Quantization (VQ), but there are two important factors, including the number of multiplications and the memory requirements which lead to complexity. For a reduction in complexity few methods have been published in literature. In this paper a new method for searching the VQ methods based upon Lloyd's Algorithm is presented. Results show the number of multiplications that could be saved is up to 74%.

Keywords: Vector Quantization, Rate-Distortion Theory and Lloyd's Algorithm

1. Introduction

Quantization is a source encoding method that converts a continuous or discrete information source for transmission over digital channel in order to reduce the rate of transmission. Because the absolute entropy of a continuous source is infinite, in quantization differential entropy is used for measuring information of a continuous source based on the following definition [1]:

$$h(X) = E[-\log_2 P(X)] = - \int_{-\infty}^{+\infty} P(X) \log_2 P(X) \quad (1)$$

where $P(X)$ is the Probability Distribution Function (PDF) and $E[g(X)]$ is the expectation value function.

Shannon's Rate-Distortion Theory [2] determines the minimum rate needed for encoding the source with a given fidelity criterion. This is the basis of continuous source quantization and a lot of efforts [3-5] have been made to achieve the minimum rate by introducing different techniques.

The Rate-Distortion function, $R(D)$, is defined as the rate based on Shannon's Rate-Distortion theory. In order to achieve $R(D)$, Vector Quantization is used instead of Scalar Quantization [2].

There are different methods used for VQ, Lloyd's Algorithm is the most popular of them [6]. The length of the vector is the most important parameter in VQ. Increasing the vector length, leads to the achievement $R(D)$. However, the encoding procedure could be more complex.

In this paper a new method, Spherical Area Allocation, is presented for design from searching in a codebook in all methods of VQ based on Lloyd's Algorithm. It is shown that the encoding procedure using this new method has less complexity than the full search and in particular multiplication operations are reduced up to 74%.

2. Rate Distortion $R(D)$ Function

Considering the source information, $\{X_k\}$, with fidelity criterion D , and conditional distribution, $Q(i|j)$. The average distortion related to Q is defined by $d(Q)$ given by:

$$d(Q) = \sum_{i,j} p(j) Q(i|j) d(i,j) \quad (2)$$

where $d(i,j)$ is difference between input and output encoder.

D-admissible conditional probability, $Q(i|j)$, is $d(Q) \leq D$, and all D -admissible $Q(i|j)$ is shown by Q_D :

$$Q_D = \{Q(i|j) : d(Q) \leq D\} \quad (3)$$

Then, the average mutual information can be calculated [1,2] by:

$$I(Q) = \sum_{j,k} P(j) Q(k|j) \log \frac{Q(k|j)}{Q(k)} \quad (4)$$

The fidelity criterion, or distortion, depends on source and VQ method and is defined by:

$$D = \int_{-\infty}^{\infty} [q(x) - x]^2 P(x) dx \quad (5)$$

And as a result signal to noise ratio is given by

$$SNR = (S/N) = 10 \times \log \sigma^2 / D \quad (6)$$

σ^2 is the variance of source.

The distortion for a vector quantizer is:

$$D = \sum_{i=1}^L \int_{C_i} (y_i - x)^2 P(x) dx \quad (7)$$

y_i , ($i=1, \dots, L$) is the codeword's vector, x is the input sequence and C_i is the i^{th} matched areas as shown in Figure 1.

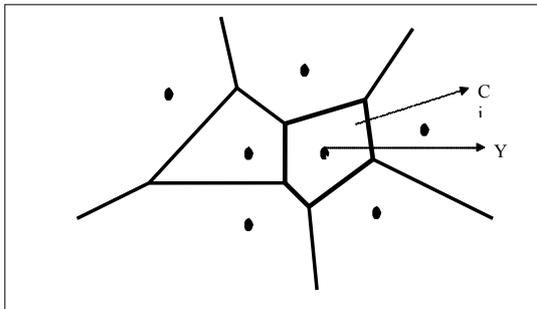


Figure 1: C_i areas and matched Codewords

Regarding to the above definitions $R(D)$ can be expressed [1, 2, 7] as

$$R(D) = \min I(Q) \quad \text{for all } Q \leq Q_D \quad (8)$$

4. Lloyd's Algorithm

Lloyd's Algorithm was introduced in 1948 [6] which is one of the most popular algorithms for VQ and close to optimal quantization. Figure 3a shows flowchart of this algorithm. There are two conditions required to use Lloyd's Algorithm including:

I – It suggests that for an input sequence, a codeword with shortest distance to the input is chosen given by:

$$x \in C_i \Rightarrow q(x) = y_i \quad (9)$$

This condition is also named as the nearest neighbor.

II – each codeword is the center of matched area respectively

$$y_i = Cent(C_i) \quad (10)$$

The most important factor in different methods of VQ based on Lloyd's Algorithm is the choice of the initial codebook. This is because codewords have to be convex for converging to the final codebook [7]. The random manner [10] and splitting method [11] are the two most well known methods for choosing the initial codebook.

5. New Method for Searching

There are two important factors in VQ including the number of multiplications and memory requirements that lead to system complexity. However, these days high capacity memory's Integrated Circuit (IC) are available and so the only problem is multiplication. Moreover increasing numbers of multiplication result in reducing VQ's speed and increasing error rounding. In recent years a few methods have been presented to reduce the number of multiplications. Gray & Bei [12-14] introduced a simple and valuable method that is shown in Figure 3b. The new algorithm is based on the allocation of a hyper sphere, a sphere in four dimensions of space or more, to any codeword. Its center is a codeword and its radius is the half of the distance between codeword and the nearest neighbor. Search algorithms that utilize this new method should determine d_{min} , and compare it with the radius of the matched hyper sphere, Figure 2. For an input sequence inside a hyper sphere, the matched codeword belonging to that sphere is selected as the encoded output sequence. If the input sequence is outside of all hyper spheres, then the Gray & Bei method [12] is used for encoding. Figure 3c shows the flowchart of the new algorithm.

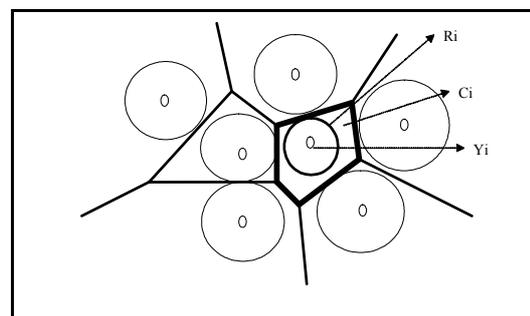
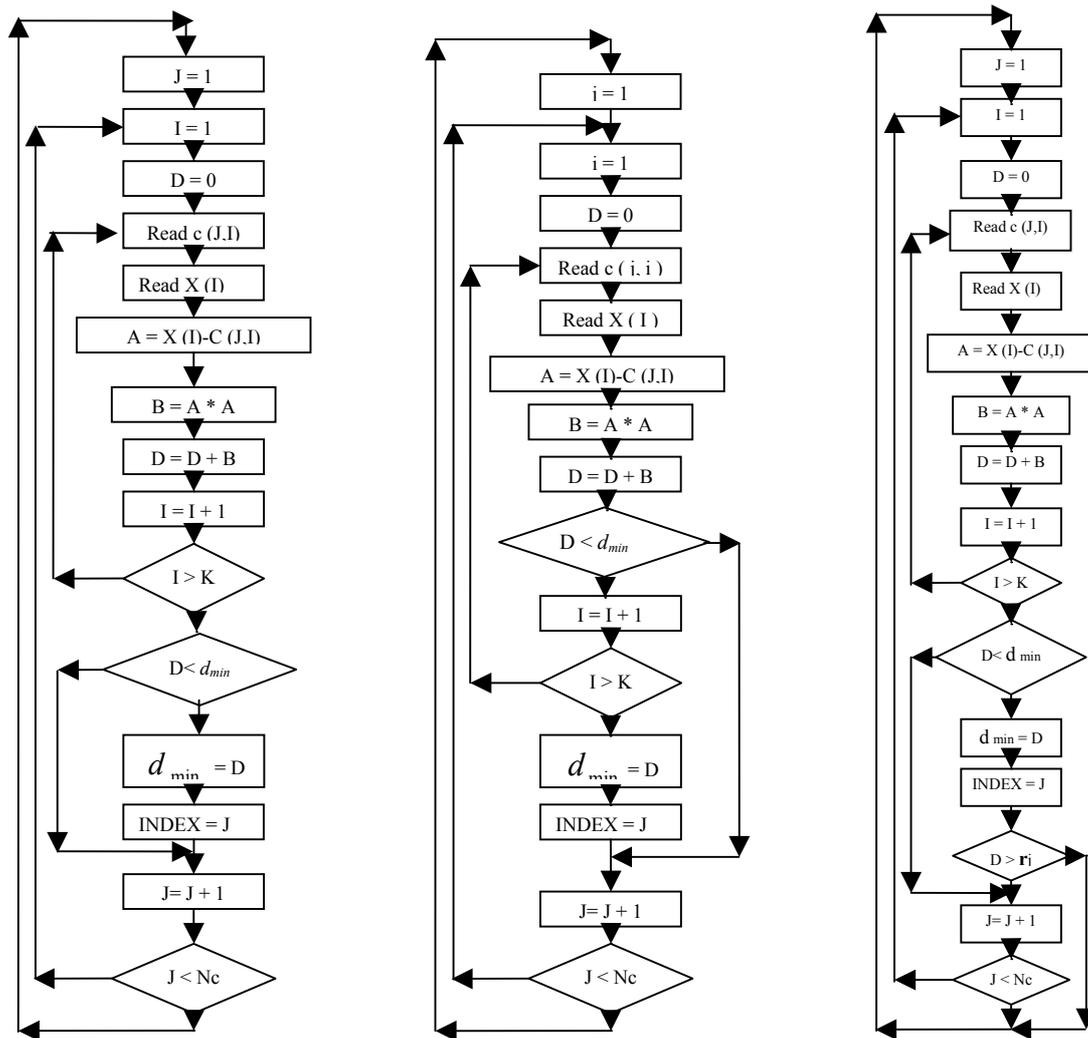


Figure 2: Spherical Area Allocation R_i , C_i and matched codewords in VQ.



a) Full search

b) Gray & Bei search

c) Spherical Area Allocation

Figure 3: a) Full search, b) Gray & Bei search and c) Improved Lloyds

19.1% to Gray & Bei. Therefore, significant savings are obtained using the new method.

6. Results

To verify the performance of the proposed new algorithm a first order Gaussian Markov input source is chosen and the number of required multiplications for different vector lengths and $R=1$ is obtained and the results using the different methods are compared. Table 1 shows the result of simulations for Full Search, Gray & Bei and spherical area allocation.

Also Figure 4 shows the savings in multiplication procedures using different methods. It can be observed that for vector length of $n=7$, specified $SNR=11.2$ dB and $R=1$ bit/symbol, the number of multiplications per symbol saved is up to 75.7% in comparison to the Full Search and

Table 1: The required number of multiplication per symbol

n	Full Search	Gray & Bei	Spherical Area Allocation
2	4	2.97	2.23
3	8	4.73	3.72
4	16	7.84	6.06
5	32	12.93	10.55
6	64	22.14	17.43
7	128	38.4	31.06

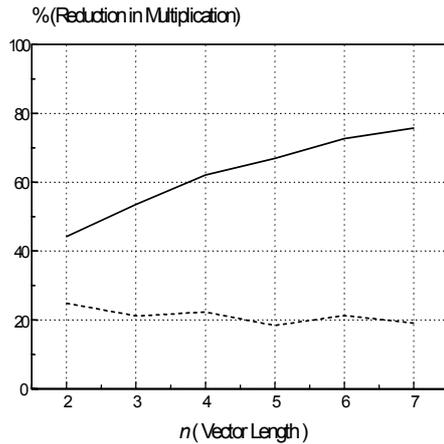


Figure 4: The percent of saving in multiplication using Spherical Area Allocation compared to Full Search and Gray & Bei

The study the effect of R on the number of multiplications, C_s leads to hyper spheres and so they cover more space, hence a number of multiplications are saved again. Table 2 shows the results for specified input sequence and $n=2$ and different methods of encoding.

Table 2: The required number of multiplication and saving for different R

R	Full Search	Spherical Area Allocation	Saving
1	4	2.23	44.2
1.5	16	3.55	77.8
2	32	5.79	81.9
2.5	64	9.62	84.9

7. Conclusion & Remark

In this paper the Spherical Area Allocation method of vector quantization based on Lloyds' algorithm was presented. For each matched codewords' area a hyper sphere was introduced. For an input sequence inside the sphere, the encoded output is the codeword of the corresponding sphere. The new algorithm was tested for a first order Gaussian Markov input source. Results show that the number of multiplications saved is up to 75% and 20% in comparison to Full Search and Gray & Bei methods respectively. Moreover, the effect of data rate was considered and results show a saving in multiplications. As a result, by reducing the complexity in computation, the vector length of the quantization procedure could be increased and better performance in achieving $R(D)$, could be obtained.

8. References:

- [1]. R.B. Ash, "Information Theory", New York, John Wiley, 1965.
- [2]. C.E. Shannon, "Coding Theorem for a discrete source with a fidelity criterion", IRE Nat. Conv. Rec. Pt.4, pp. 142-163, March 1959.
- [3]. A.B.Y. Linde A. Buzo and R.M. Gray, "An algorithm for vector quantization design," IEEE Transactions on Communications COM 28, No. 1, pp. 84-95, January 1980.
- [4]. J.G. Proakis, "Digital Communications", McGraw-Hill, New York, 1995.
- [5]. A. Gersho, and R.M. Gray, "Vector Quantization: A pattern matching technique for speech coding", IEEE Communication Magazine, December 1983.
- [6]. S.P. Lloyd, "Least squares quantization in PCM," IEEE Transaction on Information Theory, Vol. IT-28, No. 2, March 1982.
- [7]. M.M. Aldavood, "Vector Quantization," Isfahan University of Technology, Iran, Mar 1988.
- [8]. J. Makhoul, S. Roucos and H. Gish, "Vector quantization in speech coding", Proceeding IEEE Vol. 73 No.11, pp. 1551-1588, Nov 1985.
- [9]. A. Gersho, "Asymptotically optimal block quantization", IEEE Transactions on Information Theory, Vol. IT-25, No. 4, July 1979.
- [10]. R.M. Gray, "Vector quantization," IEEE Acoustics, Speech and Signal Processing Magazine, Vol. 1, No. 2, pp. 4-29, April 1984.
- [11]. R.M. Gray and A.B.Y. Linde, "Vector quantizers and predictive vector quantizers for Gauss-Markov sources," IEEE Transaction on Communications, COM-30, pp. 381-389, Feb 1982.
- [12]. R.M. Gray, C. Bei, "An improvement of the minimum distortion encoding algorithm for vector Quantization," IEEE Trans. on Comms, COM-33, No. 10, pp. 1132-1133, October 1985.
- [13]. M.R. Soleymani and S.D. Morgera, "A high speed algorithm for vector quantization," Proceedings IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 45.6.1-45.6.2, Apr 1987.

A Novel Fast Search Algorithm For MPEG Standard

S. Soongsathitanon, S.S.Dlay
School of Electrical, Electronic and Computer Engineering
University of Newcastle upon Tyne

Abstract

This paper presents a novel fast search algorithm, diagonal logarithmic search (DLS) algorithm, for block based motion estimation. The performance of the algorithm is evaluated by using standard 176×144 pixels quarter common intermediate format (QCIF) benchmark video sequences and the results are compared to a traditional well-known Full search (FS) Algorithm and a widely used fast search algorithm called the Three Step Search (3SS). The important metrics, time and Peak Signal to Noise Ratio (PSNR) are used to evaluate the novel algorithm. The results show that the strength of the algorithms lies in their speed of operation as they are much faster than the FS and 3SS. The speed of the operation is improved by 88.77 % and 40% over the FS and 3SS.

1. Introduction

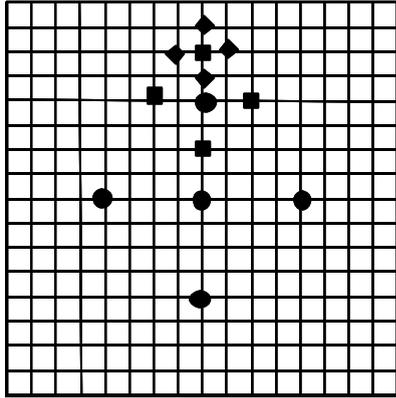
In recent years, Video compression has played a vital role in data storage and transmission as it has been the main focus in many applications such as multimedia communications, remote monitoring, videophones, videoconference etc. The effective video compression mostly employ the hybrid coding configuration which involves both intra frame and inter frame Coding. The intra frame compression is a vital means of exploiting the spatial redundancies and inter frame compression is used for exploiting the temporal redundancies. Due to the slow movement of video sequences such as 'head-and shoulders', the two consecutive frames will not dramatically change. So the current frame can be predicted from the translation of the previous frame. The method that is commonly used is Motion Compensation prediction (MC) Predictive Coding is widely used in video transmission, especially for low bit rate coding. The basic idea behind motion compensation is to estimate of the displacement of objects. The methods used are called motion estimation (ME). The ME methods can be classified into 3 types which are optical flow equation based methods[1], Pixel-recursive methods[2] and block matching methods. The block matching is widely used by the video compression

standards such as MPEG1-2, H.263 [3] because it can achieve a good balance between complexity and coding efficiency [4]. The block matching aims to find the best block from an earlier frame to reconstruct an area of the current frame. The block matching method can be categorized into frame based block matching and object based block matching method. There are many ME techniques used nowadays. The conventional frame based block matching technique is considered to be the full search or exhaustive search which searches every possible candidate block within search area in order to find the block with minimum difference. However, the implement of Full Search Algorithm requires massive computation or huge hardware cost to achieve significant speed. To alleviate these problems, many fast-searching algorithm have been developed to reduce the computation and data fetching by reducing the number of comparisons between the blocks, such as the 2-D logarithm [5], the three-step search (3SS) [6], the New three-step Search [7], the Four Step Search (4SS) [8], the orthogonal search [9] and hierarchical block matching algorithms. These proposed methods try to improve the performance of the motion estimation both in term of accuracy and speed. Nevertheless these fast searching algorithms can not improve both accuracy and speed at the same time such as the N3SS can improve the accuracy of the motion estimation but the speed of the operation is increased as it search more points than the 3SS. Having taken every factor into account, the 3SS search still remain a favorite choice for motion estimation as it still used in MPEG standard.

The novel diagonal logarithmic search (DLS) algorithm aims to improve the performance of the speed of operation and still keep the good performance in the prediction. The strategy of DLS algorithm is to reduce search points. Having reduced the search points, the computation complexity is also reduced. Therefore the speed of operation is increased. The search pattern had to be designed to keep the good performance on the prediction.

2. Diagonal Logarithmic Search (DLS) Algorithm

In video coding, the current frame is divided into equal sized non-overlapping rectangular blocks. The frame dimensions are multiples of the block size and square blocks are mostly used. Then the comparisons between the target block and candidate blocked are determined to search for the best block matching.



● Points for Stage1 ■ Points For Stage2 ◆ Points for Stage 3

Figure (1): The Example Path for convergence of Logarithmic Diagonal Search (DLS)

The novel DLS algorithm is a search algorithm to find the best block matching. The DLS has both a vertical and horizontal stage. Firstly, the method searches for the similarity in both horizontal and vertical axis initially some displacement $\pm disp$ from the centre block. The displacement is related to the size of the search area using the formula below

$$disp = \log_{10}(s + 1) / \log_{10}(2)$$

where $s \times s$ is the size of the area

From the five blocks, the block that produces the minimum distortion function becomes the center of the next step. The procedure continues until the distance ($disp$) is equal to one, and then all the surrounding blocks are tested.

The distortion function or matching criteria is used to qualify the similarity between the target block and candidate blocks. There are many kinds of matching criteria such as Mean Absolute Difference (MAD), Mean Square Error (MSE), and Pixel Different Classification (PDC) etc. The widely used distortion criterion is MAD. The MAD is the fastest matching criterion. The MAD can be found from the equation below

$$MAD_{(x,y)}(m_x, m_y) = \sum_{i=1}^m \sum_{j=1}^n |f_n(x+i, y+j) - f_{n-1}(x+i-m_x, y+j-m_y)|$$

Where m_x and m_y describe the motion vector within the search window and x and y represent the spatial positions of the MC block under consider. The m and n are the dimension of the blocks.

After the best block matching is found. The motion vector $v(x,y)$ of block (x,y) is can also be found by

$$v(x, y) = \text{argmin} MAD_{(x,y)}(m_x, m_y)$$

3. Algorithm Testing

The algorithms were implemented on a Pentium4 2.4 GHz computer. During the performance evaluation, the three QCIF video sequences, Claire, Carphone and Foreman, are selected. The novel DLS is compared to the full search algorithm and the 3SS.). As seen in figure (2) the two consecutive frames are compared. The motion vectors can be found by using DLS algorithm and then the currents frame can be predicted by adding the motion vectors information to the reference frame.



(a) Previous Frame

(b) Current Frame

Figure (2) Two Consecutive Frame of "Clair" Video Sequence

The motion vectors between two consecutive frames of "Caire" Video Sequence are shown in figure (3).

In this paper, the metrics time and PSNR (Peak Signal to Noise Ratio) are considered. The time is used as a measure of computational complexity. The PSNR is used to determine the quality of compressed images.

The PSNR for a gray scale image is defined as follows

$$PSNR = 10 \log_{10} \left(\frac{255^2}{1/M \times N \sum_{i=1}^M \sum_{j=1}^N (x_{ij} - \hat{x}_{ij})^2} \right)$$

Where are $M \times N$ the dimensions of the frames in pixels and X_{ij} and \hat{X}_{ij} are the luminance components of the original and reconstructed image respectively, at the spatial location (i, j) .

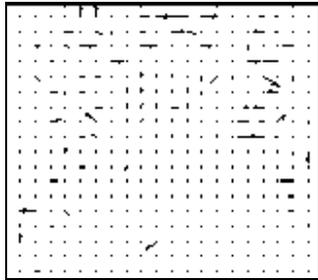


Figure (3) the motion vectors between two consecutive frames.

4. Results

The DLS algorithm is compared with the 3SS and FS algorithm. The PSNR performance of “Claire” video sequence is shown in figure (4). The PSNR values of prediction using DLS algorithm is as good as that of the 3SS and FS algorithm. As seen from the figure (4) the PSNR of all 90 frames using DLS is almost the same values as using 3SS and FS algorithm. The PSNR values of DLS algorithm lie on almost the same line with that of 3SS and FS. The PSNR values are also very high. From the figure (4) the results shown that all three algorithms, FS 3SS and DLS can successfully predict the “Claire” video sequence. The predictions are very accurate. The average PSNR over 90 frame of Claire video sequence are 41.66 dB, 41.62 dB and 41.61dB for Full Search, 3SS and DLS respectively.

The performance of DLS over “Foreman” and “Carphone” compared with 3SS and FS is shown in figure (5) and (6). The performance of DLS algorithm is slightly worse than FS algorithm in “Carphone” video sequence but it is better than 3SS. The average PSNR over 90 frame of Carphone video sequence are 33.43 dB, 32.82 dB, and 32.98 dB for Full Search, 3SS and LDS respectively.

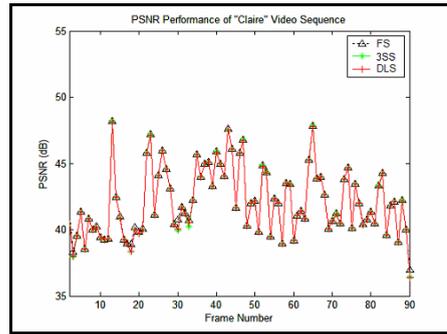


Figure (4) Comparative PSNR Performance of “Claire” Using FS, 3SS and DLS

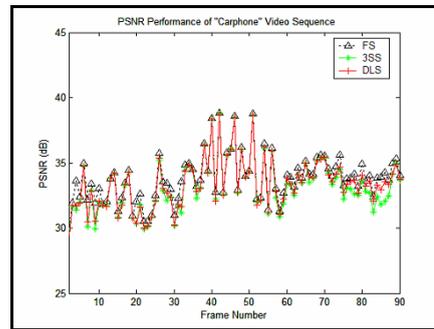


Figure (5) Comparative PSNR Performance of “Carphone” Using FS, 3SS and DLS

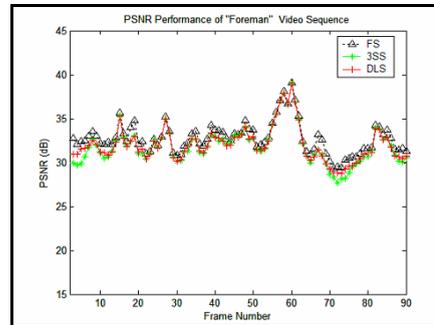


Figure (6) Comparative PSNR Performance of “Foreman” Using FS, 3SS and DLS

The results of “Foreman” video sequence come out at the same manner with the other video sequences. The DLS algorithms still remain better performance than 3SS. The average PSNR over 90 frame of Foreman video sequence are 32.39 dB, 31.53dB, and 31.68dB for Full Search, 3SS and OLS respectively.

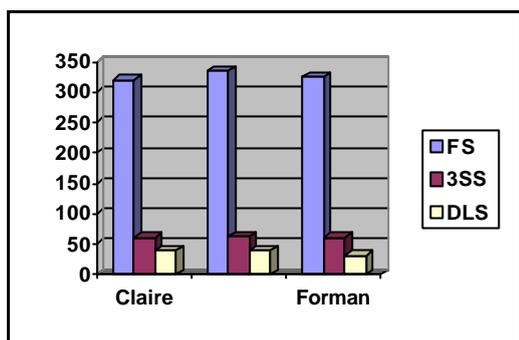


Figure (7) Comparative Speed of Operation using FS, 3SS and DLS

The advantage of DLS algorithm lies in the speed of operation. The 3SS algorithm designed for the improvement of FS in terms of processing time. The main target of DLS algorithm is also the same. The DLS aims to improve the performance of prediction both in terms of speed and quality. The results in figure (7) are the total time which the algorithm spend on the prediction over 90 frames of each video sequences. The result shows the total time of DLS is much better than FS and 3SS. Also the quality of prediction of DLS is better than 3SS.

5. Conclusion

Using the benchmark QCIF video sequences, the novel DLS is proved to be very good choice and better than the 3SS. The performance of prediction is very good both in terms of subjective and objective quality. The performance of prediction of DLS is better than 3SS as the PSNR of reconstructed frames are improved approximately 0.06 dB over 3SS. The advantage DLS algorithm lies in the speed of operation. The speed of DLS is extremely fast as it is 88.77 % faster than the FSA and over 40% faster than the 3SS. The system such as the mobile phone communications which the bandwidth and time is limited the DLS is better choice than the 3SS and FS.

References

[1] Netravali A., Haskell B., *Digital Pictures-representation and compression*, 1998

[2] Jain J., Jain A., Displacement Measurement and its Application Interframe Image Coding, *IEEE Trans. Commun.*, vol.29, no.12, 1981, pp 1799-1806

[3] Koga J., Inuma k., Hirani A., Iijima Y., Ishiguro T. (1981). "Motion Compensated Interframe Coding for Video Conferencing" Proc. National Telecommunications Conference, pp G 5.3.1-5.3.5.

[4] Puri A., hang H., Schilling D. (1987). "An efficient block-matching algorithm for motion compensated coding" proc. IEEE Conf. Acoustics, Speech and Signal Processing, pp 25.4.1-25.4.4.

[5] Chan M., Yu Y., Constantinides A. (1990). "Variable size block matching motion compensation with application to video coding" IEE Proc. I, Commun. Speech Vis., vol.137, no.4, pp 205-212.

[6] Nogaki S. and Ohta M. (1992). "An overlapped block motion compensation for high quality motion picture coding" Proc IEEE Int. Symp. Circuits, Systems, pp 184-187

[7] Yogoyama Y., Miyamoto Y., Ohta M. (1995). "Very low bit rate video coding using arbitrarily shaped region based motion compensation" IEEE Trans. Circuits, Systems for Video technology, Vol.5, no.6, pp 500-507

[8] Moon J., Kweon J., Kim H. (1999) "Boundary Block Matching technique for efficient texture coding of arbitrarily shaped object" IEEE Trans. Circuits, Systems for Video Technology, Vol.9, no1, pp35-43

[9] Pereira F. (1998). "New trends on Video and Image Coding" RECPAD 98, Lisboa Portugal.

[10] Hanzo L., Stereit S (1995). "Adaptive Low-Rate wireless videophones schemes" IEEE Trans. Circuits and Systems for video technology, vol. 5, no 4, pp 305-318

[11] Lee X., Zhang Y. (1996). "A Fast Hierarchical Motion-Compensation Scheme for video coding using Block Feature Matching" IEEE Trans. Circuits and Systems for video technology, vol. 6, no 6, pp 627-633

[12] Girod B. (1993). "Motion Compensating Prediction with fractional-pel accuracy" IEEE Trans. Commun., vol. 41, no 4, pp 604-612

[13] Chavari H., Mills M. (1990). "Block Matching Motion Estimation Algorithms - New Results" IEEE Trans. Circuits, Systems for Video Technology, vol.37, pp 649-651

[14] L. Scargall, S.S. Dlay, "Mobile Videophone Systems for Radio Speech Channels", Proc. Of SPIE'98, Jan. 24-30, 1998, San Jose, America, pp. 170-17

A Survey Paper on Regression Testing Techniques

Ajay Rana

Department of Computer Science
Amity School of Computer Sciences
Amity Campus
Noida-201303

India

ajay_rana@ascs.amity.edu

31 th January 2004

ABSTRACT

Regression testing is a necessary but expensive maintenance activity aimed at showing that code has not been adversely affected by changes. The purpose of regression testing is to ensure that changes made to software, such as adding new features or modifying existing features, have not adversely affected features of the software that should not change. Regression testing is usually performed by running some or all of the test cases created to test modifications in previous versions of the software. Many techniques have been reported on how to select regression tests so that the number of the test cases does not grow too large as the software evolves.

Keywords: Regression Testing, Test Set Minimization, Test Set Prioritization, Domain Analysis, Slicing, Scenario Based Testing, Incremental Testing, Interactive program testing.

Introduction

Software engineers often save test suites they develop so that they can reuse those test suites later as their software evolves. Such test suite reuse, in form of regression testing, is pervasive in the software industry[16]. However as a program evolves the regression test set grows larger, old tests are rarely discarded and the expense of regression testing grows. Repeating all previous test cases in regression testing after each minor software revision or patch is often impossible due to pressure of time and budget constraints. On the other hand, for software revalidation, arbitrarily omitting test cases used in regression testing is risky. In this paper, we introduce a few methods to select

small subsets of effective fault-revealing regression test cases to revalidate software.

Scenario Based

[24] discusses scenario based functional regression testing which is based on an end-to-end(E2E) integration test scenario. The test scenarios are first represented in a template model that embodies both test dependency and traceability. Traceability information shows affected components and associated test scenarios and test cases for regression testing. Further ripple effect analysis[10] can be used to identify all directly and indirectly affected scenarios and thus the set of test cases can be selected for regression testing.

Prioritization

[6] discusses test case prioritization techniques that schedule test cases for execution in an order that attempts to maximize some objective function. One such performance goal, rate of fault detection, is a measure of how quickly faults are detected within the testing process. An improved rate of fault detection can provide faster feedback on the system under test, and let software engineers begin locating and correcting faults earlier than might otherwise be possible. Such feedback can also provide earlier evidence that quality goals have not been met, allowing earlier strategic decisions about release schedules.

The author also discusses 14 test case prioritization techniques grouped under three categories, namely, Control level, Statement level and Function level. The paper also discusses different parameters that help in prioritization such as **FEP** (Fault exposing potential) based on mutation analysis[5, 18], **FI** (Fault index to represent fault proneness of a function) prioritization based on principle component analysis[21] and their variants.

Text Set Minimization

A problem with prioritization is that we do not know how much is enough while selecting test cases one by one from the top. [25] presents a test set minimization procedure that finds a minimal subset in terms of the number of the test cases that preserves the coverage with respect to a certain criterion (say@) as the original set. This implies that, if a program attribute defined by @ is covered by the original set, it will also be covered by the minimized set.

However preliminary experimental results suggests that minimization of test suites may have a significant, adverse impact on the ability to detect regression errors[20].

Incremental Regression Testing

Incremental regression testing [2] tries to find the solution to the problem of finding the test cases in the regression suite on which the modified program may differ from the original program. [4] discusses incremental regression testing and its cost reduction.

Slicing

Program Slicing[22] is an operation that identifies semantically meaningful decomposition of programs, where the decompositions consist of elements that are not textually contiguous. [17] presents the technique of using two slicing algorithms to determine directly and indirectly affected def-use(Definition –Use) associations.

The slicing algorithms are efficient in that they detect the def-use associations without requiring either the data flow history or the complete re-computation of data flow for entire program. These algorithms are based on the approach that uses the control flow graph representation of the program and only requires the computation of partial data flow information. Thus, the technique requires neither a test suite nor complete data flow information to enable selective retesting. If a test suite is maintained, fewer tests may be executed since only those tests that may execute affected def-use (Definition – Use) associations are rerun.

[4] surveys a collection of published techniques that use slicing to reduce the cost of regression testing. [3] presents an algorithm that uses language semantics to identify

program components that exhibit different run-time behavior.

Simulated Annealing and Genetic Algorithms

[15] proposes the optimal regression testing problem of determining the minimum number of test cases needed for revalidating modified software in the maintenance phase. These algorithms, based on the integer programming problem formulation and program's control-flow graph, are simulated annealing and genetic algorithms. The main advantage of these algorithms is that they do not suffer from exponential explosion for realistic program sizes.

Domain Based Testing

Domain based Testing(DBT) is a test generation method based on domain analysis and domain modeling. Domain models are used as a structure to generate tests. The domain model stores syntax and semantic information needed for test case generation. It can be configured into sub-domains for specific testing needs. [23] describes how to identify regression testing sub-domains and how to generate regression test data.

Comparison of various approaches

It is difficult to compare and evaluate these techniques because they have different goals. [9] outlines the issues relevant to regression test selection techniques, and uses these issues as the basis for a framework within which to evaluate the techniques.

[7] compares five different approaches namely slicing, incremental, firewall, genetic and simulated annealing algorithms, towards regression testing for various criteria such as execution time, number of selected retests and precision among others. The

experimental results show that the five algorithms are suitable for different requirements of regression testing. The author concludes that in most cases, the incremental algorithm seems to have more favorable properties than the other four algorithms.

Object-oriented Testing

Object oriented programming has many useful features such as information hiding, inheritance, polymorphism and dynamic binding. These features can be used to filter out test cases that are irrelevant to the modification. [1, 13, 19, 8] address the issue of object-oriented regression testing. [26] presents a selective regression technique based on a function dependence relationship.

Interactive Programs

Testing interactive programs requires interaction - usually by real people. Such testing is an expensive process and hence rarely done. Some interactive tools can be used non-interactively to a limited extent, and are often tested only this way. [14] illustrates with a couple of examples how high level scripts like Tcl (Tool Control Language) [11, 12] can be used to do testing of interactive programs.

Conclusion

Various approaches have been presented in the above papers for selecting test cases for regression testing. Most of the studies which were surveyed on selecting test cases for regression testing present techniques on traditional based software but only a couple of them seem to address the issue in the object oriented paradigm.

References

- [1] K. Abdullah and L. White. A firewall approach for the regression testing of object-oriented software. In Software Quality Week Conference, San Francisco, 1997.
- [2] Hiralal Agrawal, Joseph R. Horgan, Edward W. Krauser, Saul A. London, "Incremental Regression Testing", In Proceedings of the 1993 IEEE Conference on Software Maintenance, Montreal, Canada, September 27-30, 1993.
- [3] D. Binkley, "Reducing the Cost of Regression Testing by Semantics Guided Test Case Selection". In IEEE International Conference on Software Maintenance, 1995.
- [4] David Binkley, "The Application of Program Slicing to Regression Testing".
- [5] R. A. DeMillo, R. J. Lipton and F. G. Sayward. Hints on Test Data Selection: Help for the Practicing Programmer. *Computer*, 11(4): 34-41, Apr. 1978.
- [6] Sebastian Elbaum, Alexey G. Malishevsky, and G. Rothermel. Prioritizing test cases for regression testing. In Proceedings of the International Symposium on Software Testing and Analysis, pages 102-112, August 2000.
- [7] Ghinwa Baradhi and Nashat Mansour, "A Comparative Study of Five Regression Testing", 1997 IEEE Proceedings of the Australian Software Engineering Conference (ASWEC' 97).
- [8] G. Rothermel, M. J. Harrold and J. Dedhia. Regression test selection for C++ software. Technical Report TR99-60-01, Oregon State University, 1999.
- [9] Gregg Rothermel, Mary Jean Harrold, "Analyzing Regression Test Selection Techniques", *IEEE Transactions on software engineering*, Vol 22, No. 8, August 1996.
- [10] Hind Kabaili, Rudolf K. Keller and François Lustman, "A Change Impact Model Encompassing Ripple Effect and Regression Testing"
- [11] John Ousterhout, "Tcl: An Embeddable Command Language", Proceedings of the Winter 1990 USENIX Conference, Washington, D.C., January 22-26, 1990.
- [12] John Ousterhout, "tcl(3) - overview of tool command language facilities", unpublished manual page, University of California at Berkeley, January 1990.
- [13] D. Kung, J. Gao, P. Hsia, F. Wen, Y. Toyoshima and C. Chen. On regression testing of object-oriented programs. *The Journal of Systems and Software*, 32(1): 21-40, January 1996.
- [14] Don Libes, "Regression Testing and Conformance Testing Interactive Programs", to appear in *IEEE Transactions on Software Engineering*, New York, NY.
- [15] Mansour, N., and Fakih, K. Natural Optimization algorithms for optimal regression testing. COMPSAC' 97, Washington D.C., Aug 13-15, 1997.
- [16] K. Onoma, W.-T. Tsai, M. Poonawala, and H. Sukanuma. Regression testing in an industrial environment. *Comm. of the ACM*, 41(5):81--86, May 1988.
- [17] R. Gupta, M. Harrold, M. Soffa, "An approach to regression testing using slicing," *Conference on Software Maintenance*, 1992, pp. 299-308
- [18] R. G. Hamlet. Testing programs with the aid of a compiler. *IEEE Trans.*

- Softw.Eng., SE-3(4): 279-290, July 1977.
- [19] G. Rothermel and M. J. Harrold. Selecting regression tests for object-oriented software. In International Conference on Software Maintenance, pages 14-25, 1984.
- [20] G. Rothermel, "Efficient, Effective Regression Testing Using Safe Test Selection Techniques", PhD dissertation, Clemson Univ., May 1996.
- [21] S. G. Elbaum and J. C. Munson. Code churn: A measure for estimating the impact of code change. In Proc. Int'. Conf. Softw. Maint., pages 24 - 31, Nov. 1998.
- [22] Thomas Reps, Todd Turnidge, "Program Specialization via Program Slicing".
- [23] A. von Mayrhauser, R. Mraz, J. Walls, "Domain Based Regression Testing" , Procs. International Conference on Software Maintenance 1994, Sept. 1994, CS Press, pp. 26 -- 35
- [24] Wei-Tek Tsai, Xiaoying Bai, ray Paul, Lian Yu, "Scenario-Based Functional Regression Testing, Proceedings of the 25th Annual International Computer Software and Applications Conference (COMPSAC' 01).
- [25] W. Eric Wong, J. R. Horgan, Saul London, Hira Agrawal, "A Study of Effective Regression Testing in Practice", 8th IEEE International Symposium on Software Reliability Engineering(ISSRE'97), pp 264-274, Albuquerque, NM, November 1997.
- [26] Ye Wu, Mei-Hwa Chen, Howard M. Kao, "Regression Testing on Object-Oriented Programs"

AN EFFICIENT OF CODEBOOK DESIGN FOR SPEAKER DEPENDENT CODING SYSTEM

J.Srinonchat, S.Danaher, J.I.H.Allen
School of Engineering and Technology, Northumbria University
Newcastle upon Tyne, UK, NE1 8ST
phone: +44 191 2273624 email: jakkree.srinonchat@unn.ac.uk

A. Murray
Advanced Technology Division
Tail Electronics, Christchurch, New Zealand

Abstract

The spectral envelope and the source characteristics of speech exist in each speech waveform. However it is obviously different in each speaker. This work has been investigated and takes advantage of these characteristics of speaker to design the Speaker Dependent Coding System. Four speakers are investigated; two male and two female. The speech waveform is coded to LPC and LSP coefficients and then the K-means algorithm is employed to design an optimal codebook for each speaker. The performance of this technique will be presented. Preliminary results are shown that this technique can be reduce the bit rate and still maintains the quality of synthesis speech sound.

1. Introduction

The requirement to compress speech signal has been sought in most main speech coding research centres, which the traditional propose is to reduce the bit rate while maintain the suitable quality for storage or transmission in communication system. Recently there are a large number of results and applications in different strategies for suitably encoded speech for bandwidth-restricted [1]. The well-known Pulse Code Modulation (PCM) [2] is the earliest standard 64 kbps speech compression method for telephone network. PCM coding provides the good quality speech transmission, but uses high bandwidth for operation. Hence it has been found to be inadequate in conditions of spectrum efficiency when applied to newer bandwidth limited communication system.

Generally, speech waveform contains the information of both spectral envelope and the voice source characteristics. Hence each speaker will have individual properties speech waveform. This work has been investigated and exploited the characteristics in each speaker for design a Speaker Dependent Coding System (SDCS). The SDCS has the ability to gain knowledge from a particular speaker by training from speech samples of that

speaker. This paper is mainly concerned to exploit the repetitiveness of the attribute of a single speaker to further reduce the bit rate for wireless communication. Research on speaker dependent coding system is one of the most important future directions for achieving advanced speech information processing systems [3].

This paper is briefly introduced the Speaker Dependent Coding System. The K-means algorithm is used to design a specific codebook for an individual speaker for measure the repetitiveness of the attribute of a single speaker and also to reduce the bit rate for transmission. Resulting showing the bit rate reductions in term of transmission of codebook addresses using a database of four speakers are presented and finally suitable conclusions are made

2. SDCS VS SICS

The achievements of speech coders are growing in the array of both industrial and domestic products, by the explosion in the multi-media and personal communication industries. Speaker Independent Coding System (SICS) are the ideal method of all products that transmit speech, where they attempt to encode any speech to use in lowest bandwidth and also without degradation in the synthetic speech quality. SICS does not operate for specifically with any individual speaker. SICS can be used by any speaker or any user. Thus this is an ineffectual for secure in communication system.

The secure communications is the necessary and important in transmission system especially in military applications. Hence the SDCS, which has the ability to gain knowledge from a particular speaker, is interested to apply for the secure communications application. Normally, the data stream, which is actually transmitted from transmitter to receiver, can be reconstructed into the original speech signal, without the secrete code. If the voice description of a speaker is the secrete code to decode the data stream, therefore this offers a second layer of protection in the system. The

reconstruction would be difficult, if the voice description file of the speaker is not existed. The successful of secure transmission in the system will be completed, if both transmitter and receiver contain the voice description file that never requires for transmit. This is one of the reasons why the SDCS has been studied.

3. Codebook design.

Speech waveform is basically a continuous signal, which created by a physical system. The signal is altered in time domain upon the vocal tract vibration. The information of speech waveform exists both in the spectral envelope and in the voice source characteristics of speech. The spectral envelope, called vocal tract, is coded to LPC and LSP parameters using LPC and LSP analysis [4, 5, 6]. In this section the K-means algorithm are employed to expose voice source characteristics of a particular single speaker. To achieve gaining acknowledge of a particular speaker, the K-means algorithm is required in SDCS for design an optimal codebook.

3.1 K-means algorithm

Generally, the operation of the K-means algorithm [4, 7] is to split the training data into a set of code vectors, called K cluster. This function is like adjusting the levels of a scalar quantiser to minimise value of distortion when a signal is quantised. This set of code vectors comprises a codebook for the feature space. The process of K-means algorithm for the design of a codebook is as follows:

1. Initialisation: the set of random data has been chosen to be the initial set of code words in the codebook.
2. Nearest-Neighbour Search: For each training vector, find the code word in the current codebook that is closest (in terms of spectral distance), and assign that vector to the corresponding cell (associated with the closest code word).
3. Centroid Update: Update the code word in each cell using the centroid of the training vectors assigned to that cell.
4. Iteration: Repeat steps 2 and 3 until the average distance falls below a pre-set threshold.

The optimal codebooks were designed by using K-means quantiser process for training and grouping those coefficients by gain knowledge from a particular speaker. In each address of optimal codebook contained of a set of codevector, which represented for each group. In order of different characteristics in each speaker, the performance of

each optimal codebook also gives different effectiveness. The final 30minutes of testing speech was then fed through the optimal codebook to test the coding system in the environment of SDCS.

4. Implementation and result

In this experiment, the speech input data was generated from 2 males and 2 females. Each speaker contributed a total of 90 minutes of speech, of which 60 minutes (training data) used to extract the repetitiveness information and construct the optimise codebook, and the final 30 minutes (testing data) was used to assess coding system to quantify the error and success of the system.

To encode speech waveform, it is necessary to divided speech signal into frame. This work, speech is sampled at 8000 sampling per second and a frame size is 240 sampling per second. Each frame of speech characteristics was replicated to LPC coefficients, which only 10 LPC coefficients represented a speech frame. These coefficients were fed through LSP analysis to create the LSP coefficients. The reason to generate both LPC and LSP coefficient is to compare the effectiveness of them in the optimal codebooks. The optimal codebooks were designed by using K-means to train and classify those coefficients into groups by gain knowledge from a particular speaker. In each address of optimal codebook contains of a set of codevector, which represents for each group. The final 30 minutes of testing speech was then fed through the optimal codebook to test the coding system.

The average performances of these experiments are measured through each objective and subjective measurement. [5]

1. Signal to noise ratio (SNR). It works well as it calculates the mean square error between the original and synthesis speech waveform.
2. Mean Opinion Score (MOS). It is based on the opinion of a group of listeners for judging the quality of speech.

The results are shown in Fig 1-3. The set speech waveform of Male1, 2 and Female 1, 2 was different from the training process but still the same speaker for each codebook. It can be seem from the Fig. 2 that performance of LSP coefficients with a K-means optimal codebook is performed the best in this coding system environment and also in Fig. 3 it can seen that the size of the optimal codebook has an effectiveness to the quality of synthesis speech. Obviously the large codebook size provided the better quality than the small one. This technique can be reduce the bit rate in transmission by transmit only the address of optimal codebook

which can be seen in Fig. 1-3. It will use only 7 bits for transmit LPC and LSP parameters in codebook size 128.

5. Conclusion

This paper has been focused on exploiting the information of speech waveform for designing a Speaker Dependent Coding System. The result has been shown the performance of K-means algorithm works well with the LSP parameters especially in male voice (M1). Furthermore the quality of synthesis speech is also depending on the size of the codebook. These results were tested in the Code Excited Linear Predictive (CELP) coder that base on the FS-1016 standard. Using this technique, FS-106 can use only 7 bits per frame to transmit the LSP coefficients if the codebook size 128 is chosen. In this reason, this technique can be able to reduce the transmission rate of LSP coefficients around 20.5%.

As further work, the ability of K-means algorithm could be considered in order to define more specific performance in SDCS and the another methods are also required to more investigate to use in SDCS and compare with K-means techniques.

6. Reference

- [1]. W.T.K. Wong, et al, "Low rate speech coding for communications," *BT Technology journal*, Vol. 14, pp. 28-43, 1996.
- [2]. J. Dunlop, and D. G. Smith, *Telecommunications engineering: 2nd edition*, Chapman & Hall Publishers, 1992.
- [3]. J.Srinonchat, S. Danaher, J.I.H. Allan and A. Murray, "New bit rate CELP coder for speaker Dependent coding system," *IASTED International Conference on Artificial Intelligence and applications*, Innsbruck, Austria, Feb 2004.
- [4]. J.R. Deller, and J.G. Proakis, *Discrete-time processing of speech signals*, Macmillan Publishing, 1993
- [5]. A. Murray and S. Danaher, "Low bit rate speech coding for single speaker," *IEE Colloquium Techniques for Speech Processing and their Applications*, June 1994, pp. 2/1-2/6.
- [6]. J. Srinonchat, S. Danaher & A. Murray, "Address-Vector Quantisation applied to speech coding," *IEEE International Symposium on Signal Processing and Information Technology*, Darmstadt, Germany, Dec 14-17, 2003.

- [7]. K. Krishna, K.R. Ramakrishnan, and M.A.L Thathachar, "Vector quantisation using genetic K-means algorithm for image compression," in *Proc. International Conference of Information, Communications and Signal Processing*, Sept. 1997, pp. 1585-1587.

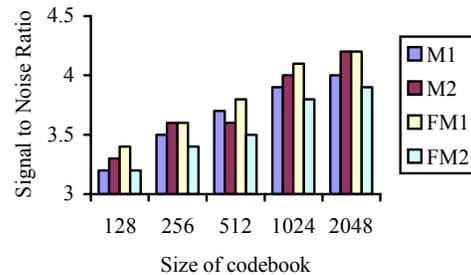


Figure 1: LPC with K-means optimal codebook

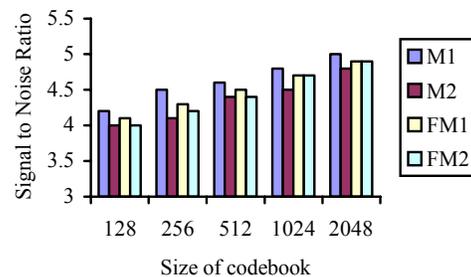


Figure 2: LSP with K-means optimal codebook

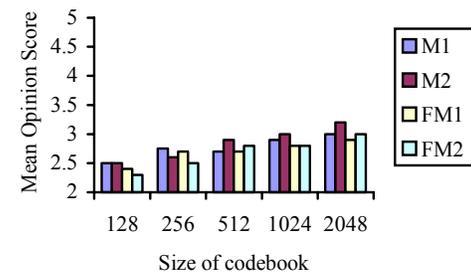


Figure 3: the quality of synthesis speech in different size of optimal codebook

Analysis of the IrFM Digital Payment System

Pi Huang and A. C. Boucouvalas

Multimedia Communications Research Group,
School of Design, Engineering and Computing,
Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, UK
{phuang, tboucouv}@bournemouth.ac.uk

ABSTRACT. IrFM, is a financial messaging specification developed by IrDA, that has great prospective for the future digital payment system. In this paper, we lay out a comprehensive analysis for IrFM. We develop a detailed IrFM transaction model which leads to derivation of the IrFM total connection time. The model allows evaluation of the impact of data rate, IrDA turnaround time and XID sending interval. Several design guidelines are calculated for achieving fast connections. Other major factors of the IrFM digital payment system are also investigated.

1. Introduction

The increase of financial products available on the market today, imposes a burden on consumers who often find themselves trying to manage several, incompatible, payment systems. In such environment, different payment details (receipts) have to be retained but are usually thrown away. However, vendors need payment receipts to accept any goods returns. Consumers are not the only losers in this scenario; vendors also have difficulties. As a group, vendors are experiencing \$500 million in annual unrecoverable loss [1] because they are unable to locate paper record receipts which must be produced as proof of purchase. Therefore there is a great desire to create a digital payment system which can reduce the costs and the risk of financial transactions dramatically.

Infrared is the perfect choice for the digital payment interface due to its low cost, security and short-range directional connection. A large number of portable devices on the market today have been equipped with IrDA ports ranging widely from mobile phones to PDAs [2]. IrDA envisioned the potential of IrDA devices in the digital payment applications and so has developed the Infrared Financial Messaging (IrFM) Point and Pay Profile specification [3]. IrFM provides a quick and seamless way for users of infrared enabled portable devices to pay for services and merchandise by beaming their “soft” credit cards, debit cards, checks, cash, or other financial instruments to a point of sale device (POS), ATM, vending machine or other compatible payment terminal.

Creating a digital payment system may not be successful unless it addresses the consumer’s needs

and provides clear benefits. Consumers will continue to use the tried and proven methods until an easier solution to their financial management requirements presents itself. A digital payment system should include the essential features in terms of security, reliability, interoperability, simplicity, as well as low transaction delay. In this paper, we are going to carry out an inclusive analysis on IrFM covering a wide range of issues of the digital payment system. First, we will have a brief introduction to the IrDA protocol stacks.

2. IrDA Protocol Stacks

The Infrared Data Association (IrDA) is an industry-based group that develops standards for indoor infrared communications. IrDA protocols are widely adopted as industry standards. By using IrDA, a proven standard which is already universally installed in many appliances could greatly benefit the digital payment applications.

Fig.1 shows the IrDA protocol stacks required for IrFM applications. A brief introduction to the IrDA stacks follows next.

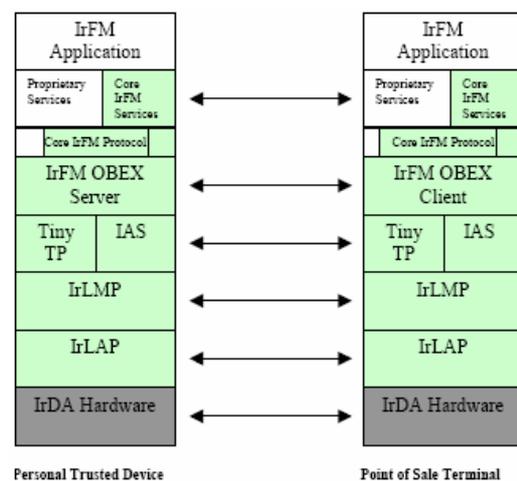


Figure 1, IrDA protocol stacks

- Physical Layer: Specifies optical characteristics, data encoding, and framing for various speeds [4].
- IrLAP: Link Access Protocol. It establishes the basic reliable connection [5].
- IrLMP: Link Management Protocol. It multiplexes

applications on the IrLAP connection [6].

- IAS: Information Access Service. It provides a “yellow pages” of services on a device [6].
- TinyTP: Tiny Transport Protocol. It adds per-channel flow control to keep things moving smoothly [7].
- OBEX: The Object Exchange protocol. It is a session protocol which provides easy transfer of files and other data objects [8].
- Core IrFM protocol performs many common commands needed for the IrFM applications [3].
- Core IrFM services support the basic IrFM transactions used in all scenarios [3].
- Proprietary services are defined by developers that provide support for different needs.

3. IrFM Connection Time

One of the most important criteria for the success of IrFM is low transaction delay. Consumers simply will not accept the digital system if they need to wait for a long time to complete their transactions. Various studies suggest that the acceptable delay should not be longer than 1 second from the time the user points the IrDA device at the POS until the price is displayed on the screen of the device [9]. As will be explained with the aid of Fig.3, after the price is displayed on the user’s mobile device, the user will be asked to choose the payment method and authenticate the payment. The system performance is mainly driven by the user thereafter. This paper will therefore focus on modelling IrFM to the point before the price is displayed and seeking the suitable parameters to achieve fast IrFM connection.

The IrFM transaction delay includes two components, the connection time and data transfer time. There have been many studies on improving IrDA throughput [10] and [11]. However, all the previous studies focus justifiably on reducing only the data transfer time by considering an infinite size packet to transmit (e.g. IrBurst application [12]). In contrast for IrFM transactions one exchanges relatively small packet sizes. Therefore, fast connection establishment plays a more important role by comparison to data transfer for IrFM applications. In the following, we are going to derive the time (T_{total}) from pointing the PTD to POS to displaying the price of the merchandise on the screen of the PTD.

In modeling IrFM, we are going to study an IrFM scenario namely ‘paying for groceries’. The scenario to be described deals payments with a credit card using a Personal Trust Device (PTD) which is an IrDA embedded device (mobile, PDA, etc.). As the groceries are being scanned, the customer approaches the register, powers on the PTD, and points it at the IR window on the register. The ‘e-wallet’ application automatically launches with a reliable IrFM

connection with the POS. The IrFM connection procedures and the detailed model for the whole scenario are elaborated in Fig.2 and 3 respectively.

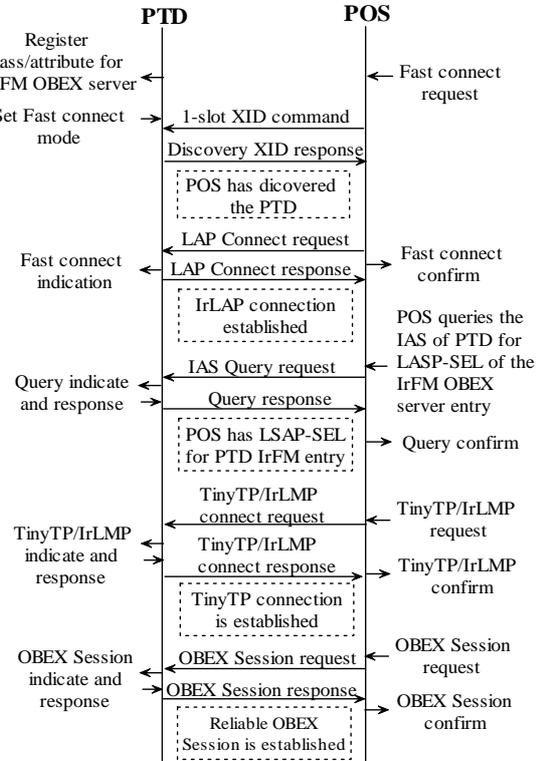


Figure 2, IrFM connection procedure

As defined in [3], all IrFM POS devices must send a ‘1-slot’ exchange station identification (XID) command at least every 300ms. PTD will respond with a 1-slot XID. We consider the longest XID interval (300ms) in this paper. A 1-slot XID command is 44bytes for both discovery and response mode. After device discovery, POS carries out the IrLAP connection. The IrLAP connection request/response (RQ/RP) command is 32bytes. POS queries IAS of PTD once the IrLAP connection is established. The query exchanges the IrLMP Link Access Service Point (LASP) of IrFM and other IrFM service information. The length of the IAS query packets are varied depending on the service name and content. To access only one object (IrFM) and assuming three attributes associated with the object, the IAS request and response packet is 250bytes and 202bytes respectively by considering integer attribute value [6] and its longest possible IAS packet length. The POS continues the IrFM connection by setting up the IrLMP/TinyTP connection. TinyTP connection parameters have the maximum length of 60bytes and lay inside the IrLMP connection packets. The IrLMP connection RQ/RP packets need 5bytes to perform its task. Therefore, it is 65bytes for each IrLMP RQ/RP packet.

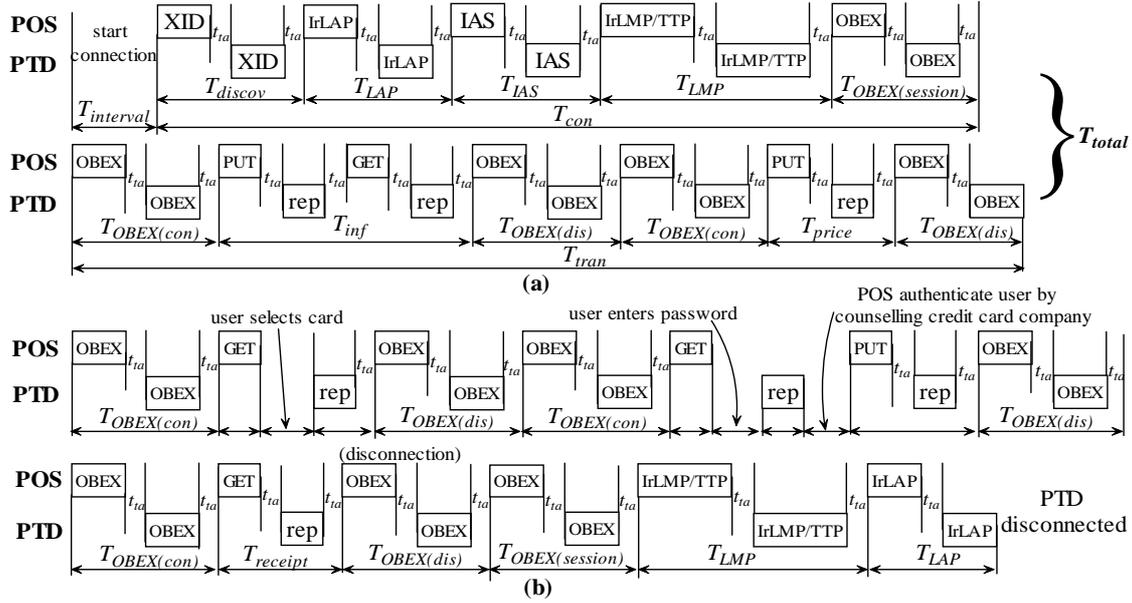


Figure 3, (a) From the beginning of the IrFM connection process to the price of the merchandise being displayed on the PTD; (b) From the payment method selection, payment authentication and payment receipt delivery to the connection termination.

The final step to establishing the IrFM connection is the creation of the OBEX session. We assume the OBEX session RQ/RP packets have the maximum allowed length of 260bytes. Once the OBEX session is set up, IrFM can start to perform its transactions between POS and PTD. IrDA physical layer adds 4 bytes overhead for each packet. For each link turnaround, the station has to wait for a ‘minimum turnaround time’ before transmitting packets. The minimum turnaround time is the time necessary for a receiver circuit to recover, following saturation by transmissions from the same device.

To display the groceries’ price onto the customer screen PTD, IrFM has to perform at least two services: Exchange IrFM Information and Put Price. These two services need to set up and tear down OBEX connection twice, as well as transmitting several OBEX operations. The length of OBEX connection and disconnection request packets are 12bytes and response packers are 8bytes. The length of the OBEX PUT and GET packets are depended on the application. Given the fact that IrFM transaction packets are very small and can always reside within one minimum OBEX packet with the length of 260bytes, we assume all the OBEX request packets (PUT/GET) are 260bytes. The OBEX response has the same length of OBEX header to perform acknowledgement which is 11bytes. The details of IrFM transactions for this scenario are given in the appendix.

Fig.3, illustrates the detailed packet exchange between POS and PTD. In this paper, we assume the processing time is small enough to be ignored. As shown in Fig.3, IrFM connection time is:

$$T_{con} = T_{discov} + T_{LAP} + T_{IAS} + T_{LMP} + T_{OBEX(session)} \quad (1)$$

The transmission time for the two necessary IrFM services is given by:

$$T_{tran} = 2T_{OBEX(con)} + T_{inf} + 2T_{OBEX(dis)} + T_{price} \quad (2)$$

The total time from pointing the PTD at the POS to displaying the price of the groceries is the sum of T_{con} and T_{tran} :

$$T_{total} = T_{interval} + T_{con} + T_{tran} \quad (3)$$

The total connection times for different data rates and turnaround times are shown in Table I. The table shows, as expected, T_{total} reduces significantly with the increasing data rate and also benefits from a smaller value of t_{ta} . T_{total} is less than 1s when $C \geq 38.4\text{Kbit/s}$.

Fig.4, is a plot of the results of the percentage time allocation of $T_{interval}$, $t_{ta}(total)$, T_{con} (excluding t_{ta}) and T_{tran} (excluding t_{ta}) for data rates from 1Kbit/s to 1Mbit/s. As all the time portions for the IrFM connection are considered, the sum of the four percentages always equals to 100%. T_{con} and T_{tran} are the two dominant factors at low data rates. When combined occupy more than 90% of the overall time at $C=1\text{Kbit/s}$. $T_{interval}$ and $t_{ta}(total)$ gradually increase their percentages and commit over 90% at $C=1\text{Mbit/s}$.

As shown in Table I and Fig.4, high data rate C can reduce the connection time but only leads to trivial improvement when $C > 100\text{Kbit/s}$. The fastest speed defined in SIR 115.2Kbit/s [5] and t_{ta} of 1ms are adequate for the IrFM applications. To further reduce the connection time, a smaller value of $T_{interval}$ should also be applied especially for the high data rate links.

Table I: T_{total} by implementing different data rate C and minimum turnaround time t_{ta} .

C (bit/s)	t_{ta}	10ms	1ms	0
9600		2.40 s	2.19 s	2.17 s
19200		1.47 s	1.26 s	1.23 s
38400		1.00 s	0.79 s	0.77 s
57600		0.84 s	0.64 s	0.61 s
115200		0.69 s	0.48 s	0.46 s

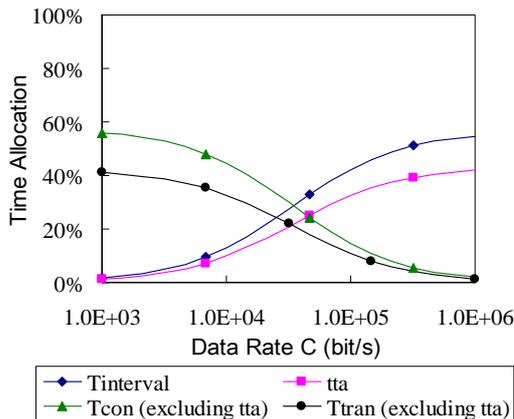


Figure 4, percentages of time allocation for four different elements of IrFM connection time

4. Security and Reliability

IrFM generally has a very high level of security. Digital signatures, password authentication, and session encryption can be easily implemented over IrFM transaction. Moreover, as the infrared communication is directed by the user's 'point and shoot' behavior and is strictly a 'peer to peer' communication, it is very difficult to intercept and naturally an ideal platform to exchange confidential information. The only concern for IrFM is the 'express payment' applications. IrFM enables 'express payment' [3] in order to reduce the payment time in small transactions (vending machines, tollbooth etc.). In these cases, user doesn't need to enter password to authenticate the payment. However, this feature is a potential vulnerable point for financial fraud but the amounts exchanged are small in such cases. The IrFM developers should take the extra precautions on designing 'express payment' applications.

IrFM applications are built around the concept of services and each service is uniquely identified to ensure the transaction reliability. A service is an entity, which performs transactions. An IrFM scenario is achieved by performing transactions between one or more services. To operate a service, the same service must appear on both PTD and POS. Services are uniquely identified by a 128bits Universally Unique Identifier (UUID). If the IrFM transaction is accidentally halted (e.g. PTD runs out of battery), the transaction can be resumed by quoting the same UUID.

5. Simplicity and Interoperability

As a digital payment system, IrFM requires very little hardware support and is very simple to use. To use IrFM, user only needs to have an IrDA embedded PTD and install the IrFM software suit in the PTD. After a few steps of adding credit card information, user can easily include all the details of their cards into a single PTD. For the POS, it is also painless to set up IrFM. POS only needs to add one IrDA port, run the IrFM software and send the IrFM transactions to the payment point where it is used to authenticate the physical credit card payments. No additional network infrastructure is required.

IrDA protocol stacks have already been installed in millions of devices over years and are proven over time. IrDA devices have excellent compatibility between different manufacturers. As the only infrared standard for digital payment, IrFM is expected to have a very high level of compatibility.

6. Conclusion

In this paper, we have carried out a systematic analysis for the IrFM 'e-payment' system. A simple model of IrFM is developed by considering the whole IrDA protocol stacks. Based on this model, the average total IrFM connection time is investigated. Several valuable design guidelines are given on data rate, IrDA turnaround time and XID sending interval to achieve fast connections. Furthermore, other major factors of IrFM including security, reliability, simplicity and interoperability are also studied. From the analysis, we show that IrFM is an excellent digital payment system by offering fast connectivity and has many desired features requiring very few hardware requirements.

References

- [1]. Maness, Terry S. & Zietlow, John T., "Short Term Financial Management", Dryden Press, 1997, p. 227
- [2]. Williams S., "IrDA: Past, Present and Future", IEEE Personal Communications, 7(1), 11-19 (2000)
- [3]. IrDA, Infrared Financial Messaging Point and Pay Profile, version 1.0 (Dec, 2003)
- [4]. IrDA, Serial Infrared Physical Layer Specification Version 1.1.
- [5]. IrDA, Serial Infrared Link Access Protocol, Version 1.1.
- [6]. IrDA, Link Management Protocol, Version 1.1.
- [7]. IrDA, Tiny Transport Protocol, Version 1.1.
- [8]. IrDA, Object Exchange Protocol (IrOBEX), Version 1.3.
- [9]. H.R. Damon, R.J. Brown, L. Faulkner, white paper "Creating an End-To-End Digital Payment System" Oct. 1999
- [10]. V. Vitsas and A.C. Boucouvalas, "Throughput Analysis of the IrDA IrLAP Optical Wireless Link Access Protocol" 3rd Conference on Telecoms, Conftele 2001, pp. 225-229
- [11]. A.C. Boucouvalas and V. Vitsas, "Optimum window and frame size for IrDA links", Electronics Letters 2001, 37(3):194-196.
- [12]. IrBurst, Version 0.6, Jan, 2004.

Appendix:

Detailed IrFM transaction for the 'paying groceries' scenario:

Exchange IrFM Information Transaction

POS creates a directed OBEX connection to the Core_IrFM_Service UUID

POS performs a GET operation for the PTD IrFM information.

POS performs a PUT operation of its IrFM information.

POS performs a disconnect of the OBEX directed connection.

Next is the Get Desired Loyalty Service Transaction (optional)

POS creates a directed OBEX connection to the Core_IrFM_Service UUID.

POS performs a get operation to get the selected loyalty service passing its own services

POS performs a disconnect of the OBEX directed connection

Next is the loyalty Transaction (optional)

POS creates a directed OBEX connection to the Loyalty service UUID

POS performs a GET operation for its loyalty information

POS performs a disconnect of the OBEX directed connection

Next is the Put Price Transaction

POS creates a directed OBEX connection to the Core_IrFM_Service UUID.

POS performs a PUT operation for the price

information

POS performs a disconnect of the OBEX directed connection.

Next is the Get Desired Payment Service Transaction.

POS creates a directed connection to the Core IrFM service UUID

POS performs a GET operation for the selected Payment service passing its own service list.

POS performs a disconnect of the OBEX directed connection

Next is the VISA Credit Payment Transaction.

POS creates a directed OBEX connection to the selected Visa Credit service UUID.

POS performs GET and PUT operations to achieve the payment transaction as defined and controlled by VISA.

POS performs a disconnect of the OBEX directed connection.

Next is the detailed receipt Transaction (optional)

POS creates a directed OBEX connection to the Core_IrFM_Service UUID

POS performs a PUT operation of the receipt object.

POS performs a disconnect of the OBEX directed connection.

At this point, the scenario is finished.

POS closes the OBEX session

POS disconnects the Tiny TP connection to the IrFM OBEX server

POS disconnects the IrLAP connection.

Arbitrarily Shaped Video Object Compression with Weighted Finite Automata

Shumu Li (S.Li@hw.ac.uk) R.J Clarke (R.J.Clarke@hw.ac.uk)
Department of Electrical, Electronic and Computer Engineering
Heriot-Watt University, Scotland, United Kingdom

Abstract An important issue in object-based video coding is the efficient description of arbitrarily shaped image segments or complete video objects (VOs). This paper outlines a solution to this task using Weighted Finite Automata (WFA). WFA is a relatively new technique for image coding, which explores image self-similarity and has been reported to be able to achieve very low bitrates. We develop a new algorithm called Shape Adaptive WFA (SA-WFA) especially for the purpose of arbitrarily shaped VO coding. The idea is that for those partitioned blocks lying within boundary areas containing both VO parts and background parts, a classification process is added in before WFA encoding. This is followed by the adaptive orthogonalization as well as a Matching Pursuit (MP) process. Experimental results prove that SA-WFA is more suitable for arbitrarily shaped VOs coding compared with existing WFA techniques.

1 Introduction

The booming computer techniques, in software as well as hardware, facilitate potentially even more enormous visual product consumption than ever before including home cinema, image communication, digital multimedia broadcasting, on site surveillance, etc. The ISO/IEC Moving Picture Experts Group (MPEG) is set up to standardize this process and has achieved great success. For the first time, the state-of-the-art MPEG-4 [1] brings object-based video compression (OBVC) into international standardization efforts. This object-based representation approach, where a scene is modeled as a composition of objects (both natural and synthetic), with which the user may interact, is at the heart of the MPEG-4 technology. Encoding/decoding of segmented video objects (termed VOs), such as talking people, moving ships, etc., with arbitrary shape and inner texture information is desirable to achieve such object-based functionality. In this paper, we address this topic.

So far existing arbitrarily shaped image segment coding techniques in the literature don't significantly differ from their predecessors. The reason is that traditional frame-based codec schemes have been highly developed and in most cases can be used for shape-variant coding; albeit certain adaptations are necessary for higher efficiency. One way of such adaptations is, for waveform coding techniques, prior to any transform, so-called padding, where the pixels located outside the VO are assigned certain luminance values, as in [2]. Another way is using shape adaptive techniques such as the shape-adaptive DCT [3]. Region-based segmented coding techniques, such as [4], are another options. However, most padding processes introduce extra unnecessary frequency coefficients as an overhead; shape-adaptive transform is very computationally expensive; region-based segmentation re-

quires too much pre-processing besides introducing the bitrate burden of shape coding. In one word, emerging techniques for arbitrarily shaped image segment coding are far from mature and need researchers to develop novel methods. In this paper we explore the Weighted Finite Automata (WFA) method for this task.

2 Image compression with WFA

Image compression using WFA was introduced by Culik and Kari in 1993 [5] and later developed by both themselves and other researchers, such as Hafner et al [6] and Katritzke [7]. The compression algorithm converts image I into automaton A , where A can be represented with fewer bits and can therefore be regarded as a compact representation of the image. The automaton graph reflects relationships between various parts of the image. According to Jiang et al [8], acyclic WFA as in [6] can always be merged into linear WFA as in [5] but not vice versa.

Formally, WFA can be specified by automaton $A = (S, \mathcal{S}, W, I, F)$, where

$S = \{s_0, \dots, s_{|S|-1}\}$, a finite set of states;

$\mathcal{S} = \{a_0, \dots, a_{|\mathcal{S}|-1}\}$, a finite alphabet;

$W: S \times \mathcal{S} \times S \rightarrow \mathbb{R}$, a weighted function;

$I: S \rightarrow \mathbb{R}$, an initial distribution;

$F: S \rightarrow \mathbb{R}$, a final distribution.

In terms of image compression application described in this paper, state s_i ($0 \leq i \leq |S|-1$) responds to image segments in current frame; Σ is a set consisting of two elements 0 and 1 (thus we use bintree segmentation); W is a weighted parameter matrix with: $(W_a)_{ij} = W(s_i, a, s_j)$, $a \in \mathcal{S}$, $s_i, s_j \in S$; I and F are $|S|$ -dimensional vectors.

Then the current frame can be calculated by formula $f_A(a_0 \dots a_{k-1}) = I^t (W_{a_0 \dots} W_{a_{k-1}}) F$, for $k \leq N$ and $a_0 \dots a_{k-1} \in \mathcal{S}^k$, $\mathcal{S}^k := \hat{E}_n \approx \mathcal{S}^n$. This is actually a theoretical decoding process since one by one we can calculate the luminance values for the pixels uniquely located by Σ^k . However, in practice this method is not efficient and thus we develop other ways for WFA image compression.

In order to construct a WFA for a given image, we have to construct the vectors I, F and the matrices defined by $W_a (a \in \mathcal{S})$. Since the vector I depends on the characteristic of the WFA adopted, we specify this vector for each algorithm. Vector F consists of the average brightness of all the state images. As for matrices W , it comes from the process where sub-segments of the state images are linearly combined from other state images using the following equation [6]:

$$f_i(a_0, \dots, a_{k-1}) = \sum_{j=0}^{|S|-1} (W_{a_0})_{i,j} f_j(a_1 \dots a_{k-1}).$$

By this method, the self-similarity within image is exploited to eliminate spatial redundancy.

It is said [7] that the first algorithm on WFA image compression proposed by Culik and Kari is as follows:

```

/* param φ is the image to be represented. */
WFA operateImage (φ);
{
  create a state which represents the whole input image;
  while (there are unprocessed states)
  {
    choose an unprocessed state;
    operate on all quadrants of the assigned state image:
    {
      if (the current quadrant can be linearly combined by
      existing states)
        store these coefficients;
      else {
        create a new state representing this quadrant;
        store a pointer to this state;
      }
    }
  }
  I=(1,0,...,0);
  return the resulting WFA;
} //end operateImage

```

Algorithm 1 Construction of the WFA

Although there is no detailed implementation description for this algorithm, it points out the basic idea for WFA image compression. In the next section, we will discuss the WFA construction in more detail, especially for arbitrarily shaped image segment processing.

3 Arbitrarily shaped image coding with WFA

3.1 Image partition

Tree structure partition is an efficient way for image segmentation representation. Two typical structures are quadtree and bintree with four and two equal-sized sons respectively for each father node. All the quadrants are denoted by a sequence of concatenated letters from alphabet Σ (See Fig. 1).

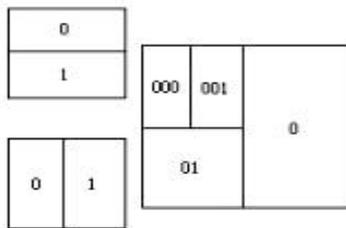


Figure 1 Locating image segments with bintree

Empirically, bintree partition generates larger approximated leaves at the bottom of the segmentation and thus is more efficient and suitable for image compression.

3.2 WFA construction

In our work, the WFA construction is done as Algorithm 2:

```

/* param φ is the whole input image to be represented and the algorithm starts with operateImage(φ)*/

```

```

WFA operateImage (e);
{
  operate on all quadrants of the assigned state image e:
  {
    if (the current quadrant can be linearly combined by existing
    states)
      {
        store these coefficients;
      }
    else {
      create a new state a representing this quadrant;
      operateImage (a);
    }
  }
  I=(0,0,...,1);
  return the resulting WFA;
} //end operateImage

```

Algorithm 2 WFA construction in the present work

The differences from Algorithm 1 are: *a)* the generated states are not put into the WFA tree until all the sub-images of this state are approximated; *b)* the tracing order of image quadrants is depth first search (DFS) rather than breadth first search (BFS). *c)* at the very beginning, the WFA tree has to be filled with some initial state basis. We choose cosine basis images.

An important part of the WFA image compression is to choose an appropriate method for state image approximation. We utilize the matching pursuit (MP) method [9], which expands a signal using an over-complete dictionary of functions and suits very low bitrate video coding. Note that for current quadrant approximation, it's not practical to use all the state images existing in WFA tree due to computational complexity. Similarly to fractal image coding, only selected state images play this role, which are termed domain images. A short introduction to MP is given below.

Let $x=\{x[n]; n=0, 1, \dots, N-1\}$ be a discrete-time signal in a finite dimensional Hilbert space $H=R^N$, with the inner product of $x, y \in H$ defined as $\langle x, y \rangle = \sum_n x[n] \cdot y[n]$, and the norm as $\|x\| = \langle x, x \rangle^{1/2}$.

Given a dictionary $D=\{\mathbf{f}_\beta; \beta \in G\}$ of basis functions in H , normally with imposed $\|\mathbf{f}_\beta\|=1$ and $span(D)=H$, the goal is to obtain an approximate decomposition of x with linear combinations of a small number of basis functions such that $x = \sum_{b \in G} a_b \mathbf{f}_b + R$ (\mathbf{f}_β are the basis

functions in D ; a_β is the coefficient of decomposition; R is the residual). This search of \mathbf{f}_β is done recursively till R is smaller than a threshold ϵ and for each search the \mathbf{f}_β with biggest inner product norm of $\|\langle x, \mathbf{f}_\beta \rangle\|$ is selected. By storing the β and corresponding a_β , which normally have smaller size, x can be approximated and compressed within some distortion

threshold. In our case x represents an image vector. It's well known that higher efficiency can be achieved if the dictionary D (or the domain) is composed of an orthogonal basis.

With respect to arbitrarily shaped video object coding, we put such a VO into an empty frame of the same size as the source video. With no need of shape coding in this way because of the prior knowledge of the empty frame background, the desired VO can be extracted from the decoded VO plane.

While Algorithm 2 is implemented for VO coding, due to the randomly arbitrarily shapes in boundary blocks, the WFA construction won't start to approximate partitioned blocks until it produces very small sized image segments. This leads to generating more unnecessary states than it should have. Our idea is to approximate these edge blocks earlier (before further partition). To do this, we add in a process to judge whether the current quadrant is an edge block or not and which type it belongs to if it is. We observe that most edge blocks can be categorized into 16 types as in Fig. 2. Only the segmentation whose bintree level is within some range is considered for this process. Accordingly, the MP process is adapted to such edge blocks.

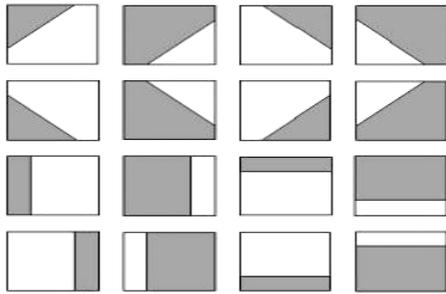


Figure 2 Sixteen types of edge blocks (shadowed area lies in VO)

constructed states is compared with that of existing WFA. Two individual samples of this comparison are shown in Fig. 5 and Fig. 6 with the precondition of preset equal decoded quality; the overall comparison is shown in Fig. 7. The improved rate-distortion performance is shown in Fig. 8. We find that for very high image quality coding, SA-WFA can not outperform WFA anymore. This is because in that case most approximation has to be done in small blocks for both and thus the improved compression ratio from the shape adaptive property in SA-WFA can not compensate for the overhead of flag coding for boundary blocks. But we think this is not a problem because normally object-based image coding techniques outperform regular frame-based image coding only for low or middle bitrate. For very high image quality coding, frame-based techniques are said to be better. Our shape-adaptive algorithm is exactly in line with this experience.



Figure 3 Original image Claire

3.3 Storage of the WFA compressed image

The information needing to be saved for future image decoding includes: bintree structure (a sequence of true/partition or false/approximation); flags for edge block types (four bits each type) for those quadrants that are considered as edge blocks; the W_a matrix containing quantized coefficients a_β and indices. All this is entropy encoded to eliminate statistical redundancy.

4 Experimental results

In this section, we show some experimental results of the proposed SA-WFA. The source image Claire (Fig. 3) is a pre-processed result using the segmentation technique we proposed in [10]. A SA-WFA decoded image with PSNR 35.92dB and compression ratio 27.62 is shown in Fig. 4. In line with what we analysed before, the proposed SA-WFA can reduce the amount of states needed to be constructed efficiently and thus improve the coding efficiency and decrease the memory consumption. The number of SA-WFA



Figure 4 SA-WFA decoded image (PSNR 35.92dB, compression ratio 27.62)

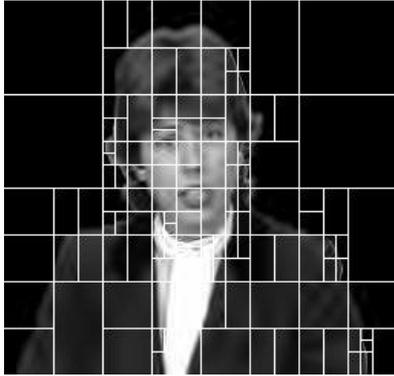


Figure 5 Coded image with marked 124 states (preset decoded quality is as same as in Fig. 6)

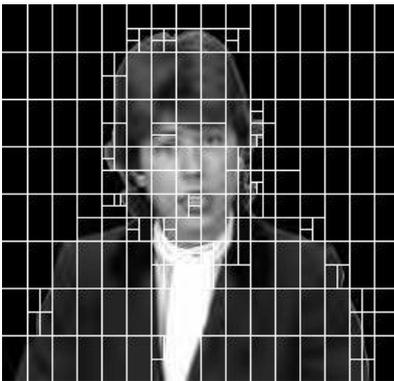


Figure 6 Coded image with marked 218 states (preset decoded quality is as same as in Fig. 5)

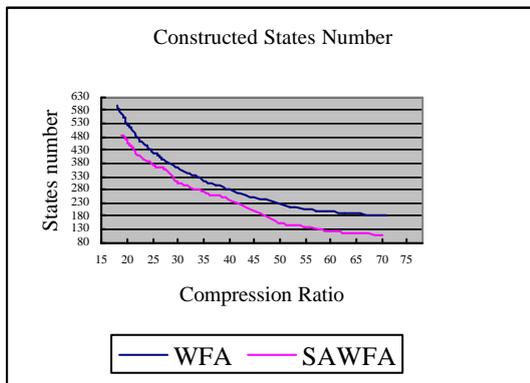


Figure 7 Constructed states number comparison

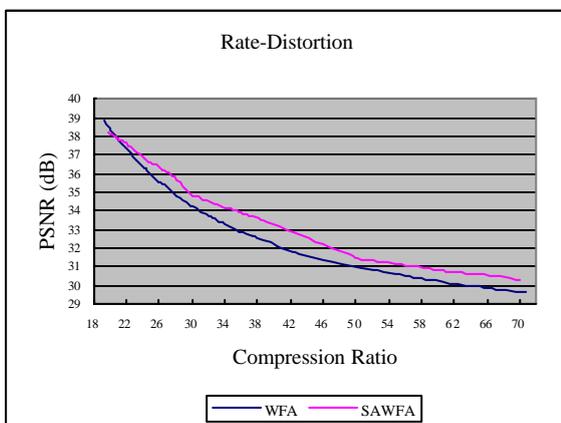


Figure 8 Rate-Distortion comparison

5 Conclusions

In this paper, we propose a new SA-WFA technique for arbitrarily shaped video object compression. We introduce a shape-adaptive process, for both partition and matching pursuits, into the existing WFA codec. Experimental results prove that for low or middle bitrate arbitrarily shaped image coding, this process improves compression efficiency and reduces the number of constructed states greatly. By exploring efficient automata methods for image coding with the unique shape adaptive characteristic, SA-WFA suits object-based video compression, towards which digital video products development heads.

Acknowledgements

We are very grateful to the support from ORS (UK) funding and a James-Watt scholarship (Heriot-Watt University). We also would like to acknowledge the help and informative advice from Dr. F. Katritzke.

References

- [1] F. Pereira, MPEG-4: Why, what, how and when?, *Signal Processing: Image Communication* 15 (2000) pp. 271-279.
- [2] G. Shen, B. Zeng and M. Liu, Arbitrarily shaped transform coding based on a new padding technique, *IEEE transactions on circuits and systems for video technology*, Vol. 11, No.1, pp. 67-79, Jan. 2001.
- [3] T. Sikora and B. Makai, Shaped-adaptive DCT for generic coding of video, *IEEE transactions on circuits and systems for video technology*, Vol. 5, No. 1, pp. 59-62, Feb. 1995.
- [4] A. Kaup and T.Aach, Coding of segmented images using shaped-independent basis functions, *IEEE transactions on image processing*, Vol. 7, No. 7, pp. 937-947, July 1998.
- [5] K. Culik and J. Kari, Image compression using weighted finite automata, *Computers and graphics*, vol. 17, no.3, pp.305—313, May/June, 1993.
- [6] U. Hafner, J. Albert, S. Frank and M. Unger, Weighted finite automata for video compression, *IEEE journal on selected areas in communications*, Vol. 16, No. 1, pp. 108-119, Jan. 1998.
- [7] F. Katritzke, Refinements of data compression using weighted finite automata, Ph.D thesis, University of Siegen, 2001.
- [8] Z. Jiang, O. Vel and B. Litow, Unification and extension of weighted finite automata applicable to image compression, *Theoretical computer science* 302, pp. 275-294, 2003.
- [9] S. G. Mallat and Z. Zhang, Matching pursuits with time-frequency dictionaries, *IEEE transactions on signal processing*, Vol. 41, No. 12, pp. 3397-3415, Dec. 1993.
- [10] S. Li, R. J. Clarke, Automatic moving object contour extraction based on 2-D spatiotemporal entropic segmentation, *Proceedings of CSNDSP 2002*, pp. 268-271.

ARTIFICIAL NEURAL NETWORK ASSESSMENT OF THE PROGNOSTIC VALUE OF TUMOUR PROMOTER GENE S100A4 AND TUMOUR SUPPRESSOR GENE NM23 IN CONJUNCTION WITH TUMOUR GRADE, TUMOUR SIZE AND STEROID RECEPTOR STATUS OF BREAST CANCER.

S.R. Grey¹, S.S. Dlay¹, B.E. Leone², F. Cajone³ and G.V. Sherbet^{1,4}

¹School of Electrical, Electronic and Computer Engineering, University of Newcastle upon Tyne UK; ²Ospedale di Monza, Monza, Italy; ³Istituto Di Patologia, University of Milan, Italy; ⁴Institute for Molecular Medicine, Huntington Beach, CA, USA.

ABSTRACT

Artificial neural network assessment of breast cancer markers can reliably tell us the role and efficacy of genes, steroid receptors and conventional tumour markers as predictors of breast cancer metastasis. The expression of tumour promoter gene S100A4, metastasis suppressor gene nm23, oestrogen and progesterone receptors, and tumour grade and size have been investigated for their potential to predict breast cancer progression. The molecular and cellular data have been analysed using artificial neural networks to determine the potential of these markers to predict the presence of metastatic tumour in the regional lymph nodes. This study of just 47 samples shows that tumour grade and size are poor predictors. The relative expression of S100A4 and nm23 genes is the single most effective predictor of nodal status. Inclusion of oestrogen- and progesterone-receptor status with tumour grade and size markers improves prediction; however, there may be some overlap between steroid receptors and molecular markers. This could aid the clinician in determining whether invasive procedures of axially node dissection can be obviated and whether conservative forms of treatment might be appropriate in the management of the patient.

Keywords: artificial neural networks, breast cancer, metastasis associated gene, metastasis suppressor gene, nm23, nodal spread prediction, S100A4

INTRODUCTION

The state of breast cancer progression is conventionally determined by assessing the dissemination of the tumour to axillary lymph nodes and the number of nodes involved, and also the criteria of tumour size, cellularity and histological grade [1], which form the basis of breast cancer management. The presence of tumour cells in the lymph nodes, as indicated by histology of lymph nodes obtained by axillary dissection, is the most reliable indicator of malignancy and prognosis. There have been many efforts to identify non-invasive surrogate methods that can predict the potential spread of cancer to the lymph nodes. Among notable molecular markers employed for this purpose are the S100A4 tumour promoter gene and the putative anti-metastasis gene nm23 [2, 3].

Oestrogens are known to stimulate the growth of a variety of target tissues. They bind to specific intracellular receptors, the oestrogen receptors (ER), which are then translocated into the nucleus where they initiate the transcription of responsive genes resulting in appropriate physiological function. The ER gene is silenced in a number of cancers. The silencing of the ER gene often correlates with disease progression. The absence of ER in breast cancer is regarded as an indicator of poor prognosis, since ER-tumours are resistant to anti-oestrogen therapy, continue rapid growth and result in a poor outcome for patients.

Statistical techniques have traditionally been employed to assess the significance of individual or a set of tumour markers in the prediction of tumour progression and prognosis. Artificial neural networks (ANN) have been demonstrated to be capable of analysing the complex relationships that subsist between the multitude of potential tumour markers currently being studied in many cancer centres and tumour progression and prognosis. Indeed, ANN has proved to be a powerful alternative to the statistical methods of Cox regression models, logistic regression and discriminate function analysis [4, 5]. In light of the intrinsic merits of ANN, here we have employed the same to analyse the influence of the expression of S100A4, nm23, ER and PgR, as well the traditionally assessed features of tumour size and histological grade, on the spread of breast cancers to axially lymph nodes.

MATERIALS AND METHODS

The clinical and molecular marker data, encompassing tumour histological grade, tumour diameter, nodal status, ER/PgR expression and the expression status of S100A4 and nm23 pertained to a series of forty-seven breast cancer specimens [2]. The present series of experiments analyses, by using artificial neural networks (ANN), the efficacy of these molecular markers to predict nodal spread of breast cancer.

ANN Design and Training

The artificial neural network (ANN) used to predict nodal metastasis is based on a feed forward multi-

layer perceptron (MLP). The network has three layers with a single output neuron in the final layer as it is proven that such layouts are universal approximators [6]. Numerical input vectors (patterns) are presented at the input layer, and activity flows through the network to the output layer. Connections have a numerical weight value associated with them, and the signal transmitted via a connection is multiplied by the weight value. Each unit computes some function of the sum of its weighted inputs, and transmits the result through its output connections. In this way an MLP implements a mapping from input space to output space.

The Back propagation algorithm (BPA) [7] provides a means of training the network to perform supervised learning tasks. Supervised learning starts with the presentation of a set of example input patterns to a system. The output is then compared with the known correct output for each pattern, and some adjustments are made so as to improve the response of the system to those patterns. In an implementation, the goal is not simply to learn the patterns in the training set (which could be accomplished by a lookup table), but to learn in some sense the characteristics of the mapping, so as to be able to generalise (i.e. produce correct responses for patterns that the network has not been trained on).

During training effort was made to prevent over training and keep the ANN as generalised as possible [8]. The Levenberg-Marquardt training algorithm [9] was used to perform gradient descent. Batch training proved to be the most efficient method of training with care being taken to avoid local minima.

Molecular and cellular marker data relating to breast cancer

From the data set of forty-seven women, fifteen are discarded due to either unknown nodal status or a lack of one of the measured elements. To efficiently analyse the performance of the network the data set of thirty-two elements is split into a training set and a testing set, both of sixteen elements each. The data elements for each set are selected randomly from the parent set.

Histological grading has been numerically represented using a scale of one through to three. ER and PgR levels are quantised into four levels due to the nature of the data and to provide the most effective input to the ANN. Nm23 and S100A4 have a threshold value of 0.01. This value is also assigned to those occurrences where the level of either marker was not detectable. This is particularly useful when calculating the ratio S100A4/nm23.

Performance of the network is analysed by considering the number of patients correctly classified out of the total number of classifications made. To prevent a particular selection of data from influencing performance due to a fortuitous selection, this method was performed three times, forming three sets of training and testing data.

RESULTS AND DISCUSSION

Each network consisted of a different selection of the inputs available to analyse the characteristics of each input. Table 1 shows the combinations of inputs for each ANN. Groups A-O analyse the effects of combining tumour grade and size with molecular marker expression. Groups P-U analyse the predictive value of S100A4/nm23 gene expression ratio and how other markers affect it.

Group comparisons:

1. [F] Tumour histological grade/diameter possess only 60% predictive value.
2. [E, F] indicates that when both S100A4 and nm23 were not detected, inclusion of ER/PgR improved nodal prediction by 13%.
3. [L, M, and P, Q] indicates that when S100A4 is detected, ER/PgR has a negative effect on prediction (-4%).
4. [J, K and J, E] Inclusion of ER/PgR with nm23 does not influence prediction, and the addition of nm23 to ER/PgR is a very small improvement (2%).
5. [F, Q and D] However, when both S100A4 and nm23 are included with ER/PgR, the prediction rate increase by 19-21%.
6. [G, M] and [G, J] show that in the absence of ER/PgR, inclusion of S100A4 improves prediction by 11% and inclusion of nm23 by 15%.
7. [O, P] show that the ratio of S100A4/nm23 expression or the expression of the individual genes is the single best predictor of nodal status. Inclusion of tumour grade [R] does not affect prediction. But inclusion of tumour size [S, T] reduces prediction by gene expression ratio.

Summary of findings

1. Tumour grade and size of the tumour were poor predictors of nodal spread. Small tumours of low grade with no detectable tumour cells in the regional lymph node status are often linked with good prognosis. But the need exists to sample a number of axillary lymph nodes for histology. Here we assess the effects of combining molecular markers with tumour grade and histology on the prediction of nodal status.
2. The expression of S100A4 and nm23 genes relative to each other or the expression of the individual genes is the best single predictor of nodal status. The inclusion of tumour grade with gene expression ratio does not influence prediction. Inclusion of tumour size has a negative effect on prediction; so does inclusion of both grade and size confirming the negative effect of tumour size.

3. ER/PgR status enables one to improve prediction of nodal involvement.
4. Inclusion of S100A4 further improves prediction. Small tumours of low grade with no detectable tumour cells in the regional lymph node status are linked with good prognosis.
5. Nm23 alone does not contribute to prediction, in the presence of ER/PgR.
6. The inclusion of both S100A4 and nm23 together with ER/PgR greatly improves prediction of nodal status. However, this performance is also seen without ER/PgR when both S100A4 and nm23 are present.
7. When ER/PgR are not assessed, S100A4 and nm23 both individually influence predictive ability. This suggests that ER/PgR influence might overlap that of nm23. This could be a reason why nm23 apparently does not contribute to prediction when ER/PgR are present [Item 5 above]. Compatible with this is the recent findings that ER α induces the expression of nm23 in breast cancer cells in culture. ER seems to enhance nm23 gene transcription via ER-responsive elements that occur in the promoter region of the ER gene [10].
8. It is uncertain why the inclusion of ER/PgR reduces predictive function of S100A4, albeit by a small margin of 4% [L, M]. ER expressing tumours have been shown to be epidermal growth factor receptor (EGFr) negative. Furthermore, breast cancer cells that are EGFr+ also express S100A4 at a high level [11]. This suggests that an inverse relationship exists between S100A4 and ER, for which there is now direct evidence [12]. Therefore it is possible that the effects of S100A4 are antagonistic to those of ER/PgR and counteract the predictive value of the steroid receptors.

CONCLUSIONS

This study using ANN suggests that relative expression of S100A4 and nm23 genes is the single most effective predictor of prognosis. Inclusion of ER/PgR with traditional markers improves prediction, however, there may be some overlap between ER/PgR and molecular markers. This study also underscores the power of ANN techniques to predict the potential of primary breast cancers to spread to axillary lymph nodes. This could aid the clinician in determining whether invasive procedures of axially node dissection can be obviated and whether conservative forms of treatment might be appropriate in the management of the patient.

REFERENCES

- [1] Fisher B, Ravdin RG, Ausman RK, Slack BNH, Moore GE, Noer RJ. Surgical adjuvant chemotherapy in cancer of the breast: Results of a decade cooperative investigation. *Ann Surgery* 1996; **168**: 337-356.
- [2] Albertazzi E, Cajone F, Leone BE, Naguib RNG, Lakshmi MS, Sherbet GV. The expression of metastasis associated genes *mts1* and nm23 in carcinoma of the breast are related to disease progression. *DNA Cell Biol* 1998; **17**: 335-342.
- [3] Sherbet GV. *Calcium signalling in Cancer: The role of calcium binding proteins in cell proliferation, invasion and metastasis*. CRC Press, Boca Raton, FL 2001.
- [4] Seker H, Odetayo M, Petrovic D, Naguib RNG, Bartoli C, Alasio L, Lakshmi MS, Sherbet GV. A fuzzy measurement-based assessment of breast cancer prognostic markers. *Proc IEEE EMBS Int Conf Information Tech Appl in Medicine, 2000*, pp. 174-178.
- [5] Seker H, Odetayo M, Petrovic D, Naguib RNG, Bartoli C, Alasio L, Lakshmi MS, Sherbet GV. Assessment of nodal involvement and survival analysis in breast cancer patients using image cytometric data: Statistical, neural and fuzzy approaches. *Anticancer Res* 2002; **22**: 433-438.
- [6] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Networks* 1998; **2**: 359-366.
- [7] Gurney K. An introduction to neural networks. UCL Press, London 1997.
- [8] Rafiq MY, Bugmann G, Easterbrook DJ. Neural network design for engineering applications. *Computers and Structures* 2001; **79**: 1541-1552.
- [9] Bishop CM. *Neural networks for pattern recognition*. Oxford University Press, Oxford 1995.
- [10] Lin KH, Wang WG, Wu YH, Cheng SY. Activation of anti-metastatic nm23-H1 gene expression by oestrogen and its alpha receptor. *Endocrinology* 2002; **143**: 467-475.
- [11] Sherbet GV, Parker C, Usmani BA and Lakshmi MS. Epidermal growth factor receptor status correlates with cell proliferation-related 18A2/*mts1* gene expression in human carcinoma cell lines. *Ann N Y Acad Sci* 1995; **768**: 272-276.
- [12] Platt-Higgins AM, Renshaw CA, West CR, Winstanley JHR, Rudland SD, Barraclough R, Rudland, PS. Comparison of the metastasis-inducing protein S100A4 (p9Ka) with other prognostic markers in human breast cancer. *Int J Cancer* 2000; **89**: 198-208.

Table 1. The structure of ANN's analysed and their average performance over three sets of data

Test	Size (mm)	Grade (1,2,3)	ER	PgR	Ratio1	S100A4 (1,2)	NM23 (1,2)	Ratio2 (1,2,3)	Performance (%)
A	*	*	*	*	*	*	*		81
B	*	*	*	*	*				75
C	*	*			*	*	*		88
D	*	*	*	*		*	*		79
E	*	*	*	*					73
F	*	*							60
G	*	*			*				77
H	*	*				*	*		81
I	*	*				*	*	*	90
J	*	*					*		75
K	*	*	*	*			*		75
L	*	*	*	*		*			67
M	*	*				*			71
N	*	*					*	*	75
O						*	*	*	88
P					*				85
Q			*	*	*				81
R		*			*				85
S	*				*				79
T	*	*			*				79
U	*	*	*	*	*				77

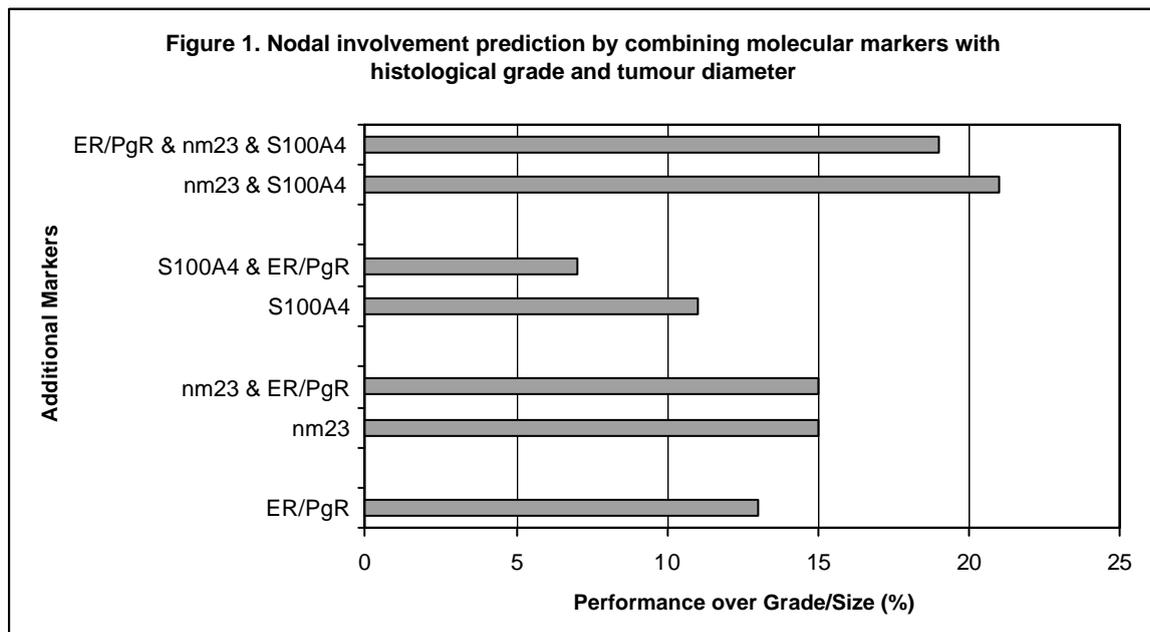
Ratio1: The ratio of the normalised values of S100A4/nm23

Ratio2: Values of ratio1 quantised into three groups

ER/PgR: Percentage values grouped into integer 1,2,3,4

A-O: Nodal prediction with reference to tumour grade/size

P-U: Nodal prediction with reference to S100A4/nm23 ratio



Biometric User Authentication

Raghav Menon, Dr. Satnam S Dlay, Dr. Wai Lok Woo

Abstract

Biometric refers to the measurement of certain biological characteristics of an individual. The most well known biometric is fingerprinting used by the law enforcing agencies for the identification of criminals. In this paper, a Fast Fourier Transform method is used to enhance the poor quality prints. Wavelet Packet Transform is used for feature extraction. Neural Networks which are algorithms that try to emulate the biological neuron are used for the purpose of identification.

It is found that the time taken for testing the fingerprints is very short and the whole process does not take more than 7 seconds. Also storage required is small. The Network was first trained for nine fingerprints and it gave an accuracy of 100%. Then the number of fingerprints was increased to twenty. It was found to give an accuracy of 84.5%.

Introduction

Biometrics can be defined as the process of identifying a person's identity uniquely based on his or her physiological or behavioural characteristic [1]. The chosen biometric should confer to certain properties which are explained in [2]. Some of the biometrics that satisfies these properties is fingerprints, iris, voice etc. Fingerprints can be considered as graphical designs that consists of ridges and valleys of every individual [2] [3]. Fingerprints are unique in case of every individual, in fact the fingerprints of identical twins are also different [4]. With the increase of crime in the society it is necessary to have a method to match the fingerprint of a person accurately. Manual method of matching takes a long time and hence there is a necessity for an automatic fingerprint verification system. A general block diagram of an automatic fingerprint verification system is shown in Figure 1.1 [5].

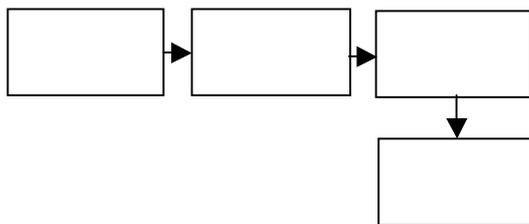


Figure 1.2 Block Diagram of Automatic Fingerprint Verification System [5]

The fingerprint identification is usually done by matching the minutiae which was introduced by Sir Francis Galton and further improved by Henry [6]. In [7] [8] [9] [10] [11] [12] [13] different methods of fingerprints recognition are explained.

In this paper fingerprint identification is done using Artificial Neural Networks. The project can be divided into two stages (i) Training stage (ii) Testing Stage. The training stage consists of the pre-processing where the low quality image is enhanced, the feature extraction during which characteristics unique to a person is extracted from the image and finally the extracted features are given as the input to the training algorithm where the Neural Network is trained. The testing stage also contains the above mentioned steps except the training stage as the trained network is now used to test the samples. The methods are explained in the following paragraphs.

Pre-Processing

Pre-processing is a very important stage in fingerprint recognition. The pre-processing method is application dependent i.e. depending on the application we can choose the pre-processing technique. The main function of the pre-processing stage is to make the fingerprint image suitable for the purpose of recognition i.e. pre-processing is done in order to remove noise and enhance the ridge pattern. Thus in the pre-processing stage the fingerprint image is processed using a series of image processing algorithms to obtain a clear, unambiguous image, clarifying smudged areas and healing cuts and scars. The fingerprint image may have distortions due to elastic deformations of the finger, cuts and abrasions on the finger, dirt, oil or moisture on the finger or scanner, partial imaging of finger tip, prints imaged with different rotations. There are various techniques to compensate these problems. There are many methods by which enhancement of fingerprints can be done. In [14] fingerprint enhancement using Gabor filters and in [15] enhancement using directional Fourier domain filtering is explained. An adaptive filtering technique is used in [16] followed by ridge enhancement based on ridge direction.

In [17] an FFT based cost-effective method is explained which is used in the project for the enhancement of fingerprints. It is known that the magnitude of the Fourier Transform contains the directional information for the block and thus in a box containing some ridges the Fourier

transform would correspond to its dominant frequencies [17]. In order to enhance a block containing the ridges thereby thickening the ridges and more finely spacing the parallel ridges, the FFT of the block is multiplied by its magnitude and hence on taking the inverse it will be found that the block is enhanced [17].

The following figures show the result of applying the above algorithm described in [17].



Figure 2.1 Block of Fingerprint before applying and after applying the algorithm

The fingerprint image was divided into blocks of 32×32 pixels and to each block the enhancement process was applied. The Figure 2.3 shows the original fingerprint and the image after the enhancement process.

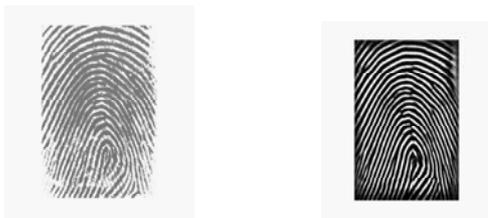


Figure 2.3 Image to be Enhanced Enhanced Image

Feature Extraction

Feature extraction is the process of acquiring higher level information for the purpose of image identification or identifying the objects in the image. Feature extraction can be done in many methods and is application dependent. In fingerprint identification systems the most commonly used feature is the minutiae details. In [10] [17] and [18] different methods for feature extraction is mentioned.

In this paper, wavelet packet based method is used to extract the features from the fingerprints. A Wavelet Packet is a generalisation of wavelet in that each octave frequency band of the wavelet spectrum is further subdivided into finer frequency bands by using the two scale relation repeatedly. In the case of Wavelet Transforms the image is decomposed into four images, the low resolution image, and an image containing the horizontal frequencies, an image containing the vertical frequencies and an image containing the diagonal frequencies. In the Wavelet case the decomposition is applied to the low

resolution image recursively while in Wavelet Packet case the decomposition is applied to all the coarse scale approximations leading to a complete wavelet tree. The Wavelet Packet transform not only the low resolution image is sub divided but also the other images are subdivided. This can be diagrammatically shown as a tree diagram as in Figure 3.1.

Here the texture information is used as the feature. In [19] it has been mentioned that texture classification can be done by calculation of the entropy of each of the sub band image. It is mentioned that different textures will have different values of entropy for each of the sub band image. It should be noted that fingerprint being unique, the texture of each fingerprint can be considered to be different and hence it can be

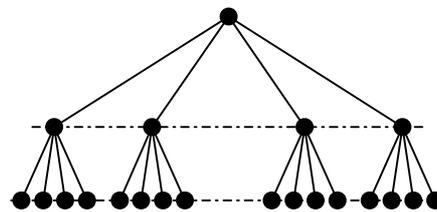


Figure 3.1 Wavelet Packet Tree

classified using the Wavelet Packet Transform. Hence each fingerprint will have a unique Wavelet Packet Signature (Entropy distribution). Some examples are shown in the figures below:

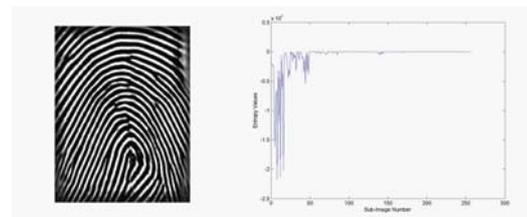


Figure 3.2 Fingerprint and its Corresponding Signature

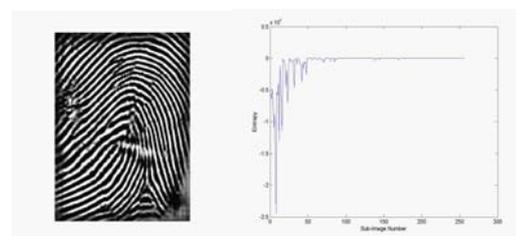


Figure 3.3 Fingerprint and its Corresponding Signature

Neural Networks

An Artificial Neural Network can be considered as an information processing system that has certain characteristics in common with the

biological neural networks. In other words, Neural Networks are computer algorithms that were developed inspired by the way in which the information is processed in the nervous system or a Neural Network is an implementation of an algorithm inspired by the research into brain [20].

The training algorithm used in here is the Backpropagation Algorithm due to the very fact that it is simple. The training of the Network involves three stages. They are i) the feed forward stage of the input ii) the calculation and backpropagation of associated errors iii) finally the adjustment of weights. A simple architecture of the Neural Network is shown in Figure 4.1[21]

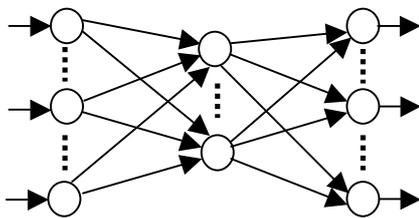


Figure 4.1 Architecture of a Neural Network [21]

After training the Network the testing algorithm is applied. The testing involves only the feedforward stage.

Results and Discussion

The Neural Networks were trained using the backpropagation algorithm first for a total of nine different fingerprints and 53 samples and the following parameters were used to train the Network. The parameters for training the Network are shown in Table 5.1 error convergence graph is shown in the Figure 5.2.

<i>Parameters</i>	<i>Values</i>
No. of Input Neurons	256
No. of Hidden Neurons	9
No. of Output Neurons	4
No. of Epochs	43466
Time Taken for Training	1.939256e+003 secs
Learning Rate	6.000000e-003
Tolerance	1.000000e-002
Momentum Factor	3.000000e-001

Table 5.1 Parameters Used for Training Neural Networks for Nine Fingerprints

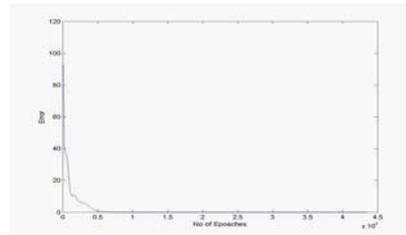


Figure 5.1 Error Convergence Graph for the Parameters in Table 5.1

It was found while testing that for nine fingerprints a 100% result was obtained. After this the network was trained for twenty different fingerprints. The parameters used are given in the table below and the error convergence graph is shown in the Figure 5.2

<i>Parameters</i>	<i>Values</i>
No. of Input Neurons	256
No. of Hidden Neurons	23
No. of Output Neurons	5
No. of Epochs	121052
Time Taken for Training	3.211222e+004 secs
Learning Rate	1.500000e-003
Tolerance	1.000000e-001
Momentum Factor	4.000000e-001

Table 5.2 Parameters Used for Training Neural Networks for Twenty Fingerprints

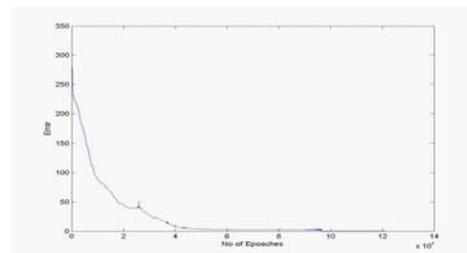


Figure 5.2 Error Convergence Graph for the Parameters in Table 5.2

While testing with twenty prints the network showed an accuracy of 84.5%. It was also found that the testing time is very small and did not exceed 7 seconds. Thus it is found from the results that Neural Networks is capable of achieving results accurately and fast. Since only the features from the image are extracted, it is not necessary to store the image and this saves the memory space. The trained Network stores the information in a weight file and it is enough that this file is stored. In the testing phase, since the features of the image can be directly read from the input image, it is only this weight file that is required.

References

- [1] Miller B, 'Vital Signs of Identity', IEEE Spectrum, February 1994, pp.22-30.
- [2] Anil K. Jain, Fellow, IEEE, Lin Hong, Sharat Pankanti, Associate Member, IEEE, and Ruud Bolle, Fellow, IEEE, 'An Identity Authentication System Using Fingerprints', Proceedings of IEEE, Vol.85, No.9, September 1997, pp.1364-1388.
- [3] Anil K. Jain, Fellow, IEEE, Salil Prabhakar, Student Member, IEEE, and Lin Hong, 'A Multichannel Approach to Fingerprint Classification', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.21, No.4, April 1999, pp.348-359.
- [4] Kalle Karu and Anil K Jain, 'Fingerprint Classification', Pattern Recognition, Vol.29, No.3, 1996, pp.389-404.
- [5] Nalini K. Ratha, Student Member, IEEE, Kalle Karu, Shaoyun Chen, and Anil K Jain, Fellow, IEEE, 'A Real Time Matching System for Large Fingerprint Databases', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.18, No.8, August 1996, pp.799-813.
- [6] Raffaele Cappelli, Alesssandra Lumini, Dario Maio, Member, IEEE, and David Maltoni, 'Fingerprint Classification by Directional Image Partitioning', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.21, No.5, May 1999, pp.402-421.
- [7] Bijan Moayer and King-Sun Fu, Fellow, IEEE, 'A Tree System Approach for Fingerprint Recognition', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.PAMI – 8, No.3, May 1986, pp.376-387.
- [8] Milos Klima, Jiri Rott, Pavez Dvorak, Daniez Gleeson, Susan McKenna-Lawlor, Jack Keating, 'Model of 2D Optical Correlator for Fingerprint Identification', IEEE Aerospace and Electronic System Magazine, Vol.12, No.7, July 1997, pp.3-8.
- [9] Emmanuel Fort and Holger Vach, 'Multiplexed Fingerprint Recognition using Holography', American Journal of Physics, Vol. 67, No. 2, February 1999, pp.116-119.
- [10] Anil K. Jain, Fellow, IEEE, Salil Prabhakar, Lin Hong, Sharat Pankanti, 'Filterbank based Fingerprint Matching', IEEE Transactions on Image Processing, Vol. 9, No.5, May 2000, pp.846-859.
- [11] Zslot Miklos Kovacs-Vajna, Senior Member , IEEE, 'A Fingerprint Verification System Based on Triangular Matching and Dynamic Time Warping', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.11, November 2000, pp.1266-1276.
- [12] Andrea R. Roddy, Member, IEEE, and Jonathan D. Stosz, 'Fingerprint Features-Statistical Analysis and System Performance Estimates', Proceedings of IEEE, Vol. 85, No. 9, September 1997, pp.1390-1421.
- [13] Louis Coetzee and Elizabeth C. Botha, 'Fingerprint Recognition in Low Quality Images', Pattern Recognition, Vol. 26, No. 10, 1993, pp.1441-1460.
- [14] Lin Hong, Yifei Wan, Anil K Jain, 'Fingerprint Image Enhancement: Algorithm and Performance Evaluation', IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 20, No. 8, August 1998, pp. 777-789.
- [15] Sherlock B, Monro D, Millard K, 'Fingerprint Enhancement by Directional Fourier Filtering', IEE Proc. Vision, Image and Signal Processing, 141 (2), 1994, pp. 87-94.
- [16] Douglas Hung D C, 'Enhancement and Feature Purification of Fingerprint Images', Pattern Recognition, Vol.26, No.11, 1993, pp. 1661-1671.
- [17] Willis A J, Myers L, 'A Cost-Effective Fingerprint Recognition System for the use with Low Quality Prints and Damaged Fingertips', Pattern Recognition, Vol.34, No.2, 2001, pp. 255-270.
- [18] Tico M, Kuosmanen P, Saarinen J, 'Wavelet Domain Features for Fingerprint Recognition', Electronic Letters, Vol.37, No.1, January 2001, pp. 21-22.
- [19] Andrew Laine, Jian Fan, 'Texture Classification by Wavelet Packet Signatures', IEEE Transactions on Pattern Analysis and Machine Vision, Vol. 15, No.11, November 1993, pp. 1186-1191.
- [20] Don Hammerstrom, 'Neural Networks at Work', IEEE Spectrum, June 1993, pp. 26-32.
- [21] Rela Nekovei, Student Member, IEEE, and Ying Sun, 'Back-Propagation Network and its Configuration for Blood Vessel Detection in Angiograms', IEEE Transactions on Neural Networks, Vol.6, No.1, January 1995, pp.64-72.

Classification of Cancerous Cells Images using Clustered Fuzzy-Neural Machine Techniques

E. NWOYE, S. S. DLAY, W. L. WOO
School of Electrical, Electronic and Computer Engineering
University of Newcastle upon Tyne
United Kingdom

Email: ephraim.nwoye@ncl.ac.uk; s.s.dlay@ncl.ac.uk; w.l.woo@ncl.ac.uk,

Abstract:

Computer assisted diagnosis of cancer has received attention in recent years. The development of automated algorithms would be a valuable tool to the Pathologist for fast verification of these cancer abnormalities. In this paper a novel method which will automatically locate differences in cancer cells images and classify cells into normal and malignant is implemented by fuzzifying image feature descriptor values and incorporating clustering paradigm into neural network to classify images. The proposed system was evaluated using 116 cancers and 88 normal colon cells images. It is more efficient, simple to implement and yields better accuracy than conventional methods.

1. Introduction

Pathologists, daily, screen large numbers of slides containing cancerous cells manually, which are similar in shape, size or cells structure. The recent advances in multimedia and image processing can be utilized to assist pathologists in this respect. Cancer of the large bowel is among the commonest of malignant tumors (third in the UK after lung and breast cancer) and represents the second largest cause of death from cancer in the western world. In UK over 30,000 people develop colorectal cancer each year [1]. The American Cancer society reported that about 105,500 people will have colorectal cancer in 2003 with about 57100 resulting to death in America. This disease is uncommon in Asia, Africa and South America due to pattern of food which explains the high incidence of colorectal carcinoma (30-60 per 100000) in USA, Western Europe and Australia [1] [2].

Several classifier algorithms have been developed. These include the traditional classifiers [3,4], and the fractal dimension algorithm [5-7] for detection and classification of cancerous colonic mucosa. Comparison shows that there was no great improvement over the discrimination already achieved. This research therefore is to investigate the possibility of classifying normal and malignant colon cells using the proposed Clustered Fuzzy-Neural Machine (CFNM) to achieve 100% discrimination. The common texture descriptors to be used are Contrast (CON), entropy (ENT), Inverse Difference Moment (IDM), Correlation (COR) and Angular Second Moment (ASM).

2. Theory and Method

The proposed system is simply presented using a two layer and one output neurons.

Assuming a training set (\bar{X}_j, \bar{T}_j)

where $1 \leq j \leq k$ and $\bar{X}_j = (\bar{x}_{j1}, \dots, \bar{x}_{jm})$ and

the actual output \bar{Y}_j , then the input to the

neuron, $\bar{I} = \sum \bar{x}_i \bar{w}_i$ using standard

fuzzy arithmetic, and the desired output

$\bar{y} = f(\bar{I})$ making \bar{y} a fuzzy subset of $[0, 1]$,

with the activation function,

$$f(x) = (1 + e^{-x})^{-1}$$

and error $\bar{E} = 1/2 [\sum_{i=1}^k (y_j - \bar{T}_j)]$, as the

weight adjusts as follows;

$$\bar{w}_i(l+1) = \bar{w}_i(l) + \eta [\Delta \bar{w}_i(l+1)] \quad (1)$$

where $1 \leq i \leq n$; $l = 1, 2, \dots$. And η is learning rate.

Equation 1 above is the fuzzified delta algorithm of the back propagation neural network. To overcome the problem of overlaps in the features that will identify the various classes in the above algorithm, probabilistic cluster technique was introduced to modify the learning rule. Consider for input pattern belonging to class q, a dominant cluster d is selected from class q based on the following criterion:

$$\|X - \bar{w}_{id}\| = \min_j \|X - \bar{w}_{ij}\| \quad (2)$$

where \bar{w}_{ij} represents the weight vector of cluster j in class q. In order to ensure that the dominant cluster can truly represent the pattern X, the similarity distance defined by

(3) between \bar{X} and \bar{w}_{id} need to be re-

examined. For the cluster \bar{w}_{id} , CFNM learning rule adjusts its weight vectors according to the following:

$$w_{id}(t+1) = w_{id}(t) + a_{id} (X - \eta w_{id}(t+1))$$

.....3

Where a_{id} is a non-negative mixing factor.

Fig. 1 shows a typical network model of proposed CFNM.

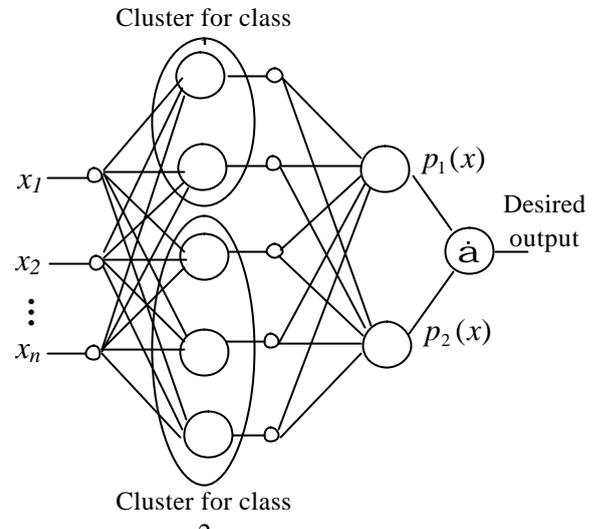


Fig. 1: Clustering Fuzzy-Neural Machine

The CFNM training Algorithm is set below in the following steps:

1. Initialize the fuzzified weights, $w_i \in [0,1]$ set all weight values 0
2. Present the fuzzified input, $x_1, \dots, x_n \in [0,1]$ and the desired output
3. Calculate the Actual outputs; using the logistic activation function,
4. Adapt weights $\bar{w}_i(t+1) = \bar{w}_i(t) + \eta [\Delta \bar{w}_i(t+1)]$
5. Modify the weights by clustering using a non-negative a probabilistic mixing factor a_{id} , $w_{id}(t+1) = w_{id}(t) + a_{id} (X - \eta w_{id}(t+1))$

We used the Fractal dimension techniques to extract the required image descriptors features. Fractals are a class of mathematical functions which have been used to characterize the geometrical properties of a set.

3. Results and Discussion

For the purpose of classification the images were divided into two sets, the training and validation. The first 44 normal images and 58 cancer images were used for training. The second 44 normal images and 58 cancer images were used for testing. All the images used were acquire as hospital samples captured with a light microscope in combination of a CCD camera with magnification of *40. Each acquired image has a spatial resolution of 256x256 pixels represented as unit8 arrays. The algorithm is implemented using MATLAB. Fig.2 shows that performance was met in training. Fig.3&4 shows respectively the normal and cancer colon cell images with their histograms. The classification rate is about 97%. In the confusion matrix given in Table1 above, the error rate is as a result of medical specimen preparation and preprocessing anomalies. All features obtained from fractal dimension were normalized.

Table 1: Classification for combined features

Feature combinations in the calculations	Classification Rate
CON	0.9625
CON & ASM	0.9667
CON, ASM & ENT	0.9676
CON,COR, ASM ENT & IDM	0.9775

Table 2: Confusion matrix

Classified as	Cancer	Normal	Classification Rate (%)	Error Rate
Correct classes				
Cancer	114	2	98.27	0.017
Normal	2	82	97.72	0.027

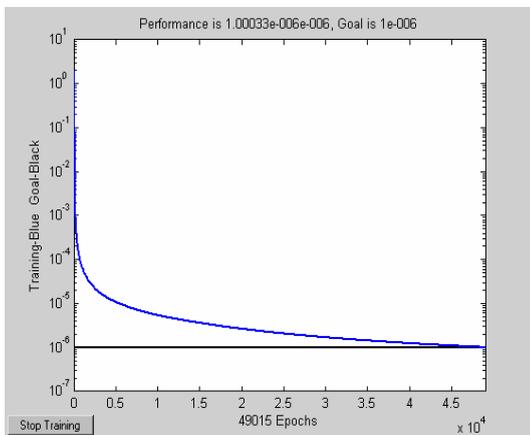


Fig.2: Training graph

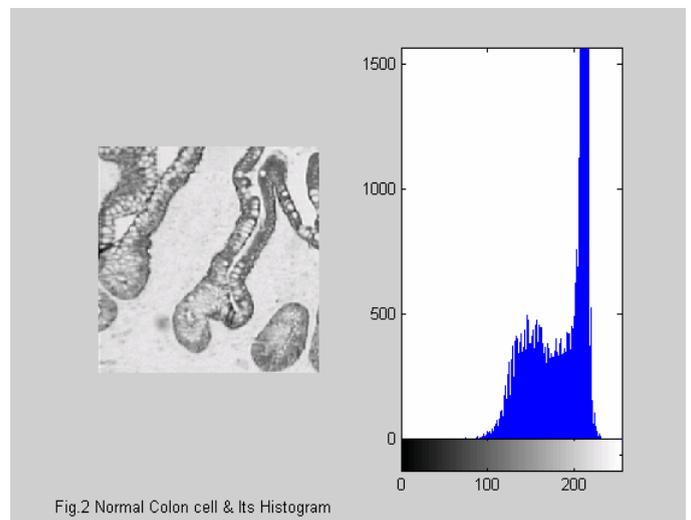


Fig. 2: Normal Colon Cell and Histogram

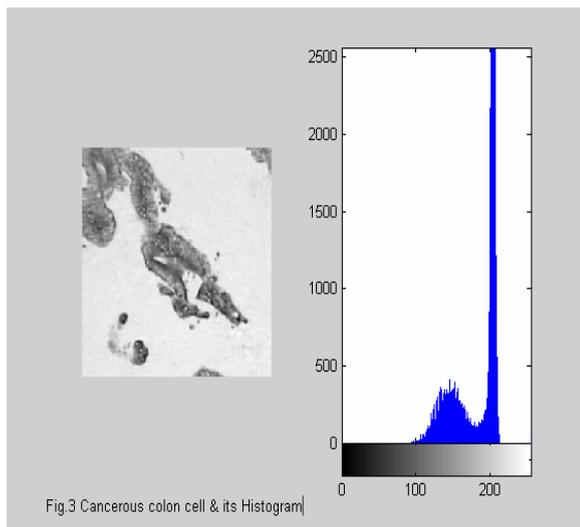


Fig. 3: Cancerous Colon Cell and Histogram

4. Conclusion

A system based on fractal geometry and fuzzy neural network was developed for colon cancer classification. The system consists of an automatic feature extraction process, image enhancement algorithm and fuzzy neural network classifier. During the process of feature extraction, colon cells were located using iterative generation of fractal values.

References:

- [1] R. Souhami and J. Tobias, *Cancer and its Management. Blackwell Pub. Co. inc., USA, pp 243-252, 2003.*
- [2] W.C. Willert, M.J. Stampfer, G.A. Golditz, B.A. Rosner and F.E. Speizer, "Relation of meat, fat and fiber intake to the risk of colon cancer in a prospective study among Women", *N Engl. J Med., pp 1664-1672, 2002.*
- [3] P.W. Hamilton, D. Thompson, J.M Sloan and P.H. Bartel, "Knowledge guided segmentation and morphometric analysis of colorectal dysplasia." *Analytical and Quantitative cytology and histology vol.17 No.3, 1995*
- [4] A.J. Wong, G.L. Kohn, H.J. Schwarz, B.H. Ruebner and M.J Lawson, "Colorectal cancer and non-cancer patients have similar labeling indices by microscopy and computed image analysis." *Human pathology vol...26, No.12, Pp1329-1332, 1995.*
- [5] A.N. Esgiar, R.N.G. Naguib, B.S. Sherif, M.K. Benneth and A.Murray, "Fractal Analysis in the Detection of Colonic Cancer Images." *IEEE Trans. Info. Tech. in Biomedicine, vol.6, No.1, pp54-58, 2001*
- [6] A. Esgiar, R. Nguib, M. Benneth and A. Murray, "Automated Feature extraction and identification of Colon Carcinoma." *Anal. Quant. Cytol., vol.20, pp297-301, 1998.*
- [7] A. Esgiar, R. Nguib, B.S. Sheriff, M. Bennett and A. Murray, "Microscopic Image Analysis for Quantitative Measurement and Feature Identification of Normal and cancerous Colonic Mucosa." *IEEE Trans. Info. Tech. in Biomedicine. No.2, pp 197-203, 1998.*

Coherence-free Photonic Notch Filter

E. H. W. Chan and R. A. Minasian

*School of Electrical and Information Engineering & Australian Photonics CRC,
University of Sydney, NSW, 2006, Australia.*

Email: r.minasian@ee.usyd.edu.au

Abstract

A new topology for a photonic notch filter that can solve the coherent interference problem in photonic signal processors is presented. It is based on a double-pass modulation technique. The key advantage of this new structure is that it completely removes the dominant phase noise limitation, which is a fundamental problem with existing incoherent processor approaches. It enables arbitrary narrow-linewidth telecommunication-type lasers to be used with stable and robust filter operation. Extension to obtain tuneable notch frequency is also demonstrated.

Introduction

The ability of photonic signal processors to process high-speed signals directly within the optical fibre has attracted significant interest. This stems from their high-time bandwidth capabilities that can overcome electronic processor limitations, particularly in future applications in high frequency optical wireless communications and in fibre-optic signal distribution for antenna remoting, where in-built signal conditioning is attractive.

Many techniques for realising photonic signal processor filters based on the optical delay line concept have been reported [1]-[4]. In order to obtain a robust transfer function irrespective of environmental perturbations, virtually all previously reported processors are based on an incoherent approach, in which the coherence of the light source is made smaller than the minimum delay time of the processor. There are two problems with this approach. First, the optical source linewidth needs to be specially designed, and normal narrow-linewidth telecommunications lasers cannot be used, in general. Second, and most importantly, the incoherent approach produces an excessive amount of phase noise at the output due to the optical interference of the delayed optical signals. This phase noise is a fundamental problem [5]. It is by far, the dominant noise source in the system, and it can severely degrade the SNR. The use of two orthogonal polarisations in hi-birefringence fibre [6] or in birefringent crystals [7] has been proposed to remove this limitation, however this approach requires very precise polarisation control and requires an additional polarisation element that can

provide the time delay between the two polarisations. In this paper, we present a new technique that can solve the coherent interference problem in photonic signal processors. It is based on a double-pass modulation technique. The technique requires only one grating after the modulator and is particularly simple. Moreover, it completely removes the phase noise, and also enables arbitrary narrow-linewidth lasers to be used while also enabling stable and robust filter operation to be realised. The new structure solves the coherence problems in photonic notch filters and can realise in-built signal conditioning in fibre-optic links. Results are presented, which demonstrate robust operation with a narrow-linewidth source, and a high SNR. Finally, we also demonstrate how this structure can be extended to tuneable notch operation.

Double-pass Modulation Photonic Notch Filter Topology

The topology of the new double-pass modulation based notch filter is shown in Fig. 1.

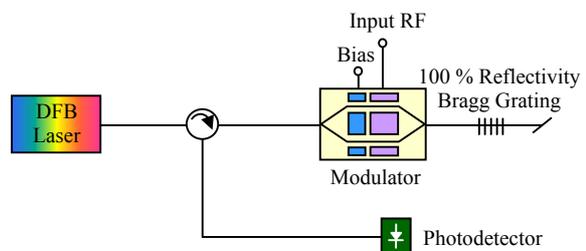


Fig. 1. The topology of the double-pass modulation based notch filter.

In this structure, CW light from a laser is modulated in a modulator in the forward direction and is then reflected back from a fibre Bragg grating to undergo a double-pass modulation in the modulator in the reverse direction. The second modulation produces notches at all frequencies where the remodulation is an odd integer multiple of 180 degrees phase difference to the returned modulated RF signal. Selecting the distance L between the modulator and the grating reflector can control the notch frequencies. It is apparent that there exists only a single optical path in the system. Hence there is no possibility of coherent interference effects. To our

knowledge, this is the first photonic notch filter structure that does not involve multiple optical signal paths. Since it is single-path and coherence free, a narrow linewidth laser source (such as a DFB laser) can be used to obtain a stable notch filter response. Also there are no source linewidth limitations to the high-frequency range of the filter operation, and only the modulator bandwidth determines this. Finally, and most importantly, there is no phase noise generated in this filter structure. The system is simple, and can be built into an existing microwave fibre-optic link, since it only requires one additional fibre Bragg grating and a circulator beyond components that already exist in a fibre-optic link, to provide filtering of unwanted RF signals.

Referring to Fig. 1, the notch filter output photocurrent is given by

$$I_o = \frac{\mathfrak{R} t_{ff}^2 P_{cw}}{4} [1 + \cos(\beta_{RF} \sin(\omega_{RF} t + \phi) + \beta_{bias}) + \cos(\beta_{RF} \sin(\omega_{RF} t) + \beta_{bias}) + \cos(\beta_{RF} \sin(\omega_{RF} t + \phi) + \beta_{bias}) \cos(\beta_{RF} \sin(\omega_{RF} t) + \beta_{bias})] \quad (1)$$

where \mathfrak{R} is the photodiode responsivity, t_{ff} is the electro-optic modulator (EOM) insertion loss, P_{cw} is the CW laser power, $\beta_{RF} = \pi V_{RF}/V_{\pi}$, $\beta_{bias} = \pi V_{bias}/V_{\pi}$, ϕ is the phase difference between remodulation and the returned modulated RF signal, V_{π} is the modulator switching voltage, V_{RF} is the RF signal amplitude and V_{bias} is the modulator bias voltage. The phase difference ϕ depends on the distance L between the modulator and the grating reflector and is given by

$$\phi = \omega_{RF} \cdot T = \omega_{RF} \cdot \left(\frac{2nL}{c} \right) \quad (2)$$

where ω_{RF} is the angular RF frequency, T is the round trip time used to set the filter centre frequency, n is the fibre refractive index and c is the speed of light.

The notch filter transfer function, defined as the ratio of the output RF electrical power to the input RF electrical power, is given by

$$H(f) = \frac{1}{32} \mathfrak{R}^2 t_{ff}^4 P_{cw}^2 K_o (U_{RF}^2 + T_{RF}^2) R_o \quad (3)$$

where U_{RF} and T_{RF} are given by

$$U_{RF} = \sin \beta_{bias} \sin \phi + \cos \frac{\phi}{2} \sin 2\beta_{bias} \sin \frac{\phi}{2}$$

$$T_{RF} = \sin \beta_{bias} \cos \phi + \sin \beta_{bias} + \cos \frac{\phi}{2} \sin 2\beta_{bias} \cos \frac{\phi}{2} \quad (4)$$

K_o is the modulator response parameter and R_o is the load resistance. The frequency response of the filter when the EOM is biased at quadrature is exactly the same as the commonly used coupler-based unbalanced Mach-Zehnder interferometer notch filter response [8].

The distance between the modulator and the grating reflection point controls the notch frequency

of the filter. Hence this structure can readily be extended to tuneable notch operation by replacing the uniform Bragg grating with a chirped grating and by tuning the source wavelength. Tuneable, narrow-linewidth laser sources are available, and enable agile notch filter frequency tuning to be obtained.

The fundamental requirement of the double-pass modulation coherence-free photonic notch filter topology is that the response of the optical modulator is the same in both directions of incident light, ie. in both the forward and reverse directions. This can cause a significant frequency limitation if travelling-wave EOMs are to be used. This is because travelling-wave EOMs are distributed waveguide structures, and while they have a very high bandwidth of response in the forward direction due to the matching of the optical and RF wave velocities travelling together in the structure, the bandwidth in the reverse direction is quite poor because in that case the optical and RF waves are travelling in opposite directions, hence the problem of velocity mismatch occurs. This effect can severely limit the notch filter frequency range and notch rejection level. To solve this problem, we propose to use an electroabsorption modulator (EAM) for the modulation device in the structure. The EAM is an integrated device and behaves as a lumped component. For example, EAMs having only 50 μm length and having bandwidths of 50 GHz have been demonstrated [9]. Being 50 μm in length means that it is truly lumped to high microwave frequencies, and there should be no velocity mismatch effect for forward or backward waves. Since the device is so short, velocity matching is not an issue up to frequencies well into the submillimetre range, which enables the photonic notch filter to operate up to very high microwave frequencies.

The coherence-free photonic notch filter topology eliminates the dominant phase induced intensity noise, which is the most important limitation in conventional photonic signal processor structures. This is a key advantage of the topology presented. The remaining noise components at the output are the usual shot noise, laser intensity noise and thermal noise, ie. there is no additional noise generated by the signal processor. Since the output RF electrical power is proportional to the square of the input optical power, and among the noises at the output, only the laser intensity noise is proportional to the square of the input optical power, hence it is possible to increase the SNR by increasing the input optical power until the system is laser intensity noise limited. This is in contrast to conventional incoherent photonic notch filter structures, in which it is not possible to increase the SNR by increasing the input optical power because the phase noise is dominant for those structures and the phase noise increases with the square of the input optical power,

just as the signal does. This means that the coherence-free notch filter topology not only has a vastly lower intrinsic noise performance than conventional incoherent unbalanced Mach-Zehnder notch filter structures, but that unlike its counterpart it enables even further improvement in noise performance to be realised by increasing the optical power, which is an important advantage.

Experimental Results

In order to verify the proof of principle for the new topology, the double-pass modulation coherence-free photonic notch filter topology as shown in Fig. 1 was set up. The laser source was a tuneable external cavity laser, which had a linewidth of less than 500 kHz and whose coherence time was much larger than the minimum delay time in the filter. The modulator was an EOM biased at the quadrature point, and had polarisation controllers at its input and output. A 1536.5 nm Bragg grating with over 99 % reflectivity was used for optical signal reflector. The distance between the EOM and the Bragg grating was 2.35 m, which corresponds to a filter FSR of 42.4 MHz. The double-pass modulated optical signal passed through an optical circulator and was detected by the photodetector, whose output was connected to a network analyser to display the filter transfer characteristic.

Results showing the measured and simulated responses of the double-pass modulation based notch filter are shown in Fig. 2. Excellent agreement between the experimental measurements and the theoretical predictions can be seen. The filter response was stable, even though the laser linewidth was significantly smaller than the filter FSR, which demonstrated that the filter operation is free of coherence interference limitations.

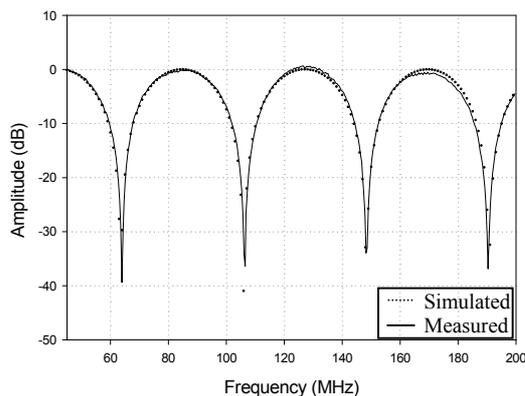


Fig. 2. Measured and simulated responses of the double-pass modulation based notch filter.

For comparison, a conventional coupler-based unbalanced Mach-Zehnder interferometer notch filter, which uses incoherent summation of delayed signals, was set up. The length difference of the two arms of the Mach-Zehnder filter was 4.48 m, which

corresponds to 44.6 MHz filter FSR and which is essentially the same as for the double-pass modulation based notch filter, and also the same tuneable laser source was used as before. Fig. 3 displays two filter responses of the unbalanced Mach-Zehnder filter measured at different times. As expected, the narrow-linewidth, highly coherent laser produced a filter response that was continuously fluctuating. More than 20 dB fluctuation on the filter passband was observed, clearly indicating optical coherence interference effects. By broadening the laser linewidth so that it is larger than the filter FSR, the fluctuation in the response could be removed, as expected and this is also shown in Fig. 3. However, even with linewidth broadening, phase noise still remains and is the dominant noise source for the incoherent Mach-Zehnder filter. By comparing the results in Fig. 2 and 3, it can be seen that both filter structures have the same transfer function. However the double-pass modulation based notch filter has the key advantage of being coherence free and of eliminating the dominant phase noise.

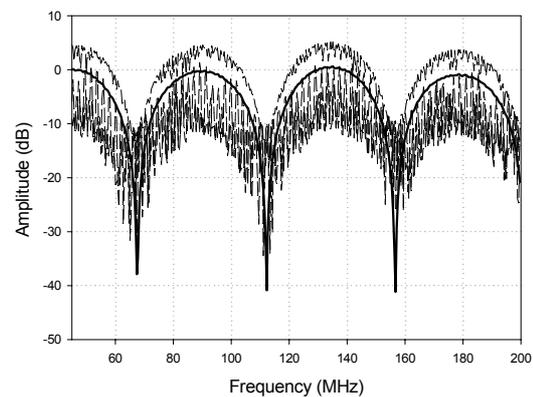
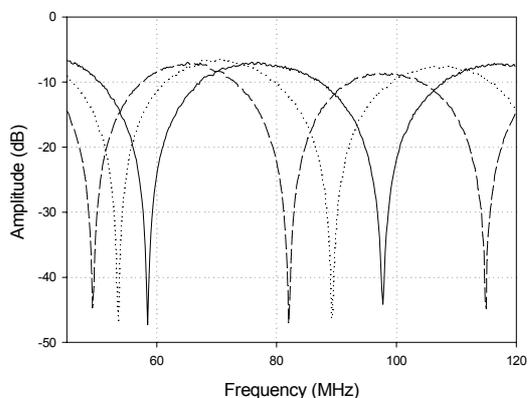


Fig. 3. Conventional unbalanced Mach-Zehnder delay line notch filter frequency responses with (solid) and without (dash) coherence control on the tuneable laser.

To demonstrate the SNR capabilities of the double-pass modulation based notch filter, a single frequency RF signal at the filter passband was applied into the EOM, and the RF signal power and the noise level at the output of the filter were measured on a spectrum analyser with 300 kHz resolution bandwidth. The tuneable laser used in this experiment had a RIN of <-145 dB/Hz for the frequencies below 1 GHz, the optical signal into the EOM was 13.4 dBm, and -10 dBm of input RF signal was used, which corresponded to 0.1 modulation depth. The measured SNR was 68 dB. The predicted SNR using the experimental parameters was 70.7 dB, which is in good agreement with the measured value. Because the RIN of the tuneable laser is relatively high, the laser intensity noise was the highest noise component in the system. For comparison, the SNR of the

conventional coupler-based unbalanced Mach-Zehnder interferometer notch filter, which uses incoherent summation of delayed signals, was measured under the same conditions. In this case, the measured SNR was only 9.4 dB. This demonstrates the vastly superior SNR performance (ie. nearly 60 dB increase) of the double-pass modulation based notch filter compared to the conventional delay line notch filter, because it eliminates the dominant phase noise. Moreover, the SNR of the Mach-Zehnder filter cannot increase with increased input optical power because the dominant phase noise increases by the same amount as the signal as the input optical power increases, whereas the SNR of the double-pass modulation based notch filter can be increased further by increasing the input optical power since phase noise is eliminated.

To demonstrate tunability of the filter structure in Fig. 1, multiple Bragg gratings written in a single fibre at different locations in the fibre was used in conjunction with a tuneable laser source. The centre wavelengths of the gratings were 1537.8 nm, 1544.2 nm, and 1550.2 nm. The separation between each grating was around 20 cm. The tuneable laser was tuned to the grating wavelength and the filter frequency responses were measured on the network analyser. The normalised notch filter responses



for different laser wavelengths of 1537.8 nm (solid), 1544.2 nm (dots) and 1550.2 nm (dash).

A small discrepancy in the notch filter response level was observed because the gratings had slightly different reflectivities. A better fabrication process should be able to avoid this non-ideality. Nevertheless, Fig. 4 shows that the notch filter frequency can be tuned over a wide frequency range.

Conclusion

A new topology for a photonic notch filter that can solve the coherent interference problem in

photonic signal processors has been presented. It is based on a double-pass modulation technique, which requires only one grating after the modulator and is particularly simple. The filtering technique has been experimentally verified. The results have demonstrated robust notch filter operation with a narrow-linewidth source. Furthermore no phase induced intensity noise was observed at the system output, yielding a vastly superior SNR performance of the double-pass modulation based notch filter compared to the conventional delay line notch filter. Tuneable coherence-free operation of the notch filter has also been demonstrated over a wide frequency range. This structure can be readily built into fibre-optic transmission systems, since it only requires a few additional components, in links for optical wireless communications and in fibre-optic signal distribution for antenna remoting. The new photonics based filter offers notch filtering to very high microwave frequencies and can be integrated into optical fibre microwave transmission systems.

References

- [1] V. Polo, F. Ramos, J. Marti, D. Moodie and D. Wake, "Synthesis of photonic microwave filters based on external optical modulators and wide-band chirped fiber gratings", *J. Lightwave Technol.*, vol. 18, no. 2, pp. 213-220, 2000.
- [2] E. H. W. Chan, K. E. Alameh, R. A. Minasian, "Photonic bandpass filters with high skirt selectivity and stopband attenuation", *J. Lightwave Technol.*, vol. 20, no. 11, pp. 1962-1967, 2002.
- [3] B. A. L. Gwandu, W. Zhang, J. A. R. Williams, L. Zhang and I. Bennion, "Microwave photonic filtering using Gaussian-profiled superstructured fiber Bragg grating and dispersive fibre", *Electron. Lett.*, vol. 38, no. 22, pp. 1328-1330, 2002.
- [4] D. Pastor, J. Capmany and B. Ortega, "Broad-band tunable microwave transversal notch filter based on tunable uniform fiber Bragg gratings as slicing filters", *IEEE Photon. Technol. Lett.*, vol. 13, no. 7, pp. 726-728, 2001.
- [5] B. Moslehi, "Analysis of optical phase noise in fiber-optic systems employing a laser source with arbitrary coherence time", *J. Lightwave Technol.*, vol. LT-4, no. 9, pp. 1334-1351, 1986.
- [6] W. Zhang, J. A. R. Williams and I. Bennion, "Optical fiber delay line filter free of limitation imposed by optical coherence", *Electron. Lett.*, vol. 35, no. 24, pp. 2133-2134, 1999.
- [7] J. Chen, Y. Wu, J. Hodiak, and P. Yu, "A novel digitally tunable microwave-photonic notch filter using differential group-delay module", *IEEE Photon. Technol. Lett.*, vol. 15, no. 2, pp. 284-286, 2003.
- [8] A. H. Quoc and S. Tedjini, "Experimental investigation on the optical unbalanced Mach-Zehnder interferometers as microwave filters", *IEEE Microwave Guided Wave Lett.*, vol. 4, no. 6, pp. 183-185, 1994.
- [9] T. Ido, S. Tanaka, M. Suzuki, M. Koizumi, H. Sano and H. Inoue, "Ultra-high-speed multiple-quantum-well electro-absorption optical modulators with integrated waveguides", *J. Lightwave Technol.*, vol. 14, no. 9, pp. 2026-2033, 1996.

Combined Spatial and Temporal Domain Wavelet Shrinkage Algorithm for Video Denoising

Eric J. Balster and Yuan F. Zheng, Department of Electrical Engineering, The Ohio State University
Columbus, OH 43210 USA, balstere, zheng@ee.eng.ohio-state.edu

Robert L. Ewing, Embedded Information Systems Engineering Branch, Air Force Research Laboratory
Wright-Patterson AFB, OH 45433, robert.ewing@wpafb.af.mil

Abstract—A combined spatial and temporal domain wavelet shrinkage algorithm for video denoising is presented in this paper. The spatial domain denoising technique is a selective wavelet shrinkage method which uses a two-threshold criteria to exploit the geometry of the wavelet sub-bands of each video frame, and each frame of the image sequence is spatially denoised independently of one another. The temporal domain denoising technique is a selective wavelet shrinkage method which estimates the level of noise corruption as well as the amount of motion in the image sequence. The amount of noise is estimated to determine how much filtering is needed in the temporal domain, and the amount of motion is taken into consideration to determine the degree of similarity between consecutive frames. The similarity affects how much noise removal is possible using temporal domain processing. Using motion and noise level estimates, a video denoising technique is established which is robust to various levels of noise corruption and various levels of motion.

Keywords—video denoising, combined spatial and temporal domain processing, selective wavelet shrinkage, motion estimation.

I. INTRODUCTION

The recent advance in multimedia technology has promoted a large amount of research in the area of image and video processing. Image and video processing applications which include compression, enhancement, and target recognition, require preprocessing functions for noise removal to improve performance. Noise removal is one of the most common and important processing steps in many image and video systems.

Because of the commonality of noise removal functions in most image and video systems, there has been an large amount of research dedicated to the subject of image denoising over the past several decades, and many different mathematical tools have been proposed. Various established denoising methods using variable coefficient linear filters [10, 13], adaptive non-linear filters [9, 17], DCT based solutions [5], cluster filtering [16], genetic algorithms [15], fuzzy logic [6, 14], etc., have all been proposed in the literature.

The wavelet transform has also been used to suppress noise in digital images. It has been shown that the reduction in absolute value of wavelet coefficients is successful in signal restoration [8]. This process is known as wavelet shrinkage. Other denoising techniques select or reject wavelet coefficients based on their predicted contribution to reconstructed image quality. This process is known as *selective* wavelet shrinkage, and many works have used it as the preferred method of image denoising [1, 4, 7, 8, 11].

However until recently, the removal of noise in video signals has not been seriously studied. Cocchia, et. al., developed a three dimensional rational filter for noise removal in video signals [3]. The 3D rational filter is able to remove noise, but preserve important edge information. Also, the 3D rational filter uses a motion estimation technique. Where there is no motion detected, the 3D rational filter is applied in the temporal domain. Otherwise, only spatial domain processing is applied.

Later, Zlokolica, et. al. uses two new techniques for noise removal in image sequences [17]. Both these new techniques show improved results upon the method of [3]. The first method is an alpha-trimmed mean filter of [2] extended to video signals, and the second is the K nearest neighbors (KNN) filter. Both alpha-trimmed and KNN denoising methods are based on ordering the pixel values in the neighborhood of the location to be filtered, and averaging a portion of those spatially contiguous pixels. Each of these methods attempts to average values which are close in value, and avoid averaging values which are largely dissimilar in value. Thus, the image sequence is smoothed without blurring edges.

However, because the success of the wavelet transform over other mathematical tools in denoising images, some researchers believe that wavelets may be successful in the removal of noise in video signals as well. Pizurica, et. al. uses a wavelet-based image denoising method to remove noise from each individual frame in an image sequence and then applies a temporal filtering process for temporal domain noise removal [12]. The combination of wavelet image denoising and temporal filtering outperforms both wavelet based image denoising techniques [1, 7, 8, 11] and spatial-temporal filtering techniques [2, 3, 17].

In this paper, we develop a noise removal algorithm for video signals. This algorithm uses selective wavelet shrinkage in all three dimensions of the image sequence and proves to outperform the few video denoising algorithms given in the relevant literature. First, the individual frames of the sequence are denoised by the method of [1], which we had developed earlier. Then a new selective wavelet shrinkage method is used for temporal domain processing.

Following the Introduction, Section II describes the temporal domain wavelet shrinkage method and explores the proper order of temporal and spatial domain processing functions. Section III provides the proposed motion estimation index used in the temporal domain processing. Section IV develops the parameters for temporal domain processing, and Section V gives the experimental results of the proposed method as well

as other established methods. Section VI concludes the paper.

II. TEMPORAL DENOISING AND ORDER OF OPERATIONS

In this section, we develop the principal algorithm for video denoising. Additional mechanisms required by this algorithms will be discussed in latter sections.

A. Temporal Domain Denoising

Let us define f_l^z as a pixel of spatial location l and frame z in a given image sequence. The non-decimated wavelet transform applied in the temporal domain is given by

$$\lambda_{k+1}^{3D}[l, z] = \sum_p g[p] \alpha_k^{3D}[l, 2^{k+1}p - z], \quad (1)$$

and

$$\alpha_{k+1}^{3D}[l, z] = \sum_p h[p] \alpha_k^{3D}[l, 2^{k+1}p - z], \quad (2)$$

where

$$\alpha_{-1}^{3D}[l, z] = f_l^z. \quad (3)$$

$\lambda_k^{3D}[l, z]$ is the high-frequency wavelet coefficient of spatial location l , frame z and scale k . Also, $\alpha_k^{3D}[l, z]$ is the low-frequency scaling coefficient of spatial location l , frame z and scale k . Thus, multiple resolutions of wavelet coefficients may be generated from iterative calculation of Equations 1 and 2.

The wavelet function used in the temporal domain denoising process is the Haar wavelet given by

$$h[n] = \begin{cases} \frac{1}{\sqrt{2}}, & n = 0, 1 \\ 0, & \text{else} \end{cases} \quad g[n] = \begin{cases} \frac{-1}{\sqrt{2}}, & n = 0 \\ \frac{1}{\sqrt{2}}, & n = 1. \end{cases} \quad (4)$$

The decision to use the Haar wavelet is based on experimentation with several other wavelet functions and finding the greatest results with the Haar. The compact support of the Haar wavelet makes it a suitable function for denoising applications. Because of its compact support, the Haar coefficients represent least number of original pixels in comparison to other types of wavelets. Thus, when a coefficient is removed because of its insignificance, the result affects the smallest area of the original signal in the reconstruction.

Significant wavelet coefficients are selected by their magnitude with a threshold operation.

$$L_k^{3D}[l, z] = \begin{cases} \lambda_k^{3D}[l, z], & \text{when } |\lambda_k^{3D}[l, z]| > \tau_z, \\ 0, & \text{else} \end{cases}, \quad (5)$$

where $L_k^{3D}[l, z]$ are the thresholded wavelet coefficients used in signal reconstruction, and τ_z is the threshold value. The resulting denoised video signal is computed via the inverse non-decimated wavelet transform

$$\hat{\alpha}_k^{3D}[l, z] = \frac{1}{2} \sum_p h[p] \hat{\alpha}_{k+1}^{3D}[l, z - 2^{k+1}p] + \frac{1}{2} \sum_p g[p] L_{k+1}^{3D}[l, z - 2^{k+1}p], \quad (6)$$

where

$$\hat{f}_l^{z, 3D} = \hat{\alpha}_{-1}^{3D}[l, z]. \quad (7)$$

$\hat{f}_l^{z, 3D}$ is the temporally denoised video signal.

B. Order of operations

With a spatial denoising technique of [1] and a temporal denoising technique established above, there still remains the question of the order of operations. The highest quality may occur with temporal domain denoising followed by spatial domain denoising (TFS), or spatial denoising followed by temporal denoising (SFT). There are two thresholds used in the denoising process; one corresponding to each the spatial and temporal denoising techniques. The spatial threshold is labeled τ , and the temporal threshold is labeled τ_z , described above.

Thus, a simple test is conducted with two test video signals. The first video signal is one which contains little motion, and the other contains a great deal of motion. The selected image sequences are the "CLAIRE" sequence from frame #104-167 and the "FOOTBALL" sequence from frame #33-96.

In order to determine the correct order of operations, both of the image sequences are denoised with τ and τ_z ranging from 0–30 for both TFS and SFT denoising operations. The number of resolutions of the non-decimated wavelet transform used in both the spatial and temporal denoising methods is $k = 1 \dots 5$. The average PSNR of each trial is recorded.

Figure 1 gives the results of testing. As shown in Figure

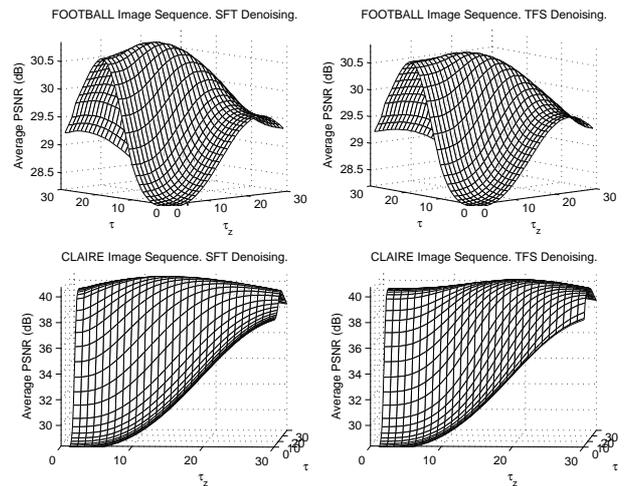


Fig. 1. Test Results Both TFS and SFT Denoising Methods. Upper Left: FOOTBALL Image Sequence, SFT Denoising, Max. PSNR = 30.85, $\tau = 18$, $\tau_z = 12$. Upper Right: FOOTBALL Image Sequence, TFS Denoising, Max. PSNR = 30.71, $\tau = 18$, $\tau_z = 12$. Lower Left: CLAIRE Image Sequence, SFT Denoising, Max. PSNR = 40.77, $\tau = 19$, $\tau_z = 15$. Lower Right: CLAIRE Image Sequence, TFS Denoising, Max. PSNR = 40.69, $\tau = 15$, $\tau_z = 21$.

1, the highest average PSNR is achieved by SFT denoising; first spatially denoising each frame of the sequence followed by temporal domain denoising. Thus, for the proposed denoising method, spatial domain denoising occurs prior to temporal domain denoising, exclusively.

III. PROPOSED MOTION INDEX

A motion index is important in the success of a denoising method in order to discriminate between large temporal variances in the video signal which are caused by noise and large temporal variances which are caused by motion in the original (noiseless) signal. A motion index is able to aid temporal denoising algorithms to eliminate the large temporal variances caused by noise while preserving the temporal variances

caused by motion in the original image sequence, creating a higher quality video signal. That is, the motion index is used to determine τ_z

A. Motion Index Calculation

Several works have developed a motion estimation index to determine the amount of temporal domain processing to perform, i.e., the amount of information that can be removed from the original signal to improve the overall quality [3, 12]. However, none of these proposed indexes are robust to noise corruption, which is an important feature in a motion index. There are a few characteristics that a motion index must possess. One, a motion index should be a localized value. The reasoning behind a localized motion index is because the amount of motion may vary in different spatial portions of the image sequence, thus the motion index should be able to identify those differences. Two, a motion index needs to be unaffected by the amount of noise corruption in a given video signal. A motion index should be robust to noise corruption to aptly determine the proper amount of temporal domain processing.

Thus, a localized motion index is developed which is relatively unaffected by the level of noise corruption in the original image sequence. A spatially averaged temporal standard deviation (SATSD) is used as the index of motion. Spatial averaging is used to remove the noise inherent in the signal, and the standard deviation is used to detect the amount of activity in the temporal domain.

Let us define $\hat{f}_l^{z,2D}$ as pixel value in the spatial location l of the z^{th} frame of an image sequence already processed by the 2D denoising method of [1]. The spatial averaging of the spatially denoised signal is given by

$$A_l^z = \frac{1}{B^2} \sum_{i \in I} \hat{f}_i^{z,2D}, \quad (8)$$

where I is the set of spatial locations which form a square area centered around spatial location l , and B^2 is the number of spatial locations contained in I ; typically, $B = 15$. The value of B must be an odd value to allow for the square area to set centrally around spatial location l . This average is used to find the standard deviation in the temporal domain.

$$\mu_l = \frac{1}{F} \sum_{i=1}^F A_l^i, \quad (9)$$

and

$$M_l = \sqrt{\frac{1}{F} \sum_{i=1}^F (A_l^i - \mu_l)^2}. \quad (10)$$

M_l is the localized motion index, F is the number of frames in the image sequence, and μ_l is the temporal mean of the spatial average at location l .

IV. TEMPORAL DOMAIN PARAMETER SELECTION

The amount of temporal denoising which is beneficial to an image sequence is dependent upon the amount of noise corruption as well as the amount of motion. Thus, the threshold τ_z is given by

$$\tau_z = \alpha \widetilde{\sigma}_n + \beta M_l \quad (11)$$

where the two parameters α and β are determined experimentally using test image sequences. In order to obtain an estimate of the noise level, $\widetilde{\sigma}_n$, an average is taken from the noise estimates of each frame in the image sequence, given by the noise estimates of [1, 11]. If we assume an IID (independent, identically distributed) noise model, we are able to accurately estimate noise in the temporal domain by estimating the noise in the spatial domain. The motion index, M_l is calculated by equation 10.

Again, the sequences CLAIRE and FOOTBALL are used for α and β selection. Each of the image sequences are corrupted with differing levels of noise corruption ($\sigma_n = 10, 20$) and denoised with the SFT denoising method where Equation 11 is used as the temporal domain threshold. Values of α and β are used ranging from $\alpha = 0$ to 3.0 and $\beta = -0.3$ to 0.3. The results of this testing is given in Figure 2. As shown in Figure

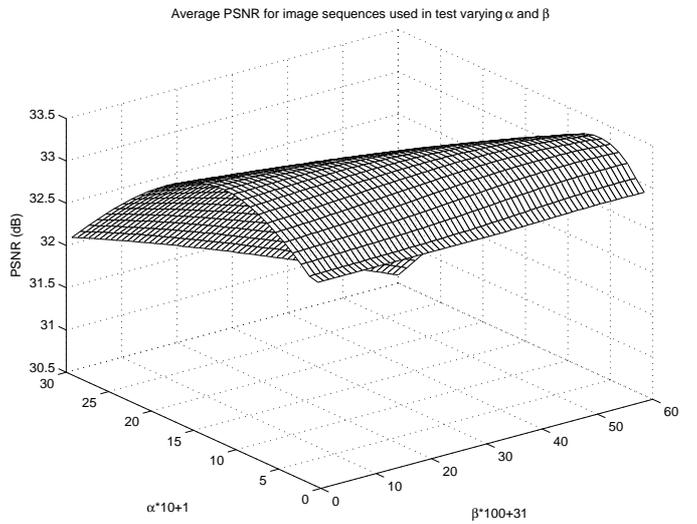


Fig. 2. α and β Parameter Testing for Temporal Domain Denoising.

2 the maximum average PSNR is achieved when $\alpha = 1.0$ and $\beta = -0.16$. The result is reasonable, of course, because as the motion increases in an image sequence the redundancy between frames decreases, and the benefits of temporal domain processing decrease. Thus, as the testing has shown, the temporal domain threshold decreases as the motion increases.

V. EXPERIMENTAL RESULTS

The proposed video denoising algorithm first is applied to each of the video frames individually and independently. The method of [1] was developed earlier by our previous research to denoise images, and is used as the spatial denoising portion of the proposed video denoising algorithm.

The video signal is then denoised in the temporal domain by the method developed in Sections II and IV. The temporal denoising algorithm is a selective shrinkage algorithm which uses a proposed motion estimation index to determine the temporal threshold, τ_z . The temporal threshold is modified by the motion index to effectively eliminate temporal domain noise while preserving important motion information.

The SALESMAN image sequence is used to determine the effectiveness of the proposed video denoising method. It is corrupted with various levels of noise and denoised with the

methods of [1, 3, 12, 17] as well as the proposed method. The results are given in Figures 3 and 4. As shown in Figures

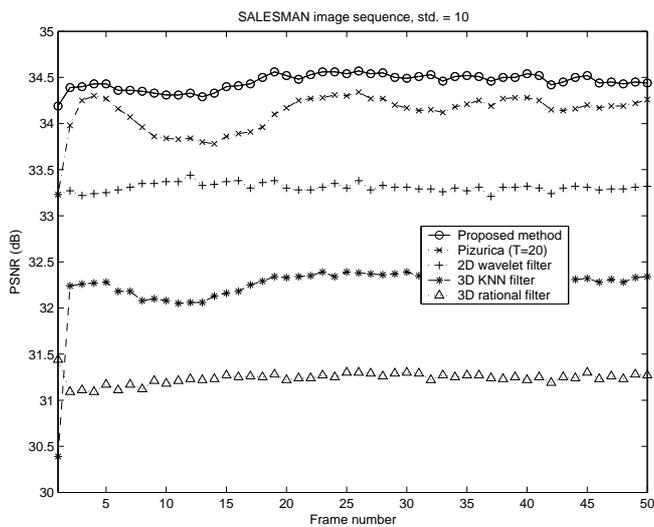


Fig. 3. Denoising Methods applied to the SALESMAN Image Sequence, std. = 10

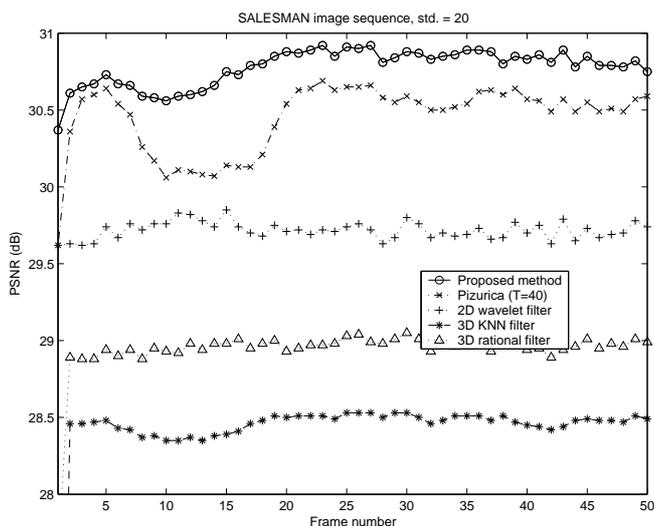


Fig. 4. Denoising Methods applied to the SALESMAN Image Sequence, std. = 20

3 and 4, the proposed method outperforms the other methods presented. In both cases, the proposed denoising method has a higher average PSNR than the other denoising methods tested. Also, note that in the method of [12], the threshold T is manually changed due to video content and noise level to obtain the highest average PSNR using that particular method. In the proposed method, the temporal domain threshold is automatically calculated due to estimates of the noise level and motion.

VI. CONCLUSIONS

In this paper, a new combined spatial and temporal domain wavelet shrinkage method is developed for the removal of noise in video signals. The proposed method uses a geometrical approach to spatial domain denoising to preserve edge information, and a newly developed motion estimation index for selective wavelet shrinkage in the temporal domain.

The spatial denoising technique is a selective wavelet shrinkage algorithm developed in [1] and is shown to outperform other wavelet shrinkage denoising algorithms given in the literature both in denoised image quality and computation time.

The temporal denoising algorithm is also a selective wavelet shrinkage algorithm which uses a motion estimation index to determine the level of thresholding in the temporal domain.

The proposed motion index is experimentally determined to be more robust to noise corruption than other methods, and is able to help determine the threshold value for selective wavelet shrinkage in the temporal domain. With the motion index and temporal domain wavelet shrinkage, the proposed video denoising method is experimentally proven to outperform other methods given in the literature for various levels of noise corruption applied to video signals with varying amounts of motion.

REFERENCES

- [1] E. J. Balster, Y. F. Zheng, and R. L. Ewing. "Fast, Feature-based Wavelet Shrinkage Algorithm for Image Denoising". In *Proc. IEEE Int. Conf. Integration of Knowledge Intensive Multi-Agent Systems*, volume 1, pages 722–728, 2003.
- [2] J. B. Bednar and T. L. Wat. "Alpha-Trimmed Means and Their Relationship to Median Filters". *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-32:pages 145–153, Feb. 1984.
- [3] F. Cocchia, S. Carrato, and G. Ramponi. "Design and Real-Time Implementation of a 3-D Rational Filter for Edge Preserving Smoothing". *IEEE Trans. Consumer Electronics*, vol. 43:pages 1291–1300, Nov. 1997.
- [4] M. Ghazel, G. H. Freeman, and E.R. Vrscay. "Fractal-Wavelet Image Denoising". In *Proc. IEEE Int. Conf. Image Processing*, volume 1, pages 1836–1839, 2002.
- [5] S. D. Kim, S. K. Jang, M. J. Kim, and J. B. Ra. "Efficient Block-Based Coding of Noise Images by Combining Pre-Filtering and DCT". In *Proc. IEEE Int. Symp. Circuits and Systems*, volume 4, pages 37–40, 1999.
- [6] W. Ling and P. K. S. Tam. "Video Denoising Using Fuzzy-connectedness Principles". In *Proc. 2001 Int. Symp. Intelligent Multimedia, Video, and Speech Processing*, pages 531–534, 2001.
- [7] M. Malfait and D. Roose. "Wavelet-Based Image Denoising Using a Markov Random Field A Priori Model". *IEEE Trans. Image Processing*, vol. 6:pages 549–565, Apr. 1997.
- [8] S. Mallat and W. L. Hwang. "Singularity Detection and Processing with Wavelets". *IEEE Trans. Information Theory*, vol. 38:pages 617–623, March 1992.
- [9] M. Meguro, A. Taguchi, and N. Hamada. "Data-dependent Weighted Median Filtering with Robust Motion Information for Image Sequence Restoration". *IEICE Trans. Fundamentals*, vol. 2:pages 424–428, 2001.
- [10] O. Ojo and T. Kwaaitaal-Spassova. "An Algorithm for Integrated Noise Reduction and Sharpness Enhancement". *IEEE Trans. Consumer Electronics*, vol. 46:pages 474–480, May 2000.
- [11] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy. "A Joint Inter- and Intrascale Statistical Model for Bayesian Wavelet Based Image Denoising". *IEEE Trans. Image Processing*, vol. 11:pages 545–557, May 2002.
- [12] A. Pizurica, V. Zlokolica, and W. Philips. "Combined Wavelet Domain and Temporal Video Denoising". In *Proc. IEEE Int. Conf. Advanced Video and Signal Based Surveillance*, volume 1, pages 334–341, July 2003.
- [13] P. Rieder and G. Scheffler. "New Concepts on Denoising and Sharpening of Video Signals". *IEEE Trans. Consumer Electronics*, vol. 47:pages 666–671, Aug. 2001.
- [14] L. Shutao, W. Yaonan, Z. Changfan, and M. Jianxu. "Fuzzy Filter Based on Neural Network and Its Applications to Image Restoration". In *Proc. IEEE Int. Conf. Signal Processing*, volume 2, pages 1133–1138, 2000.
- [15] C. Vertan, C. I. Vertan, and V. Buzuloiu. "Reduced Computation Genetic Algorithm for Noise Removal". In *Proc. of IEEE Conf. Image Processing and Its Applications*, volume 1, pages 313–316, July 1997.
- [16] Y. F. Wong, E. Viscito, and E. Linzer. "PreProcessing of Video Signals for MPEG Coding by Clustering Filter". In *Proc. IEEE Int. Conf. Image Processing*, volume 2, pages 2129–2133, 1995.
- [17] V. Zlokolica, W. Philips, and D. Van De Ville. "A New Non-linear Filter for Video Processing". In *Proc. IEEE Benelux Signal Processing Symposium*, volume 2, pages 221–224, 2002.

Fiber Optic Interferometric Sensing via Digital Signal Processing

Barry Griffin, Michael Connelly

Optical Communications Research Group, Department of Electronic and Computer Engineering,
University of Limerick, Limerick, Ireland.

e-mail: *Barry.Griffin@ul.ie*

Abstract – This paper presents a digital based acquisition system, concerning fiber optic sensing. The aforementioned system transmits a demodulated signal to the sensing interferometer, demodulates the returning interferometric signal, while simultaneously relaying control and signal data via PC based software.

1 Introduction

Fiber-based interferometry has been employed in the sensing of an array of different physical phenomena for a number of years now. These include the sensing of strain, pressure, temperature, and flow [1]. Presented in this paper is an approach to interferometric pressure sensing which is conducted using a Digital Signal Processor (DSP). The DSP is capable of modulating the laser transmitter and demodulating the photodetected signal, as well as engaging in duplex communication with a Windows PC application. The application facilitates the user in controlling the various parameters involved when obtaining the interferometric measurand signal. This software also permits the user to view and record time and spectral domain plots concerning the demodulation of the sensed signal. The system outlined in this paper is targeted primarily at industrial applications requiring the general purpose measurement of dynamic pressure, via fiber optic interferometric sensors; and in addition, lends itself as a generic design for more specific applications. Such applications might well involve sharing the processor with operations not pertinent to pressure sensing. At the crux of sensor-based interferometric detection systems, are the demodulation schemes they employ when converting the photodetected signal to a signal proportional to that being sensed. Demodulation schemes fall into two main categories, active and passive [2]. Active schemes, while more sensitive, require electrical feedback, negating one of the primary advantages associated with fiber sensors. Passive schemes are therefore more popular than their active counterpart, and for this reason the authors have decided to concentrate their design efforts on the former. In particular, we are investigating the Synthetic-Heterodyne demodulation technique [3][4].

From a commercial perspective, this system is attractive for three major reasons when compared

with more traditional analog systems. The first reason is with respect to the flexibility provided by digital hardware, over analog. For example, if the demodulation or control algorithm requires a modification, this can be achieved by simply modifying the software, which runs on the DSP. It is also possible to select different demodulation algorithms depending on the type of sensor used, while under such circumstances, an analog design would require additional hardware. The second advantage relates to the fact that DSPs are now readily available in all sorts of environments. Again as an example, many aircraft now rely heavily on DSP processors for control and monitoring purposes. Therefore for such cases, it makes sense to fully utilize the available resources by opting for an acquisition method based around digital signal processing. The final advantage granted concerns integrating the primary function of obtaining the pressure signal, with the auxiliary functions of logging and control. Logging is handled by simply storing digital samples of the acquired signal in storage memory, while control is handled by augmenting the software to flag the required events.

2 System

Fig. 1. outlines the various components and processes involved when sensing a pressure signal, and in presenting this signal to the user in an appropriate form. This diagram can be abstracted into three physical parts, the PC, the Acquisition Unit and the Sensor/Interferometer.

Generally, passive demodulation techniques require that the laser output is frequency modulated, usually via amplitude modulation of the laser drive current [5]. To achieve this, a Digital-to-Analog Converter (DAC) interfaces to the laser driver, allowing the DSP to dictate the type of modulation applied, whether that be sine wave or saw-tooth modulation, etc. An Analog-to-Digital Converter (ADC) is also utilized in converting the

photodetected signal to the digital domain. This digitalized signal is then demodulated by the DSP in order to extract the measurand signal. The laser is a Distributed Feedback (DFB) fiber laser, operating with a very narrow linewidth, at a wavelength of 1550nm. A laser driver was selected to provide both a d.c. bias and a modulation current to the laser. This is accomplished by means of voltage to current conversion, applied to the laser driver voltage input. A clipper circuit between the DAC and the voltage input is also necessary in order to protect the laser from excessive currents. Mounting the laser on a copper block provides satisfactory temperature stabilization. A p-i-n photodiode is used in converting the returning lightwave signal from the interferometer to an electrical signal. The photodetector circuitry is concerned with the suitable amplification of this signal, and in insuring that capacitances within the circuit do not forward bias the photodiode at any stage of its operation. The Codec Daughterboard consists of two 16bit DACs and two 16bit ADCs, operating with a maximum sample rate of 44.1kHz. The ADCs include an amplifier stage, which can be controlled by the DSP software in real-time. As a result, the user has the option, via the control software, of adjusting the amplification applied to the photodetected signal. A 32bit floating-point processor, the Texas Instruments TMS320C6711, was chosen to serve as the motherboard for the Acquisition Unit. The PC and Acquisition Unit communicate with each other via a parallel port interface. With respect to the Synthetic-Heterodyne demodulation technique, this system is capable of demodulating signals with a maximum frequency of 4.41kHz. A correlation algorithm is employed by the DSP in order to synchronise the incoming modulated signal with the local oscillator signals employed by the detection scheme, if required.

3 Software

PC based software has been specially designed and written to assist the user in both the control of the acquisition system, as well as the display and storage of the measured signal. From a control perspective, the user selects a demodulation scheme and adjusts the parameters relating to this scheme to suit their particular application. These parameters, in the case of Synthetic-Heterodyning, include the frequency and depth of laser modulation, the amplification applied to the photodetected signal, the filter lengths, as well as the sensor transfer function, which relates the demodulated signal with the physical phenomenon under investigation. The user also has the option of saving the parametric settings to file. A number of demodulation schemes are currently covered, including Synthetic

Heterodyne [3], Hmodyne with Phase Generated Carrier [6] and Pseudo-Heterodyne [7]. With regard to signal presentation, real-time plots of the interferometric signal, the demodulated signal, as well as the magnitude and phase spectra of the detected and demodulated signals, are available, as seen in Fig. 2. The user may also save the demodulated signal to file for subsequent analysis.

4 Experimental Results

In order to establish the sensitivity and reliability of the system, a low-finesse extrinsic Fabry-Perot cavity has been constructed. This sensor is diagrammatised in Fig. 3. and includes a reflecting mirror mounted on a piezoelectric cylinder, which moves in response to an applied voltage. The PZT is capable of inducing a maximum phase shift of π radians within the cavity, which corresponds to a displacement of 775nm. For experimental purposes, the PZT was driven at 10Volts, producing a phase shift of $\pi/10$ radians. The results are depicted in Fig. 4. below. The distortion of the demodulated signal is believed to result primarily from multiple reflections within the cavity.

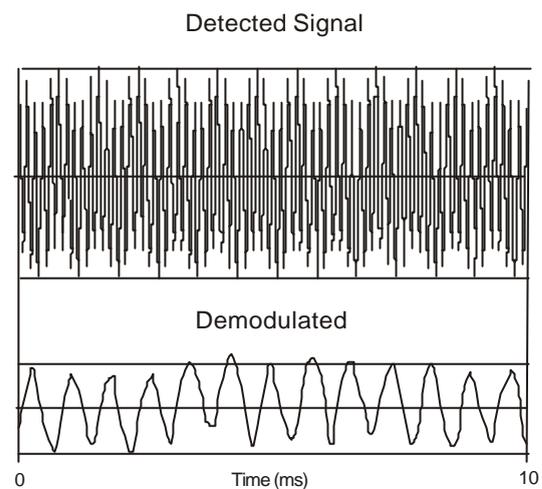


Figure 4: Above: signal prior to demodulation, Below: Signal after demodulation.

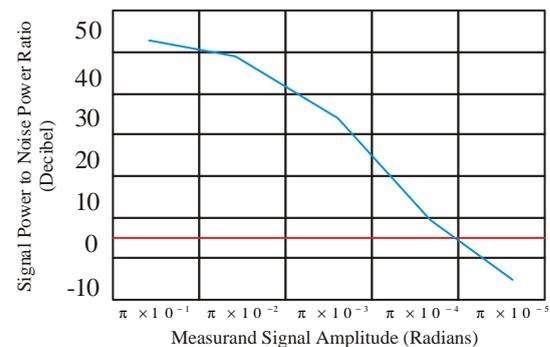


Figure 5: Minimum Detectable Phase.

The sensitivity of the system was also established as follows -

$$\frac{1 \times 10^{-4} \text{ rads}}{\sqrt{44100 \text{ Hz}}} = 5 \times 10^{-7} \text{ rads} / \sqrt{\text{Hz}}$$

where 1×10^{-4} is approximately the minimum detectable phase for the bandwidth of 44.1kHz.

5 Future Work

Active demodulation schemes are DC based. To facilitate these schemes, it would be necessary to remove the DC coupling stages from both the ADCs and DACs. The current system is designed specifically for two-arm interferometers. An increase in the number of arms would require an increase in ADCs and DACs. A future design could include converters with a higher resolution than that currently available. Replacing the 16bit ADCs and DACs would result in a decrease in quantization noise and therefore an increase in sensitivity. This system also has the potential to sense physical properties other than dynamic pressure.

6 Conclusion

We have designed and developed a digital interferometric sensor interrogation system and its associated software, which enables the user to configure a wide variety of demodulation schemes and also carry out sophisticated real-time analysis on the demodulated signal.

Acknowledgement

Supported by Enterprise Ireland Research Innovation Fund, Project IF/2002/332N.

REFERENCES

- [1] A.D. Kersey, A. Dandridge, "Applications of Fiber-Optic Sensors, IEEE Transactions on Components and Manufacturing Technology", Vol. 13, No. 1, March 1990.
- [2] T.G. Giallorenzi, J.A. Bucaro, A. Dandridge, G.H. Siegel, J.H. Cole, S.C. Rasleigh, R.G. Priest, "Optical Fiber Sensor Technology", IEEE J. Quantum Electron., Vol. QE-18, No. 4, 1992.
- [3] J.A. Cole, B.A. Danver, J.A. Bucaro, "Synthetic-Heterodyne Interferometric Demodulation", IEEE J. Quantum Electron., Vol. QE-18, No. 4, 1982.
- [4] M. J. Connelly, "Digital Synthetic-Heterodyne Interferometric Demodulation", J. Opt. A: Pure Appl. Opt. 4, 2002.
- [5] A. Dandridge, L. Goldberg, "Current Induced Frequency Modulation in Laser Drivers", Electron. Letters, Vol. 18, 1982.
- [6] A. Dandridge, A. B. Tveten, T.G. Giallorenzi, "Homodyne Demodulation Scheme for Fiber Optic Sensors with Phase Generated Carrier", IEEE J. Quantum Electron., Vol. QE-18. No. 4, 1982.
- [7] H. Sing, J.S. Sirkis, J. Andrews, R. Pulfrey, "Evaluation of Integrated Optic modular-Based Detection Schemes for In-Line Fiber Etalon Sensors", J. Lightwave Technol., Vol. LT-13, 1995.

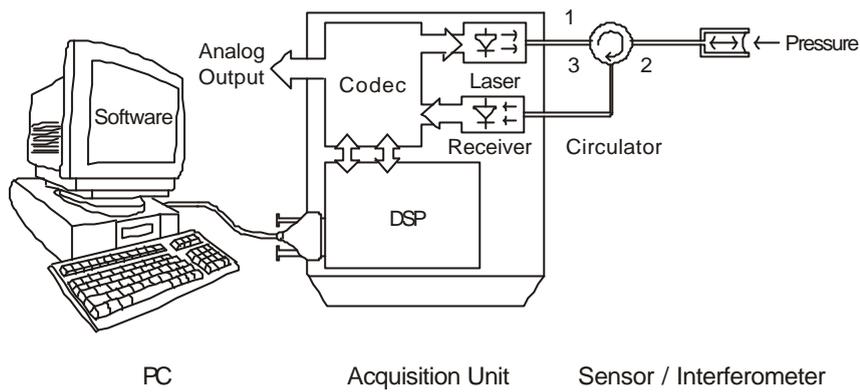


Figure 1: System Overview.

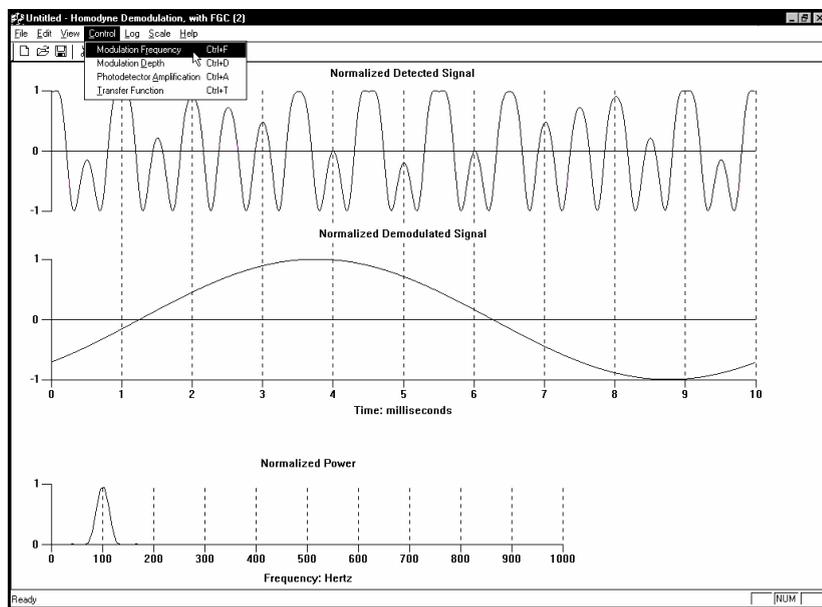


Figure 2: Top: Detected signal; Middle: Demodulated signal; Bottom: Magnitude spectrum of demodulated signal.

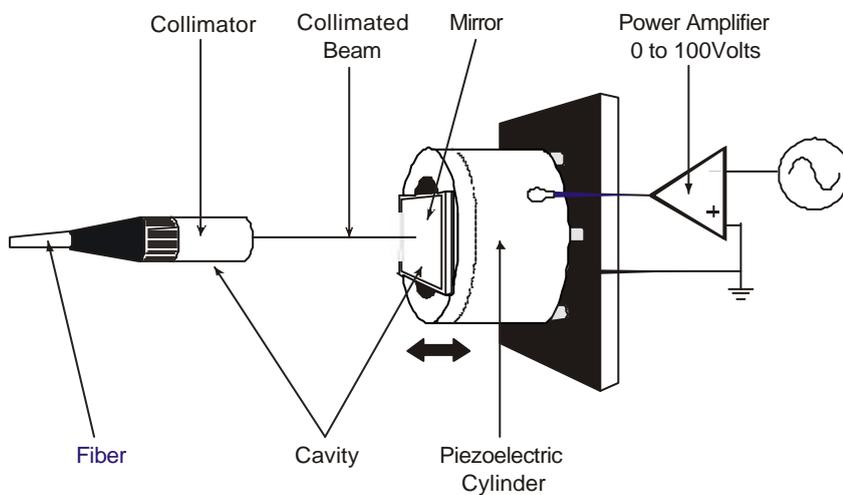


Figure 3: Fabry-Perot Cavity.

FUZZY IMAGE SEGMENTATION CONSIDERING OBJECT SURFACE SIMILARITY

M. Ameer Ali, Laurence S Dooley and Gour C Karmakar

Gippsland School of Computing & Information Technology
Monash University, Australia

Email: {Ameer.Ali, Laurence.Dooley and Gour.Karmakar}@infotech.monash.edu.au

ABSTRACT

Clustering algorithms are very sensitive to the feature used and the type of an object. To reduce this, a *fuzzy image segmentation using location and intensity information* (FSLI) algorithm was developed by considering connectivity and similar surface variations (SSV) of an image. However, FSLI did not consider SSV of each object and was unable to segment images containing multiple objects with SSV and at least one object with dissimilar surface variations. To improve the effectiveness of FSLI for multiple regions, this paper introduces a new *fuzzy image segmentation considering object surface similarity* (FSOS) algorithm, which directly considers object SSV and the arbitrariness of the fuzzy c-means (FCM) algorithm for pixel location. It also perceptually selects the threshold within the range of human visual perception. Quantitative results confirm the improved segmentation performance of FSOS compared with FSLI, FCM, possibilistic c-means (PCM), and suppressed fuzzy c-means clustering (SFCM) for many different images.

1. INTRODUCTION

Object-based image segmentation is a most challenging task because of a myriad number of objects and the huge variations among them that make it very complex to approximate all the objects within a general framework [1], [2].

Clustering algorithms [3]-[11], use many different feature types, such as brightness (e.g., pixel *intensity*¹ of a gray-scale image) and geometric information (e.g., location) but their effectiveness is solely dependent on the types of the feature used and the information about the objects in an image. This raises the question about which type of feature produces better results for which type of image, thereby limiting the generalization of the clustering algorithm [3]. For instance, FCM [3], PCM [5] and SFCM [12] cannot separate image regions (objects) having similar pixel intensities by considering only their PI. They may however be able to, by exploiting PL information or a CIL. In the same way, clustering cannot segment asymmetrically oriented adjacent regions having different intensities by only considering PL, but may well be able to do so by considering respective PI. It has been found that even clustering algorithms using both features, i.e. CIL, do not necessarily produce the

¹ The terminology *PL*, *PI* and *CIL* refer respectively to *pixel location*, *pixel intensity* and a *combination* of pixel intensity and normalized pixel location.

expected results for all images. This motivates to produce a final result by fusing the initial segmented results produced by any clustering algorithm separately using different feature sets. To address these factors, Ameer et al. [13] recently proposed an algorithm called *fuzzy image segmentation using location and intensity information* (FSLI) which considered connectivity and SSV of the entire image. This algorithm however failed to segment all objects well if an image contained multiple objects having SSV and at least one other object having dissimilar surface variations (DSV). It is also sensitive to the perceptually selected thresholds.

This paper introduces a new algorithm called *fuzzy image segmentation considering object surface similarity* (FSOS) by considering each object SSV and the arbitrariness of FCM for location to improve the effectiveness of FSLI for multiple regions. In order to reduce the sensitivity of the thresholds, an algorithm for automatically selecting the feature sets is introduced in the bedrock of FSOS. The threshold value is perceptually selected from a range of values by considering human visual perception. This paper also includes a numerical analysis of PCM and SFCM for all feature sets and the proposed FSOS in addition to FSLI and FCM for all feature sets using one of the objective segmentation evaluation methods [1].

The paper is organized as follows: In Section 2, issues relating to the identification of SSV are discussed, while the theoretical underpinning of the FSOS algorithm is presented in Sections 3 and 4. A detailed qualitative and quantitative performance analysis of the segmentation results of the new algorithm is provided in Section 5, with some conclusions presented in Section 6.

2. IDENTIFICATION OF SIMILAR SURFACES

A fundamental drawback of FCM using PL is that it arbitrarily divides objects, irrespective of the object SSV and DSV. To reduce this arbitrariness², it is important to try and identify similar and dissimilar object surfaces in an image, though this represents an intractable challenge. Two possible scenarios exist by which surfaces may be considered similar from a brightness perspective. These are that the surfaces have: (i) similar intensity and (ii) possess SSV. In the former, FCM using CIL performs a function similar to FCM using PL. Note, when objects have SSV i.e., comprise repeated patterns of bright and dark pixels, FCM using a combination is unable to separate them. This provided

² A mathematical proof of this is not included due to page limitation.

the motivation to explore FCM using a combination to determine the type of object surface variations. This implies that when FCM using a combination is unable to separate a group of objects, all these objects will have SSV^2 .

3. THE MODIFIED FCSI ALGORITHM

In the FCSI algorithm, the rationale for the merging of the segmented results independently produced by FCM using only PL and PI had already been presented. However, the FCSI algorithm is very sensitive to the threshold and higher level features of an object are not considered. To address these issues, the original FCSI algorithm was modified by incorporating a connectedness feature [13]. The overlapping pixels in the modified FCSI (MFCSI) algorithm are distributed using 8-connected objects and 8-connectivity. The remaining pixels are redistributed by FCM using CIL and this reduced the sensitivity of the perceptual threshold.

4. THE FSOS ALGORITHM

The original FSLI algorithm [13] was unable to correctly segment images comprising objects containing two or more similar objects with SSV and at least one other object having DSV. To address this limitation and also reduce the arbitrariness of FCM using PL (Section 2), this paper proposes a new algorithm (Algorithm 2) called *fuzzy image segmentation considering object surface similarity* (FSOS) which directly considers object surface similarity.

Given any image may contain objects with both similar and dissimilar surface variations; any strategy that is able to identify between them in the segmentation process has the potential to offer superior results. A detailed description of how this is achieved is presented in Algorithm 1. This uses the initially segmented regions (R^C) produced by FCM using CIL and the area of the segmented region. Area $A^{R_i^C}$ of region R_i^C can be defined as,

$$A^{R_i^C} = \text{Area}(\text{Convexhull}(R_i^C)) \quad (1)$$

where $\text{Area}(\bullet)$ and $\text{Convexhull}(\bullet)$ determine the area and the vertices of the convex hull of a region respectively. The merging of the two regions of R^C to form a new region R_k^M is expressed by:

$$R_k^M = \{P(x, y) | P(x, y) \in R_i^C \vee P(x, y) \in R_j^C\} \quad (2)$$

where M is the number of merged regions in R_k^M , $2 \leq M \leq \mathfrak{R}$, $1 \leq k \leq \mathfrak{R}/2$, $i \neq j$ and $1 \leq i, j \leq \mathfrak{R}$, \mathfrak{R} is the number of segmented regions. The difference between the maximum area of all merged regions and the area $A^{R_k^M}$ of the merging region R_k^M represents a measure of the shape distortion (Step 5 in Algorithm 1). If this is within the tolerance (≈ 0.5 dB) of the human visual perception, the merged regions will have SSV. This hypothesis is applicable to all possible merging regions. Finally, objects having DSV are separated from those objects having SSV which are represented by region R^D , where D is the number of objects having DSV and $1 \leq D \leq \mathfrak{R}$. For objects having SSV, FCM using PL must be used for segmentation

because they cannot be separated by FCM using CIL and PI (Step 2 in Algorithm 2). Those multiple objects having DSV are segmented using the MFCSI algorithm since they have distinctive PIs. For \mathfrak{R} regions, (since the degree of arbitrariness of FCM using PL increases in $O(\mathfrak{R}^2)$), for more than two objects with DSV, FCM using CIL produces better segmented results than FCM using only PL. In this case therefore, MFCSI is used with R^C and R^I where R^I is the initial segmented regions by FCM using PI. However, for two regions it is important to make a decision about which feature (either CIL or PL) is used in conjunction with PI. This decision has been considered by taking account of the superiority of using CIL and PI over PL. This is illustrated in Figure 1 (a) and (b), where angles θ_1 and θ_2 represent the discrepancy between the two decision boundaries separately produced by FCM, using *PL and CIL* and *PL and PI* respectively.

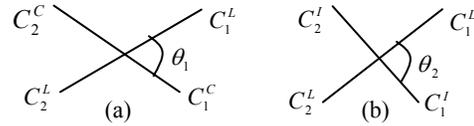


Figure 1: Angle between two decision boundaries produced by FCM separately using (a) *PL and CIL*, (b) *PL and PI*.

Algorithm 1: Identification of objects having SSV and DSV

Precondition: Initially segmented regions R^C and \mathfrak{R} .

Post condition: A list of objects having similar and dissimilar surface variations are R_i^M and R^D respectively.

1. Set $M = 1$ and $k = 1$.
2. Form region R_k^M by combining R_i^C and R_j^C using (2).
3. Calculate areas $A^{R_k^M}$, $A^{R_i^C}$ and $A^{R_j^C}$ using (1).
4. Find the maximum area $A_i^{R^C}$ of $(M+1)$ regions in R_k^M .
5. IF $\left(\left| A^{R_k^M} - A_i^{R^C} \right| \leq T_{\max} \right)$ THEN the two regions R_i^C and R_j^C have SSV and increment M
6. Repeat Steps 2-5 for forming R_k^M by merging R_k^M and another region from R^C except the earlier merged regions.
7. IF $(M \geq 2)$ THEN increment k and repeat Steps 1-6.
8. Separate region R^D having D objects with DSV from R^C .

In Figure 1, $\theta_1 > 45^\circ$ indicates the existence of two different objects with quite different brightness (PI values for gray-scale image) and CIL outweighs PL. In these circumstances, there is a strong argument for using CIL with PI in MFCSI, without connectivity. The reason for this is that all regions have distinctive PI, but one particular region may possess similar PIs to those of other regions, which may be connected to it. To avoid the possibility of misclassification, connectivity is not applied (Step 4 in Algorithm 2). When $\theta_1 \leq 45^\circ$ and $\theta_2 > 45^\circ$, it implies that CIL produces comparable results to PL, while PI differs with PL. This motivates the use of CIL with PI in MFCSI using connectivity because of the potential impact of PL rather than PI in CIL. For

all other cases, PL and PI are used in MFCSI using connectivity. Note, that the value of the threshold T_{\max} in Algorithm 1 is determined by considering human visual perception.

Algorithm 2: The FSOS algorithm

Precondition: The foreground region f to be segmented, \mathfrak{R} , θ_1 and θ_2 .

Post condition: The final segmented regions R .

1. Segment f by FCM using a combination into \mathfrak{R} regions represented by R^C .
2. Find R_k^M and R^D applying

Algorithm 1 for R^C .

3. IF ($k \geq 1$) THEN FOR $i = 1, \dots, k$
 Segment R_i^M into M regions by FCM using PL.
4. IF ($D \geq 2$) THEN
 IF ($D = 2$) THEN
 IF ($\theta_1 > 45^\circ$) THEN segment R^D into D regions by MFCSI without using connectivity for R^1 and R^C .
 ELSE IF ($(\theta_1 \leq 45^\circ)$ AND ($\theta_2 > 45^\circ$)) THEN
 Segment R^D by MFCSI for R^1 and R^C
 ELSE segment R^D by MFCSI for R^1 and R^L where R^L is the initial segmented regions by FCM using PL.
 ELSE segment R^D by MFCSI for R^1 and R^C .

5. EXPERIMENTAL RESULTS

The new FSOS, FSLI, fuzzy c-means (FCM) [3], PCM [5] and SFCM [12] algorithms were implemented using Matlab 6.1 (The Mathworks Inc.). For FCM, PCM and SFCM, three feature sets: PI, PL, and CIL were used. A total of 146 different types of natural and synthetic 8-bit gray-scale images were used in the experiments. These comprised different regions (objects) having similar and dissimilar PI and surface variation³. In segmenting the foreground objects, the background was manually removed from all images, and since the background of an object is filled with zero, all foreground zeros were replaced by 1 to differentiate them. PL in the form of the (x, y) coordinates of a pixel were normalized within the range $[0, 255]$ in order to keep them within the same range as the pixel intensities in order to reduce the effect of image size. All the results of the new FSOS algorithm presented were produced using a perceptually selected threshold T_{\max} set to 5% shape distortion of the largest merged region (see Section 4). The quantitative analysis was conducted using *discrepancy based on the number of misclassified pixels* [1]. Type I, $errorI_i$ is the error percentage of all i^{th} region pixels that are misclassified in the other regions, while Type II, $errorII_i$ is the error percentage of all region pixels that are misclassified into i^{th} region. The original images and their manually segmented reference regions are shown in Figures

³ IMSI’s Master Photo Collection, 1895 Francisco Blvd. East, San Rafael, CA 94901-5506, USA.

2(a)-2(b) and 3(a)-3(b). Note, that the manually referenced and segmented regions are displayed by different gray levels instead of the original region intensities, to provide a better visual interpretation of the segmentation results. For the space limitations, the best two results of FCM, PCM and SFCM with the results of the FSLI and FSOS algorithms are provided here.

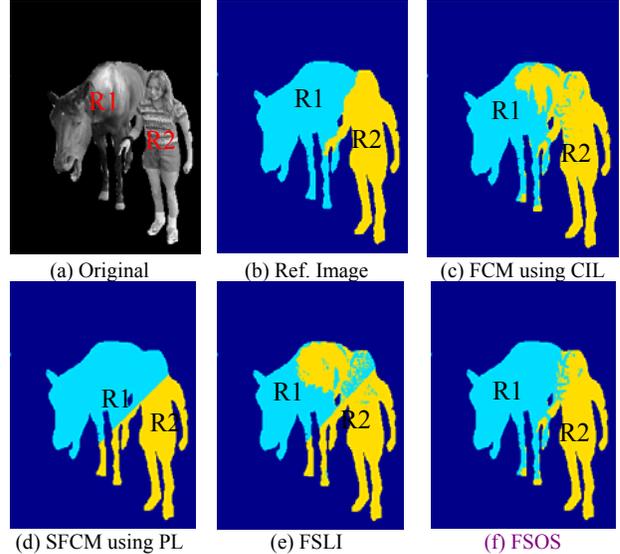


Figure 2: (a) Original horse image, (b) Manually segmented reference of (a). (c)-(f) Segmented results of (a).

Table 1: Average error percentages and the number of images for the superior results of different algorithms.

Algo-rithm	FCM			PCM			SFCM			FSLI	FSOS	
	PL	PI	CIL	PL	PI	CIL	PL	PI	CIL			
Average Error	Horse	14.4	35.4	13	48.6	48.6	46.4	17	35.4	35.4	25.1	5.3
	Snake	17.6	41.3	14.2	24.9	39.1	19.7	20.1	41.1	41.7	18.2	4.3
	146 Im	24.8	30.1	20.5	36.8	33.7	33.9	24.7	29.8	30.1	20.2	17.5
#Image	146	22	5	27	14	11	7	23	5	6	32	50

The experiments were performed using a horse image (Figure 2(a)) having different two regions: the horse (R_1) and the woman (R_2). The four best segmented results of FCM, PCM, SFCM, FSLI and FSOS taken one from each algorithm for this image are shown in Figure 2(c)-(f). If the segmented results in Figure 2(c)-(e) are compared with the manually segmented reference regions in Figure 2 (b), it is visually apparent that a considerable number of pixels of the horse (R_1) region were misclassified into the woman (R_2) region and vice versa since both the horse and the woman regions contain a certain range of PI variations. Most of these pixels were correctly classified by

the FSOS algorithm (Figure 2 (f)) because of mainly selection of appropriate feature sets. The numerical results (the average of Type I and Type II errors) of FCM, PCM and SFCM for all feature sets (PI, PL and CIL), FSOS and FSLI for the horse image are shown in Table 1. Table 1 shows that the average error percentage is 5.3 for FSOS, while the second best average error percentage achieved by FCM using CIL is 13. This also proves the superiority of FSOS over FSLI and FCM, PCM and SFCM using any of the three feature sets.

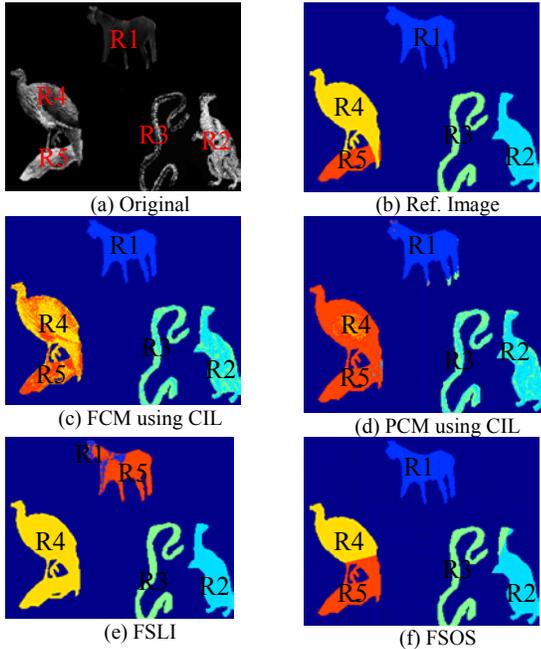


Figure 3: (a) Original snake image, (b) Manually segmented reference of (a). (c)-(f) Segmented results of (a).

Another sample image in Figure 3(a) used in the experiment called the snake image having five different regions: the horse (R_1), the kangaroo (R_2), the snake (R_3), the peacock (R_4) and the branch of a tree (R_5). Note, the kangaroo (R_2) and the snake (R_3) have SSV; the peacock (R_4) and the branch of a tree (R_5) have SSV and the remaining horse (R_1) has only DSV. The four best segmented results taken one from each algorithm for the snake image (Figure 3 (a)) are shown in Figure 3(c)-(f) respectively. For objects having SSV, FCM and PCM using CIL could not separate the peacock (R_4) from the branch of a tree (R_5) and the kangaroo (R_2) from the snake (R_3) (Figure 3(c)-(d)). FSLI was unable to separate the peacock (R_4), the branch of a tree (R_5) and the horse (R_1) because of not considering SSV for each object (Figure 3 (e)). The result shown in Figure 3(f) shows that FSOS completely separated the horse (R_1). FSOS also clearly separated the snake (R_3) from the kangaroo (R_2) and the branch (R_5) from the peacock (R_4) except a few amount of misclassified pixels of the kangaroo (R_2) and the peacock (R_4). This confirms that FSOS produced better results than FCM, PCM, SFCM and FSLI because of utilizing the concept of SSV and DSV for individual object and selection of proper feature sets. FSOS also produced a

considerable less average error (4.3%) compared with the second best average error (14.2%) produced by FCM using CIL given in Table 1.

Experiments were conducted on total 146 real and synthetic images, including up to 5 regions. Overall, the proposed FSOS algorithm produced better results for 50 images compared with the second best 32 produced by FSLI shown in Table 1. Table 1 also shows that the average error percentages of the proposed FSOS and FSLI algorithms for 146 images are 17.5 and 20.2 respectively, while the best average error percentages of FCM, PCM and SFCM for 146 images are 20.5, 33.7 and 24.7 respectively. From all the experimental results, it can be concluded that the proposed FSOS algorithm outperforms the FSLI and FCM, PCM and SFCM for all three feature sets.

6. CONCLUSIONS

This paper has introduced a new algorithm called *fuzzy image segmentation considering object surface similarity* (FSOS) which has exhibited superior segmentation performance compared to FSLI, FCM, PCM and SFCM both qualitatively and quantitatively. The value of T_{max} is perceptually selected from a range of values by considering shape distortion based on the human visual perception.

7. REFERENCES

- [1] Gour C Karmakar, "An Integrated Fuzzy Rule-Based Image Segmentation Framework," PhD, Thesis Dissertation, 2002
- [2] I. Gath, and A. B. Geva, "Unsupervised Optimal Fuzzy Clustering," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 2(7), pp. 773-781, 1989.
- [3] J.C. Bezdek, "Pattern Recognition with fuzzy objective function algorithm," (New York: Plenum Press, 1981).
- [4] Jian Yu, Houkuan Huang and Shengfeng Tian, "An Efficient Optimality Test for the Fuzzy C-Means Algorithm," IEEE Intern. Con. on Fuzzy Systems, pp. 98-103, 2002.
- [5] R. Krishnapuram and J. M. Keller, "A Possibilistic Approach to Clustering," Fuzzy Systems, Vol. 2(2), pp. 98-110, 1993.
- [6] Carl G. Looney, "Interactive clustering and merging with a new fuzzy expected value," Pattern Recognition Letters, Vol. 35(11), pp. 2413-2423, 2002.
- [7] Y. A. Tolias and M. Panas, "On applying Spatial Constraints in Fuzzy Image Clustering Using a Fuzzy Rule-Based System," IEEE Intern. Con. on Signal Processing Letters, Vol. 5(10), pp. 245-247, 1998.
- [8] A. W. C. Liew, S. H. Leung and W. H. Lau, "Fuzzy Image Clustering Incorporating Spatial Continuity," IEE Proc.-Vis on Image Signal Process, Vol. 147(2), pp. 185-192, 2000.
- [9] Gerardo Beni and Xiaomin Liu, "A least biased fuzzy clustering method," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 16(9), pp. 954-960, 1994.
- [10] Z. Chi, H. Yan, and T. Pham, "Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition," (Singapore: World Scientific Publishing Co. Pte. Ltd. 1996).
- [11] Frank Hoppner and Frank Klawonn, "A new approach to fuzzy partitioning," IEEE, 2001(C), pp. 1415-1424.
- [12] Jiu-Lun Fan, Wen-Zhi Zhen and Wei-Xin Xie, "Suppressed fuzzy c-means clustering algorithm," Pattern Recognition Letters, Vol. 24, pp. 1607-1612, 2003.
- [13] M. Ameer Ali, L. S. Dooley, and G.C. Karmakar, "Fuzzy image segmentation using location and intensity information," IASTED Intern. Con. on Visualization, Imaging and Image processing, Benalmadena, Spain, 8-10 September, 2003.

Information Theoretic Approach to Blind Deconvolution under Impulsive Noise

C. Wei, L.C. Khor, W.L. Woo and S.S. Dlay

School of Electrical, Electronic and Computer Engineering,

University of Newcastle upon Tyne,

NE1 7RU, Newcastle, United Kingdom

Email: chen.wei@ncl.ac.uk, l.c.khor@ncl.ac.uk, w.l.woo@ncl.ac.uk and s.s.dlay@ncl.ac.uk

Abstract: *In this paper, the problem of blind deconvolution of sub-Gaussian signals under impulsive noise is investigated. Combining Maximum Negentropy optimisation and Expectation Maximisation algorithm with Myriad statistics, the investigated scheme shows high performance over conventional FIR filter under the condition of appropriate selection of linearity parameter k in proportion to the degree of impulsiveness of the noise. Simulations which have been undertaken to study the performance of the investigated approach in mitigating impulsive noise are presented.*

Key Words — *Myriad statistics, Blind deconvolution, adaptive estimation.*

1. Introduction

Blind deconvolution has attracted attention for its diverse applications such as communication channel equalisation, seismic deconvolution and image deblurring. The reason it is called ‘blind’ is that we assume that no prior knowledge about the source and its consequent convolution process are known. Therefore, it is impossible to recover the source signal by, like a traditional method, given a train sequence. However, this weak assumption makes it a powerful tool in signal processing. In this paper, blind deconvolution of sub-Gaussian signals under impulsive noise is the main focus, and this investigation can be utilised in many fields. Previous literature [1] has provided the solutions for blind deconvolution under noiseless or Gaussian distributed noise. However, impulsive noise such as atmospheric noise in radio links, ocean acoustic noise, industrial activity and atmospheric noise [2] could severely corrupt observed signals and may lead to the failure of these techniques. Therefore, new feasible algorithm for blind deconvolution under impulsive noise in particular is becoming essential and paramount in both theoretical and practical levels. A technique is investigated and studied alongside FIR filter, which combines Maximum Negentropy (MN) optimisation for blind deconvolution and Myriad statistics for suppressing the impulsive noise.

Generally, each observation is from the convolution of the source signal and channel with additive noise, and can be modelled as

$$\begin{aligned} d_i &= s_i * h_i + \mathbf{n}_i \\ &= \sum_{l=0}^L h_l s_{i-l} + \mathbf{n}_i \end{aligned} \quad (1)$$

where d_i , s_i , h_i and v_i represent the observation sequence, source signal, the convolutive channel, and noise at time i respectively, and $*$ represent the convolution function.

2. Myriad statistics

It is well known that linear filters perform poorly when outliers exist in the data. Therefore, several techniques are proposed to counter this difficulty. Most of these techniques are based on robust statistics which perform well under or a little away from nominal conditions. Median filters, based on order statistics, are used to counter Laplacian noise model, whose distribution is more heavy-tail than the Gaussian distribution. However, its applications are limited mainly in the image processing due to its selection property. Myriad statistics, a branch of robust statistics to reduce the distortion caused by outliers, is derived from the M-estimation for the a -stable distribution.

Given a set of $L+1$ data samples $\{d_{i-l}\}_{l=0}^L$ at time i , the M-estimation locates the optimal selection y by solving the following cost function

$$\tilde{y}_i = \operatorname{argmin}_{y_i} \sum_{l=0}^L \mathbf{r}(d_{i-l} - y_i), \text{ where } \mathbf{r}(x) = -\log f(x)$$

and $f(x)$ is the distribution of x . here, we use a -stable distribution to model $f(x)$, which vary from highly impulsive distribution to Gaussian distribution. In this distribution range, a -stable distribution has no closed form except $a=2$, corresponding to Gaussian distribution, and $a=1$, corresponding to Cauchy distribution. We consider a variant of Cauchy distributed noise $f(\mathbf{n}) = \frac{k}{p k^2 + \mathbf{n}^2}$ which is

characterised by a long tail distribution. The output of the Myriad statistics is given by

$$\tilde{y}_i = \underset{y_i}{\operatorname{argmin}} \prod_{l=0}^L [k^2 + |\mathbf{w}_l| (\operatorname{sign}(\mathbf{w}_l) d_{i-l} - y_i)^2] \quad (2)$$

where k represents the Myriad linearity parameter and w_l is a real weight. Since $\log(\cdot)$ is a mono-increasing function, finding extreme value of y_i can be replaced by finding extreme value of $\log(y_i)$. Under this substitution, we can define the following:

$$P(y_i) = \log(Q(y_i)) = \sum_{l=0}^L \log(k^2 + |\mathbf{w}_l| (\operatorname{sign}(\mathbf{w}_l) d_{i-l} - y_i)^2) \quad (3)$$

To find the roots of $P(y_i)$, we differentiate $P(y_i)$ with respect to y_i and let $P'(y_i)$ be equal to zero, which is expressed by

$$P'(y_i) = -2 \sum_{l=0}^L \frac{|\mathbf{w}_l| (\operatorname{sign}(\mathbf{w}_l) d_{i-l} - y_i)}{k^2 + |\mathbf{w}_l| (\operatorname{sign}(\mathbf{w}_l) d_{i-l} - y_i)^2} = 0 \quad (4)$$

The weighted myriad is one of the real roots of the function $P'(y_i)$ of (4). An iterative procedure based on the fixed point search algorithm [3] to compute the sample myriad is performed until the estimated output y_i converges, i.e. $\|y_i(n+1) - y_i(n)\|^2 < \epsilon$, where ϵ is a small constant threshold. We introduce new functions:

$$\mathbf{x}_l(y_i(n)) = \frac{|\mathbf{w}_l|}{k^2 + |\mathbf{w}_l| (\operatorname{sign}(\mathbf{w}_l) d_{i-l} - y_i(n))^2} \quad (5)$$

Substituting (5) into (4), we get:

$$P'(y_i) = \sum_{l=0}^L \mathbf{x}_l(y_i(n)) (\operatorname{sign}(\mathbf{w}_l) d_{i-l} - y_i(n+1)) = 0 \quad (6)$$

Rewrite equation (6) and combining with (5), we derive an iteration algorithm by:

$$y_i(n+1) = \frac{\sum_{l=0}^L \mathbf{x}_l(y_i(n)) \operatorname{sign}(\mathbf{w}_l) d_{i-l}}{\sum_{l=0}^L \mathbf{x}_l(y_i(n))} \quad (7)$$

In particular, the Myriad linearity parameter k plays a major role in controlling the robustness of the output against outliers. Specifically, the smaller the k , the stronger the effect on impulsiveness. When $k \gg 8$, myriad filter performs like a linear filter; when $k \rightarrow 0$, myriad filter output mode-myriad [4].

3. Blind Deconvolution by Maximum Negentropy Optimisation

Gaussian mixture model (GMM) is widely used for modelling real phenomena since they are composed by many independent distributed effects. It can be used to model speech signals whose power distribution with respect to frequency is time variant. One typical model is the Bernoulli-Gaussian model. In this paper, the probability density function (pdf) of the source wavelet is modelled by GMM:

$$p(s_i) = \sum_{j=1}^M \frac{\mathbf{a}_j}{\sqrt{2\pi\mathbf{s}_j}} e^{-\frac{s_i^2}{2\mathbf{s}_j^2}} \quad (8)$$

where a_j and s_j are the percentage of contribution and variance of the j^{th} component in the mixture respectively. It has been shown that any continuous smooth enough function could be approximated with arbitrary accuracy by Gaussian Mixture Model with a sufficient number of Gaussian basis functions [5].

For noiseless blind deconvolution, a commonly used method is based on the maximisation of a global likelihood function [6]. This algorithm, however, suffers from high computational complexity. Based on the central limit theorem, we investigate to maximise the marginal Negentropy which is formulated as follows:

$$J = E \left[\log \frac{p(d_i)}{f(d_i)} \right] \quad (9)$$

$$= E[\log p(d_i)] - E[\log f(d_i)]$$

$$= E \left[\log \sum_j p(g_j) p(d_i | g_j) \right] - E[\log f(d_i)]$$

$$p(d_i | g_j) = \frac{1}{\sqrt{2\pi\mathbf{s}_j}} e^{-\frac{d_i^2}{2\mathbf{s}_j^2}} \quad \text{and} \quad p(g_j) = \mathbf{a}_j \quad (10)$$

where $p(d_i | g_j)$ is the pdf of d_i given the j^{th} Gaussian component and $f(\cdot)$ represents the corresponding Gaussian distribution with the same variance as s_i , i.e. $\mathbf{s}^2 = \sum_j \mathbf{a}_j \mathbf{s}_j^2$. The maximisation of the cost

function J can be adaptively implemented by using the gradient ascent algorithm shown as

$$\frac{\partial J}{\partial \mathbf{w}_l} = \frac{\partial J}{\partial y_i} \frac{\partial y_i}{\partial \mathbf{w}_l} = \left(\frac{y_i}{\mathbf{s}^2 \sum_j \frac{p(g_j | y_i)}{\mathbf{s}_j^2}} - y_i \right) \frac{\partial y_i}{\partial \mathbf{w}_l} \quad (11)$$

The derivative of the output with respect to the Myriad filter coefficients [7] is

$$\frac{\partial y_i}{\partial \mathbf{w}_i} = \frac{-\left[\frac{k^2 \text{sign}(\mathbf{w}_i)(y_i - \text{sign}(\mathbf{w}_i)d_{i-1})}{(k^2 + |\mathbf{w}_i|(y_i - \text{sign}(\mathbf{w}_i)d_{i-1})^2)^2} \right]}{\sum_{j=0}^L |\mathbf{w}_j| \frac{k^2 - |\mathbf{w}_j|(y_i - \text{sign}(\mathbf{w}_j)d_{i-j})^2}{(k^2 + |\mathbf{w}_j|(y_i - \text{sign}(\mathbf{w}_j)d_{i-j})^2)^2}} \quad (12)$$

Inserting (12) into (11), we can update the coefficients of Myriad filter at time i as follows:

$$\mathbf{w}_i(i+1) = \mathbf{w}_i(i) + \mathbf{m} \frac{\partial J}{\partial \mathbf{w}_i} \quad (13)$$

For a linear FIR filter, (12) is replaced with $\frac{dy_i}{d\mathbf{w}_i} = d_{i-1}$.

To complete the optimisation procedure, the Expectation-Maximisation (EM) algorithm is applied to estimate the parameters in the GMM as follows:

$$\mathbf{s}_j^2(i+1) = \mathbf{g}_j^2(i) + (1 - \mathbf{g})y_i \quad (14)$$

$$\mathbf{a}_j(i+1) = \mathbf{g}\mathbf{a}_j(i) + (1 - \mathbf{g})p(\mathbf{g}_j|y_i) \quad (15)$$

where \mathbf{g} is a constant within the range of $0.8 < \mathbf{g} < 0.95$ and

$$p(\mathbf{g}_j|y_i) = \frac{p(y_i|\mathbf{g}_j)p(\mathbf{g}_j)}{\sum_{j=1}^M p(y_i|\mathbf{g}_j)p(\mathbf{g}_j)} \quad (16)$$

4. Results

Simulation has been carried out with 1000 samples sub-Gaussian sequence to investigate the performance of the Myriad filter and linear FIR filter. Different types and levels of noise (modelled by α -stable distribution) were added to the convolved signals. In order to test the performance of these two filters, they are applied to a chirp-type binary signal. This test source signals are given in Fig. 1(a)

Firstly, we compare the filters performance under different Generalised Signal-to-Noise Ratio (GSNR). The source signal is convolved through a channel whose response is given by $H(z) = 1 + 0.8z^{-1}$ and contaminated by moderate impulsive signal, where $\alpha = 1.75$, GSNR=8dB. It is obvious that the outputs of different filters are quite different from the desire signals. The outputs of Myriad filter using algorithm based on equations (11)-(15) are shown in Fig 1(b), which are visually similar to the source signal. On the other hand, the outputs of linear FIR filter as shown in Fig 1(c) are greatly affected by the impulsive noise. When the magnitude of GSNR is decreased to 20dB, it is seen that the performance of the FIR filter increases

(Fig 2(b)), but Weighted Myriad filter results in higher improvement (Fig 2(c)) than the FIR filter

Secondly, the performance under different degree of impulsive noise is compared. The impulsive parameter is changed to $\alpha = 1.3$ with GSNR=8dB. Due to highly impulsive noise, the performances of both filters degrade substantially (Fig 3(b), (c)) and the linear FIR filter fails to extract the source signal (Fig 3(c)).

The performance of both conventional FIR filter and weighted Myriad filter under GSNR from 0dB to 30dB and a value from 1.3 to 2 has also been evaluated using Mean Absolute Error (MAE) defined as

$$E_{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - s_i|$$

which is plotted in Fig. 4. As

can be seen from plot, the performance shows that the weighted Myriad filter outperforms the conventional FIR filter under any degree of impulsive noise when GSNR is below 12dB. On the other hand, both performances become equal under Gaussian noise as the weighted Myriad filter simply reduces to a FIR filter. However, the performance of the linear FIR filter significantly drops as soon as α changes from 1.75 to 1.3.

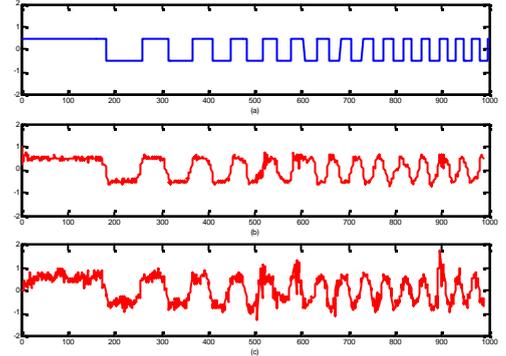


Figure 1: (a) Chirp-type binary signal. (b) Output of Weighted Myriad filter. (c) Output of linear FIR filter

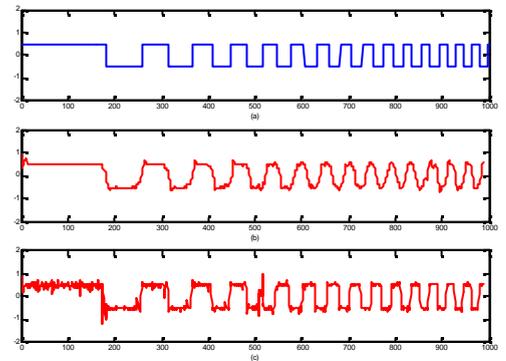


Figure 2: (a) Chirp-type binary signal. (b) Output of Weighted Myriad filter. (c) Output of linear FIR filter

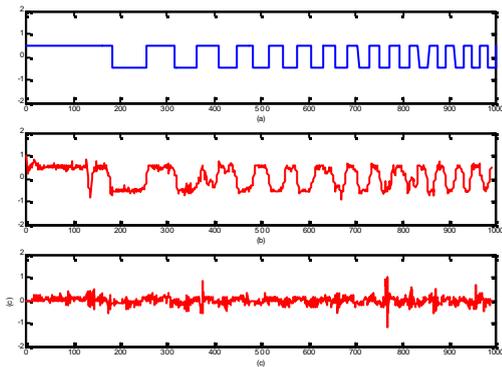


Figure 3: (a) Chirp-type binary signal s_i . (c) Output of Weighted Myriad filter. (d) Output of linear FIR filter.

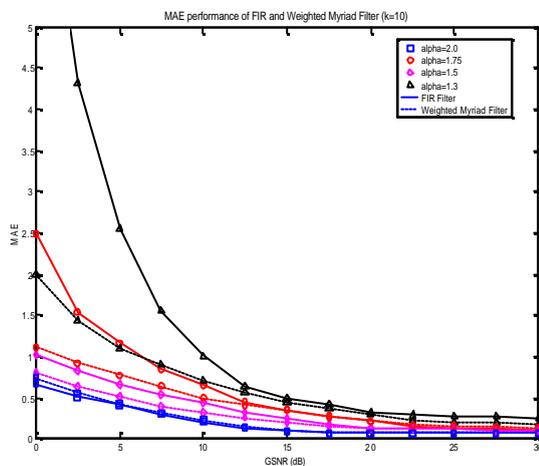


Figure 4: MAE performance of FIR and Weighted Myriad Filter

5. Conclusion

This paper presents an investigation of a possible solution to combat the adverse effect caused by impulsive noise under convolutive environment in which no prior information about the source or/and the channel is given. The investigated scheme is derived from the Maximum Negentropy framework and the use of Myriad statistics shifts the scope from a Gaussian noise model to an impulsive noise model which is more realistic to data analysis. The scheme shows promising applicability under moderate degree of impulsiveness of the noise. Further research is proposed on the adaptive estimation of the parameter k to improve the performance.

6. References

- [1] Luo, H. and Y. Li (Oct. 1998), "The application of blind channel identification techniques to prestack seismic deconvolution," *Proceedings of the IEEE* 86: 2082 – 2089.
- [2] M.P. Shinde and S.N. Gupta (1974), "Signal detection in the presence of atmospheric noise in tropics," *IEEE Transactions communications* 22: 1055-1063
- [3] Kalluri, S. and G. R. Arec (2000), "Fast Algorithms for Weighted Myriad Computation by Fixed Point Search." *IEEE Transactions on Signal Processing* 48: 159-171.
- [4] Kalluri, S. and G. R. Arec (1998), "Adaptive Weighted Myriad Filter Algorithms for Robust Signal Processing in a-Stable Noise Environments." *IEEE Transactions on Signal Processing* 46: 322-334.
- [5] Haykin, (1999). *Neural Networks: a comprehensive foundation*. New Jersey, Prentice-Hall, Inc.
- [6] Santamaria, I., C. J. Pantaleon, et al. (1999), "Deconvolution of Seismic Data Using Adaptive Gaussian Mixtures." *IEEE Transactions on Geoscience and Remote Sensing* 37: 855-860.
- [7] Kalluri, S. and G. R. Arec (Nov. 2001), "Robust Frequency-selective Filtering using Weighted Myriad Filter Admitting Real Valued Weights." *IEEE Transactions on Signal Processing* 49: 2721-2733.

INTERACTIVE SPEECH RECOGNITION SYSTEMS USING COMMITTEE MACHINES

M.SRINIVASAN, S.S.DLAY, W. L.WOO

School of Electrical, Electronic and Computer Engineering, University of
Newcastle upon Tyne.

srinivasan.meenakshi-sundaram@ncl.ac.uk

Abstract: Research to develop the speech recognition based interactive systems is an active area applicable to various fields. This paper presents a unique architecture namely committee machine (CM) to construct such a system. The proposed novelty contributes to research by analysing the existing techniques in developing speech recognition based systems before presenting the theory of committee machine architecture to build an interactive system. A performance of 74.86% is obtained for speech samples from TIMIT database substantiating its contribution to systems using the Committee Machine architecture.

Keywords:

Man Machine Systems, Committee Machines, Radial Basis Functions, Hybrid Connectionist, and Lip Synchronisation.

I. INTRODUCTION

The interaction between man and machines is currently undergoing rapid transformation from traditional mode of communication through external devices to natural modes such as speech, gesture etc. The focus towards this justifies the need to develop efficient architectures. This is addressed by introducing a novel architecture for man machine systems involving speech recognition and lip synchronisation. The utilization of speech recognition and lip synchronisation for man machine interaction has been galvanised by the introduction of statistical HMM. Most the methods successfully employed HMM for statistical modelling. [1,9] Alternatively the speech recognition solutions were built with the Artificial Neural Networks (ANN) with limited success due to complex training schemes. Interestingly some of researchers looked at adding visual features along with acoustic features to enhance recognition.[4] Other dominant methods have included use of non-geometric and geometric facial feature vectors for speech recognition front end . However the above methodologies suffer due to the following: HMMs reliance on the statistical assumption that

emission and transition probabilities depend only on current state pose problems in real time issues where the linear probability trade-off is evident. ANN on the other hand is undermined by its inability to model long term dependencies. These drawbacks led to hybrid paradigm of HMM and ANN utilising their advantages upon statistical, sequential modelling and classification respectively. As a result a number of different hybrid paradigms for ASR were proposed [2,3][5,6]. However the utility of such hybrid systems is limited due to highly complex training schemes and computationally intensive solutions. These drawbacks present a real challenge to introduce a new architecture and this paper addresses the above issues with its novel methodology.

II. COMMITTEES OF RADIAL BASIS FUNCTIONS ARCHITECTURE

The motivation for committee machines is induced by the need for improving the generalization performance under noisy conditions. This is an important issue in applications where the existing methods can work at local region of interests in feature space. However their performance is limited for global feature vector space consideration. Committee machines address this issue by partitioning the set of input data into different sub regions and feeding into to the set of neural networks. Combining these estimators' predictions to a single output would improve the generalization performance to a great extent than those where input data is treated and trained globally resulting in poor generalised solution. [7]

The important aspect with respect to committees is the modularity they can bring while analysing the input feature space. Secondly committees allow the reduction in complexity as training the whole feature space with a single estimator and the type of estimator network chosen is critically important for committee machines performance. Through analysis of performance the

Radial Basis Function (RBF) networks stand apart from the rest of networks for their clustering and classifying properties and ensemble of RBF is an ideal unit that takes care for the variability of data presented before the committee for classification. In the following section the above idea is presented in the form of a method: The input feature space for classification is divided into a number of subspaces and each RBF focuses on a particular region. The modular processes can be treated as two phases namely *independent* and *combination*. During the *independent phase* the parameters used with each RBF's include the central mean, the width associated with the Gaussian functions and the weights. These parameters are adjusted and network is trained estimator for classification then passing through the *combination phase* where the different estimators are combined to produce the response for a specific input space using a gating network mechanism that combines the estimator output.

Training the architecture:

The issues to be addressed when committee machines are trained are associated with respect to individual or global training of networks and in incorporating gating network and finding an optimised algorithm in estimating the target. It is important to have a for a scheme provision for weighing the network members. For a real time system it is desired to be able to train continuously which is not possible with a static system. Thus a gating network is introduced to adjust the weight of networks and it is useful in attaining the target with a feedback between the input and the output from members. The next section gives an analytical approach with respect to training.

Step 1-Training of RBF member networks:

The RBFs are seen as a two layer feed-forward neural network with nonlinear input and output of RBF is linear. Typical RBF network has an input layer consisting of set of data to be trained I , a hidden layer k made of neurons each basis functions and an output layer O which has weighed combination of hidden layer output. Lets us consider such a topology with input represented as x_n and the output of networks is represented as $y_1, y_2, y_3, \dots, y_n$, characterised by the center of Gaussian, the width of the activation surface and the weighting factors w_i . RBF training can be done in two stages.

RBF hidden layer stage: Here the parameters of hidden units' e.g. Gaussian parameters using clustering algorithms are estimated. For training of RBF we have chosen to use the popular k-means algorithm and least means squares for location of center, width and weighting factors calculation respectively. With the identification of the individual centers of the Gaussian functions and their common width using k-means

algorithm the RBF learning continues with calculation of weights in the *Output Layer Stage*.

RBF Output Layer Stage: This stage involves the calculation of the weights and final output calculation depends on weighted sum of inputs using Least Mean Squares (LMS) algorithm. The weights w_i are updated using the *least mean square (LMS)* algorithm during online training. There are alternative methods such as *orthogonal least square (OLS)*. We have chosen LMS to update the weights as OLS lacks a generalization learning algorithm and it uses windowed data and contains information for particular window.

Step 2-Training of Committee Machines:

For illustration of training this architecture let us consider the individual RBFs emit output represented by y_i . Let us consider such a system where the output is represented as Z . Representing the output in terms of individual networks output y_i we can write

$$Z = \sum_{i=1}^N y_i$$

Where N denotes the number of RBF networks used within committee. With the weighed scheme applied the above equation can be written as

$$Z = \sum_{i=1}^N g_i y_i$$

The gating network parameters g_i are chosen with respect to the subjected to training and previous state output. Eg. $0 \leq g_i \leq 1$ Such that $\sum_i g_i = 1$

For classification problems the generalisation performance is critical where the networks output is compared with the desired output classes with prior knowledge. We introduce a self maintaining mechanism involving a gating network that adjusts the weights associated with individual member networks using the feedback from output. This gives better generalisation compared to individual networks. The significance of gating networks lies in its ability to combine the networks output according to the statistical significance and networks contribution analysis that differs in input space. The error function e_r can be defined as the difference between the committee machines output and the desired response during classification problem. If we denote the desired response by d the aim is to reduce the error function which can be written as

$$e_r = d - Z$$

The architecture could be optimal if the error function is kept at minimum and from the committee architecture figure 2 we can write the error function e_r to be

$$e_r = d - \sum_{i=1}^N y_i g_i \quad (1)$$

The squared error function denoted by J can be given as

$$J = E[e_r^2]$$

The weights can be defined by the following equation

$$g_i = \frac{\exp[-\mathbf{m}_i]}{\sum_{i=1}^M \exp[-\mathbf{m}_i]} \quad \text{where } i = 1, 2, 3, \dots, M$$

Where $\mathbf{m} = a_k^T x$ and M denotes no of sub-spaces. Differentiating the error function with respect to the variant a_k

$$\frac{dJ}{da_k} = -2 e_r \left(\sum_{i=1}^M \frac{dg_i}{da_k} y_i \right) \quad (2)$$

Substituting $\mathbf{m} = a_i^T x$ in the weighting factors

$$\text{equation } g_i = \frac{e^{-a_i^T x}}{\sum_{i=1}^M e^{-a_i^T x}} \quad (3)$$

$$\frac{dg_i}{dak} = \frac{d}{dak} \left[\frac{e^{-a_i^T x}}{\sum_{i=1}^M e^{-a_i^T x}} \right]$$

Using the rules for division we get the expansion of (3) as

$$\left[\frac{\sum_{i=1}^M e^{-a_i^T x} \left(-x.e^{-a_i^T x} \right) - \left(e^{-a_i^T x} \right) \left(e^{-a_i^T x} \right) x}{\left(\sum_{i=1}^M e^{-a_i^T x} \right)^2} \right] \quad (4)$$

$$\frac{dg_i}{dak} = \left(-x.e^{-a_i^T x} \right) \left[\frac{\sum_{i=1}^M e^{-a_i^T x} + e^{-a_i^T x}}{\left(\sum_{i=1}^M e^{-a_i^T x} \right)^2} \right] \quad (5)$$

Substituting the equation (5) in (2) the state equation can be written as

$$a_k^{(t+1)} = a_k^{(t)} + 2.m.e_r \left[\frac{\sum_{i=1}^M y_i . x . e^{-a_i^T x} \cdot \left[\frac{\sum_{i=1}^M e^{-a_i^T x} + e^{-a_i^T x}}{\left(\sum_{i=1}^M e^{-a_i^T x} \right)^2} \right]}{\sum_{i=1}^M y_i . x . e^{-a_i^T x}} \right] \quad (6)$$

With the optimal solution arriving by the proper choice of the parameter \mathbf{m} we can train it sequentially for the set of training data. During training the error function is minimised using the above state equation and every successive state is corrected using the weighting factors of gating networks. With the above divide and rule using the RBFs we can find an optimal solution for any input space containing set of data that are trained using RBFs and committee machines. The gating parameters g_k are trained for the minimisation of the error cost function. $E[e_r^2] \rightarrow 0$. The parameters are adjusted till the steady state conditions are achieved and the output class of committee machine matches the desired class.

III. EXPERIMENTS AND RESULTS

The novel architecture proposed is analysed using TIMIT speech database consisting of 6300 utterances, 10 sentences spoken by each of 630 speakers from 8 dialect regions in United States. The input speech data sampled at 16 KHz divided into 20ms frames with 10ms overlap and applied Hamming windowing process. The frames of speech data subjected to MFCC analysis resulting in a 39 dimension feature vector. These features are then fed into the recognition engine. HMM model is a three state 39 classes model and resulting in 117 states involved in recognition engine.

The vector set emitted by HMM are placed as input nodes for RBF with ONE hidden layer for each RBF. The input nodes are grouped according to the data orientation in the space and for RBF parameters are initialised and trained. These results are weighted and compared with the desired class.

First the recognition performances of the existing systems are compared indicating the scope for improvement (Table 1). The analysis showed that the existing methods can yield an accuracy of app 56 %. Thus a suitable network is selected with analysis on different neural networks (Table 2). Here the Multi Layered Perceptron and radial basis functions are compared showing the simplicity of RBF motivating their utility for the committee machine architecture. RBF with lesser training time and computations could achieve the required performance. Then the RBFs are combined with their output nodes weighted, weights chosen at different ranges and performances are analysed. Table 3 and 4 indicate the performance results for committee machine and its significance when

compared to using a single network for recognition. The results clearly confirm the superior performance of the committee architecture using RBFs.

IV .CONCLUSION

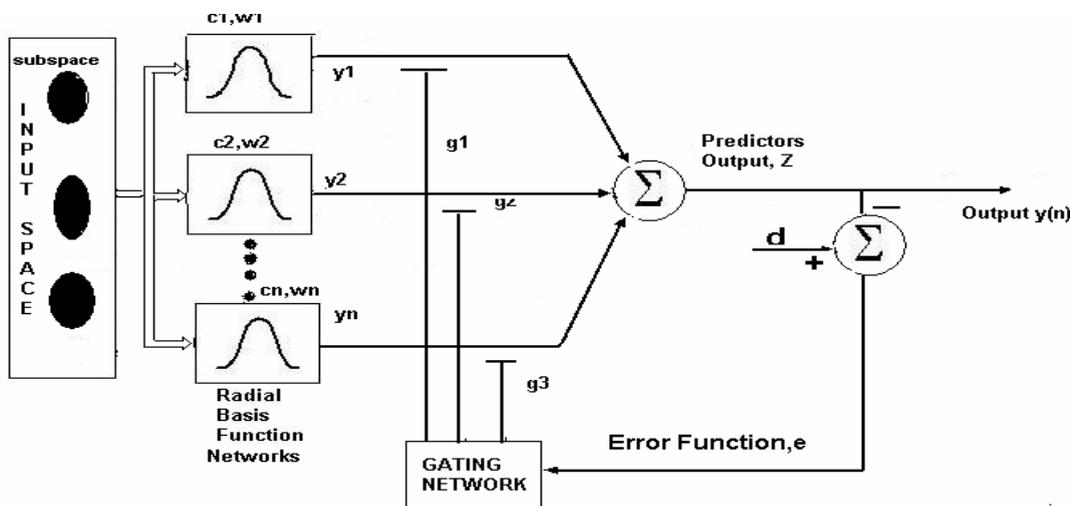
This paper has presented a novel architecture for the man machine systems. The performance analysis through tests and the subsequent results for a speech database have confirmed the theory that this novel architecture can perform efficiently for generalization with human modes and can be trained easily compared to existing systems. The architecture could act as building block in many real-world applications including natural man-machine interfacing; digital intelligent agent for hearing impaired community. With the research community's vision firmly focussed on the man machine interaction this method could be a catalyst for further research in the area of speech recognition architectures and ignite the path towards bridging the man machine interaction to an impressive level.

REFERENCES

1. Ajit V.Rao and Kenneth Rose, "Deterministic Annealed design of Hidden Markov Speech Recognisers" , IEEE Trans. on Speech an

- Audio Processing, Vol. 9,No.2 pp. 111-125, Feb 2001.
2. Bourlard.H and N. Morgan, Connectionist Speech Recognition, Kluwer Academic Publ., Massachusetts, 1994.
3. Dupont.S, "Missing Data Reconstruction for Robust Automatic Speech Recognition in the Framework of Hybrid HMM/ANN Systems", Proc. ICSLP'98, pp 1439-1442.Sydney, Australia, 1998.
4. Eric Vatikiotis-Bateson and Hani.C.Yehia, "Speaking Mode variability in Multimodal Speech Production", IEEE Transactions on Neural Networks, Vol. 13 No.4 pp.895-899, 2002.
5. Gong.Y, "Speech recognition in noisy Environments: A survey", Speech communication Vol. 12, No. 3, pp. 231--239, June, 1995.
6. Nelson Morgan and Hervé Bourlard, "An Introduction to Hybrid HMM/Connectionist Continuous Speech Recognition", IEEE Signal Processing Magazine, pp. 25-42, May 1995
7. S.Haykin, Neural Networks, Prentice Hall, 1999.

PERFORMANCE ANALYSIS TESTS RESULTS:



COMMITTEE MACHINE - DIAGRAM

TABLE 1: Recognition results of existing recognition systems and the Committee machine

Recognition Configuration	Recognition rate (%)	Subs. (%)	Del. (%)	Ins. (%)	Total Error rate(%)
Baseline HMM	56.3	27.1	10.1	6.5	43.7
HMM + MLP	59.09	26.24	8.91	5.76	40.91

TABLE 2: Neural Networks analysis: MLP vs. RBF performances for the committee machine architecture.

S.No.	Feature	MLP (%)	RBF (%)
1	Recognition Accuracy	59.09	58.45
2	Error rate	40.91	41.55

TABLE 3: Speaker-independent, continuous speech – Recognition results using committee machine architecture and comparison with single network.

S.No.	Feature	Single RBF Network (%)	Committee machine (%)
1	Recognition Accuracy	58.45	65.87
2	Error Performance	41.55	34.13

TABLE 4: Speaker-independent, continuous speech -Human lip synchronization performance of system

Configuration	Rec. Accy. %	Sub. Err%	Del. Err %	Ins. Err %	Error Rate %
CM performance	65.87	23.59	6.15	4.39	34.13

MORPHOLOGICAL ANALYSES OF HISTOLOGICAL IMAGES USING FRACTAL

¹Khaled Marghani, ¹Satnam Dlay, ¹Bayan Sharif, ²Andrew Sims

¹k.a.s.marghani@ncl.ac.uk, ¹s.s.dlay@ncl.ac.uk, ¹byan.sharif@ncl.ac.uk, and ²a.j.sims@newcastle.ac.uk

¹School of Electrical, Electronic and Computer Engineering, Merz Court, University of Newcastle upon Tyne, NE1 7RU, United Kingdom

²Medical Physics Department, Freeman Hospital, NE7 7DN, UK

ABSTRACT

Objectives: In order to assist the decision of the pathologist in cancer diagnosis, a new algorithm using morphological analyses based on fractal geometry is investigated. Samples from two different colorectal tissue types were used, consisting of 24 microscopic images represent moderately differentiated adenocarcinoma, and 22 images of normal tissues. **Methods:** morphological examinations based on box-counting method were applied and fractal dimension (FD) were estimated. For validation, statistical analyses for assessing the identification of abnormalities (adenocarcinoma vs. normal mucosa) were performed. **Results:** Obtained results show a very strong significance of 5.57013×10^{-10} using analysis of variance for comparing the means of two different populations of the independent FD observations. **Conclusion:** The potential for applying morphological analyses of histological microscopic images based on fractal features is established. In brief, quantitative measurements, based morphological features, provide useful information can help decision-making. Further research of combining advanced feature based morphology for pathological diagnosis is required.

Key words: Medical Diagnosis, Histological Images, Segmentation, Morphology, and Fractal.

1. INTRODUCTION

1.1 Benefits of quantitative analyses.

Cancer is a leading cause of fatality amongst all diseases for human beings in the world. However, there is still no known way of preventing cancer, but early detection allows treatment before it spreads to other parts of the body and is the key to its cure. Accurate and reliable decision making in cancer prognosis, made by pathologists in histological laboratories, can help in the planning of appropriate surgery and therapy and, in general, optimize patient management through the different stages of the disease. This may be a demanding target as analysis of microscopic images of stained tissues is often adversely affected by poor acquisition and tissue preparation. This is because of the effect of manual functions required to prepare tissue and acquire images. In addition, the volume and complexity of these data have been increased [1,2]. Analysis of microscopic images of cells and tissues of such cancers is still assessed manually, which is arduous, time consuming, and consequently may effect the histologists' decision. On the other hand, during the last few decades, there has been a rapid development of image processing and pattern recognition techniques, in addition to computerization of standard statistical analysis. Therefore, the use of modern technology to support data analysis based on quantitative

measurement instead of manual methods has become imperative.

1.2 The importance of morphological analysis.

Although considerable research, in the last two decades, has looked into different methods based on morphometric, densitometry, statistical and fractal techniques, to derive treatment or therapy for different types of cancer, rather less attention has been paid to the analysis of colon cancer. Different approaches for the purpose of the automation of computer based methods of cell and tissue analysis have become available; for instance flow cytometry [3], or image analysis including morphometry [4] and texture analysis. More interest was considered using the last approach as many studies have been performed to identify and analyze isolated cells or gland tissue according to parameters measured by image analysis systems such as cell size, shape, density, color hue, etc. [5, 6]. However, in the case of comprehensive cells, problems remained for image segmentation and shape description and analysis. The estimated parameters such as size and density of the cells occupied surface usually do not afford exact morphological information. Therefore, using mathematical morphology in the area of image analysis has been of great interest to researchers.

1.3 Automatic image segmentation and shape description.

Aziz, [7] pointed out that the histopathological features such as nuclear size, shape, and polymorphism must be converted to image features such as area, shape factor, and variance of the area; this feature vector must be correlated with the Pathologist's expert opinion or diagnosis. In 1996, a new method by Thiran [8] was proposed for automatic recognition of microscopic cancer images. He pointed out that mathematical morphology provides good efficiency for classification of digital images of cancer. Hamilton, et al [9] in 1997 used texture features, extracted using features of the co-occurrence matrix and the number of low optical density pixels in the image, to classify colon cancer. Similar results using texture features have been obtained by Esgiar et al[10], when they investigated texture and geometric features, and their initial result in the geometric approach showed promise as an indicator for classification.

1.4 The aim of this study

As evidence mentioned above shows the strength of morphological shape analysis, in this paper we present a new automated algorithm for the identification between cancer and normal colon tissue glands, based on morphological segmentation and shape analysis using fractals morphology.

2. METHODS

1.2 Image acquisition

Two main parts are considering image acquisition related to medical preparation and image capturing. Medical preparation was started as samples of colon tissue, obtained from sequential resections and diverticulosis, which received within 30 minutes after resection. They were fixed in formalin or alcohol. After a period of 24hours, the samples were placed into fixative for a further 24-hour. Next, all samples were sliced thinly (3~4 mm) and implanted in paraffin (tissue processing), and sectioned to between 4 to 6 μm . In worm temperature glass microscopic slides were stickled to sectioned tissues. Finally, they were immunohistochemically stained for cytokeratins. This technique enables tissue configuration and regions to be identified.

As about image capturing the histological samples, several steps were required to allow further examination and analysis. Firstly, using a light microscope and JVC camera attached to a Leica Q500 image analysis system, stained samples were captured. Only exact white image level can be regulated by the arranged system. Afterwards, under low magnification of x40, images were sampled and digitised in 8bit format with grey level range 0-255, and stored on a

personal computer in 512x512 size format. A number of samples from two colon tissue lesions types were digitally acquired consisting of 24 microscopic images represent moderately differentiated adenocarcinoma samples, and 22 images of normal colon tissues.

As a result of some limitations relating to tissue preparation, another process known as *Image sampling* become necessary. This issue concerns the tissue distribution on images since some images did not contain tissue, and other image samples did not have enough tissue compared with the whole image size. Therefore, all images have been divided into four small images of 256x256 formats. 102 samples from 184 original images, representing 44 cases of normal tissues and 58 of malignant image samples, were selected and used for our examinations. Only samples that more than 10% of its image size occupied by object tissue were selected. An example from the final outcome for both cases can be illustrated in fig (2).

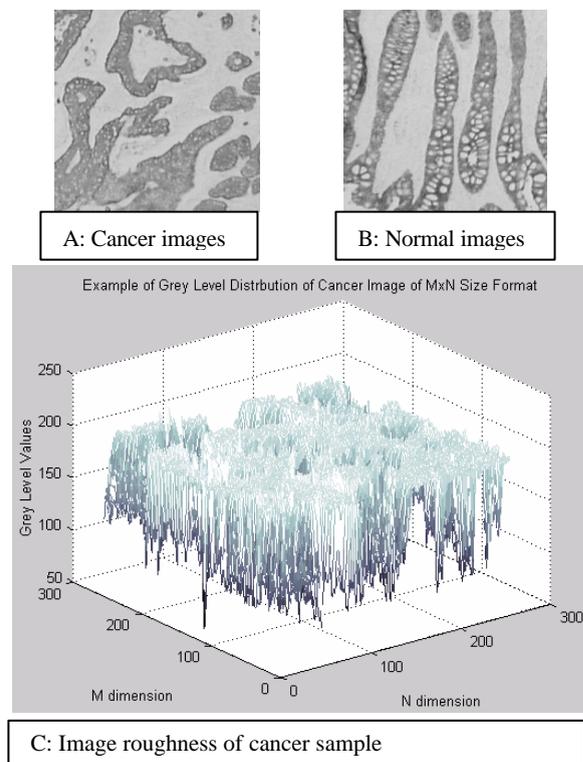


Figure (2): Samples set of different type of colon tissue in 2D gray-level digital image format: (A) - Cancer tissue images, (B) - Normal sample images, and (C): Example of intensity distribution of the cancer sample.

2.4 Image segmentation

Our method of geometric analysis requires isolating the nuclear shape of normal and malignant colon tissues.

To achieve this aim, images were segmented into two dissimilar regions, background and foreground, and

stored in a binary image format. Segmentation was performed with a threshold method and the best thresholding level was selected using Otsu's method [12] that chooses the threshold to minimize the interclass variance of the black and white pixels.

After the segmentation process, gland objects of histological tissues were affected in their configuration, and thus the morphological image processing operations such as filling, dilation and erosion are required [13, 14]. Firstly, in order to avoid object disorientation, the hole-filling process was applied to smooth the shape. Also, filtering the image using the opening operation with a disk structure element removed superfluous overlapped objects. This selection made the shape object smoother at the same time filtering redundant objects. Dilation was directly applied in small squarely structured element to recover any loss of the original shape. Finally, the opening operation was applied again to ensure that object tissues are filtered promptly. Data sets are now in binary image format and ready for further analysis to be made under geometric approach.

2.5 Fractal measures for estimating a set's dimension

A brief definition of well-known fractal measures, using the box-counting method, of fractal sets is required. The definition of the box-counting method is a constrained model of the Hausdorff measurement of fractal dimension. In contrast, under approaches such as the covering blanket method, features can be classified as a simple example of Hausdorff measures. The application of the fractal geometry in pathology has been extensively reviewed [17]. Kerenji et al [18] carried out research of using FD to distinguish between normal liver and hepatocellular carcinoma. Their proposed method was based on identifying quantitative measurements of nuclear shape of cell objects. The reason for keeping special interest on fractal theory to identify abnormalities between different texture objects is the high correlation between the human perception of surface roughness fig (2C) and its geometric dimension estimated by fractal models [20]. Our experience [11] of investigating the potential of morphological analysis based on fractal geometry to examine colon dataset under different resolution levels shows evidence of identifying the difference between histological microscopic images.

2.6 Box-counting method

There are several well-known methods in which the box counting has been implemented to estimate the fractal dimension and the mass local distribution [15,16,19, and 21]. The self-similarity property of the fractal set with dimension D implies that it should be

the same over a range N of non-overlapping copies of the origin, each of which has been scaled by a ratio r in both axis. Under these circumstances the relationship $I = N r^D$ has to be satisfied. Converting the above notion for our algorithm is explained as; traditional box counting method approach analyzes gland objects by counting the number of grid squares $N(r)$ that cover the whole object as a function of the grid size $(1/r)$ without considering how many pixels in the grid size. The formula, which shows how box counting methods run and the relationship between both variables can be defined as:

$$D_f = \lim_{r \rightarrow 0} \frac{\log N(r)}{\log(1/r)}$$

We estimated the fractal dimension for every image by plotting the relationships between box size, and the number of boxes covering the whole image, when Log-log plot was used as $\log(1/r)$ in the x-axis, against $\log N(r)$ in the y-axis.

3. RESULTS AND ANALYSIS

3.1 Fractal feature

Very strong significant of 5.57013×10^{-10} using traditional box counting method is achieved as it shows a distinct separation between normal and malignant colon tissue images using geometric approach. As can be seen from fig (2) regularity was lost in cancer tissues, and therefore, its estimated fractal dimension is high. The mean value and standard deviation for the cancer images was 1.418 ± 0.025 , while its 1.389 ± 0.059 for the normal glands. For all data sets, the sensitivity and specificity of 96.6% and 56.8%, respectively, with an overall accuracy of 79.4% is achieved.

3.2 Discussion

Box-counting method can be used in such different approaches, therefore, its mean values can be different. For example, multiresolution box counting method [11] is depend on changing the image resolution by reconstructing the image in different image size and counting the number of pixels representing glands objects. In this paper, only traditional concept of counting the number of boxes required to cover such objects at different box size were applied. In both cases, estimated fractal dimension (FD) are exceeded the topological values, and the mean FD for cancer samples are higher than for normal samples. Very strong significant of 5.57013×10^{-10} using traditional box counting method is achieved compared with only 2.2343×10^{-6} with reconstructed images algorithm. Improvement of these results compared with previous work [11, 19] can primarily be explained by the use of the morphological approach. Our approach proposes a

hole-filing operation, in order to identify border objects. This process enables the object tissue to be isolated from the background, using threshold operation based on segmenting the image into two main classes. Furthermore, two important operations, known as erosion and dilation, are also used in a unique way so that the glands become clear and ready to be examined. The importance of these operations was noted in [11, 13, and 14]. However, promising results for identifying abnormalities of colon cancer using the fractal approach alone is still being investigated.

4. CONCLUSIONS

Achieved results show that fractal geometric approach using traditional box-counting method can be a good descriptor feature to distinguish between different tissue categories but it's not sufficient to explore in clinical trial yet. Further investigation in this task on combining other morphological features may reach our desirable target.

5. REFERENCES

1. Expert Advisory Group on Cancer: A policy Framework for Commissioning Cancer Services. London, Department of Health, 1995.
2. Cancer Research UK, <http://www.cancerresearchuk.org/>
3. L.G.Koss, B. Czerniak, F. Herz, and R.P. Wersto, "Flow cytometric measurement of DNA and other cell components in human tumors. A critical appraisal, " *Human Pathol.*, vol. 20, pp. 258-548, 1989.
4. J. P. A. Baak and J. Oort, *Morphometry in Diagnostic Pathology*. Berlin: Springer Verlag, 1983.
5. C. Garbay, J. M. Chassery, and G. Brugal, "An interactive region-growing process for cell image segmentation based on local color similarity and global shape criteria," *Anal. Quantum Cytol. Histol.*, vol. 8, pp25-34, 1986.
6. C. Gauvain, D. Seigneurin, and G. Brugal, " A quantitative analysis of the human bone marrow erythroblastic cell lineage using the SAMBA 200 cell image processor," *Anal. Quantume Cytol. Histol.*, vol. 9, pp. 253-262, 1987.
7. Aziz DC, Barathur RB, 1994, 'Quantitation and morphometric analysis of tumours by image analysis. *J Anal Biochem*; 19: 120-125.
8. Thiran J P, 1996,"Morphological feature extraction for the classification of digital images of cancerous tissues", *IEEE Tran. On Biomedical_Eng.*, VOL. 43, NO. 10,1011-1019.
9. P. W. Hamilton, P. H Bartels, D. Thompson, N. H. Anderson, and R. Montironi, 1997,

"Automated location of dysplastic fields in colorectal histology using image texture analysis", *J. Pathol.*, VOL. 182, pp. 68-75.

10. Esgiar A N, Naguib R N G, Sharif B S, Bennett M K, and Murray A, 1998, "Automated feature extraction and identification of colon carinoma", *Anal. and Quant. Cyto. and hist.*, VOL. 20, No. 4, 279-301.
11. K. Marghani, S. Dlay, B. Sharif, and A. Sims, Morphological and texture features for cancers tissues microscopic images, *Medical Imaging 2003, SPIE* Vol. 5032, Part Three, pp 1557-1564.
12. N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, VOL. 9, No. 1, pp. 62-66, 1979.
13. Robert M Haralick, Stanley R Sternberg, and Xinhua Zhuang, 1987, "Image Analysis Using Mathematical Morphology", *IEEE Transactions on pattern analysis and machine intelligence*, VOL. PAMI-9, No. 4, July, p 532-550.
14. Serra J. *Image analysis and mathematical morphology*, volume one. Academic Press Inc., London u.a, 1982.
15. Cross S S, and Cotton W K, 1992, "The Fractal dimension may be a useful morphometric discriminate in histopathology", *J of Pathology*, VOL. 166, 409-411.
16. Cross S S, Bury J P, Silcocks P B, Stephenson T J, and Cotton D W K, 1994, "fractal geometric analysis of colorectal polyps" , *J. of Pathology*, VOL.172, 317-323.
17. Cross S. S., 1997, "FRACTALS IN PATHOLOGY" *JOURNAL OF PATHOLOGY*, VOL. 182: 1-8.
18. Kerenji AS, Bozovic ZL, Tasic MM, Budimlija ZM, and Polzovic AF, 2000, 'Fractal dimension of hepatocytes' nuclei in normal liver vs hepatocellular carcinoma (HCC) in human subjects -preliminary results', *Archive of Oncology*: 8(2): 47-50.
19. Esgiar, Abdelrahim Nasser; Naguib, Raouf N. G.; Sharif, Bayan S.; Bennett, Mark K.; Murray, Alan, 2002, "Fractal analysis in the detection of colonic cancer images", *IEEE Transactions on Information Technology in Biomedicine*, v 6, n 1, March , p 54-58.
20. D. C. Knill, D. Field, D. Kersten, 1990, "Human discrimination of fractal images", *J. Opt. Soc. Amer.*, .7, No.6, pp. 1113.
21. R. Voss, *Random, Scaling Phenomena in Disordered Systems*, R. Pynn and Skjeltorop., Eds., Plenum, N. Y., 1986.

NESTED WIGNER DISTRIBUTIONS AND THEIR APPLICATION TO SIGNAL SEPARATION

Saad A. Qazi and Lampros K. Stergioulas
 Dept. of Information Systems and Computing
 Brunel University, Uxbridge
 Middlesex UB8 3PH

Abstract - The nested Wigner distribution is introduced and its practical discrete version derived. Numerical results for the case of two-tone separation demonstrate the potential of this method for signal detection and separation.

1. Introduction

The Wigner Distribution (WD) has proved to be a powerful tool for Time-Frequency analysis over the last two decades. It is the most popular bilinear time-frequency representation amongst the Cohen's class of representations [1], as it is considered to be superior than other TF distributions, firstly due to its fairly straightforward calculation [1] and secondly because its time and frequency resolutions are the highest achievable by any distribution – these resolutions are not determined by a window, but rather by the intrinsic resolution of the signal itself [2]. The WD however has also its shortcomings, the most important one from a practical point of view being the presence of usually significant cross-terms, which are undesirable and a source of error.

The WD has been mainly used for one-dimensional signals, but is not restricted to this case. There has been some work involving the application of the WD to optics and image processing, in which cases, the signals are two-dimensional and the distribution then is four-dimensional [3,4,5]. In this paper, we propose the use of a new, more complex representation for 1-D signals, namely the Second-Order Nested Wigner distribution (NWD), which is derived as the Wigner distribution of a Wigner distribution, resulting in a quartic 4-D function, in other words we employ the Wigner operator on a 2-D Wigner function. It is obvious that the same process can be invoked for a theoretically infinite number of times, yielding higher-order (multi-dimensional) nested distributions. In the following sections, the mathematical formulation of the NWD is presented, a practically applicable discrete version is derived, and its usefulness is demonstrated by its application to a simple example of tone separation.

2. Definition of the Second-Order Nested Wigner Distribution.

The auto-Wigner distribution [6] of a signal $s(t)$ is defined as

$$W(t, \omega) = \int s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau \quad 1$$

$$W(t, \omega) = \int S\left(\omega - \frac{\theta}{2}\right) S^*\left(\omega + \frac{\theta}{2}\right) e^{-j\theta t} d\theta \quad 2$$

where $S(\omega)$ is the Fourier transform of $s(t)$. By applying this Wigner function on itself, the NWD therefore can be derived as

$$W(t, \omega, \tau, \theta) = \iint W\left(t + \frac{\alpha}{2}, \omega + \frac{\beta}{2}\right) \quad 3$$

$$W^*\left(t - \frac{\alpha}{2}, \omega - \frac{\beta}{2}\right) e^{-j(\alpha\theta + \beta\tau)} d\alpha d\beta$$

$$W(t, \omega, \tau, \theta) = \iint A\left(\tau - \frac{\phi}{2}, \theta - \frac{\lambda}{2}\right) \quad 4$$

$$A^*\left(\tau + \frac{\phi}{2}, \theta + \frac{\lambda}{2}\right) e^{+j(\phi\omega + \lambda t)} d\phi d\lambda$$

The expressions (3) and (4) can be considered to be simple extensions to (1) and (2); however it will be shown that this new, more complex representation offers a greater capability for signal processing applications. Since $s(t)$ and $S(\omega)$ are Fourier Transform pairs, we can conclude that the WD is a joint time-frequency distribution. Building an analogy from this, we also know that $W(t, \omega)$ and the Ambiguity function

$$A(\tau, \theta) = \iint W(t, \omega) e^{j\theta t + j\tau\omega} dt d\omega \quad 5$$

form a 2-D Fourier Transform pair, where 'τ' and 'θ' are time and frequency lags respectively. Thus the new four-dimensional space relates time and frequency with time and frequency lags.

Because of this analogy, the 2-D marginal over the whole time-frequency or time lag-frequency lag planes is due to return $|W(t, \omega)|^2$ and $|A(\tau, \theta)|^2$ respectively, as indeed given in the equations below:

$$|W(t, \omega)|^2 = \iint W(t, \omega, \tau, \theta) d\tau d\theta \quad 6$$

$$|A(\tau, \theta)|^2 = \iint W(t, \omega, \tau, \theta) dt d\omega \quad 7$$

Four other 2-D distributions can also be derived from various combinations of time, frequency and their lags, which can offer different specialised approaches and may be more powerful in certain applications of signal detection and separation.

For practical signal processing, there is a need to redefine (3) and (4) in a discrete-time and discrete frequency context, leading to the formulation of the discrete version of the NWD of a discrete-time signal $s(n)$ of length N :

$$W(n, m, \eta, \mu) = \frac{1}{N^2} \sum_{k_2=-\frac{N}{2}}^{N/2} \sum_{k_1=-\frac{N}{2}}^{N/2} W(n+k_1, m+k_2)$$

$$W^*(n-k_1, m-k_2) e^{-j2\pi(\frac{\eta k_2}{N} + \frac{\mu k_1}{N})}$$

8

where n , m , η and μ stand for the discrete counterparts of time (t), frequency (ω), time lag (τ) and frequency lag (θ), respectively.

A significant problem related to the discrete form of the WD, and consequently also for the NWD, is aliasing. This aliasing problem will naturally get worse, as we move to higher orders/dimensions of nested Wigner functions. From previous work [2,7] on this problem, it is known that $W(n, m)$ is periodic in m with period π . It is easy to show that the NWD $W(n, m, \eta, \mu)$ is similarly periodic in μ with period $\pi/2$ as well as in η with period $N/2$, leading to dilation in the time- and frequency-lags.

3. Application of the Nested Wigner distribution for two-tone separation

In this section, we discuss the case of a signal which is a sum of two tones with frequencies ω_1 and ω_2 .

$$s(t) = e^{j\omega_1 t} + e^{j\omega_2 t} \quad 9$$

The WD of (9) will be

$$W(t, \omega) = \delta(\omega - \omega_1) + \delta(\omega - \omega_2) + 2\delta[\omega - \frac{1}{2}(\omega_1 + \omega_2)] \cos(\omega_1 - \omega_2)t \quad 10$$

The WD and the Ambiguity function (AF) of the superposition of these two tones are shown in figures 1(a,b). Now, using Eq. (3) to calculate the NWD yields the following:

$$\begin{aligned} W(t, \omega, \tau, \theta) &= \delta(\theta)\delta(\omega - \omega_1) + \delta(\theta)\delta(\omega - \omega_2) + \\ &\delta(\omega - \frac{Y}{2})[\delta(\theta - \lambda) + \delta(\theta + \lambda) + 2\delta(\theta)\cos 2\lambda t] + \\ &2\delta(\omega - \frac{A}{4})[\delta(\theta + \frac{\lambda}{2})\cos \lambda(t - \frac{\tau}{2}) + \delta(\theta - \frac{\lambda}{2})\cos \lambda(t + \frac{\tau}{2})] + \\ &2\delta(\omega - \frac{B}{4})[\delta(\theta + \frac{\lambda}{2})\cos \lambda(t + \frac{\tau}{2}) + \delta(\theta - \frac{\lambda}{2})\cos \lambda(t - \frac{\tau}{2})] + \\ &2\delta(\theta)\delta(\omega - \frac{Y}{2})\cos \lambda \tau \end{aligned} \quad 11$$

where, $\lambda = \omega_1 - \omega_2$, $Y = \omega_1 + \omega_2$, $A = 3\omega_1 + \omega_2$ and $B = \omega_1 + 3\omega_2$. In case where the corresponding frequency of each tone is unknown, one can always derive this information from the marginal $|G(\omega, \theta)|^2$. It should be noted that there are more than one ways to do this, but here our preference is to illustrate how this can be done in the new NWD approach. For a noise-free case, $|G(\omega, \theta)|^2$ is given by

$$\begin{aligned} |G(\omega, \theta)|^2 &= \delta(\theta)[\delta(\omega - \omega_1) + \delta(\omega - \omega_2)] + \\ &\delta\left(\omega - \frac{1}{2}(\omega_1 + \omega_2)\right)[\delta(\theta - (\omega_1 - \omega_2)) + \delta(\theta + (\omega_1 - \omega_2))] \end{aligned} \quad 12$$

In the frequency and frequency-lag distribution plane, the NWD exhibits four spikes (maxima) at locations that can be easily guessed from (12). These locations are shown in figure 2. Furthermore, if both ω_1 , ω_2 are now known, one can extract the square-magnitude ambiguity function for a single tone (ω_1 or ω_2) (Figure 3a) using one of the following substitutions:

$$|A(\tau, \theta)|_{\sigma_1}^2 = \int W(t, \omega_1, \tau, \theta) dt \quad 13$$

$$|A(\tau, \theta)|_{\sigma_2}^2 = \int W(t, \omega_2, \tau, \theta) dt \quad 14$$

The two tones can then be separately recovered from eqs. 13 and 14 with straightforward inverse transforms.

The square-magnitude WD $|W(t, \omega)|^2$ for a tone of frequency $\omega_1 - \omega_2$ (Figure 3b) can be calculated in a similar way. From (12), one can determine the value $\omega_1 - \omega_2$ from the distance of the corresponding spikes along the time-lag axis, so the substitution $\tau = \omega_1 - \omega_2$ and integration along t yields

$$|W(t, \omega)|_{\omega_1 - \omega_2}^2 = \int W(t, \omega, \omega_1 - \omega_2, \theta) dt \quad 15$$

It is subsequently demonstrated that even in the presence of high levels of noise, this method for tone detection is still able to work successfully. One such situation is illustrated in figure 4, where the sum of two tones is subjected to additive white Gaussian noise (AWGN) of SNR level -2 db. In figure 4(a), the WD of the contaminated signal is

shown, and in *figure 4(b)*, the reconstruction of the marginal $|G(\omega, \theta)|^2$ can still determine the frequencies of these tones with a very high accuracy. In the presence of a single tone, frequency detection is possible up to -10 db SNR.

The presented results compare favourably with previous work on single-tone frequency estimation

(eg. [8] reports a best-case RMS error for -2 dB SNR, which is higher than 0.1 Hz).

The successful use of NWD for two-tone separation, as evidenced in the presented results, goes some way to suggest that efficient solutions for many difficult signal processing problems, such as time-frequency filtering and cross-term suppression, may be provided by (higher order) nested distributions.

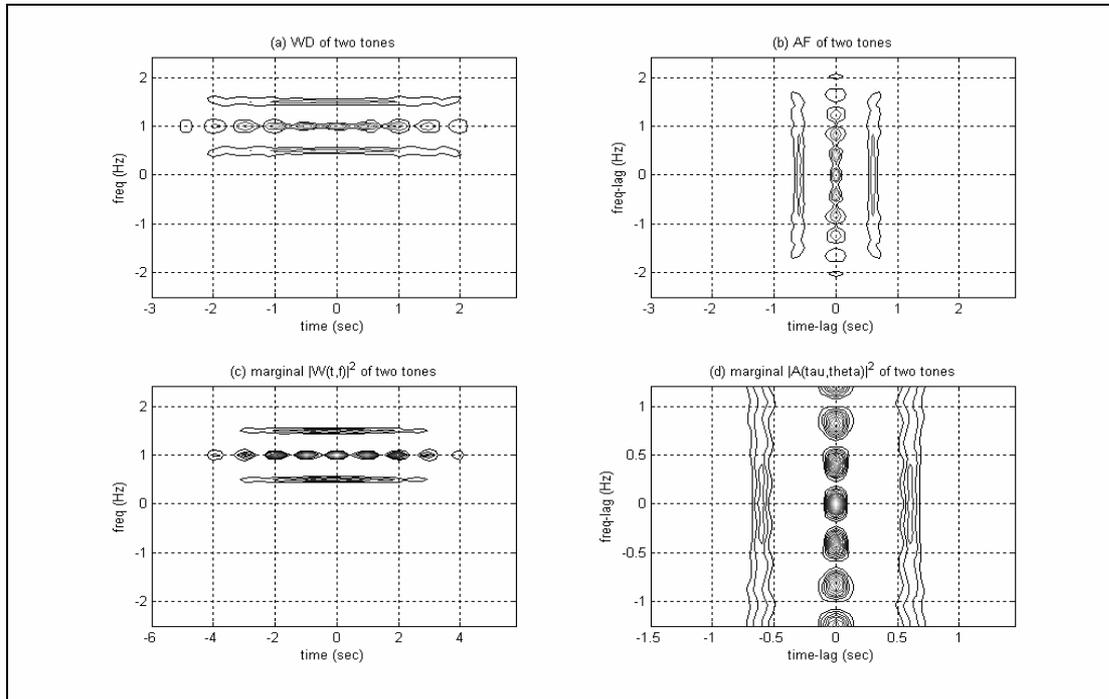


Figure 1 (a) WD of a two-tone signal with frequencies 0.5 and 1.5 Hz. Cross-terms can be seen at the location of the average frequency of the two tones, i.e. 1.0 Hz. (b) The ambiguity function of the same signal. (c) The 2-D Marginal of the 4-D NWD along τ and θ , equivalent to $|W(t, f)|^2$. (d) The 2-D Marginal of the 4-D NWD along time and frequency, equivalent to $|A(\tau, \theta)|^2$. The effects of dilation are visible in τ and θ .

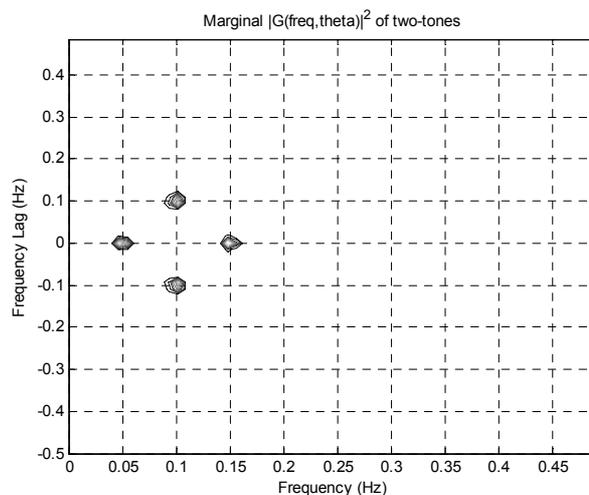


Figure 2 The 2-D Marginal $|G(\omega, \theta)|^2$ of 4-D NWD along time (t) and time-lag (τ), from which the exact frequencies of the two tones can be determined.

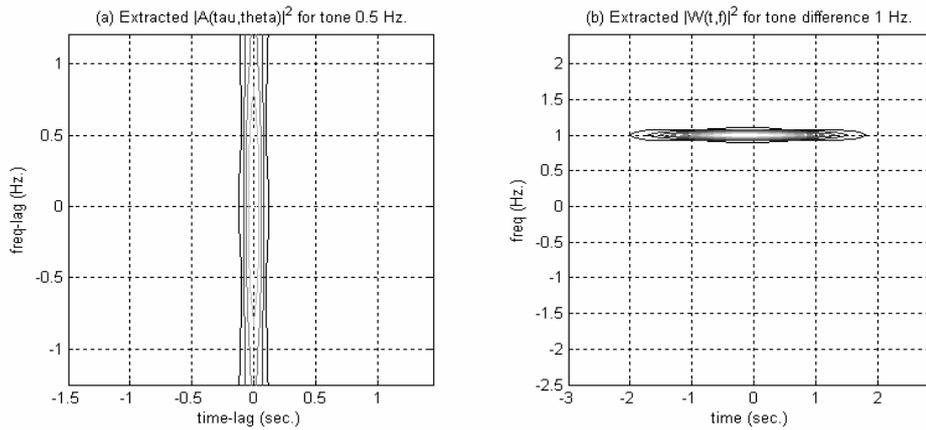


Figure 3 (a) The 2-D marginal representing the squared magnitude Ambiguity function $|A(\tau, \theta)|^2_{\omega_1}$, extracted by double-integrating along ω_1 (0.5 Hz). (b) The 2-D marginal representing the squared magnitude WD function $|W(t, f)|^2_{\omega_1 - \omega_2}$, for the tone of 1Hz. ($\omega_1 - \omega_2$).

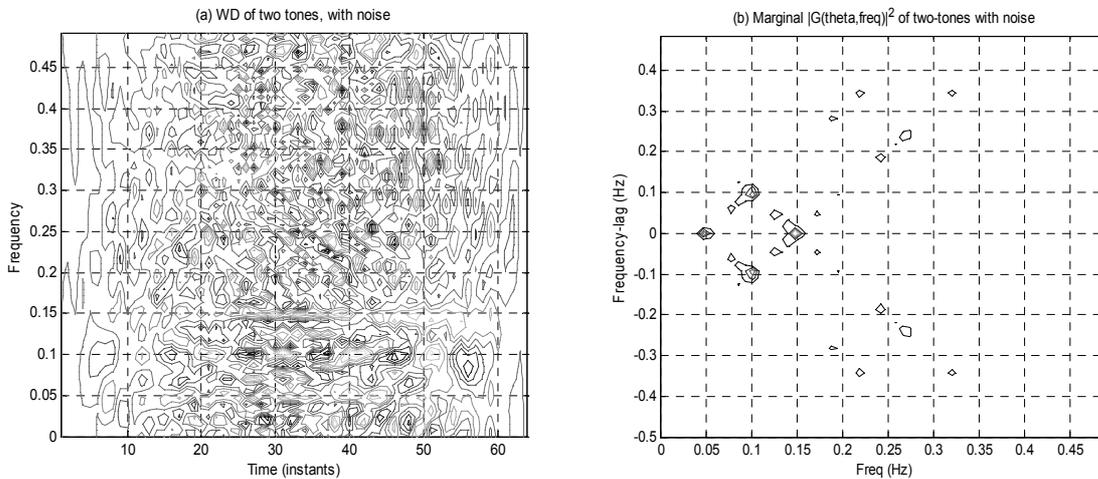


Figure 4 (a) The WD of the sum of two tones exposed to -2db AWGN. (b) The 2-D Marginal $|G(\omega, \theta)|^2$ gives the exact frequencies of the two tones, despite the presence of noise.

4. Conclusion

The second-order Nested Wigner Distribution was introduced for continuous- and discrete- time signals. The NWD is a quartic function, which essentially relates the time-frequency distribution to the Ambiguity Function space, i.e. to time and frequency lags. The nested Wigner function is closely related to the quantum mechanical formalism of the Wigner function and the extended phase space [9,10]. This new approach is shown to provide substantial benefits for signal separation, making it highly promising for signal processing and communications.

References

[1] L. Cohen, Time frequency Distributions: A review, *Proceedings of IEEE*, Vol.77, No. 7, 1989.
 [2] A.C.M. Claasen, W.F.G. Mecklenbrauker, The aliasing problem in Discrete-Time Wigner Distributions, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 31, No. 5, 1983.

[3] R. Balmer, H. Glunder, the Wigner Distribution function of two-dimensional signals, Coherent –optical generation and display, *Optica Acta*, Vol. 30, No. 12, 1983
 [4] M.J. Bastiaans, The Wigner Distribution function applied to Optical signals and systems, *Optics communications*, Vol. 25, No. 1, 1978.
 [5] Y.M. Zhu, Y. Gao, Analysis of Chest CT images using Wigner Distribution, Engineering in Medicine and Biology Society, 1992. Vol.14. *Proceedings of the Annual International Conference of the IEEE*, Volume: 5, October 29- November 1, 1992, Pages:1926 – 1927
 [6] A.C.M. Claasen, W.F.G. Mecklenbrauker, The Wigner Distribution-A tool for time-frequency signal analysis, Part I and II, *Phillips J. Research*, Vol. 35, pp. 276-350, 1980
 [7] A. V. Oppenheim, Speech spectrograms using the fast Fourier transform, *IEEE Spectrum*, Vol. 7, pp. 57-62, 1970.
 [8] Rife, D., Boorstyn, R, Single tone parameter estimation from discrete-time observations, *Information Theory, IEEE Transactions on*, Volume: 20, Issue: 5, Sep 1974, Pages: 591 - 598
 [9] S. Chountasis, L.K. Stergioulas and A. Vourdas, “Quantum filtering of noise in the Wigner function”, *Journal of Modern Optics*, vol. 46, No. 15, pages 2131-2141, 1999.
 [10] S Chountasis and A Vourdas, The extended phase space: a formalism for the study of quantum noise and quantum correlations, *J. Phys. A: Math. Gen.*, Vol. 32, pp. 6949-6961, 1999.

New Method of Comparison NS-FRLS and NLMS Algorithms for the Acoustic Echo Cancellation

M.Arezki (*), A.Benallal (*), F.Ykhlef(*), A.Guessoum (*), D.Berkani (**)

(*) :LATS Laboratory –Department of Electronic, Faculty of Engineering
Saad Dahlab University, BLIDA, ALGERIA. md.arezki@caramail.com

(**) : Signal &Communication Laboratory -(ENP), ALGIERS, ALGERIA.

Key words: NLMS, FRLS, adaptive Filtering, adaptation algorithm , adaptation gain, least squares, speech signal, echo.

Abstarct

In this paper, we present a comparative study of numerically stable fast recursive least squares (NS-FRLS) algorithms and the normalized least mean squares (NLMS) for the acoustic echo cancellation. We propose a methodology of comparison which consists in choosing equivalent adaptation gains for the two types of algorithms. We thus develop a new stabilisation method of the algorithms NS-FRLS with speech signal. The tests were carried out on speech signals convoluted with impulse responses of real acoustic environment with and without the addition of noise. This study confirms the superiority of the algorithms NS-FRLS on those of type NLMS.

1 Introduction

Communication means like the telephone and the data-processing networks are essential elements of telecommunications with long distance. The telephone conversation allows already a high level of communication by the choice of the words and the tone of the voice; the visiophonie adds new dimensions to this communication: face expression, gestures, body language, visible environment. Documents with texts, images or other data can be used to come to support the face to face, in order to have a better comprehension.

The audioconference is a technique making it possible to establish the communication between two conference rooms. The sound recording in each room regards the speaker as the useful source. Unfortunately, all the traditional equipment free hands [1] has an acoustic reaction problem which depends on the loudspeaker and the microphone disposition, on their immediate environment, and the place where they are used.

The microphone free hands is not only exposed to the sounds diffused directly by the closer interlocutor. Part of the sound outlet of the loudspeaker, returns directly to the microphone through the air, due to the reflection of the sound waves by the surrounding

objects and then is returned to the distant interlocutor. It is what one called acoustic echo.

The objective of this paper is a comparative study of the acoustic echo cancellation using two types of different adaptation algorithms : the algorithm FRLS and algorithm NLMS [2].

We compare the initial convergence speed and the filtering error energy (EQM) between FRLS and NLMS algorithms with speech signal input in presence or not of the additive noise at output.

2 The acoustic echo cancellation model

The acoustic echo takes source in the sound propagation since the loudspeaker towards the terminal microphone ("hand-free"); this propagation includes the reflexions on the room walls and the obstacles present in this room. We represent it by an acoustic channel of coupling. The microphone collects in more, the ambient noise in the room (figure 1).

The acoustic echo is unpleasant for the distant correspondent as soon as the delay of transmission in the telecommunications network exceeds 30 ms ; thus it must be attenuated as much as possible [3].

The echo cancellation represented schematically in figure 1 consists in cutting off to the signal y_t coming from the microphone, the "echo" \hat{s}_t estimated from the received signal x_t by filtering through the filter $H_{N,t-1}$.

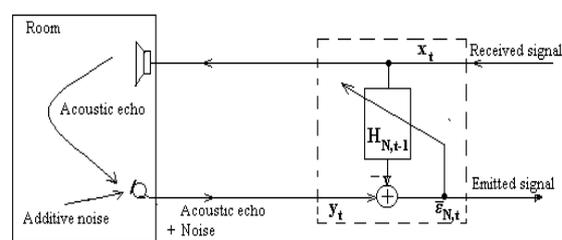


Figure 1: Audio terminal with acoustic echo model. The framed part schematizes the principle of the echo cancellation.

Classically $H_{N,t-1}$ is a transverse finite impulse response filter (FIR) of N coefficients [4] directly represents the "useful part" (beginning) of the impulse response, it is obtained by an adaptive identification from the signal received x_t as input and from the output signal y_t where t indicates the discrete time index.

3 The NS-FRLS and NLMS algorithms comparison for the acoustic echo cancellation

3.1 Comparison methodology

An acoustic echo canceller is summarised by the following equation:

$$\bar{\epsilon}_{N,t} = y_t - H_{N,t-1}^T X_{N,t} \quad (1)$$

with

$$H_{N,t} = H_{N,t-1} + G_{N,t} \bar{\epsilon}_{N,t} \quad (2)$$

where $\bar{\epsilon}_{N,t}$ is the weakened echo, $X_{N,t}$ is a vector which summarizes the past of x_t on N samples and $H_{N,t}$ (vector of order N) is the impulse response of the way of echo estimated by an adaptive algorithm. The exponent T symbolises the transposition operator.

The adaptation gain $G_{N,t}$ is:

For NLMS case:

$$G_{N,t} = G_{NLMS}(t) = \frac{\delta}{X_{N,t}^T X_{N,t}} X_{N,t} \quad (3.a)$$

with the convergence condition on δ data given by [5]:

$$0 < \delta < 2$$

And for RLS case [6]:

$$G_{N,t} = G_{RLS}(t) = R_{N,t}^{-1} X_{N,t} \quad (3.b)$$

$$R_{N,t} = \sum_{i=1}^t \lambda^{t-i} X_{N,i} X_{N,i}^T = \lambda R_{N,t-1} + X_{N,t} X_{N,t}^T$$

The quantity $R_{N,t}$ is the N-by_N sample covariance matrix of the input signal x_t . The adaptation gain of this algorithm is calculated, independently of the filtering part $H_{N,t}$, by a fast recursive least squares algorithm (FRLS) by using a prediction forward/backward analysis on the signal x_t . The adaptation gain is written then [7]:

$$G_{N,t} = G_{FRLS}(t) = \gamma_{N,t} \tilde{K}_{N,t} \quad (3.c)$$

where the variables $\gamma_{N,t}$ and $\tilde{K}_{N,t}$ respectively denote the likelihood variable and normalized Kalman gain vector.

A numerical stabilisation algorithm version (NS-FRLS) [8] is listed on table 1, for our simulations, we took the choice ($\mu^B = \mu^\beta = \mu^\gamma = 1$ and $\mu_s = 0.75$).

To be able to put the two algorithms NS-FRLS and NLMS under the same adaptation conditions, we sought an equivalence between the adaptation gains. For a large order N, and by supposing that the input signal x_t is stationary and slightly correlated, the asymptotic gains are written:

$$G_{NLMS}(t) \approx \frac{\delta}{N\sigma_x^2} X_{N,t} \quad (4.a)$$

$$G_{RLS}(t) \approx \frac{1-\lambda}{\sigma_x^2} X_{N,t} \quad (4.b)$$

σ_x^2 is the variance of the input signal x_t . The exponential forgetting factor of the algorithm NS-FRLS ($0 < \lambda \leq 1$), is calculated generally by the following equation [8], [9]:

$$\lambda = 1 - \frac{1}{pN} \quad (5)$$

where the parameter p must be a real strictly greater than 2 to ensure the numerical stability of the algorithm [8], [10].

The comparison of the equations (4.a) and (4.b), by taking account of (5), gives the following equation:

$$G_{NLMS}(t) = \delta p G_{NS-FRLS}(t) \quad (6)$$

In all our simulations, we will take the same gain;

$$\delta p = 1 \quad (7)$$

For these adaptive algorithms, we will have always to make, in asymptotic mode, a compromise between a weak EQM (near to 1 (NS-FRLS); weak in front of 2 (NLMS)) and a better convergence speed (weak in front of 1 (NS-FRLS); near to 2 (NLMS)).

3.2 Technique of stopping the adaptation during silences.

Numerically stable fast recursive least squares (NS-FRLS) algorithms with stationary signals at the input diverge with a strongly non-stationary signals as speech. To ensure the working without divergence of these algorithms, we propose a technique which allows an execution without degradation of the algorithms NS-FRLS [10]. It is based on the stopping (freezing) of the adaptation during the silence zones of the speech signal.

Silences are quite simply intervals where the useful signal is missed. These silent periods are necessary for the intelligibility of the vocal signal, they occupy a significant part time of locution.

In practice, it is question of noises, of various origins, and negligible energy in front of that of the useful signal.

During the sequences of silence, the variances of the forward/backward predictions errors in the algorithm tend towards zero, and on arrival of the useful signal the algorithm is badly initialized what causes significant discontinuities in certain variables of the algorithm, so this last diverge.

We propose a technique, which consists in benefiting from the silence intervals to block the adaptation, since the signal is absent, thus the information is not lost and to take again the adaptation as soon as there is useful signal. For that, we permanently estimate the energy of the speech signal x_t :

$$E_x(t) = \lambda_c E_x(t-1) + (1 - \lambda_c) x_t^2 \quad (8)$$

The only energy level to be fixed is the threshold energy E_{seuil} beyond which there is adaptation i.e.:

If $E_x(t) \leq E_{seuil} \Rightarrow$ no adaptation

When $E_x(t)$ becomes higher than E_{seuil} , we reinstate the algorithm on the filter parameters already identified and we continue the adaptation. In our simulations, we chose a forgetting factor $\lambda_c < 1$.

4 Simulation results

We have an impulse response of acoustic coupling measured in a videoconference room in stationary regime. It contains 4096 points (figure 2).

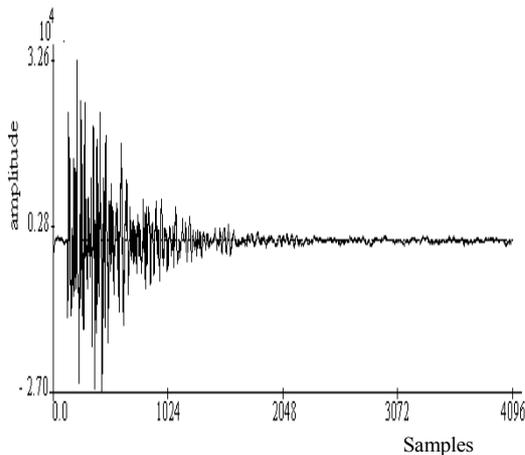


Figure 2: Impulse responses of a videoconference room

The input signal x_t , is a speech signal sampled at 16kHz. The output signal y_t is the signal x_t transformed by the coupling acoustic impulse response between the loudspeaker and the microphone.

The impulse response energy of the coupling acoustic channel decreases exponentially.

The reverberation time, measures this decreasing rate [1], [4]: it is the time which puts the sound wave energy to decrease of 60 dB; it gives the order of the significant temporal length magnitude of the impulse response.

The classically objective criterion used in echo cancellation is the weakening of the echo signal between the sound recording microphone (signal y_t) and the signal at exit of the echo canceller (signal $\bar{\varepsilon}_{N,t}$). With speech signals this criterion is not very significant in the silences zones.

For this reason, we will use primarily for the comparison the energy of the filtering error (EQM). For the curves, we calculated the EQM on 128 consecutive samples. The final EQM of the tables below is calculated on the last 5 blocks (of 256 points per block) of the error signal.

$$EQM(t) = 10 \log(\sigma_{\bar{\varepsilon}_{N,t}}^2) \quad (9)$$

$$\sigma_{\bar{\varepsilon}_{N,t}}^2 = \frac{1}{L} \sum_{i=1}^L \bar{\varepsilon}_{N,i}^2 \quad (10)$$

4.1 Order influence of the transverse filter

The stationary channel is formed of the first 2048 points of a coupling acoustic impulse response measured in a videoconference room. The results of this application are summarized on table 2 and figure 3 below.

Filter order N	Final EQM (dB)	
	NS-RLS	NLMS
Underestimated 1024	-20.81	-20.48
Real 2048	-91.95	-44.52
Over-estimated 4096	-84.22	-40.27

Table 2: Comparative table of the final EQM between NS-RLS and NLMS algorithms for given orders N

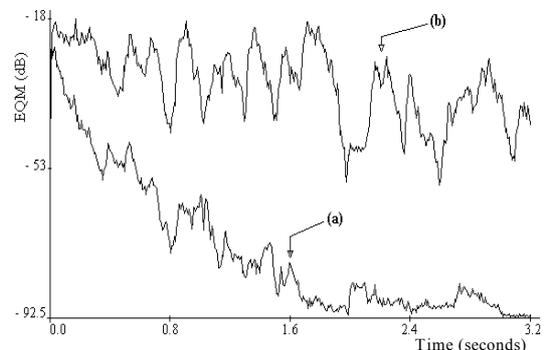


Figure 3: Temporal evolution of the EQM
(a) NS-RLS, (b) NLMS:
(N = 2048; no noise at exit)

Table 2 shows well that the NS-RLS algorithm is better than the NLMS from the viewpoint of the final filtering error energy.

Figure 3 corresponds to the case where there is no added noise at exit. We check that the NS-FRLS is much higher than the NLMS in convergence speed. We will also note that towards the end of the experiment of duration 3.2 seconds (52500 iterations approximately) the NLMS is still in transitional stage. It is the principal disadvantage of the NLMS: the convergence speed is slow and depends on the input signal.

4.2 Additive noise influence

In this section, we simulated the effect of an additive noise on the echo signal. This last is written as follow:

$$y_t/n = y_t + n_t \quad (11)$$

where n_t is a stationary white Gaussian noise. The signal-to-noise ratio SNR_{dB} is calculated by the following equation :

$$SNR_{dB} = 10 \log \frac{\sigma_y^2}{\sigma_n^2} \quad (12)$$

where σ_y^2 is the total energy of y_t and σ_n^2 is the total energy of the additive noise n_t .

For our simulations, the real size of the acoustic channel to identify is equal to that of the transverse adaptive filter. The identification is about the impulse response of a videoconference room of $N=2048$. The results are given as follows:

SNR _{dB}	Final EQM (dB)	
	NS-RLS	NLMS
50	-57.27	- 44.11
40	-46.76	- 41.83
30	-37.34	- 35.57
20	-26.90	- 25.81
10	-17.70	- 16.70

Table 3: Comparative table of the final EQM between NS-FRLS and NLMS algorithms for $N=2048$ with an output noise of a given SNR_{dB} .

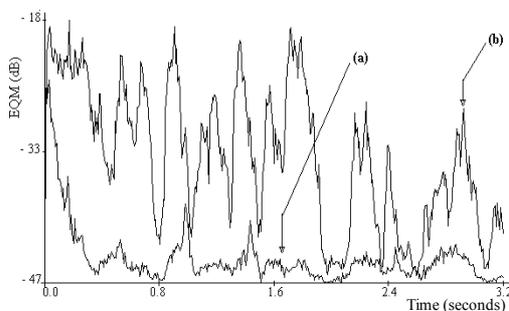


Figure 4: Temporal evolution of the EQM
(a) NS-RLS, (b) NLMS
($N=2048$, $SNR_{dB} = 40$ dB)

By comparing the results of the obtained tables, we note that the sound effects of the echo signal introduce an inevitable degradation into the EQM.

The filtering EQM is located slightly below the disturbing noise energy. This is in conformity with the adaptive filtering theory since the added noise is not taken into account by the adaptive filter.

Figure 4 represents the temporal evolution of the EQM for a signal to noise ratio about 40 dB at exit. We observe a degradation of the NS-FRLS asymptotic EQM compared to the case without noise but the convergence of this algorithm is definitely visible whereas for NLMS algorithm, degradation is considerable. The algorithm is strongly disturbed; in fact, we can say that it does not converge any more.

This is explained by the fact that the filter adaptation is strongly disturbed by the additive noise. It results a random disturbance from the adaptive filter coefficients. This last starts to converge during strong energies time of the speech signal, but this convergence is slow and the filtering error cannot reach suitable asymptotic values before the arrival of weak energies zones of the speech signal where the noise energy is almost equal to the useful signal energy.

5 Conclusion

We proposed a new stabilisation technique without interruption of the fast recursive least squares algorithms (NS-FRLS) with a speech signal in input. This technique consists in stopping the adaptation during silences. We compared it thereafter with the normalized least mean squares (NLMS) algorithm according to equivalent adaptation gains.

The comparative study shows that the NS-FRLS algorithm is more robust in the case of the echo acoustic cancellation (speech in entry) and with the additive noise at exit than the NLMS algorithm. This robustness is due primarily at its high convergence speed which does not depend on the nature of the input signal.

Bibliography

- [1] A.Gilloire and J.P.Julien, "Rooms acoustics in telecommunications", the echo of research, No 127, 1st half-year 1987.
- [2] B.Frahang-Boroujeny, "Fast LMS/Newton algorithms based on auto regressive modelling and their application to echo acoustic annulment", IEEE Transactions on signal processing, Vol.45, N°8, pp.1987-2000, August 1997.
- [3] Y.Itoh, Y.Maruyama, N.Furuya and T.Araseki, "An Echo Acoustic Chancellor For Teleconference", IEEE - ICC 85, CHICAGO, 1985, pp.1498-1502.
- [4] A.Gilloire, "Identification of transverse structures: applications to the acoustic answers", Signal processing, volume 6, N° 5, 1989, pp.347-351.
- [5] B Widrow and S. D. Stearns, Adaptive Processing Signal. Englewood Cliffs, NJ: Prentice-hall, 1985.
- [6] S.Hayking, Adaptive Filter Theory, 3rd ED. Englewood Cliffs, NJ: Prentice-hall, 1996
- [7] J.CIOFFI, T.KALATH "Transverse Fast RLS Filters for adaptive filtering" IEEE press. One ASSP 1984.
- [8] M.Arezki, A.Guessoum, D.Berkani, "New version of numerically stable algorithm MCR for adaptive filtering", national Conference on the engineering of electronics, Oran, Algeria, pp.53-55, December 2002.
- [9] A.BENALLAL, A.GILLOIRE "A New method to stabilize fast RLS algorithms based on a first-order model of the propagation of numerical errors" Proc. ICASSP88, New York 1988.
- [10] Mr. Arezki " Study of the fast recursive least squares least Algorithms in the presence of the speech signal " Magister. Blida's university 1997.

- Prediction part

Variables available at the moment t :

$$A_{N,t-1}; B_{N,t-1}; \tilde{K}_{N,t-1}; \gamma_{N,t-1}; \alpha_{N,t-1}; \beta_{N,t-1}; H_{N,t-1}$$

New information : x_t and x_{t-N}

Modeling of x_t and x_{t-N}

$$\bar{e}_{N,t} = x_t - A_{N,t-1}^T X_{N,t-1}$$

$$\alpha_{N,t} = \lambda \alpha_{N,t-1} + \gamma_{N,t-1} \bar{e}_{N,t}^2$$

$$\tilde{K}_{N+1,t} = \begin{bmatrix} 0 \\ \tilde{K}_{N,t-1} \end{bmatrix} + \frac{\bar{e}_{N,t}}{\lambda \alpha_{N,t-1}} \begin{bmatrix} 1 \\ -A_{N,t-1} \end{bmatrix}$$

$$\begin{bmatrix} \tilde{K}_{N,t} \\ 0 \end{bmatrix} = \tilde{K}_{N+1,t} - \tilde{k}_{N+1,t,N+1} \begin{bmatrix} -B_{N,t-1} \\ 1 \end{bmatrix}$$

$$A_{N,t} = A_{N,t-1} + (\gamma_{N,t-1} \tilde{K}_{N,t-1}) \bar{e}_{N,t}$$

$$\bar{r}_{N,t}^c = x_{t-N} - B_{N,t-1}^T X_{N,t}$$

$$\bar{r}_{N,t}^{f_0} = \lambda \beta_{N,t-1} \tilde{k}_{N+1,t,N+1}$$

$$\bar{r}_{N,t}^{f_1} = \lambda^{-N+1} \gamma_{N,t-1} \alpha_{N,t-1} \tilde{k}_{N+1,t,N+1}$$

$$\xi_{N,t} = \bar{r}_{N,t}^c - [(1 - \mu_s) \bar{r}_{N,t}^{f_0} + \mu_s \bar{r}_{N,t}^{f_1}]$$

$$\bar{r}_{N,t}^B = \bar{r}_{N,t}^c + \mu^B \xi_{N,t}$$

$$\bar{r}_{N,t}^\beta = \bar{r}_{N,t}^c + \mu^\beta \xi_{N,t}$$

$$\bar{r}_{N,t}^\gamma = \bar{r}_{N,t}^c + \mu^\gamma \xi_{N,t}$$

$$\gamma_{N+1,t} = \frac{\lambda \alpha_{N,t-1}}{\alpha_{N,t}} \gamma_{N,t-1}$$

$$\gamma_{N,t} = \frac{\gamma_{N+1,t}}{1 - \gamma_{N+1,t} \bar{r}_{N,t}^\gamma \tilde{k}_{N+1,t,N+1}}$$

$$B_{N,t} = B_{N,t-1} + (\gamma_{N,t} \tilde{K}_{N,t}) \bar{r}_{N,t}^B$$

$$\beta_{N,t} = \lambda \beta_{N,t-1} + \gamma_{N,t} (\bar{r}_{N,t}^\beta)^2$$

- Filtering Part

$$\bar{e}_{N,t} = y_t - H_{N,t-1}^T X_{N,t}$$

$$H_{N,t} = H_{N,t-1} + (\gamma_{N,t} \tilde{K}_{N,t}) \bar{e}_{N,t}$$

Table 1: Numerically stable fast recursive least squares NS-FRLS algorithm

New method to count vehicles in highway by using some pixels and threshold

Dr. Heydat Toossian Shandiz Assistant professor at Shahrood university of technology hshandiz@shahrood.ac.ir, Seyed Saeid Mirsharifi, Shahrood university of technology, 7th Tir Square, Po Box 36155-316, Shahrood, Iran

Abstract

There are many methods in counting vehicles in order to control the traffic, such as, mosaic method, gradient Chebychev filtering and so on. The new method is used some pixels and threshold gray level. This method has easy algorithm, high speed and not depend to type and width of vehicle. Furthermore it is independent of type of standing the vehicle.

Keywords: Digital image processing, Traffic control, Vehicles counting

Introduction

Continued growth in traveling, make government to investment a lot of money to use intelligent systems in controlling traffic. The reasons for improving these systems are reducing the people death and injury in car accidents and wasted fuel costs.

The first step in controlling traffic or improving traffic routes is measuring the number of passing vehicle. There are many sensors for controlling the traffic. However, these devices can't detect obstruction and dangerous conditions such as vehicles blocking sensor and person standing on them. Moreover non-visual information is not widely accepted on the courts. Other reasons for using camera and image processing algorithms are high price, fixing and low accuracy of sensors [1]. Moreover the camera which is used for counting may be used for other tasks, such as recording the speed, car number and type of vehicle.

There are three problems in processing and making decision based on captured image from a highway. First problem is huge amount of data which are processed, especially in real time processing. For example, if the image has 512 by 512 pixels for 30 frames in each second and 256 gray levels, then about 246 Mega bit/sec have to be processed. Second problem is effect of environment over the captured image, such as, rainy weather, fog, shadow, reflection and changing in brightness of cars pluses the difference shape of them. The third problem is the price of hardware and software.

Although distortion due to recording, changing in weather, reflection, and so on can't be removed completely [2], but system based on machine vision is preferred.

There are different methods for controlling traffic, such as MOSAIK method which is based on the idea of interacting between detail recognition, fast tracking and supervising attention control[3] [4]. Speed estimation method [5], gradient, chebychev filtering, in which finding edge of image is first task, are other methods. Edge detection algorithms are time consuming process. Pattern recognition and tracking and speed estimation [6] [7] is another method which is used edge detection and some pattern for comparison.

The new method

This method is based on three steps: choosing the pixels, taking the reference image and detection rule. Reference image is an empty image from the road.

The reference images are image for different time of a day. They can be taking by camera in different time of a day or by changing the gray level of an empty image in computer. They are taken or calculating one time and stored in computer memory. In this algorithm traffic image changed to black and white. In each image some lines are chosen. These lines contain the chosen pixel. Number of lines in pictures depends on average length of cars. In this work it is 4 m plus 0.5 m for dead band. The same lines are chosen from reference image. Then the corresponding line are subtracted and saved in an X array. In order to reduce the effect of noise and avoid use of edge detection algorithms a threshold level U is chosen. If the result array has gray level higher than U, then corresponding pixel change to 1 otherwise it set to be zero. The result array X is tested to recognize the number of changing from 0 to 1 and vice versa. The number of cars in each image is count from following formula

$$N = \sum_{i=1}^n B_i$$

In which N is total number of cars, n is the number of line in each image, B_i is the number of cars in i th line. If the number of changing level in line i is C_i , then

$$B_i = C_i/2$$

This method theoretically overcomes three main problems better than other methods. The algorithm is easy. By choosing U for each camera in calibrating system for first time, there is no difficulty to find the edge of cars. There is no need to use gradient algorithm. As choosing some lines in each image the total processed pixels are reduced extremely. The result of reducing the pixels is low memory higher processing speed and finally reducing the price of system. In this method the counting is independent of type and width of car. Moreover how the car aligns in image can't affect the counting. Standing over the street line still has no effect the counted number.

Testing the algorithm

The algorithm was tested based on real images taking by Tehran control traffic center cameras. Some sample images are in figure 1. The results of running program are in table 1.

Table 1 Testing results

Figure 1	Real number	Calculated number	Error
a	2	2	0
b	8	6.5	1.5
c	1	1	0
d	5	5	0
e	6	6.5	0.5
f	6	6	0
g	4	5	1
h	1	2	1
i	2	2	0
j	2	2	0
k	3	3	0

The errors are in figure 1 (b), (e), (g) and (h). The first error is due to three cars which are not completely in image. Each of them is counted as half. The second error is because of existence of half of a car in other street. The third and fourth errors are due to length of vehicle. If the number of car is important they are errors but if the flow of car and controlling of traffic lights are the objective of counting, they are not errors.

Conclusion

The new algorithm reasonably works well. It is easy to implement. As the algorithm takes some line in image for processing the amount of pixels are too low. It is easy to run program on line with cheap instruments.

Acknowledgment

Authors want to express their thanks to Tehran control traffic center in Iran.

References

- [1] Inigo, R. Application of machine vision to traffic monitoring & control. *IEEE transactions on vehicular technology*, vol.38, pp.112-122, 1989
- [2] Hoose, N. "Computer image processing in traffic engineering". Research Studies Press, Taunton, 1991.
- [3] A. Benschrair, M. Bertozzi, A. Broggi, P. Mich'e, S. Mousset, and G. Toulminet, "A cooperative approach to vision-based vehicle detection". In *Proc. IEEE Intl. Conf. on Intelligent Transportation Systems*, pages 209-214, 2001
- [4] Sobottka, K. & Wetzel, D. Attention control integrated in a system to Autonomous driving & collision avoidance under ego motion. Bavarian Research Center. *IEEE*, pp. 796-800, 1997
- [5] Abut, H. & Bilgin, A. & Bernardi, and R.M. A case studies in IVHS implementation an image processing application for 1-15 hov lanes. E.C.E department, San Diego State University, San Diego, CA92182, *IEEE*, pp.2979-2982, 1995
- [6] Zhimin Fan, Jie Zhou, Dashan Gao and Zhiheng, Li, "Contour Extraction and Tracking of Moving Vehicles for Traffic Monitoring", In *Proc. IEEE International Conference on Intelligent Transportation Systems (ITSC 2002)*, pp. 84-87, 2002
- [7] Zhimin Fan, Jie Zhou, Dashan Gao and Gang Rong, "Robust Contour Extraction for Moving Vehicle Tracking", In *Proc. IEEE International Conference on Image Processing (ICIP 2002, Oral)*, vol. 3, pp. 625-628, 2002



(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)



(i)



(j)



(k)

Figure 1 Real images from Tehran traffic center

On nonuniform sampling of one-dimensional signals

K. Busawon, R. Binns and B. Livingstone

Northumbria University,

School of Engineering, Ellison Building,

Newcastle upon Tyne NE1 8ST, U.K.

Phone: (0) 191 227 3103, Fax: (0) 191 227 3684

email: krishna.busawon@unn.ac.uk

Abstract

This paper investigates the possibility of extending the Shannon interpolation formula to the nonuniform sampling case for one-dimensional signals. It is shown that such extension is possible for particular types of nonuniform sampling sequences which possess some specific properties. In addition, a special sampling sequence which is dependent on the rate of change of signal with time is given. This has the benefit of obtaining more sample points when the signal is fast varying and less sample points when the signal is slow varying. As a result, a better representation of the signal, from a linear interpolation viewpoint, is obtained when the irregular sampling process is employed. Simulations are provided to show the performance and characteristics of the sampling sequence.

1 Introduction

This paper considers the problem of nonuniform sampling of a continuous-time one-dimensional signal according to a special sampling sequence. First, we start by setting some notations and giving some definitions. The set of natural numbers is denoted by N , and the set of integers is denoted by Z . The process of sampling a continuous-time one-dimensional signal $x(t)$ consists in selecting or measuring instantaneous values of its amplitude according to a divergent monotone increasing sequence of time $\{\tau(n)\}_{n \in Z}$. The sequence $\{\tau(n)\}_{n \in Z}$ is called a *sampling sequence*. In general, when $\{\tau(n)\}_{n \in Z}$ is a divergent monotone increasing arithmetic sequence, the sampling is said to be *uniform* or *regular*. However, it is commonly understood that a uniformly sampled signal is one that is sampled according to the special arithmetic sequence $\tau(n) = nT$ where $T > 0$ is known as the *sampling interval*. As a result, we obtain the following set of regularly spaced samples $\{\dots, x(-2T), x(-T), x(0), x(T), x(2T), \dots\}$ which can be written in a compact form as

$$\hat{x}(t) = x(t)\Pi(t, nT) \quad (1)$$

where $\Pi(t, nT) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$ is the *uniform comb function* and $\delta(t)$ is the dirac function.

On the other hand, if the sampling sequence $\{\tau(n)\}_{n \in Z}$ is a divergent monotone increasing nonlinear function of n , the sampling is said to be *nonuniform* or *irregular* and the resulting sampled signal, $\tilde{x}(t)$, can be written as

$$\tilde{x}(t) = x(t)\Pi(t, \tau(n)) \quad (2)$$

where $\Pi(t, \tau(n)) = \sum_{n=-\infty}^{+\infty} \delta(t - \tau(n))$ is the *nonuniform comb function*.

Going back to uniform sampling, it is also worthwhile noting that the sequence $\tau(n) = nT$ is not dependent on the signal that is being sampled and that the operation of uniform sampling is, in fact, a linear operation. Uniform sampling has led to the establishment of the uniform sampling theorem also called the Shannon theorem (or sometimes the Shannon Whittaker Kotelnikov theorem) and which is indisputably most famous theorem in information theory. It is stated as follows:

Theorem 1 (Shannon Sampling Theorem): *If $x(t)$ is a real valued band-limited signal such that its fourier transform $X(\omega) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t} dt = 0$ for $|\omega| > \omega_{\max}$, then $x(t)$ can be uniquely determined from its sample values $x(nT)$ sampled at uniform intervals if $\omega_s = \frac{2\pi}{T} \geq 2\omega_{\max}$. More precisely, $x(t)$ can be recovered by the following interpolation formula:*

$$x(t) = \sum_{n=-\infty}^{+\infty} x(nT) \frac{\sin \omega_{\max}(t - nT)}{\omega_{\max}(t - nT)}.$$

The above formula is known as the *Shannon interpolation formula*. One of the main shortcoming of uniform sampling is that it is not suited for nonstationary signals; that is signals that have different frequency components at different interval of time. For such signals, if regular sampling is employed, valuable data can be lost if the sampling theorem is not obeyed. On the other hand, large amount of unnecessary data will be stored

if the signal is constant or slowly varying within a given interval of time. Consequently, it would be more judicious to adopt an irregular sampling strategy to deal with such signals.

In recent years, a number of research works has been conducted on irregular sampling, for both bandlimited and nonbandlimited signals, using various types of nonuniform sampling sequences (see [1, 2] and references therein). In [1] for example, the problem of nonuniform sampling of nonbandlimited signal was treated for a continuous-time signal with bounded variation. A particular sampling sequence was employed where each sampling instant was assumed to be inside a well defined regular interval.

In this paper, the possibility of extending the Shannon interpolation formula to the nonuniform sampling case is investigated. It is shown that such extension is possible for special types of nonuniform sampling sequences satisfying some specific properties. This naturally leads to the establishment of a new complete system of orthogonal exponential functions through which the proposed nonuniform interpolation result is derived. In addition, a special sampling sequence which depends on the rate of change of the signal with time is proposed. This has the benefit of obtaining more sample points when the signal is fast varying and less samples points when the signal is slow varying. As a result, a better representation of the signal is obtained when the irregular sampling process is employed. Simulations are provided to show the performance and characteristics of the sampling sequence.

Notations: Throughout the paper, ξ^* will denote the conjugate of a complex number ξ . The closed interval of length T_0 is denoted by $T_0 = [t_0, t_0 + T_0]$ where t_0 is a real number. Consequently, $\int_{T_0} f(t)dt = \int_{t_0}^{t_0+T_0} f(t)dt$.

We denote by $L^2(T_0)$ the set of square integrable functions on the interval T_0 . The scalar or inner product of two functions f and g belonging to $L^2(T_0)$ is denoted by $\langle f, g \rangle = \int_{T_0} f(t)g^*(t)dt$ and the norm of f is denoted by $\|f\| = \sqrt{\langle f, f \rangle}$.

2 A nonuniform interpolation result

Let $\{u(n)\}_{n \in Z}$ be a divergent monotone increasing sequence of integers (i.e. $u(n) \in Z$) with the following properties:

P1) $u(0) = 0$

P2) There exists a finite positive number $L \geq 1$ such that $n \leq u(n) \leq nL$ for $n = 1, 2, \dots$

P3) $u(-n) = -u(n)$.

Proposition 1: Let $\{u(n)\}_{n \in Z}$ be a divergent monotone increasing sequence of integers satisfying properties P1-P3. Set $\omega_0 = \frac{2\pi}{T_0}$ with $T_0 > 0$. Then, the following statements hold:

a) The set of exponentials $\{e^{ju(n)\omega_0 t}\}$ ($n = 0, \pm 1, \pm 2, \dots$) is orthogonal over any interval of duration T_0 ; that is

$$\int_{T_0} e^{ju(n)\omega_0 t} (e^{ju(m)\omega_0 t})^* dt = \begin{cases} 0 & m \neq n \\ T_0 & m = n \end{cases}$$

b) If for some function $f(t) \in L^2(T_0)$ there exist real constants c_n such that $f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{ju(n)\omega_0 t}$ then:

i) $c_n = \frac{1}{T_0} \langle f, e^{ju(n)\omega_0 t} \rangle$

ii) If $s_n(t) = \sum_{k=-n}^n \lambda_k e^{ju(k)\omega_0 t} \in L^2(T_0)$, then the integral $E = \|f(t) - s_n(t)\|^2$ is minimum if $\lambda_n = c_n$

iii) $\|f\|^2 \geq T_0 \sum_{k=-n}^n c_k^2$

iv) $\lim_{n \rightarrow +\infty} c_n = 0$.

c) The set of exponentials $\{e^{ju(n)\omega_0 t}\}$ ($n = 0, \pm 1, \pm 2, \dots$) is complete.

The proof of Proposition 1 is fairly straightforward since $u(n)$ is an integer. In effect, one has simply to replace $u(n)$ by some integer \bar{n} and follow the same line of proof as in the uniform case. A detailed proof of Proposition 1 is given in [6].

It is worthwhile noting that the function $\varphi_n(t) = e^{ju(n)\omega_0 t}$ is periodic with period T_0 for every $n \in (-\infty, +\infty)$ since $u(n)$ is an integer. Also, the completeness of $\{e^{ju(n)\omega_0 t}\}_{n \in Z}$ implies that any function $f(t) \in L^2(T_0)$ can be expressed as

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{ju(n)\omega_0 t}.$$

This leads to the following lemma.

Lemma 1: Consider the nonuniform comb function $\Pi(t, u(n)T_0) = \sum_{n=-\infty}^{+\infty} \delta(t - u(n)T_0)$ where $u(n)$ is as above. Then,

i)

$$\Pi(t, u(n)T_0) = \frac{1}{T_0} \sum_{n=-\infty}^{+\infty} e^{ju(n)\omega_0 t}.$$

ii) and the Fourier transform of $\Pi(t, u(n)T_0)$ is given by

$$\mathcal{F}\{\Pi(t, u(n)T_0)\} = \omega_0 \sum_{n=-\infty}^{+\infty} \delta(\omega - u(n)\omega_0)$$

$$\text{where } \omega_0 = \frac{2\pi}{T_0}.$$

For the proof of this lemma see [6].

We can now state the following:

Theorem 2. Consider a band-limited real valued signal $x(t)$ such that its Fourier transform $\mathcal{F}\{x(t)\} = X(\omega) = 0$ for $|\omega| > \omega_{\max}$. Let $\{u(n)\}_{n \in \mathbb{Z}}$ be a divergent monotone increasing sequence of integers satisfying properties P1-P3. Let $x(t)$ be sampled according to the sampling sequence $\{\tau(n)\}_{n \in \mathbb{Z}}$ where $\tau(n) = u(n)T_0$ with $T_0 > 0$. Then $x(t)$ can be uniquely determined from its sampled values $x(\tau(n))$ if $\omega_0 = \frac{2\pi}{T_0} \geq 2\omega_{\max}$. More precisely, $x(t)$ can be recovered from its sampled values $x(\tau(n))$ by the following interpolation formula:

$$x(t) = \sum_{n=-\infty}^{+\infty} x(\tau(n)) \frac{\sin(\omega_{\max}(t - \tau(n)))}{\omega_{\max}(t - \tau(n))}.$$

Proof: First, note that from Lemma 1, we have

$$\mathcal{F}\{\Pi(t, \tau(n))\} = \omega_0 \sum_{n=-\infty}^{+\infty} \delta(\omega - u(n)\omega_0).$$

Next, the Fourier transform of the sampled signal $\tilde{x}(t) = x(t)\Pi(t, \tau(n))$ is given by:

$$\begin{aligned} \tilde{X}(\omega) &= \mathcal{F}\{\tilde{x}(t)\} \\ &= \frac{1}{2\pi} \mathcal{F}\{x(t)\} * \mathcal{F}\{\Pi(t, u(n)T_0)\} \\ &= \frac{1}{T_0} \sum_{n=-\infty}^{+\infty} X(\omega) * \delta(\omega - u(n)\omega_0) \\ &= \frac{1}{T_0} \sum_{n=-\infty}^{+\infty} X(\omega - u(n)\omega_0) \end{aligned} \quad (3)$$

Now, to recover $X(\omega)$ from $\tilde{X}(\omega)$, there should be no overlap between successive cycles of $\tilde{X}(\omega)$. This requires that

$$2\omega_{\max} \leq u(n)\omega_0 - u(n-1)\omega_0 \quad (4)$$

for all values of n . On the other hand, since $u(n)$ is monotone increasing, there exists a positive integer $1 \leq a_n \leq L$ such that $u(n) - u(n-1) = a_n$. Therefore, the no overlap requirement (4) amounts to:

$$2\omega_{\max} \leq a_n\omega_0. \quad (5)$$

On the other hand, the no overlap requirement (4) will be satisfied for all values of n if:

$$2\omega_{\max} \leq \omega_0 \quad (6)$$

since $\omega_0 \leq a_n\omega_0$. Now, from equation (3) and under the condition $X(\omega) = 0$ for $|\omega| > \omega_{\max}$ and $2\omega_{\max} = \omega_0$, we have

$$T_0\tilde{X}(\omega) = X(\omega) \quad \text{for } |\omega| < \omega_{\max}. \quad (7)$$

On the other hand,

$$\tilde{x}(t) = \sum_{n=-\infty}^{+\infty} x(u(n)T_0)\delta(t - u(n)T_0).$$

Therefore,

$$\tilde{X}(\omega) = \sum_{n=-\infty}^{+\infty} x(u(n)T_0)e^{-j\omega u(n)T_0}.$$

Substituting in (7), we get

$$X(\omega) = T_0 \sum_{n=-\infty}^{+\infty} x(u(n)T_0)e^{-j\omega u(n)T_0} \quad \text{for } |\omega| < \omega_{\max}.$$

Taking the inverse Fourier transform of the above equation yields,

$$\begin{aligned} x(t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega)e^{j\omega t} d\omega \\ &= \frac{T_0}{2\pi} \int_{-\omega_{\max}}^{\omega_{\max}} \sum_{n=-\infty}^{+\infty} x(u(n)T_0)e^{j\omega(t - u(n)T_0)} d\omega \\ &= \frac{T_0}{2\pi} \sum_{n=-\infty}^{+\infty} x(u(n)T_0) \int_{-\omega_{\max}}^{\omega_{\max}} e^{j\omega(t - u(n)T_0)} d\omega \\ &= \sum_{n=-\infty}^{+\infty} x(\tau(n)) \frac{\sin(\omega_{\max}(t - \tau(n)))}{\omega_{\max}(t - \tau(n))} \end{aligned}$$

This completes the proof of Theorem 2. \square

In the next section, we give a nonuniform sampling sequence that satisfy properties P1-P3.

3 A special nonuniform sampling sequence

Consider a continuous real valued bandlimited causal signal $x(t)$. Recall that $x(t)$ is causal if $x(t) = 0$ for $t < 0$. Here, we shall consider only causal signals since in practice we generally deal with such signals. It is assumed that the time derivative of $x(t)$ exists and is continuous and bounded. More precisely, there exists a positive constant M such that $\left| \frac{d}{dt}x(t) \right| \leq M$ for all $t \geq 0$.

Let $x(t)$ be sampled according to the following sequence:

$$\tau_x(n+1) = \tau_x(n) + T\beta(n) \quad \text{for } n = 1, 2, 3, \dots \quad (8)$$

where $\beta(n)$ is defined as

$$\beta(n) = \frac{1}{\left| \frac{x(\tau_x(n)) - x(\tau_x(n-1))}{\tau_x(n) - \tau_x(n-1)} \right| + 1} \quad (9)$$

with initial conditions $\tau_x(0) = 0$ and $\beta(0) = 1$ so that $\tau_x(1) = T$. The value of T is chosen such that if $x(t)$ is sampled uniformly with the sampling interval T , then the Shannon sampling theorem is obeyed. More precisely, $T \leq \frac{\pi}{\omega_{\max}}$ where ω_{\max} is the highest frequency component of $x(t)$.

3.1 Some characteristics of the sampling sequence $\tau_x(n)$

1) The sampling sequence (8) is inversely proportional to the modulus of the derivative of the signal $x(t)$. This implies that whenever the signal is varying quickly (high derivative) the shorter is the sampling step. Conversely, the slower the variation of the signal (low derivative) the larger is the sampling step. In this respect, it can be said that the sampling sequence $\{\tau_x(n)\}_{n \in \mathbb{N}}$ is an adaptive sequence.

2) If we set

$$\alpha(n) = \frac{x(\tau_x(n)) - x(\tau_x(n-1))}{\tau_x(n) - \tau_x(n-1)}$$

then, $\frac{1}{M+1} \leq \beta(k) \leq 1$ since $|\alpha(k)| \leq M$. This means that the smallest step size is $\frac{T}{M+1}$ and the largest step size is T . The largest step size T is reached whenever the signal is constant. Recall that T is chosen such that the Shannon sampling theorem is obeyed; that is $0 < T \leq T_{\max} = \frac{\pi}{\omega_{\max}}$.

3) The main benefit of the above sampling algorithm can be appreciated when a nonstationary signal is considered. Indeed, Figure 3 shows the following nonstationary signal

$$\begin{aligned} x(t) = & \sin(t) [u(t) - u(t - \pi)] \\ & + \sin(2t) [u(t - \pi) - u(t - 2\pi)] \\ & + \sin(3t) [u(t - \pi) - u(t - 2\pi)] \end{aligned}$$

when it is sampled according to $\{\tau_x(n)\}_{n \in \mathbb{N}}$ with $T = 0.628s$. The function $u(t)$ is the unit step function. Figure 4 shows the same signal when it sampled regularly with a sampling interval $T_s = T =$

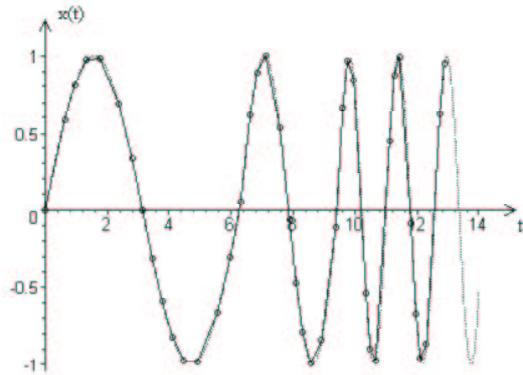


Figure 1: Nonuniform sampling of a nonstationary signal

0.628s. It is clear that the highest frequency component of $x(t)$ is 3rads^{-1} . Consequently, the sampling frequency, in the regular sampling case, is 1.6 times the Nyquist frequency. By comparing both figures, it can be observed that the nonuniform sampling process provides a better approximation of the signal, from a linear interpolation viewpoint. On the other hand, in the regular sampling case, even though the sampling theorem is obeyed, the sample points do not provide a good representation of the signal. This means that the sampling interval T_s should have been chosen smaller compared to T in order to obtain an 'equivalent' linearly interpolated signal as obtained via the nonuniform sampling sequence.

4) It is easy to check that

$$\tau_x(n) = \sum_{k=0}^{n-1} T\beta(k) \quad (10)$$

On account Remark 2, we have

$$\frac{nT}{M+1} \leq \tau_x(n) \leq nT. \quad (11)$$

5) Finally, it is important to note that the amplitude of $\tau_x(n)$ is analog in nature. This makes it particularly difficult to establish an interpolation formula from which $x(t)$ can be recovered from its samples values $x(\tau_x(n))$. Consequently, it is imperative to perform a quantisation or discretisation of the sampling sequence. In addition, this discretisation is necessary for practical implementation purposes.

3.2 Discretisation of $\tau_x(n)$

To discretise $\tau_x(n)$, we simply need to discretise the value of $\beta(n)$. Since $0 < \frac{1}{M+1} \leq \beta(n) \leq 1$, we can discretise the interval $[0, 1]$ up to L quantisation steps.

We shall therefore choose a quantisation step Δ such that $L\Delta = 1$ for some positive integer $L \geq 1$. In general, a value of Δ that is sufficiently close to the value of $\frac{1}{M+1}$ is chosen. The quantise value $\hat{\beta}(n)$ of $\beta(n)$ is then defined as follows:

$$\hat{\beta}(n) = m\Delta \text{ if } \beta(n) \in [(m-1)\Delta, m\Delta]; m \in \{1, 2, \dots, L\} \quad (12)$$

It is now clear that if we defined the new sampling sequence

$$\hat{\tau}_x(n+1) = \hat{\tau}_x(n) + T\hat{\beta}(n) \quad \text{for } n = 1, 2, 3, \dots \quad (13)$$

where $\hat{\beta}(n)$ is the quantised version of $\beta(n)$ as defined above, and, with initial conditions $\hat{\tau}_x(0) = 0$ and $\hat{\beta}(0) = 1 = L\Delta$, then

$$\hat{\tau}_x(n) = T \sum_{k=0}^{n-1} \hat{\beta}(k) = u(n)T\Delta = u(n)T_0 \quad (14)$$

where $u(n)$ is a positive integer such that $n \leq u(n) \leq nL$ and $T_0 = T\Delta$. This means that $\hat{\tau}_x(n)$ is a discrete version of $\tau_x(n)$. The sequence $\{u(n)\}_{n \in \mathbb{N}}$ satisfies properties P1 and P2 of Section 2. It is immaterial whether $\{u(n)\}_{n \in \mathbb{N}}$ satisfies property P3 since we are dealing only with causal signal. Consequently, the signal $x(t)$ can be recovered, for $t \geq 0$, by the following formula:

$$x(t) = \sum_{n=0}^{+\infty} x(\hat{\tau}(n)) \frac{\sin(\omega_{\max}(t - \hat{\tau}(n)))}{\omega_{\max}(t - \hat{\tau}(n))}.$$

Note that, as far as computer simulation is concerned, the above discretization scheme will generally amount to a rounding off of the values of $\hat{\tau}_x(n)$ up to some decimal places. Finally, it is important to realise that several variations of the sampling sequence $\{\tau_x(n)\}_{n \in \mathbb{N}}$, can be employed. More specifically, $\beta(n)$ can be replaced by $\beta(n) = \frac{1}{f(\alpha(n)) + 1}$ where $f(\alpha(n))$ is a positive function in $\alpha(n)$.

4 Conclusions

In this paper, an extension of the Shannon interpolation formula to the case of nonuniform sampling for one-dimensional signals has been given. It has been shown that such extension is possible for particular types of nonuniform sampling sequences satisfying some specific properties. A special sampling sequence which depends on the rate of change of signal with time has also been proposed. This has the benefit of obtaining more sample points when the signal is fast varying and less sample points when the signal is slow varying. It is shown via simulation that the proposed nonuniform sampling process performs better than the uniform one in the sense that more valuable information are captured via the nonuniform sampling algorithm and a better representation of the signal is provided.

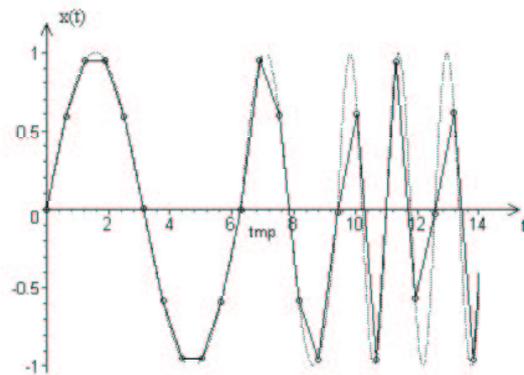


Figure 2: Uniform sampling of a nonstationary signal

References

- [1] Ferreira, P. J., Nonuniform sampling of nonbandlimited signals, *IEEE Signal Processing Letters*, vol. 2, no.5, 1995, pp. 89-91
- [2] Jerri, A. J., The Shannon sampling theorem-Its various extensions and applications: A tutorial review, *Proceedings of the IEEE*, vol. 65, no. 11, Nov. 1977.
- [3] Ortigueira, M.D., What about the FT of the comb signal?, *Proceedings of SAMPTA-97*, Aveiro, Portugal, 1997, pp28-33.
- [4] Howell, K. B., *Principles of Fourier Analysis*, Studies in Advanced Mathematics, Chapman and Hall/CRC, 2001.
- [5] Browder A., *Mathematical Analysis - An introduction*, Springer-Verlag, New York, 1996.
- [6] Busawon, K., R. Binns and B. Livingstone, *A nonuniform sampling sequence*, Internal report, no. 1203, 2003.
- [7] Frazier, M. W., *An introduction to wavelets through linear algebra*, Springer-Verlag, New York, 1999.

PARAMETER ESTIMATION & DIGITAL PEAK LOCALIZATION ALGORITHMS FOR GAMMA RAY SPECTROSCOPY

M.W.Raad¹ J.M..Noras² and M. Deriche³

¹Computer Engineering Department,
King Fahd University of Petroleum and Minerals,
Dhahran, Saudi Arabia,
mwraad@ccse.kfupm.edu.sa

²Electrical Engineering Department,
University of Bradford, Bradford, UK,
J.M.bradford@ac.uk

³Electrical Engineering Department
King Fahd University of Petroleum and Minerals,
Dhahran, Saudi Arabia,
mderiche@kfupm.edu.sa

ABSTRACT

In this paper, we discuss the problem of pulse pile up in Gamma-Ray spectroscopy. The aim is to estimate the different parameters of the basic pulse. A number of parameter estimation techniques for online peak localization are discussed. In particular, 3 and 4 point deconvolution algorithms and polynomial fitting of the pulse model are implemented. The set-up was tested with random signals from a ¹³⁷CS test source. Gamma pulses from a ³Na(Tl) scintillation detector were captured as single and double pulses for testing purposes. The *deconvolution* using a 3 and 4 coefficient filters resulted in a resolution of pile-up to an average percentage of 93% *pileup* free. A polynomial of degree 8 and higher has been found to fit the Gamma Ray pulse accurately.

1. PULSE PILE UP IN GAMMA-RAY SPECTROSCOPY

A common problem in nuclear spectroscopy is pulse pileup caused by the non-zero response time of the detection system. For germanium detectors, the time required to collect all the ionization current associated with an event ranges from 0.5 to 6.0 μ s [1]. The fact that pulses from a radiation detector are randomly spaced in time can lead to interfering effects between pulses when counting rates are not low. These effects are generally called pileup and can be minimized by making the total width of the pulses as small as possible.[2]. Pileup phenomena are of two types. The first type is known as tail pileup and involves the superposition of pulses on the long duration tail from a preceding pulse (see Fig 1). Tails can persist for relatively long periods

of time so that tail pileup can be significant even at relatively low counting rates. A second type of pileup is the peak pileup, which occurs when two pulses are sufficiently close together so that they are treated as a single pulse by the analysis system [6]. From a practical point of view, we have found this pulse spacing to be best taken as approximately equal to the peaking time of the Gamma event. The peak pileup available in real Gamma pulses amounts to approximately 16% of the Gamma pulses captured. Researchers often simply reject pileup, when it is recognized [3]. In our case, we do not wish to lose the information associated with such events, so we use a 3-point deconvolution to separate the pulses.

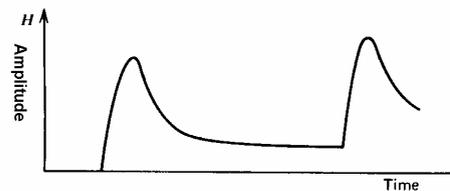


Figure 1. Pileup effect on a pulse peak from the tail of a preceding pulse

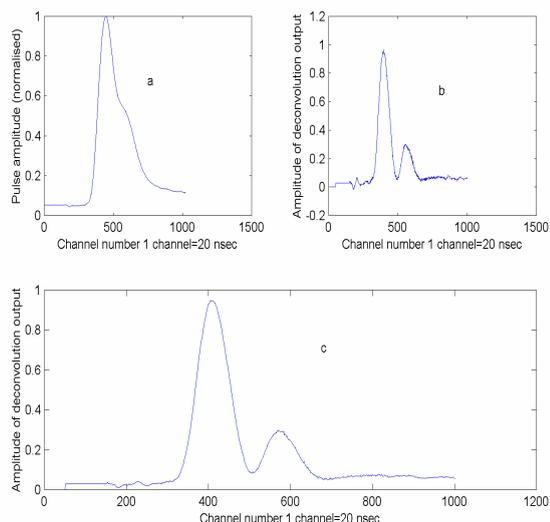


Figure 2. a) Shows a real pileup of two gamma ray events captured at ERL. B) Result of 3-point deconvolution. C) Result of 4 point deconvolution.

2. Peak detection using deconvolution

In many applications we are given an output signal from a system whose characteristics are known and require to determine the input signal. In nuclear spectrometry for instance, the semi-Gaussian shaping

amplifier acts as a convolving filter on the signal coming from preamplifier and causes an elongated tail

on the shaped signal, which makes it prone to pile-up. In such a case the problem is to design a corrective system which, when cascaded with the shaping system, produces an output that, in some sense, removes the effect of the shaper. The corrective system is called an inverse system and the operation is called deconvolution [2].

The deconvolution algorithm uses knowledge of the pulse shape to help locate and separate overlapping pulses. The pulse shape or the impulse response of the system has the following formula:

$$v(t) = t \cdot \exp(-t/\tau) \quad (1)$$

where τ is the shaping time constant. We have confined our system to a CR-RC to make the transfer function as simple as possible. It has been found that the inverse system of this preamplifier CR-RC shaper has the following three weights [4]

$$w1 = (1/x) \cdot \exp(x-1) \quad (2a)$$

$$w2 = (-2/x) \cdot \exp(-1) \quad (2b)$$

$$w3 = (1/x) \cdot \exp(-x-1) \quad (2c)$$

where $x = \Delta t/\tau$ and Δt is the sampling interval. The deconvoluted output has the form :

$$S_k = w1 \cdot v_k + w2 \cdot v_{k-1} + w3 \cdot v_{k-2} \quad (3)$$

This means that a filter performing this operation can be constructed by forming the weighted sum of three consecutive voltage samples in time. We have verified this by computing the inverse matrix of the impulse response matrix, which showed the weights are non-zero. A robust algorithm which maximises signal at the expense of some loss in time resolution is to make a sum of two deconvoluted samples. This is equivalent to using an algorithm with 4 weights which are easily calculated from the three original weights.

The 3-point deconvolution technique applied to tail pile-up events digitized at the ERL facility, succeeded in resolving pile-up by 93% in average, which means that 93% of the events after deconvolution were pileup free. However, deconvolution was not able to recover peak pileup, which is better rejected. Gamma Ray records captured at the ERL, showed a total pulse width of approximately $16\mu s$, or $16t$, where t is the pulse shape time constant. This means that the next event should not arrive within $16t$, otherwise a signal pileup will occur. Hence for this particular case the maximum count rate or the maximum throughput of any pileup classifier is $1/(16t)$. For Na(Tl) detector this translates to a maximum throughput $.06 \text{ Mcps}$ [5].

We have used synthesised simulated data based on the impulse response of the CR-RC amplifier/shaper explained previously, and added 500,000 samples of white gaussian noise of variance of $3.55e-006$. Applying the 3-point deconvolution on single Gamma pulses has resulted in the reduction of pulse width from $16t$ to approximately $4t$. This amounts to enhancing the maximum throughput of a pulse analyzer by almost 4 fold.. See figure 2 for the result of applying 3-point and 4-point deconvolution on real pile-up. The sharp fluctuations in the background shown in the deconvoluted pulse do not degrade the performance because they are outside the pulse window under study. See figure 3 for result of 3-point deconvolution on a simulated Gamma pulse after applying a 9-point moving average smoother to smooth out the noise added by deconvolution.

3. Polynomial fitting

Researchers have found that polynomial fitting of higher orders was a very attractive solution to the parameter estimation modeling problem since it exhibited linearity and followed their pulse shapes very well especially for slowly varying pulse shapes [7]. We have used simulation to investigate using polynomial filters to fit both the

deconvoluted results and the real Gamma event by using the χ^2 goodness of fit and sum of error square criteria from statistics. Due to the high computational cost of least square fit to the Gamma Ray established model, we found that a polynomial fit of 8th degree approximates the Gamma Ray event and can be applied on the *deconvoluted* events to give a χ^2 of approximately 1.2 only. Figure 4 shows the relation between the degree of polynomial fit and the sum of error square. We observe from figure 4 that the least sum of error square corresponds to degrees 8-10. Fitting to degrees less than 6 is not acceptable at all. This result shows that a polynomial of degree higher than 8 does not only fit well to the Gamma Ray pulse model, but fits extremely well to the deconvoluted events. It is also clear that the sum of error square gives better results in the case of fitting a Gamma ray which is expected since the polynomials do not fit well near the narrow peaks. For polynomial of degree 8, the sum of error square was estimated to be less than 1% in the case of fitting a Gamma Ray.

4. CONCLUSIONS

A number of parameter estimation and digital online peak localisation algorithms are being developed for the purpose of Gamma Ray pulse identification in absence of noise and *pileup*. A 3-point and 4-point *deconvolution* techniques have been successfully tested and compared and found to resolve pile-up up to 93%. The 3-point deconvolution gave a maximum throughput of the pulse analyzer of almost 0.25 MCPS, amounting to an enhancement by 4 fold compared to the Gamma Ray. The pulse shape of the Gamma ray was found to approximate very well to a polynomial of degree 8 and above, leading to a χ^2 goodness of fit of approximately 1.2. Work is under progress to implement the parameter estimation algorithms in real time using DSP processor.

Acknowledgments: Support by the EE department University of Bradford UK and Computer Engineering Department and the Research Institute of King Fahd University of Petroleum and Minerals is gratefully acknowledged.

REFERENCES

[1] R.E.Chrien and R.J.Sutter, Nucl. Instr. and Meth., A249(1986)421-425.
 [2] John G. Proakis and Dimitris G. Manolakis, Digital Signal Processing, Macmillan Publishing Company, New York, 1992.
 [3] R.A.Boie and K.R.Wildnauer, IEEE Trans. Nucl. Sci. NS-24, NO.1, February 1977.
 [4] S.Gadomski, G.Hall, T.Hough, P.Jalocha,

E.Nygaard and P.Weillhammer, Nucl.Instr. and Meth., A320(1992)217-227.

[5] W. H. Wong and H. Li, IEEE Transactions on Nuclear Science, Vol 45, No. 3, June 1998.

[6] Glenn F. Knoll, Radiation Detection and Measurement, John Wiley & Sons, 2000.

[7] Paulo C.P.S.Simoes, Jose C. Martins and Carlos M.B.A. Correia, IEEE Trans. Nucl. Sci. Vol.43, NO.3, June 1996.

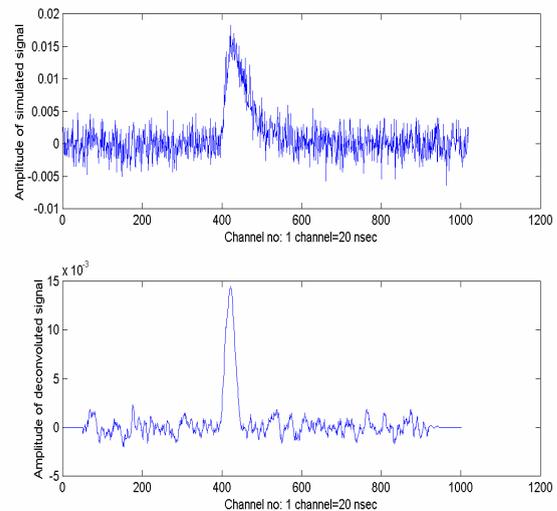


Figure 3. Result of applying 3-point deconvolution on simulated Gamma Ray after filtering with a 9-point moving average filter.

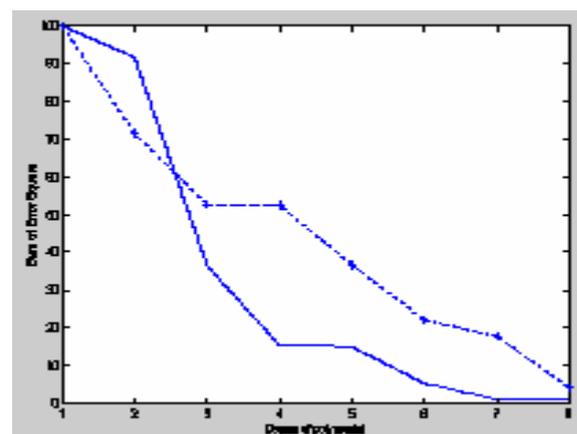


Figure 4. Effect of degree of polynomial fit to 3-point deconvolution output on the sum of error square. The solid line shows fit to a typical Gamma Ray, and dotted line fit to the result of 3point Deconvolution.

Pseudo PCA for Cluster Analysis of Time-Series Data

Daewon An, and K. Wendy Tang
Electrical and Computer Engineering
Stony Brook University
Stony Brook, NY 11794-2350

Abstract

In this paper, we proposed the Pseudo Principal Component of a single dimensional data as a similarity measure in Cluster Analysis or Data Mining, although PCA requires multi dimensional data set and a principal component of a single dimensional data cannot be computed directly. In order to get the principal component for the single dimensional data, we took advantage of re-sampling with time-delay embedding. We applied this method to the cluster analysis of the time-series data, and the results showed that the method performs a better clustering accuracy than the common Euclidian distance method.

Keywords: Cluster Analysis, time delay embedding, time series data; PCA; eigenvector; eigenvalue;

1. Introduction

Cluster Analysis often known as Data Mining can be considered as an unsupervised non-parametric statistical pattern recognition [1], i.e., a study on how to organize given data into groups of meaningful and similar items without human feedbacks.

Principal Components Analysis (PCA) is a well known statistical method to reduce the dimension of a data set or to find features in the data of high dimension. PCA is to analyze the correlations or covariances between multiple variables, and hence requires the data to be two or higher dimensional. However, time-series data are not always multi dimensional and PCA is not applicable for time-series data of a single dimension.

In the following sections, we will shortly review the Cluster Analysis, the time-series data particularly focused on the single dimensionality, and Principal Component Analysis (PCA). Then we describe the proposed method to compute a pseudo principal component from the data of a single dimension by re-sampling with time-delay embedding. We will show how this method works for a practical time-series data set (the Synthetic Control Chart data [6], [7]).

2. Cluster Analysis

Cluster Analysis often known as Data Mining is a study on how to organize given data into meaningful and similar groups without human feedbacks, and is one of the most difficult problems in classification system design [1]. In general, the number of groups or clusters may not even be specified, while exceptionally K-Means algorithm specifies the number of clusters. The key role in Cluster Analysis is how to measure the similarity or dissimilarity between data. The most common measure of the similarity is the Euclidian distance. The correlation coefficients and some other coefficients may be used to measure the similarity

or dissimilarity [4]. In this paper, we propose the principal component as the measure of the similarity.

Once we decide how to measure the similarity we can group together the data depending upon the measured similarity. There are many strategies about how to group the similar data, and a Monothetic Divisive Hierarchical (MDH) method is one among others [8]. We took the MDH as our strategy for our work described in this paper. We start with the given data as one big cluster and divide it into two (binary) smaller groups at each level of the cluster tree as illustrated in Fig. 1.

3. Time-series data - Synthetic Control Chart data

Time-series or time-sequence data are a collection of the sequential measurements of some physical system over a certain period of time [5]. More examples and study can be found in [5].

Let's consider an arbitrary variable, say, $x(t)$ generated from a certain machine or processor. We may measure the values of $x(t)$ over time $[t_1, t_2]$ using tools like a seismograph. We can then get a time sequence of N samples, using the Analog to Digital Converter (ADC) technique. The details about how to get these discrete samples from a continuous signal can be found in many textbooks like [11] and [12].

For our practical experiment, we generated the Synthetic Control Chart data of length 60, according to the equations given in [2]. The data sets consist of 6 different predetermined pattern classes, namely A, B, C, D, E, and F. They are described as normal, cyclic, increasing trend, decreasing trend, upward shift and downward shift. The characteristics of each class can be more understandable if it is plotted and Fig. 2 shows the examples of the data of those 6 classes.

4. Principal Component Analysis

PCA is a statistical tool to identify the variability of the multi-dimensional data, and PCA analyzes the correlations between variables. More details and advanced materials about PCA can be found in [3] and [9].

PCA starts with a collection of multi-dimensional data. We then compute the covariance matrix from the data, and compute the eigenvalues and corresponding unit eigenvectors of the covariance matrix. Schur-Factoring with QR algorithm and similarity transformation (e.g. Householder transformation) can be used iteratively to compute the eigenvalues and corresponding unit eigenvectors of the covariance matrix. As an eigenvector of a larger eigenvalue is more related to the data and contains more information about the data, the largest eigenvalue called the Principal Component of the data is important for feature extraction and pattern recognition. Also, as eigenvectors of small values contains less information about the data, we can ignore some of those with a certain threshold. The reduction of the dimension (or compression) of a data set is accomplished by transforming the data into a lower dimensional space without the components of less important eigenvectors. We may lose the information contained in the eigenvectors ignored, but we can tolerate if the amount of loss is little. This is how data compression is achieved by PCA.

5. Pseudo Principal Component Analysis

PCA requires multi-dimensional square covariance matrix, and we cannot directly apply PCA to the data of a single dimension. In order to apply PCA, we take advantage of re-sampling from the original data with time-delay embedding [10] and setting up a new data set of multi dimension. In this section, we will discuss about the re-sampling with time-delay embedding and how to extract the principal components from the new data set.

We, first, re-sample from the original data with a sampling filter kernel, and then set up a new set of data of multi dimension with a certain time delay. We used a non-causal MA (moving average) filter as a sampling kernel, however theoretically there is no limitation for the kernel. Let's consider the single dimensional data $x[n]$ and its time series or time sequence $\{x[n]\}$ with $n=1,2,3,\dots,N$ as shown in Fig. 3. We set up data $\{d_1[n]\}$ and $\{d_2[n]\}$ of the same length N using MA of kernel size K as shown in Fig. 3.

Now we get 2-dimensional data $\{d_1[n]\}$ and $\{d_2[n]\}$ with the same length N from the single dimensional data $\{x[n]\}$. We compute covariance

matrix $CM^{2 \times 2}$ of the new data of 2^{nd} order as follows;

$$CM^{2 \times 2} = \begin{bmatrix} \text{cov}(d_1, d_1) & \text{cov}(d_1, d_2) \\ \text{cov}(d_2, d_1) & \text{cov}(d_2, d_2) \end{bmatrix}$$

$$\text{cov}(d_i, d_j) = \frac{1}{N-K} \sum_{m=1}^{m=N-K} (d_i[m] - \bar{d}_i)(d_j[m] - \bar{d}_j)$$

$$\bar{d}_i = \frac{1}{N-K} \sum_{m=1}^{m=N-K} d_i[m]$$

From $CM^{2 \times 2}$, we compute the principal component (PC) of the new data and this PC represents the principal component of the original times-series data $\{x[n]\}$. Fig. 4 conceptually depicts each step to compute the principal components of a single dimensional data set $\{x[n]\}_i$ for $i=1,2,3,\dots,M$ with re-sampling and time-delay embedding. We illustrated how to construct a new 2-dimensional data set in Fig. 4. However, we may construct a new data set of more than 2 dimensions, if necessary.

6. Monothetic Clustering with PCA

We start with not knowing the number of clusters nor the number of members of each of the cluster. The accuracy of the clustering can only be verified by comparing the clusters with actual classes of the data sets, only if the class information is known.

See the distribution of the pseudo principal components of the Synthetic Control Chart data shown in Fig. 5-(a) and (b). The vertical axis is the pseudo principal components and the horizontal axis is the number of data. The values from the same class are marked with same symbol and same color. As we can see in Fig. 5-(a), the pseudo principal components from the same class are localized around a certain area and look like a stripe. Some stripes are overlapping each other and, in particular, the Class A is buried in the Class B. Fortunately, the localization of each Class varies, as we change the kernel size K . With the kernel size $K=1$, we can separate the Class A from the others, although the rests are all overlapped, as we can see in Fig. 5-(b).

Now, let's discuss how the principal component of each data can be used for clustering of data. In Fig. 1, each level of hierarchical tree corresponds to each pair of kernel size K . We divide the cluster into two groups at each level; one whose principal component is lower than the middle point of the minimum and maximum of all principal components of the cluster, and the other whose principal component is larger. We can summarize as:

- 1) Start with the given data as a one big cluster.

- 2) Set up the new data set by re-sampling with kernel size K and time-delay embedding.
- 3) Compute the principal components from the new data set.
- 4) Compute middle point of the principal components for each of the clusters.
- 5) Divide each of the clusters into two groups as new clusters; one whose principal component is lower than the middle point and the other otherwise.
- 6) Stop if one of the new clusters contains smaller number of data than a threshold value μ , otherwise repeat from 2), if necessary with a new kernel size K .

We verified the results by comparing the clusters with actual classes of the data sets. The performance was measured in % by

$$P = \frac{100(C - E)}{M},$$

where C is the number of items clustered together correctly which are actually in the same class, E is the number of items clustered together incorrectly which are not in the same class and M is the total number of items in the given data.

7. Experimental Results

We generated a total of $M = 600$ sequences (100 sequences per Class) of the Synthetic Control Chart data with length $N = 60$ as an initial cluster, according the equations given in [2]. The threshold value μ was set as 10% of the initial cluster, i.e., the clustering processes stops at the level where one of the new clusters contains less items than $\mu = 60$. We used 3 different pairs of sampling kernel size as $K_1 = 1, K_2 = 60$ and $K_3 = 58$ in order, and K_3 was for the last level where one of the new clusters contains less items than the threshold μ .

It turned out that the new clustering method is superior to the common Euclidian distance method. The new method grouped 79.83% of items correctly. This performance is much better than the Euclidian distance method. Table 1 shows the performance comparison between the new method and the Euclidian distance method, where CO is co-occurrence feature based, SO is second-order feature based from [2] and PPC is pseudo principal component based.

Table 1 Performance comparison

Method	Euclidian distance		PCA
Feature	CO	SO	PPC
Correct	45.2 %	47.3 %	79.83 %

(The results for Euclidian distance are from [2].)

8. Conclusion

We proposed the Pseudo Principal Component of a single dimensional data as a similarity measure in a

Cluster Analysis or Data Mining, despite Principal Component Analysis (PCA) requires multi dimensional data set. We illustrated how PCA can be performed for a single dimensional data by re-sampling with time-delay embedding. Experimenting with the Synthetic Control Chart data, we showed that the new clustering method is superior to the common Euclidian distance method.

9. References

- [1] A.K. Jain, P.W. Duin and J. Mao, "Statistical Pattern Recognition: A Review", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, January 2000.
- [2] R. J. Alcock and Y. Manolopoulos, "Time-Series Similarity Queries Employing a Feature Based Approach", in *7th Hellenic Conference on Informatics*, Ioannina, Greece, August 27-29. (1999)
- [3] Michael E. Tippong and C. M. Bishop, "Probabilistic Principal Component Analysis", *Journal of the Royal Statistical Society*, Series B, 61, Part 3, pp. 611-622. (1999)
- [4] Qin He, *A Review of Clustering Algorithms as Applied in IR*, Master's Thesis, University of Illinois at U-C (1999)
- [5] A. S. Weigend and N. A. Gershenfeld et al., *Time Series Prediction: Forecasting the future and understanding the past. Proc. of the NATO advanced research workshop*, Santa Fe, New Mexico, May 14-17. (1998)
- [6] D. T. Pham and A. B. Chan, Control Chart Pattern Recognition Using a New Type of Self-Organising Neural Network. *Proceedings of Institution of Mechanical Engineers*, Vol. 212, No. 1, pp. 115 – 127. (1998)
- [7] D.T. Pham and E. Oztemel, "Control chart pattern recognition using combinations of multi-layer perceptrons and learning-vector-quantization neural networks", *Proceedings of Institution of Mechanical Engineers*, vol. 207, Part I, pp 113-118. (1993)
- [8] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York. (1990)
- [9] R. W. Preisendorfer, *Developments in Atmospheric Science 17: Principal Components Analysis in Meteorology and Oceanography*, Elsevier Science Publishers B. V. (1988)
- [10] F. Takens, "Detecting strange attractors in turbulence," *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, Vol. 898, Springer-Verlag, Berlin, pp. 366-381. (1981)
- [11] C. T. Chen, *Analog & Digital Control System Design: Transfer-function, State-space, & Algebraic Methods*, Saunders College Publishing. (1993)

[12] Alen V. Oppenheim and Ronald W. Schaffer, *Discrete-Time Signal Processing*, Englewood-Cliffs, New Jersey, Prentice-Hall. (1989)

Appendix: Figures

Fig. 1 Monothetic Cluster Analysis with pseudo principal components

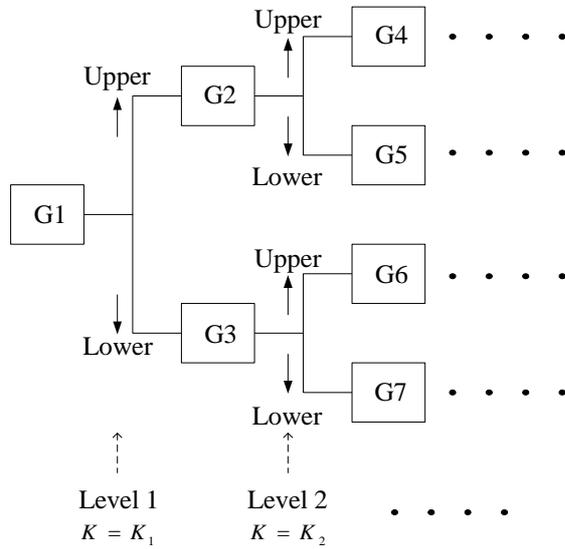


Fig. 2 Examples of the Synthetic Control Chart data of length 60

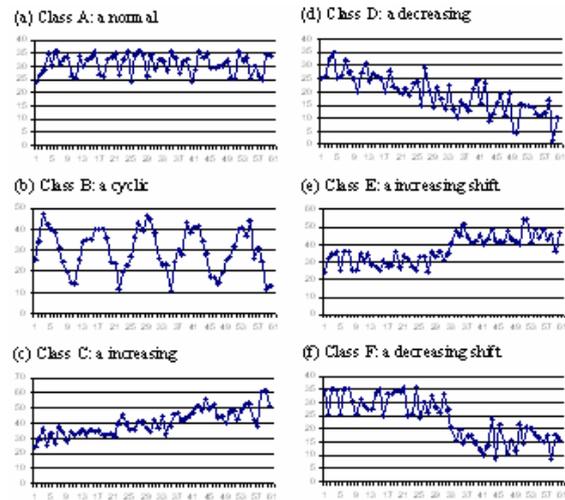


Fig. 3 Re-sampling with time-delay embedding

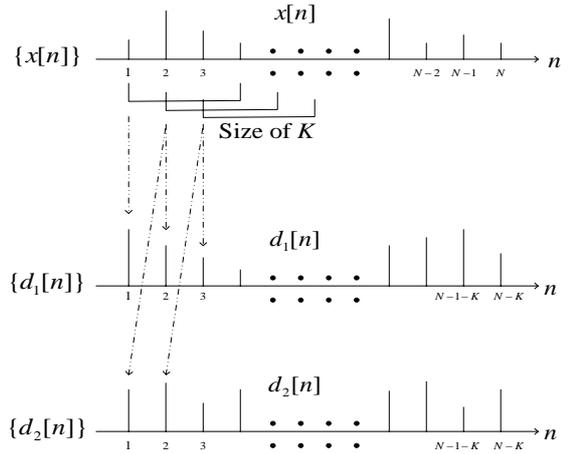


Fig. 4 Steps for computing principal components

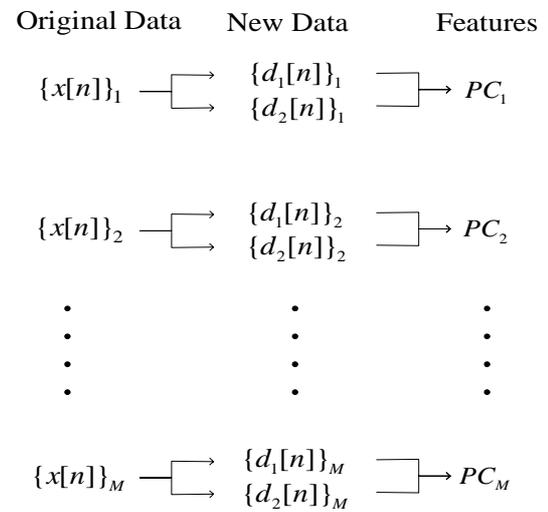
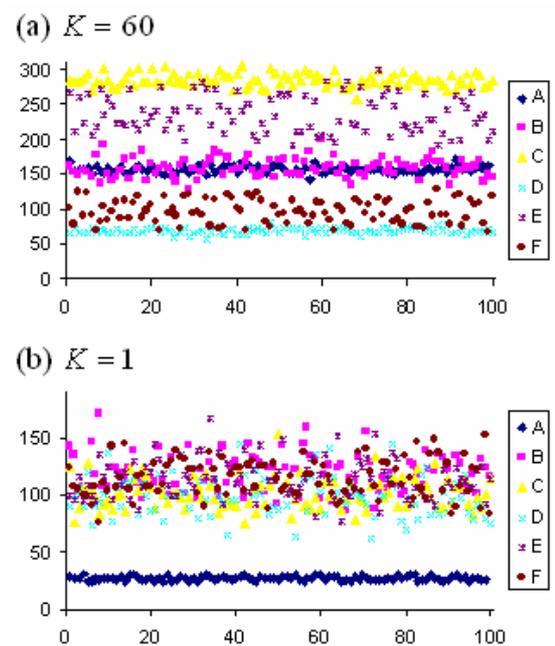


Fig. 5 The distribution of the principal components of the Synthetic Control Chart data.



Sampling-Reconstruction Procedure of Gaussian Processes with a Finite Number of Samples with Jitter

VLADIMIR A. KAZAKOV AND DANIEL RODRIGUEZ S.

*Department of Telecommunications of the School of Mechanical and Electrical Engineering,
National Polytechnical Institute of Mexico, U. Zacatenco, C.P. 07738, Mexico City
vkazakov41@hotmail.com
dannyrसान@hotmail.com*

Abstract: The Sampling-Reconstruction Procedure (SRP) of some Gaussian processes with an arbitrary number of samples is investigated. All or some samples can have the independent jitter. The probability density function of jitter is the Beta-distribution. The method of the investigation is based on the conditional mean rule. The types of the covariance functions and the types of the Beta distributions are changed. The optimal error reconstruction functions are calculated for some new examples.

1. Introduction

The recent important book [1] is devoted to the consideration of the different aspects of nonuniform Sampling. The problem of the Sampling-Reconstruction Procedure (SRP) of stochastic processes with jitter is one of the very significant cases of nonuniform Sampling.

The problem of SRP with jitter was investigated since 60 years [2-4]. During the last time the interest to this problem is also increased (see for instance papers [5-6]). We notice here some specific features of all mentioned works [1-6]: 1) the probability density functions of the sampled stochastic processes are not determined; 2) the sampled stochastic processes are described by the power spectrum, in the majority cases this spectrum is limited by a boundary frequency; 3) the investigations have carried out in the frequency domain, consequently the main results have been formulated by the spectral characteristics; 4) the quantity of the samples are equal to infinity; 5) the jitter effect is described by the same features independently of the number of the samples.

In the present paper we used the conditional mean rule in order to describe the SRP of the Gaussian processes with jitter. This approach has been productively applied in the statistical SRP description of different types stochastic processes [7-12]. The application of the conditional mean rule for the SRP of Gaussian processes with jitter allows us to investigate some new aspects in this problem: 1) the number of samples is arbitrary and limited; 2) the parameters of the jitter distribution can be the same or the different for different samples; 3) in

particularly some samples can be jittered and other samples are fixed.

We apply the mathematical derivations exceptionally in the time domain. Owing to this approach we calculate exact curves for the minimum error reconstruction functions as the functions of time. We demonstrate how the type of the covariance function of the sampled processes, the sampling interval and the jitter characteristics influence on the error reconstruction functions.

The results of the present investigation are obtained by the simple mathematical method with comparison the methods [1-6].

2. General Expressions

Let us consider Gaussian processes $x(t)$ with the mathematical expectation $m(t)$, the variance $\mathbf{s}^2(t)$, and the covariance function $K(t_1, t_2)$. We can write the multidimensional probability density function (pdf) for such process. We fix an arbitrary set of samples in this process $X, T = \{x(T_1), x(T_2), \dots, x(T_N)\}$. Then instead of the given process we have the conditional process $\tilde{x}(t)$. One can find its main statistical characteristics [13, 14]:

$$\tilde{m}(t) = m(t) + \sum_{i=1}^N \sum_{j=1}^N K(t, T_i) a_{ij} [x(T_j) - m(T_j)], \quad (1)$$

$$\tilde{\mathbf{s}}^2(t) = \mathbf{s}^2(t) - \sum_{i=1}^N \sum_{j=1}^N K(t, T_i) a_{ij} K(T_j, t), \quad (2)$$

where $\tilde{m}(t) = \tilde{m}(t | X, T)$ is the conditional mean and $\tilde{\mathbf{s}}^2(t) = \tilde{\mathbf{s}}^2(t | X, T)$ is the conditional variance, and a_{ij} - is the element of the inverse covariance function.

Following this rule we choose the reconstruction function $\hat{x}(t)$ as the conditional mean function (1), i.e. $\hat{x}(t) = \tilde{m}(t)$, and the quality of the reconstruction procedure is estimated by the conditional variance (2), i.e. $\mathcal{Q}^2(t) = \tilde{\mathbf{s}}^2(t)$.

The expressions (1)-(2) and their multidimensional generalizations were used in the statistical SRP description of many types of Gaussian

processes and their transformations [7-12] when the set of samples X, T is fixed.

Now we consider the situation when the instant time T_i of any sample can have a jitter. In the present paper we consider the case of the independent jitter of samples times, i.e. $\tilde{T}_i = T_i + \mathbf{e}_i$, where \mathbf{e}_i is the random variable with known pdf $w(\mathbf{e}_i)$. There must be the restrictions on the type of $w(\mathbf{e}_i)$ determined in the interval (a_i, b_i) ($b_i > a_i$):

$$(b_{i-1} - T_{i-1}) + (T_i - a_i) \leq \Delta T_i. \quad (3)$$

Below we pay our attention on the error reconstruction function $\tilde{\mathbf{s}}^2(t)$. In (1) and (2) one can see three fixed time arguments: (t, T_i) , (T_i, T_j) , and (T_j, t) . In the case of jitter we need to change them: (t, \tilde{T}_i) , $(\tilde{T}_i, \tilde{T}_j)$ and (\tilde{T}_j, t) . In the stationary given process we need to write the differences of the corresponding arguments, i.e. the duration of the random intervals: $[\tilde{T}_i - t]$, $[\tilde{T}_i - \tilde{T}_j]$ and $[\tilde{T}_j - t]$.

The first and the third intervals depend on only one random variables \mathbf{e}_k (for $k=i, j$), because $\tilde{T}_k = T_k + \mathbf{e}_k$ and T_k is the constant value. Hence pdf of the random variables \tilde{T}_k can be determined on the base pdf $w(\mathbf{e}_k)$ as:

$$W(\tilde{T}_k) = w(\tilde{T}_k - T_k), \quad k=1, \dots, N. \quad (4)$$

Now let us consider pdf of the interval between any two samples:

$$\Delta \tilde{T}_{j-i} = \tilde{T}_j - \tilde{T}_i = (T_j - T_i) + (\mathbf{e}_j - \mathbf{e}_i) = \Delta T_{j-i} + \mathbf{h}_{j-i}, \quad (5)$$

here $\Delta T_{j-i} = T_j - T_i$ is the deterministic value and $\mathbf{h}_{j-i} = \mathbf{e}_j - \mathbf{e}_i$ is the random variable.

Putting $w(\mathbf{e}_j) = p(\mathbf{e}_j) = p(\mathbf{h}_{j-i} + \mathbf{e}_i)$ we can find the pdf $q(\mathbf{h}_{j-i})$ of the difference between two random variables (see, for instance, [14], [15]):

$$q(\mathbf{h}_{j-i}) = \int w(\mathbf{e}_i) p(\mathbf{h}_{j-i} + \mathbf{e}_i) d\mathbf{e}_i. \quad (6)$$

After that we find pdf of the sum (5)

$$W(\Delta \tilde{T}_{j-i}) = q(\Delta \tilde{T}_{j-i} - \Delta T_{j-i}). \quad (7)$$

Knowing the distributions $W(\tilde{T}_k)$ ($k=1, \dots, N$) and $W(\Delta \tilde{T}_{j-i})$ ($i, j=1, \dots, N$) one can obtain the general average expression of the error reconstruction function $\langle \tilde{\mathbf{s}}^2(t) \rangle$:

$$\begin{aligned} \langle \tilde{\mathbf{s}}^2(t) \rangle = & \int \dots \int \mathbf{s}^2(t) W(\tilde{T}_1) \dots W(\tilde{T}_N) w(\Delta \tilde{T}_{1-1}) \\ & \times w(\Delta \tilde{T}_{1-2}) \dots w(\Delta \tilde{T}_{1-N}) w(\Delta \tilde{T}_{2-1}) \dots w(\Delta \tilde{T}_{N-N}) \\ & \times d\tilde{T}_1 \dots d\tilde{T}_N d(\Delta \tilde{T}_{1-1}) \dots d(\Delta \tilde{T}_{N-N}) \end{aligned} \quad (8)$$

where $\tilde{\mathbf{s}}^2(t)$ is determined by (2).

Below the complex general expression (8) will be concretized in some examples.

3. Models of the Jitter and the Sampled Processes

In this paper, we use the Beta-distribution [16] in order to describe the jitter:

$$w(\mathbf{e}_k) = \frac{(\mathbf{e}_k - a_k)^{\gamma-1} (b_k - \mathbf{e}_k)^{\beta-1}}{B(\gamma, \beta) (b_k - a_k)^{\gamma+\beta-1}}, \quad a_k \leq \mathbf{e}_k \leq b_k \quad (9)$$

where γ, β are the parameters of the Beta distribution.

Following (9) the jitter distribution can be represented by different ways depending on the values of the parameters γ and β . For example, if $\gamma=\beta=1$ the form of the jitter distribution is uniform; if $\gamma=\beta>1$ the density of the jitter is of type Gaussian truncated; if $\gamma=\beta<1$ the jitter has a distribution in form of U; if $\gamma \neq \beta$ and both are bigger than the unit, we have an asymmetric distribution; and if $\gamma=1$ and $\beta=2$, specifically, we have a distribution in the ramp form.

We emphasize that any sample may have its own type of the distribution among all types of Beta-distribution.

Now we have to choose some Gaussian models of the stochastic processes. We restrict our consideration by some stationary Gaussian processes with low band spectrums. It is rather convenient to use some processes on the output of some linear filters driven by white noise. Namely, on the output of the integrated RC circuit we have the Gaussian Markov process with the normalized covariance function in the stationary regime:

$$R(\mathbf{t}) = \frac{K(\mathbf{t})}{\mathbf{s}^2} = \exp[-\mathbf{a} | \mathbf{t} |] \quad (10)$$

where \mathbf{a} is the parameter of RC circuit.

The output process on the two series integrated RC circuits with the equal parameters has the next normalized covariance function:

$$R(\mathbf{t}) = (1 + \mathbf{a} | \mathbf{t} |) \exp[-\mathbf{a} | \mathbf{t} |], \quad (11)$$

and the output process on the three series integrated RC circuits has the next normalized covariance function:

$$R(\mathbf{t}) = (1 + \mathbf{a} | \mathbf{t} | + \mathbf{a}^2 \mathbf{t}^2 / 3) \exp[-\mathbf{a} | \mathbf{t} |]. \quad (12)$$

We can carry out the correct comparisons of the obtained results if all processes have the same covariance time \mathbf{t}_c . In order to do this we have to multiply the parameters \mathbf{a} in (11) and (12) by the coefficients 2 and 8/3 correspondingly.

4. Examples

A. The Case of Two Samples

This case is the simplest among other interpolation procedures. The expressions (5)-(8) will be naturally simplified:

$$\Delta\tilde{T}_i = \tilde{T}_{i+1} - \tilde{T}_i = (T_{i+1} - T_i) + (\mathbf{e}_{i+1} - \mathbf{e}_i) = \Delta T_i + \mathbf{h}_i, \quad (13)$$

$$q(\mathbf{h}_i) = \int w(\mathbf{e}_i) p(\mathbf{h}_i + \mathbf{e}_i) d\mathbf{e}_i, \quad (14)$$

$$W(\Delta\tilde{T}_i) = q(\Delta\tilde{T}_i - \Delta T_i), \quad (15)$$

$$\langle \tilde{\mathbf{s}}^2(t) \rangle = \int \dots \int \tilde{\mathbf{s}}^2(t) w(\tilde{T}_i) w(\tilde{T}_{i+1}) \times w(\Delta\tilde{T}_i) d\tilde{T}_i d\tilde{T}_{i+1} d(\Delta\tilde{T}_i). \quad (16)$$

The inverse matrix in the stationary regime is determined by the formula:

$$A = \begin{bmatrix} a_{i,i} & a_{i,i+1} \\ a_{i+1,i} & a_{i+1,i+1} \end{bmatrix} = \begin{bmatrix} K(0) & K(\Delta\tilde{T}_i) \\ K(\Delta\tilde{T}_i) & K(0) \end{bmatrix}^{-1} = \begin{bmatrix} \frac{K(0)}{K^2(0) - K^2(\Delta\tilde{T}_i)} & \frac{-K(\Delta\tilde{T}_i)}{K^2(0) - K^2(\Delta\tilde{T}_i)} \\ \frac{K(\Delta\tilde{T}_i)}{K^2(0) - K^2(\Delta\tilde{T}_i)} & \frac{K(0)}{K^2(0) - K^2(\Delta\tilde{T}_i)} \end{bmatrix}. \quad (17)$$

We consider a Markov process with two samples. The normalized covariance function of this processes is given by the expression (10) and this process have a covariance time $\mathbf{t}_c = 1/\mathbf{a}$.

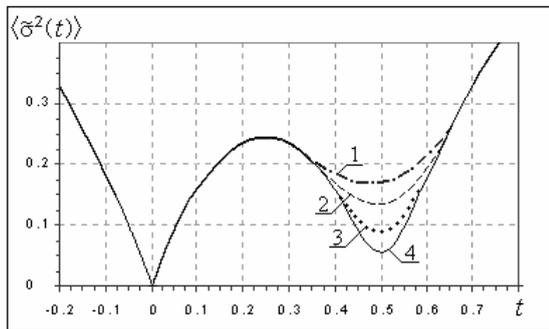


Fig. 1. The error reconstruction functions of the Markov process for four different cases.

In this case, the first sample is fixed in the instant $T_1=0$ and the second sample $T_2=0.5$ has jitter and it is represented by four different symmetrical Beta distributions: 1) the distribution is uniform with the parameters: $\mathbf{g}=\mathbf{b}=1$; 2) the distribution is the semicircle with $\mathbf{g}=\mathbf{b}=2$; 3) the

distribution is the truncated Gaussian when $\mathbf{g}=\mathbf{b}=5$; 4) the distribution is the truncated Gaussian with small \mathbf{s}^2 determined by $\mathbf{g}=\mathbf{b}=15$. One can mention that one can represent the jitter distribution with an asymmetric Beta distribution too.

All these distributions are determined in the interval (0.3, 0.7). The results of the average error reconstruction functions $\langle \tilde{\mathbf{s}}^2(t) \rangle$ are shown in the Fig. 1 for each case.

In Fig. 1 we can see the next features. When a sample does not have the jitter, the error reconstruction function is zero, as it is observed in $T_1=0$. When a sample has jitter, the error reconstruction function is different to zero and this error depend on the type of pdf and its variance especially. For instance, the values of the variance are decreased from the first case to the fourth case.

B. The Case of Four Samples

It is possible to consider any arbitrary number of samples to describe the SRP of the random processes with jitter using the general expressions (5)-(8). As one example, we consider four samples.

These samples can or can not have jitter, the jitter distribution $w(\mathbf{e}_k)$ of each sample can be the same or different and these jitter distributions can have symmetrical or asymmetrical types. For example, let us consider the following cases:

1) The first sample has a symmetrical Beta distribution with parameters $\mathbf{g}=\mathbf{b}=2$. The interval of this distribution is $(-0.1, 0.1)$, i. e. $b_i - a_i = 0.2$.

2) The second sample has the same characteristics like the first sample. The average sampling interval among these samples is $\Delta T_{2-1} = T_2 - T_1 = 0.5$. The interval of the Beta distribution of this second sample is $(0.4, 0.6)$.

3) The third sample has an asymmetric Beta distribution with parameters $\mathbf{g}=1, \mathbf{b}=2$. Therefore, the average sampling interval between the second and third samples is smaller than 0.5. The interval of the jitter distribution of this sample is twice wider $(0.8, 1.2)$.

4) The fourth sample is free of jitter. The last two samples have the average sampling interval bigger than 0.5.

The calculation results of the error reconstruction functions are shown in the Fig. 2, when $T_1=0, T_2=0.5, T_3=1, T_4=1.5, \Delta T_{j-i}=0.5$.

In Fig. 2 we can see the error reconstruction functions for three different types of processes: 1) Markov process; 2) at the output of two series integrated RC circuits and 3) at the output of three series integrated RC circuits. The normalized covariance functions of these three types of processes are determined by the formulas (10)-(12).

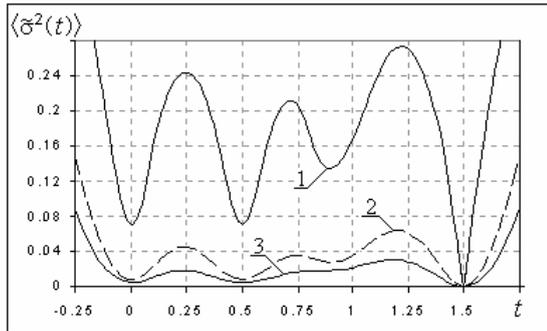


Fig. 2. The error reconstruction functions for different types of processes and considering four samples.

The curve (3) is the error reconstruction function for the third type and it is smaller in comparison with the curves (1) and (2), because the realizations of this process are very smooth.

5. Conclusions

The statistical description of the Sampling-Reconstruction Procedure with the jitter of some Gaussian processes is given on the base of the conditional mean rule. The quantity of samples is arbitrary. Generally the location of samples is arbitrary as well. The Beta distribution is used for the jitter statistical description. Any sample can have its own jitter distribution. Particularly, some samples can be fixed. Some nontrivial examples are considered.

References

- [1] F. Marvasti, Editor. "Nonuniform Sampling: Theory and Practice", Kluwer Academic / Plenum Publishers, New York, p. 924, 2001.
- [2] A. V. Balakrishnan, "On the problem of time jitter in sampling", *IRE Trans. on Information Theory*, vol. IT-8, pp. 226-236, April 1962.
- [3] W. M. Brown, "Sampling with random jitter", *J. SIAM*, vol. 11, pp. 460-473, June 1963.
- [4] B. Liu and T. P. Stanley, "Error bound for jittered sampling", *IEEE Transactions on Automatic Control*, vol. 10, no. 4, October, 1965.
- [5] B. Lacaze, "A note about stationary process random sampling", *Statistics & Probability Letter*, vol. 31, pp. 133-137, 1996.
- [6] B. Lacaze, "Stationary clock changes on stationary processes", *Signal Processing*, Vol. 55, pp. 191-205, July 1996.
- [7] V.A. Kazakov, "Regeneration of samples of random processes following nonlinear inertialess conversions", *Telecommunication and Radioengineering*, vol. 43, pp. 94-96, no. 10, 1988.
- [8] V.A. Kazakov, S.A. Afrikanov, M.A. Beliaev, "Comparison of two algorithms of the realization restorations using random numbers of counts", *Radioelectronics and Communication Systems*, vol. 37, pp. 43-45, no. 4, 1994.
- [9] V.A. Kazakov, M.A. Beliaev, "Reconstruction of realizations of Gaussian processes from readings of the process and linear transformation of it" *Telecommunications and Radioengineering*, vol. 49, pp. 97-102, no. 9, 1995.
- [10] V.A. Kazakov, M.A. Beliaev, "Sample-reconstruction procedure of some nonstationary processes", *Radioelectronics and Communication Systems*, Vol. 40, no. 9, pp. 43-49, 1997.
- [11] V.A. Kazakov, M.A. Belyaev, "Sampling reconstruction procedure for Non-Stationary Gaussian processes based on conditional expectation rule", *Sampling Theory in Signal and Image Processing*, vol. 1, no. 2, pp. 135-153, May 2002.
- [12] V.A. Kazakov, S. Sánchez, "Sampling-reconstruction procedure of random processes at the output of exponential Non-Linear converters", *Electromagnetic Waves and Electronic Systems*, vol. 8, no. 7-8, , pp.77-82, 2003.
- [13] G. Cramer, "Mathematic Methods of Statistic", Princeton, N. J.: Princeton University Press, 1946.
- [14] R. L. Stratonovich, "Topics in the theory of random noise", New York: Gordon and Breach, 1963.
- [15] A. Papoulis, "Probability, Random Variables, and Stochastic Processes", 3er. Edition, McGrawHill, 1991.
- [16] A. Stuart, K. Ord, "Kendall's advanced theory of statistics", 6th. Edition, New York: Arnold, 1994.

SIGNAL SEPARATION OF NONLINEAR MIXTURES USING FUZZY-GENETIC ALGORITHM

L.C. Khor, W.L. Woo and S.S. Dlay

School of Electrical, Electronic and Computer Engineering,

University of Newcastle upon Tyne,

NE1 7RU Newcastle, United Kingdom

Email: l.c.khor@ncl.ac.uk, w.l.woo@ncl.ac.uk and s.s.dlay@ncl.ac.uk

Abstract: This paper proposes a new technique of nonlinear blind signal separation algorithm where benefits of Fuzzy Logic and Genetic Algorithm are incorporated into the natural gradient-based algorithm. Fuzzy Logic provides a simple and computationally efficient method for controlling the learning rate and Genetic Algorithm provides an assurance of global solution which is inherent in nonlinear separation. The effectiveness of the proposed algorithm is compared with linear and nonlinear gradient algorithms in terms of convergence speed and accuracy of separation. We provide simulations to corroborate the superiority of the proposed algorithm.

1. Introduction

Blind Source Separation (BSS) via Independent Component Analysis (ICA) is a technique of recovering source signals based on information obtained from the observed signals. The signals are recovered by finding a set of demixing matrix that maximises the statistical independence of the signals. This technique has gained recognition from various fields like speech recognition, biomedical, telecommunications and astrophysical imaging [1]. Classical linear mixing models have recently been extended to more pragmatic nonlinear models [1-2] in its general form of:

$$\mathbf{x}(t) = f(\mathbf{s}(t)) \quad (1)$$

where $f(\mathbf{s}) = \mathbf{A}_1(g(\mathbf{A}_2 \dots g(\mathbf{A}_L \mathbf{s} + \boldsymbol{\varepsilon})))$ is a smooth nonlinear function for a multilayer perceptron network, $\mathbf{s}(t) = [s_1(t) \dots s_n(t)]^T$ are original source signals, $\mathbf{x}(t) = [x_1(t) \dots x_n(t)]^T$ are the output of the mixing model, $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_L$ are mixing matrices and $\boldsymbol{\varepsilon}$ the bias. Superscript ' T ' denotes the vector transposition.

The statistical independence of the signals is optimised in the form of an information theoretic cost function using a natural gradient approach [3]. Gradient descent optimisation is however fraught with issues pertaining to convergence rates and local minima. It is inevitable that algorithms employing this optimisation procedure may produce less favourable solutions which are not

only local minima but also slow in the process. Issue of speed afflicts both linear and nonlinear models but local minima remain an ever-present conundrum associated with nonlinear models. It is therefore imperative and crucial to reduce if not eliminate the risk of a local solution for nonlinear mixing models. In our proposed algorithm, we propose a blend of techniques to increase convergence rate using Fuzzy Logic and improve upon a global solution via Genetic Algorithm (GA) for gradient descent based algorithms for nonlinear blind signal separation (NLFGA).

2. Proposed Algorithm

A three-layer perceptron network for nonlinearly mixed signal separation was adopted for the incorporation of the proposed algorithm comprising the gradient descent based update of parameters with Fuzzy Logic control of learning rates and Genetic Algorithm search.

2.1. Gradient Descent Based Update

We propose optimising a multilayer Perceptron (MLP) demixer in Fig. 1 based on gradient descent with self-adaptive learning rates and a more global search.

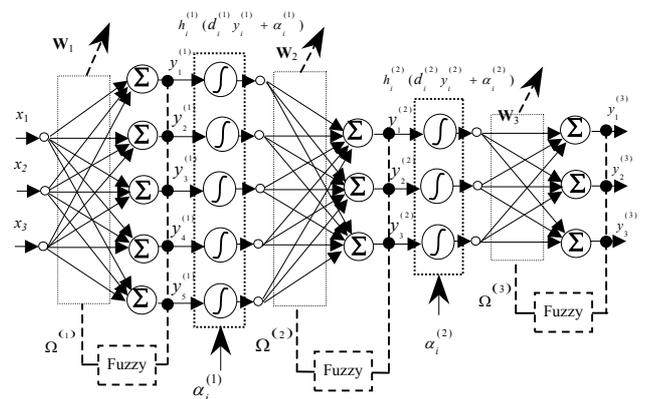


Figure 1: MLP Demixer with Fuzzy Logic and GA

The output of a L -layer perceptron network is $\mathbf{y}^{(l)} = \mathbf{W}^{(l)} h^{(l-1)} (\text{diag}[\mathbf{D}^{(l-1)}] \mathbf{y}^{(l-1)} + \boldsymbol{\alpha}^{(l-1)})$; $l = 1, \dots, L$

where $\mathbf{y}^{(l)}$, $\mathbf{W}^{(l)}$, $h(\cdot)$, $\mathbf{D}^{(l-1)}$ and $\boldsymbol{\alpha}^{(l-1)}$ represent the output of the l^{th} layer, demixing matrix, nonlinear

function, gradient of nonlinear function and bias weights of the $(l-1)^{th}$ layer respectively. Based on Maximum Likelihood [4] the proposed algorithm aims to minimise the cost function:

$$J = -\sum_{i=1}^L \log |\det \mathbf{W}^{(l)}| - \sum_{i=1}^N \log p_i(y_i^{(L)}) - \sum_{l=1}^{L-1} \left\{ \sum_{i=1}^N \log \hat{h}_i^{(l)}(d_i^{(l)} y_i^{(l)} + \alpha_i^{(l)}) + \log |\det \text{diag}[\mathbf{D}^{(l)}]| \right\} \quad (2)$$

where $p_i(y_i^{(L)})$, d and $\hat{h}_i^{(l)}(\cdot)$ are the marginal probability density function of the hypothesized distribution, diagonal element of \mathbf{D} and derivative of nonlinear function of the i^{th} sensor at l^{th} layer respectively. The update equations are obtained by differentiating (2) with respect to $G = \left\{ \left\{ \mathbf{W}^{(l)} \right\}_{l=1}^L, \left\{ \text{diag}[\mathbf{D}^{(l)}] \right\}_{l=1}^L, \left\{ \alpha^{(l)} \right\}_{l=1}^L \right\}$:

$$\frac{dJ}{d\mathbf{M}^{(l)}} = -\mathbf{I} + \text{diag}[\mathbf{D}^{(l)}] \boldsymbol{\tau}^{(l)}(\mathbf{y}^{(l)})^T \quad (3)$$

$$\boldsymbol{\tau}^{(l)} = \begin{cases} \varphi(\mathbf{y}^{(L)}) & ; l = L. \\ \boldsymbol{\psi}^{(l)} + \text{diag}[\dot{\mathbf{h}}^{(l)}] \mathbf{W}^{(l+1)T} \varphi(\mathbf{y}^{(l+1)}) & ; l = 1, \dots, L-1. \end{cases} \quad (4)$$

where

$$\varphi(\mathbf{y}^{(L)}) = \left[\frac{d}{d\hat{y}_1} \log p(y_1^{(L)}) \dots \frac{d}{d\hat{y}_N} \log p(y_N^{(L)}) \right]^T, \quad d\mathbf{M}^{(l)} = d\mathbf{W}^{(l)} \mathbf{W}^{(l)-1},$$

$$\boldsymbol{\alpha}^{(l)} = d\text{diag}[\mathbf{D}^{(l)}] \text{diag}[\mathbf{D}^{(l)}]^{-1}, \quad \boldsymbol{\psi}^{(l)} = \begin{bmatrix} \ddot{h}_1^{(l)} & \dots & \ddot{h}_N^{(l)} \\ \dot{h}_1^{(l)} & & \dot{h}_N^{(l)} \end{bmatrix}, \quad \dot{h}_i^{(l)} \text{ and}$$

$\ddot{h}_i^{(l)}$ represent the first and second order derivative of $\hat{h}_i^{(l)}$ respectively. This leads to the update equation for the demixing matrix:

$$\mathbf{W}^{(l)} \leftarrow \mathbf{W}^{(l)} + \begin{cases} \eta_w^{(l)} \left\{ \mathbf{I} - \varphi(\mathbf{y}^{(l)}) (\mathbf{y}^{(l)})^T \right\} \mathbf{W}^{(l)} & ; \text{for } l = L. \\ \eta_w^{(l)} \left\{ \mathbf{I} - \text{diag}[\mathbf{D}^{(l)}] \boldsymbol{\tau}^{(l)}(\mathbf{y}^{(l)})^T \right\} \mathbf{W}^{(l)} & ; \text{for } l = 1, \dots, L-1. \end{cases} \quad (5)$$

where $\eta_w^{(l)}$ represents the learning rate for $\mathbf{W}^{(l)}$. The following section lays out the formulation of update equations for the parameters $\mathbf{D}^{(l)}$ and $\alpha^{(l)}$.

$$\frac{dJ}{d\zeta^{(l)}} = \boldsymbol{\tau}^{(l)}(\mathbf{y}^{(l)})^T \text{diag}[\mathbf{D}^{(l)}] - \mathbf{I} \quad (6)$$

$$\frac{dJ}{d\alpha^{(l)}} = \boldsymbol{\psi}^{(l)} + \text{diag}[\dot{\mathbf{h}}^{(l)}] (\mathbf{W}^{(l+1)})^T \varphi(\mathbf{y}^{(l+1)}) = \boldsymbol{\tau}^{(l)} \quad (7)$$

$$\text{diag}[\mathbf{D}^{(l)}] \leftarrow \text{diag}[\mathbf{D}^{(l)}] + \eta_D^{(l)} \left\{ \mathbf{I} - \boldsymbol{\tau}^{(l)}(\mathbf{y}^{(l)})^T \text{diag}[\mathbf{D}^{(l)}] \right\} \text{diag}[\mathbf{D}^{(l)}] \quad (8)$$

$$\alpha^{(l)} \leftarrow \alpha^{(l)} - \eta_\alpha^{(l)} \boldsymbol{\tau}^{(l)} \quad (9)$$

where $\eta_D^{(l)}$ and $\eta_\alpha^{(l)}$ represent the learning rate for $\mathbf{D}^{(l)}$ and $\alpha^{(l)}$ respectively.

2.2. Fuzzy Logic

The performance of any algorithm using gradient descent to optimise the cost function is heavily dependent on initialised η . If η is set too large, the algorithm can become unstable and diverge from the solution. Otherwise, slow convergence speed may occur if η is set too small and the initialised values are too far from solution. Fuzzy Systems injects some degree of intelligent control over η . The learning rates

$U^{(l)} = \left\{ \eta_w^{(l)}, \eta_D^{(l)}, \eta_\alpha^{(l)} \right\}$ for $l=1, 2, \dots, L$ designated for

every parameter $G = \left\{ \left\{ \mathbf{W}^{(l)} \right\}_{l=1}^L, \left\{ \text{diag}[\mathbf{D}^{(l)}] \right\}_{l=1}^L, \left\{ \alpha^{(l)} \right\}_{l=1}^L \right\}$ will

be fine-tuned if necessary until the optimum solution is obtained. The system will monitor the progress and determine the amount to adjust $U^{(l)}$ based on the error and change of error $\Upsilon = \{\theta, \lambda\}$ in current and previous iterations. Crisp input values $\Upsilon = \{\theta, \lambda\}$ are defined as:

$$\begin{aligned} \theta_w^{(l)}(k+1) &= \sum_i \sum_j |F_{ij}(\mathbf{W}^{(l)}; k+1)| \\ \theta_D^{(l)}(k+1) &= \sum_i |F_{ii}(\mathbf{D}^{(l)}; k+1)| \\ \theta_\alpha^{(l)}(k+1) &= \sum_i |F_i(\alpha^{(l)}; k+1)| \end{aligned} \quad (10)$$

where

$$\begin{aligned} F(\mathbf{W}^{(l)}; k+1) &= \mathbf{W}^{(l)}(k+1) - \mathbf{W}^{(l)}(k) \\ F(\mathbf{D}^{(l)}; k+1) &= \text{diag}[\mathbf{D}^{(l)}(k+1)] - \text{diag}[\mathbf{D}^{(l)}(k)] \\ F(\alpha^{(l)}; k+1) &= \alpha^{(l)}(k+1) - \alpha^{(l)}(k) \end{aligned} \quad (11)$$

and $F_{ij}(\mathbf{W}^{(l)}; k+1)$ is the $(i, j)^{th}$ element of $F(\mathbf{W}^{(l)}; k+1)$ where k and $k+1$ represent the current and next iteration, respectively. Similarly $F_{ii}(\mathbf{D}^{(l)}; k+1)$ is the diagonal element of $F(\mathbf{D}^{(l)}; k+1)$ and $F_i(\alpha^{(l)}; k+1)$ is the i^{th} element of $F(\alpha^{(l)}; k+1)$. Let $\theta^{(l)} = \{\theta_w^{(l)}, \theta_D^{(l)}, \theta_\alpha^{(l)}\}$ and the change of error $\lambda^{(l)} = \{\lambda_w^{(l)}, \lambda_D^{(l)}, \lambda_\alpha^{(l)}\}$ is then:

$$\lambda^{(l)}(k+1) = |\theta^{(l)}(k+1) - \theta^{(l)}(k)| \quad \text{for } l = 1, 2, \dots, L \quad (12)$$

The degree of membership of (10) and (12) is evaluated based on the membership function in Fig. 2. A fuzzy value is calculated based on evaluation rules (Table 1). This fuzzy value is converted into crisp values of $\Omega^{(l)} = \{\gamma_w^{(l)}, \gamma_D^{(l)}, \gamma_\alpha^{(l)}\}$. The fuzzy output

$$z(\Upsilon^{(l)} = \{\theta^{(l)}, \lambda^{(l)}\}) = \Omega^{(l)} \quad (13)$$

is incorporated into next iteration as :

$$U^{(l)}(k+1) = U^{(l)}(k) - \Omega^{(l)}(k) \quad (14)$$

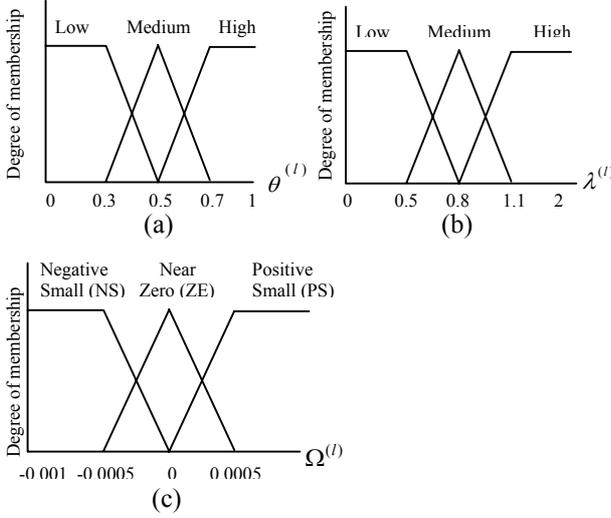


Figure 2: Membership Function of Learning Rate

$\lambda_{(l)}$ \ $\theta_{(l)}$	Low	Medium	High
Low	PS	PS	ZE
Medium	PS	PS	ZE
High	ZE	ZE	NS

Table 1: Rule Matrix for Proposed Algorithm

2.3. Genetic Algorithm

The use of gradient descent to minimise the cost function may also lead to a local minimum solution. This is undesirable if the contour is composed of multiple local minima, typical in nonlinear models. The aim is to find a global optimum solution

$$\hat{q} = \arg \max_q \{J(q); q \in \Theta\} \quad (15)$$

where $\hat{q} = \{\mathbf{W}^{(l)}\}_{l=1}^L$, $\mathbf{W}^{(l)}$ is the $(i \times j)$ demixing matrix of the l^{th} layer and Θ is the region of solutions. We propose to maximise the probability of a global solution via GA [5]. In our proposed algorithm, we implement a modified GA alongside Fuzzy control of the learning rates and the gradient descent based update. A population of M chromosomes are initialised:

$$\Gamma(k) = \{\omega_p(k) : \omega_p \triangleq \text{vec}([\mathbf{W}_p^{(1)} \ \mathbf{W}_p^{(2)} \ \dots \ \mathbf{W}_p^{(L)}])\}_{p=1}^{M_l} \quad (16)$$

where k , M_l , $\text{vec}(\cdot)$ represent the number of iteration, population size of $\mathbf{W}^{(l)}$ and function converting $\mathbf{W}^{(l)}$ into column vectors respectively and inserted into the gradient descent update.

The performance of each chromosome is evaluated using the fitness function:

$$E = -\sum_{i=1}^N \log p(y_i^{(l)}) - \log \left| \prod_{l=1}^L \mathbf{W}^{(l)} \text{diag}[\mathbf{D}^{(l)}] \right| - \sum_{l=1}^{L-1} \sum_{i=1}^N \log h_i^{(l)} \quad (17)$$

and ranked in descending order:

$$\bar{\Gamma}(k) = \{\omega_{p(r)}(k)\} \quad (18)$$

where $r=1, \dots, R$, Γ and $\bar{\Gamma}$ is the rank of ω_p , the population before and after ranking respectively. The chromosome with the lowest entropy corresponds to the fittest chromosome and thus ranked highest. The ranked chromosomes are categorised into sets A and B:

$$A(k) = \left\{ \omega_{p(r)}(k) : 1 \leq r \leq \frac{R}{2} \right\} \quad (19)$$

$$B(k) = \left\{ \omega_{p(r)}(k) : \frac{R}{2} + 1 \leq r \leq R \right\} \quad (20)$$

where set A is retained as part of the next generation. In the parent selection procedure, randomly selected $M/2$ chromosomes from set A and B are selected as parents. To generate offsprings, these parent chromosomes undergo a crossover procedure at crossover points $(l-1)(i \times j) + \left(\frac{i \times j}{2}\right)$ where $l=1, \dots, L$.

The final mutation process involves a small random change to the offsprings that replaces set B:

$$B \leftarrow B + v \quad (21)$$

where $v \in \mathbb{R}$ is a Gaussian random variable from the normal distribution. Set A (19) and mutated offsprings (21) form $\Gamma(k+1) = \{\omega_p(k+1)\}_{p=1}^{M_l}$ of the next iteration.

This is repeated with $U^{(l)} = \{\eta_w^{(l)}, \eta_D^{(l)}, \eta_a^{(l)}\}$ and $G = \left\{ \{\mathbf{W}^{(l)}\}_{l=1}^L, \{\text{diag}[\mathbf{D}^{(l)}]\}_{l=1}^L, \{\mathbf{a}^{(l)}\}_{l=1}^L \right\}$ iteratively updated until convergence is achieved.

3. Results

A 3-layer perceptron network for nonlinearly mixed signal separation was adopted to incorporate the proposed algorithm. A logistic function is used in the hidden neuron. The algorithms MLP with delta-bar-delta learning rule (MLP-DBD) [6] and Linear algorithm (Linear) are compared against the proposed NLFGA. Recorded speech signals were mixed nonlinearly via $\mathbf{A}_3 \tanh\{\mathbf{A}_2 \tanh(\mathbf{A}_1 \mathbf{s} + \boldsymbol{\varepsilon})\}$ with additional noise $\boldsymbol{\varepsilon}$ and the 4×4 mixing matrices $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ generated randomly. Sensors were perturbed with white Gaussian noise of 30dB signal-to-noise ratio. The score function $\varphi(y_i^{(L)}) = k_i \tanh(y_i^{(L)})$ where

$$k_i = \text{sign} \left[E \left\{ \sec h^2 \left(y_i^{(L)} \right) \right\} E \left\{ y_i^{(L)} \right\}^2 - E \left\{ \tanh \left(y_i^{(L)} \right) y_i^{(L)} \right\} \right]$$

customises $\varphi(y_i^{(L)})$ to cater for sub-gaussian and super-gaussian signals [4] was used. To compare the accuracy of the algorithms, we measured the performance error using (17).

The proposed algorithm was initialised with population size of $M = 8$. The learning rate of MLP-DBD was adapted according to the following equation:

$$\Delta\beta(k) = \begin{cases} \kappa & \overline{\delta(k-1)} \delta(k) > 0 \\ -\zeta\beta(k) & \overline{\delta(k-1)} \delta(k) < 0 \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

where κ is a constant, $\delta(k)$ is the gradient of the cost function, $w(k)$ is the value of a single weight, $\overline{\delta(k)}$ is the exponential average of the current and past gradients and $\beta(k)$ is the learning rate corresponding to $w(k)$. If current and past gradients are of opposing signs, $\beta(k)$ is decreased by a proportion of ζ . Montecarlo experiments were carried out to choose the best values for MLP-DBD (22) which required careful selection as the performance was heavily dependent upon κ . This weakness of DBD techniques are further compounded by the lack of appropriate method to determine the optimum value for κ and ζ . This inconvenience of ascertaining ‘good’ values of $U^{(i)} = \{\eta_w^{(i)}, \eta_D^{(i)}, \eta_a^{(i)}\}$ afflicts both the MLP-DBD and Linear algorithm, which is essentially the nature of classic gradient descent algorithms. The effectiveness of the proposed algorithm in the automated manoeuvring of $U^{(i)} = \{\eta_w^{(i)}, \eta_D^{(i)}, \eta_a^{(i)}\}$ is clearly demonstrated in Fig. 3 of the speech signals and entropy convergence. Fig. 3a and Fig. 3b are plots of the original and mixed speech signals whereas Fig. 3c, Fig. 3d and Fig. 3e present the signals recovered by NLFGA, MLP-DBD and Linear Algorithm respectively. Fig. 3f indicates the convergence of entropy achieved by NLFGA (continuous line), MLP-DBD (dotted line) and Linear (dash-dot) which clearly shows the efficiency of the proposed algorithm in speeding up the convergence and achieving the most optimum solution.

4. Conclusion

The proposed algorithm for nonlinear blind source separation displayed a performance in speed and accuracy surpassing that of the compared techniques. This is ultimately due to the classic MLP gradient based learning algorithm incorporated with Fuzzy Logic to ameliorate the learning speed of the algorithm and Genetic Algorithm based search to minimise the risk of arriving at a local solution. Issues of

convergence speed and convergence to local minima have been aptly addressed. This paper has proposed a blend of techniques that produces an impressive and attractive solution to nonlinear blind signal separation.

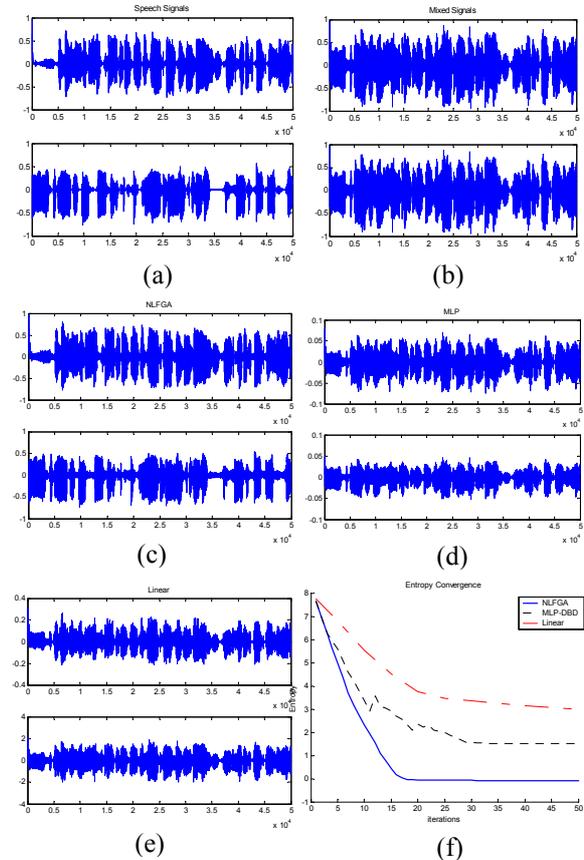


Figure 3: Speech Signals and Entropy Convergence

5. References

- [1] C. Jutten, and J. Karhunen, 'Advances in nonlinear blind source separation'. Proc. of ICA2003, 2003, pp. 245-256.
- [2] W. L. Woo and L. C. Khor, "Blind restoration of nonlinearly mixed signals using multilayer polynomial neural network", *IEE Proc. Vision Image and Sig. Proc.*, Feb 2004.
- [3] Y. Tan, et. al., "Nonlinear blind source separation using a radial basis function Network," *IEEE Trans. on Neural Networks*, vol. 12, pp. 124-135, 2001.
- [4] T.-W. Lee, et. al., "Independent component analysis using an extended infomax algorithm for mixed sub-gaussian and super-gaussian sources," *Neural Computation*, vol. 11, pp. 417-441, 1999.
- [5] F. Olivier and C. Lavergne, "Design of evolutionary algorithms — A statistical perspective", *IEEE Trans. on Evolutionary Computation*, vol. 5, no. 2, pp. 129-148, 2001.
- [6] R. A. Jacobs, "Increased rates of convergence through learning rate adaptation," *Neural Networks*, vol. 1, pp. 295-307, 1988.

Simple expressions for the locations and amplitudes of the extrema of the $(\sin x/x)^r$ function

Jean LE BIHAN

Ecole Nationale d'Ingénieurs de Brest, RESO, CS 73862, 29238 BREST cedex 3, France
Tel. +33 298 05 66 35, Fax +33 298 05 66 06, E-mail Jean.Le.Bihan@enib.fr

Abstract. *The locations and amplitudes of the extrema of the $(\sin x/x)^r$ function, when expressed under the form of series expansions, can be calculated very fastly through a straightforward recursion formula. Moreover, very simple accurate algebraic expressions are given for evaluating these locations and amplitudes.*

1. Introduction

The $\sin x/x$ function plays a key role in the fields of signal processing and communications [1, 2], so its properties have been extensively studied. Its Fourier transform, the rectangular function, is also widely used. On the other hand, the convolution operator and its counterpart regarding the Fourier Transform, the product operator, are of great importance. The purpose of this paper is to present the computation of the extrema of the function $(\sin x/x)^r$, where r is a positive integer. This function can be viewed as the result of the Fourier transform of a product of r identical rectangular pulses. A useful case is that of the function $(\sin x/x)^2$ which represents the Fourier transform of a triangular function.

For example, a window is frequently used in the time domain to limit the duration of a signal before calculating its Discrete Fourier

Transform. Many kinds of windows can be considered. Generally, the window is chosen with a smooth return to zero at its ends, in order to avoid discontinuity between two successive replicas of the periodized signal. This leads to enhanced properties of its Discrete Fourier Transform.

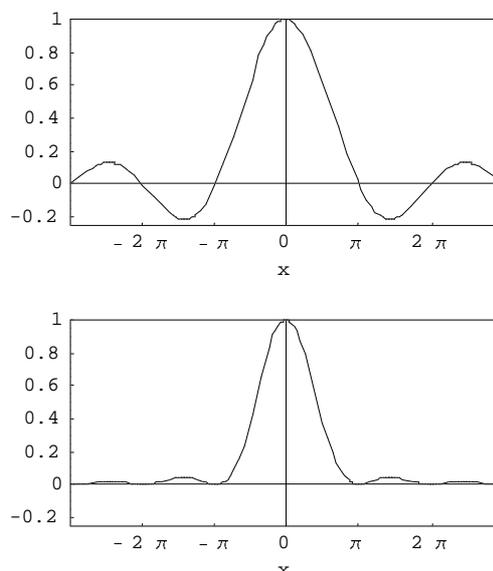


Fig. 1. The $\sin x/x$ (upper plot) and $(\sin x/x)^2$ (lower plot) functions

The well known rectangular window does not satisfy this criterion, but its multiple con-

volution do. The smoothness at the ends of the so-called polynomial windows [3] increases with the number of convolutions. So the frequential properties of a polynomial window, which corresponds to the convolution of r rectangular windows of duration T , are given by its Fourier Transform proportional to $(\sin x/x)^r$, with $x = \pi T f$. In particular, the triangular window and the parabolic window are relative to $r = 2$ and $r = 3$, respectively.

The $(\sin x/x)^r$ function may appear in other applications, for instance when studying random processes [4]. Let us consider a simple random binary wave, with pulses of amplitude $+A$ and $-A$ and duration T . In this case, the autocorrelation function has a triangular waveform and the expression of the power spectral density, is $A^2 T (\sin x/x)^r$, with $x = \pi T f$.

It is clear that the zeros of the function $(\sin x/x)^r$ are at $\pm k\pi$ ($k = 1, 2, \dots$) and the global maximum, which equals 1, at $x = 0$. The minima of $(\sin x/x)^r$ for an even r are at $\pm k\pi$ ($k = 1, 2, \dots$) and equal 0. But the local extrema $(\sin x/x)^r$ for any r are not at $\tilde{x}_m = \pm(2m + 1)\pi/2$ ($m = 1, 2, \dots$). Their locations x_m are, for all r , the solutions of the following equation

$$x = \tan x \quad (1)$$

Once the x_m obtained, the extrema amplitudes $y_{m,r}$ can be deduced from

$$y_{m,r} = \cos^r x_m \quad (2)$$

or $y_{m,r} = (-1)^{mr} (1 + x_m^2)^{-r/2}$.

In some applications, accurate determination of the extrema is needed [5]. In [5], a closed-form approximation and a recursion formula were proposed to determine the locations of the local extrema, with a good accuracy. In a more recent paper [6], methods were presented and used for the computation of both the locations and the amplitudes of the $\sin x/x$ function's extrema. Especially, series expansions were introduced to express the extrema locations and amplitudes, as well as simple algebraic approximations, allowing to reach an excellent accuracy. In this paper, we extend these results to the function $(\sin x/x)^r$. In particular, we show how to compute the coefficients of the $y_{m,r}$ series expansions and give approximate algebraic expressions for $y_{m,r}$.

2. Locations of the extrema

2.1. Series expansion

Let us denote $e_m = \tilde{x}_m - x_m$ the difference between each exact extremum location x_m and its nearest odd multiple of $\pi/2$, i.e. \tilde{x}_m . Then let us write e_m under the form of a series expansion in successive powers of $(\tilde{x}_m)^{-1}$: $e_m = \sum_{n=0}^{+\infty} e_m^n (\tilde{x}_m)^{-n}$. Therefrom the expansion coefficients of the extrema locations $x_m = \sum_{n=-1}^{+\infty} x_m^n (\tilde{x}_m)^{-n}$ can be immediately deduced: $x_m^{-1} = \tilde{x}_m$, $x_m^n = -e_m^n$ ($n \geq 0$).

We show now that using basic z -Transform techniques allows to derive the coefficients of the expansions. Changing \tilde{x}_m into z_m for this purpose and leaving out the index m for the coefficients are independent of m , eqn 1 takes the form

$$\tan e = 1/(z - e) \quad (3)$$

The first derivative can be expressed as

$$\frac{de}{dz} = -1/(z - e)^2 \quad (4)$$

A second derivation leads to the following relation

$$(z - e) \frac{d^2 e}{dz^2} - 2 \left(\frac{de}{dz} \right)^2 + 2 \frac{de}{dz} = 0 \quad (5)$$

Interpreting

eqn 5 by using basic z -Transform properties (among them $Z^{-1}\{-z \frac{de}{dz}\} = ne^n$) allows the straightforward computation of the expansion coefficients of e through the following relation (where n in e^n denotes an upper index)

$$\begin{aligned} e^n &= \{1/[n(n-1)]\} \\ \sum_{j=1}^{n-2} j \text{ odd } j(2n-j-1)e^j e^{n-j-1}, & \\ n \geq 3, n \text{ odd} & \\ e^n = 0, n \geq 0, n \text{ even} & \end{aligned} \quad (6)$$

with $e^1 = 1$.

Therefore the expansion of x_m begins as follows

$$\begin{aligned} x_m &= \tilde{x}_m - (\tilde{x}_m)^{-1} - \frac{2}{3}(\tilde{x}_m)^{-3} \\ &\quad - \frac{13}{15}(\tilde{x}_m)^{-5} - \frac{146}{105}(\tilde{x}_m)^{-7} - \dots \end{aligned} \quad (7)$$

2.2. Algebraic approximation

By limiting this expansion, simple closed-form approximations are obtained for x_m . Table 1 shows the values obtained for x_1, x_2, x_3

using the successive terms of the series expansion (the table starts with $x_m \simeq \tilde{x}_m$ at rank $p = 0$ ($n = 2p - 1$)). It can be observed for example that after only two ranks, the maximum relative error of the very simple algebraic approximation $x_m \simeq \tilde{x}_m - (\tilde{x}_m)^{-1} - \frac{2}{3}(\tilde{x}_m)^{-3}$ is 8.96×10^{-5} and it occurs for $m = 1$. The relative error is less than 3.86×10^{-6} for the other values of m (see Fig. 2). This very simple formula gives results comparable to those of eqn 1 in [5].

Table 1. Successive values for x_1, x_2, x_3 (series expansion)

p	x_1	x_2	x_3
0	4.712388980	7.853981634	10.99557429
1	4.500182390	7.726657680	10.90462861
2	4.493811716	7.725281614	10.90412712
3	4.493438770	7.725252614	10.90412173
4	4.493411825	7.725251859	10.90412166
Exact	4.493409458	7.725251837	10.90412166

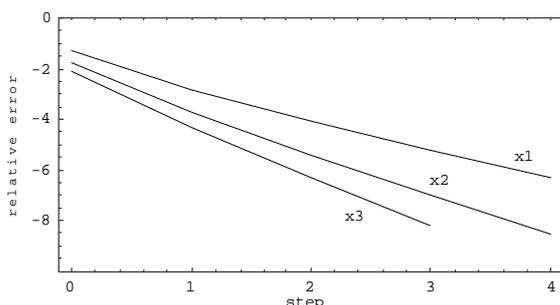


Fig. 2. Relative error (log scale) for successive values of extrema locations x_1, x_2, x_3

3. Amplitudes of the extrema

3.1. Series expansion

Using the same approach as for the locations, the extrema themselves can be directly calculated. From eqn 1, their amplitudes are $y_{m,r} = (-1)^{mr} \sin^r e_m$. Assuming that $y_{m,r}$ can be expanded under the form $y_{m,r} = (-1)^{mr} \sum_{n=0}^{+\infty} y_{m,r}^n (\tilde{x}_m)^{-n}$, and changing \tilde{x}_m into z_m , eqns 3 and 4 lead to the simple following equation (m is omitted for the same reason as above)

$$-(z - e) \frac{d|y_r|}{dz} = r |y_r| \quad (8)$$

Interpreting eqn 8 by using basic z-Transform properties allows here the straightforward computation of the expansion coefficients of $(-1)^{mr} y_{m,r}$ through the following relation (where n in y_r^n or e^n denotes an upper index)

$$y_{m,r}^n = [1/(n-r)] \sum_{j=r}^{n-1} j e^{n-j-1} y^j, \quad n > r, (n-r) \text{ even} \quad (9)$$

$$y_{m,r}^n = 0, \quad n \geq 0, (n < r, (n-r) \text{ odd})$$

with $y_{m,r}^r = 1$.

Therefore the expansion of $y_{m,r}$ begins as follows, for $r = 1$ and $r = 2$ as examples

$$y_{m,1} = (-1)^m \left[(\tilde{x}_m)^{-1} + \frac{1}{2}(\tilde{x}_m)^{-3} + \frac{13}{24}(\tilde{x}_m)^{-5} + \frac{61}{80}(\tilde{x}_m)^{-7} + \dots \right] \quad (10)$$

$$y_{m,2} = (\tilde{x}_m)^{-2} + (\tilde{x}_m)^{-4} + \frac{4}{3}(\tilde{x}_m)^{-6} + \frac{31}{15}(\tilde{x}_m)^{-8} + \dots \quad (11)$$

3.2. Algebraic approximation

By limiting this expansion, simple close-form approximations are obtained for $y_{m,r}$.

Table 2 shows the values obtained for y_1^1, y_2^1, y_3^1 using the successive terms of the series expansion (the table starts with $y_{m,1} \simeq (-1)^m (\tilde{x}_m)^{-2}$ at rank $p = 0$ ($n = 2p + 1$)). It can be observed for example that after only two ranks, the maximum relative error of the simple algebraic approximation $y_{m,1} \simeq (-1)^m \left[(\tilde{x}_m)^{-1} + \frac{1}{2}(\tilde{x}_m)^{-3} + \frac{13}{24}(\tilde{x}_m)^{-5} \right]$ is 7.34×10^{-5} and it occurs for $m = 1$. The relative error is less than 3.31×10^{-6} for the other values of m (see Fig. 3).

Table 2. Successive values for y_1^1, y_2^1, y_3^1 (series expansion)

p	y_1^1	y_2^1
0	-0.2122065908	0.1273239545
1	-0.2169845959	0.1283560036
2	-0.2172176874	0.1283741288
3	-0.2172324632	0.1283745424
4	-0.2172335359	0.1283745532
Exact	-0.2172336282	0.1283745535

p	y_3^1
0	-0.09094568177
1	-0.09132179296
2	-0.09132516305
3	-0.09132520229
4	-0.09132520282
Exact	-0.09132520282

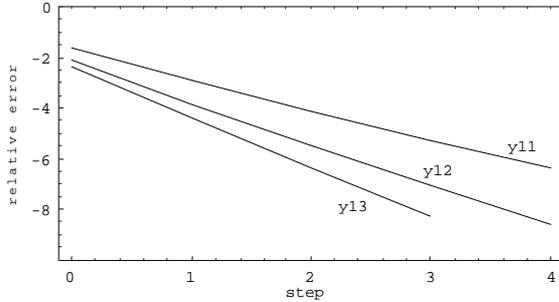


Fig. 3. Relative error (log scale) for successive values of extrema amplitudes y_1^1, y_2^1, y_3^1

Table 3 shows the values obtained for y_1^2, y_2^2, y_3^2 using the successive terms of the series expansion (the table starts with $y_{m,2} \simeq (\tilde{x}_m)^{-2}$ at rank $p = 0$ ($n = 2p + 2$)). It can be observed for example that after only two ranks, the maximum relative error of the simple algebraic approximation $y_{m,2} \simeq (\tilde{x}_m)^{-2} + (\tilde{x}_m)^{-4} + \frac{4}{3}(\tilde{x}_m)^{-6}$ is 1.95×10^{-4} and it occurs for $m = 1$. The relative error is less than 8.91×10^{-6} for the other values of m .

Table 3. Successive values for y_1^2, y_2^2, y_3^2 (series expansion)

p	y_1^2	y_2^2
0	0.04503163717	0.01621138938
1	0.04705948552	0.01647419853
2	0.04718124196	0.01647987920
3	0.04718974044	0.01648002194
4	0.04719039121	0.01648002587
Exact	0.04719044923	0.01648002599

p	y_3^2
0	0.008271117032
1	0.008339528409
2	0.008340282860
3	0.008340292533
4	0.008340292669
Exact	0.008340292671

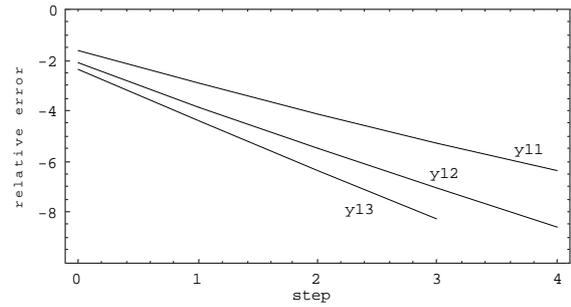


Fig. 4. Relative error (log scale) for successive values of extrema amplitudes y_1^2, y_2^2, y_3^2

4. Conclusion

A method for the computation of the extrema of the function $(\sin x/x)^r$ is presented here. The extrema locations and amplitudes are expressed under the form of series expansions. It is shown that the expansions coefficients can be obtained through straightforward recursive relations which can be easily implemented on a computer or a DSP. Very simple algebraic expressions are derived which give accurate values of the extrema locations and amplitudes, using only a very few coefficients.

Both methods (recursive relations or algebraic expressions) can be useful for various applications in different domains such as signal processing and communications.

References

- [1] H. Stark, F. B. Tuteur, and J. B. Anderson, Modern electrical communications, Englewood Cliffs : Prentice Hall, 1988
- [2] J. G. Proakis, and M. Salehi, Communications systems engineering, Englewood Cliffs : Prentice Hall, 1994
- [3] M. Kunt, Traitement numérique des signaux, Lausanne : Presses Polytechniques et Universitaires Romandes, 3e éd., pp. 107-123, 1999
- [4] S. Haykin, Communication systems, New York : John Wiley & Sons, 3rd ed., ch. 4, pp. 242-264, 1994
- [5] N. C. Beaulieu, "Extrema of $\sin x/x$ function", Electron. Lett., vol. 31, 15, 1995, p. 1215
- [6] J. Le Bihan, "Straightforward recursive relations for computing the locations and amplitudes of the $\sin x/x$ function's extrema", Electron. Lett., vol. 34, 25, 1998, pp. 2385-2386

Symbol Synchronizers Operating at Sub Data Rate

António D. Reis^{1,2}, José F. Rocha¹, Atilio S. Gameiro¹, José P. Carvalho²

¹Dep. de Electrónica e Telecomunicações / Instituto de Telecomunicações, Universidade de Aveiro, 3810 Aveiro, Portugal

²Dep. de Física, Universidade da Beira Interior Covilhã, 6200 Covilhã, Portugal

Abstract

In this paper, we present four types of symbol synchronizers, namely the analog, the hybrid, the combinational and the sequential.

Then, we explore the potentialities of the sequential synchronizer in order to obtain synchronizer versions operating at sub data rates.

The main objective is to create sequential synchronizer versions operating internally at very low speed (frequency) and transmit externally at very high data rate.

Key words: Synchronism in Digital Communications

I. INTRODUCTION

In this work we present four types of symbol synchronizers namely, the analog, the hybrid, the combinational and the sequential.

All this normal synchronizers need one or more delay adjust to work perfectly.

Fig.1 shows the general block diagram of a closed loop synchronizer.

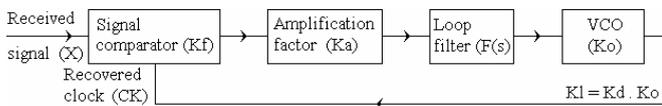


Fig.1 General closed symbol synchronizer

The closed loop synchronizer is composed by the signal comparator, amplification factor, filter and VCO.

The great difference between them is inside the signal comparator. The others blocks are similar.

At low SNR, the sequential synchronizer has the worst performance, but this can be minimized with a prefilter.

However at high SNR, where they normally operate, all synchronizers have similar jitter - noise curves.

The sequential synchronizer has more potentialities to create sub data rate versions, due to its signal comparator memory that guards previous data transition information, when posteriorly a transition clock occur.

Next, we present the four synchronizers (analog, hybrid, combinational and sequential). Referring the main difference between them.

After, we show the sequential versions operating at sub data rates.

Posteriorly, we show the design, tests and results with some comparisons [1, 4].

Finally, we present the main conclusions.

II. SYMBOL SYNCHRONIZERS TYPES

We introduce the appropriate signal comparator obtaining

the four closed loop symbol synchronizer types, namely the analog, the hybrid, the combinational and the sequential [3].

A. Analog synchronizer

Fig.2 shows the analog type, in which the signal comparator is composed by three analog multipliers (ideal multiplier) two for data transitions and one for comparison.

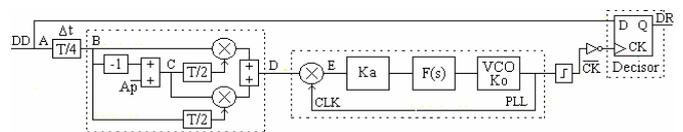


Fig.2 Analog synchronizer

The input signals (main input and VCO output) are both analog.

B. Hybrid synchronizer

Fig.3 shows the hybrid type, in which the signal comparator is composed by an exor with delay followed of an analog multiplier or switch (real multiplier).

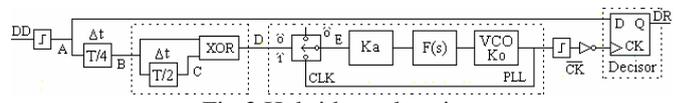


Fig.3 Hybrid synchronizer

The main input signal is already digital but the VCO output is still analog.

C. Combinational synchronizer

Fig.4 shows the combinational type, in which the signal comparator is composed by an exor with delay followed of a digital demux.

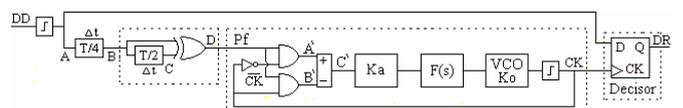


Fig.4 Combinational synchronizer

The two signal comparator inputs are both digital and the output is only function of the present inputs.

D. Sequential synchronizer

Fig.5 shows the sequential type, in which the signal comparator is composed by a flip flop with exor (variable

^{1,2}UA-UBI

pulse P_v) and an exor with a delay (fixed pulse P_f) [2].

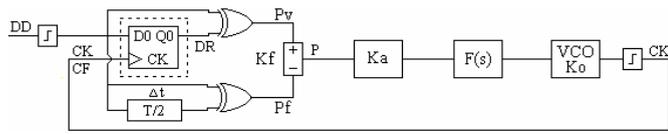


Fig.5 Sequential synchronizer

The two signal comparator inputs are both digital but now its output is simultaneously function of the two entries and signal comparator state (memory).

III. SYNCHRONIZERS OPERATING AT SUB RATES

The sequential synchronizer has potentialities that permit obtain versions operating at sub data rates.

The difference is inside of the signal comparator. So initially at the rate the dash outlined is only a flip flop, at half rate is two flip flops with mux, at a quarter rate is four flip flops with mux, etc.

A. Synchronizer operating at the rate

To understand the synchronizers operating at sub data rates, we consider the start model of Fig.6.

The dash outlined is the circuit that varies from a model to another. Here the dash outlined is only a flip flop (Fig.5).

Fig.6 shows the synchronizer operating internally at the same of the external transmission rate.

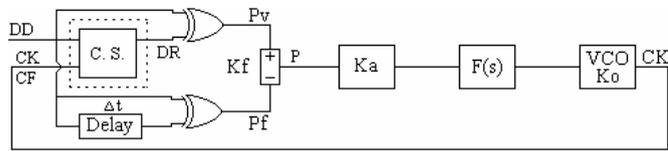


Fig.6 Synchronizer operating at the rate

Posteriorly, the CS dash outlined rectangle must be substituted by a more complex circuits.

B. Synchronizer operating at half rate

The synchronizer operating at half rate consists in the substitution of the CS dash outlined by two flip flops with a mux.

Fig.7 shows the synchronizer operating internally at half of the external transmission rate.

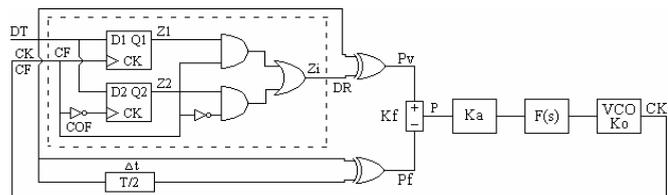


Fig.7 Synchronizer operating at half rate

The difference is inside of the signal comparator block, the others blocks are equal.

C. Synchronizer operating at a quarter rate

The synchronizer operating at half rate consists in the substitution of the CS dash outlined by four flip flops with a mux.

Fig.8 shows the synchronizer operating internally at a quarter of the external transmission rate.

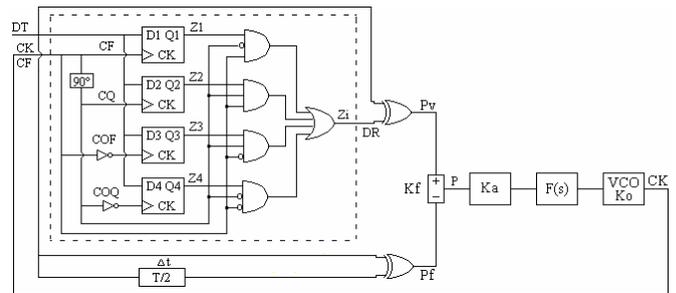


Fig.8 Synchronizer operating at a quarter rate

The difference is inside of the signal comparator block, the others blocks are equal.

D. Generalization for submultiples of the rate

The synchronizer operating at half rate consists in the substitution of the CS dash outlined by N flip flops with a mux.

Fig.8 shows the synchronizer operating internally at 1/N of the external transmission rate.

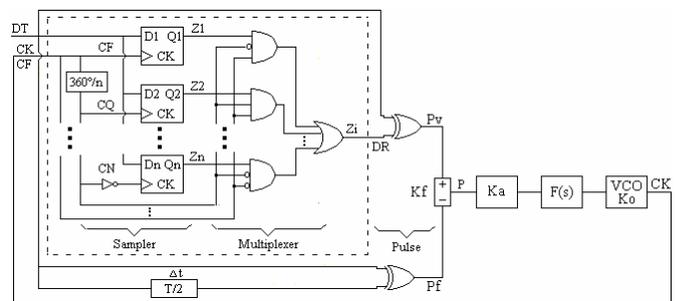


Fig.9 Synchronizer operating at 1/N rate

The characteristic curve of this phase detectors, although operating at different data rates, is equal for all them and its gain is given by $K_f = 1/2\pi$.

IV. DESIGN, TESTS AND RESULTS

A. Test setup

Fig.10 shows the setup that we used to get the jitter-noise curves of the various synchronizers [5].

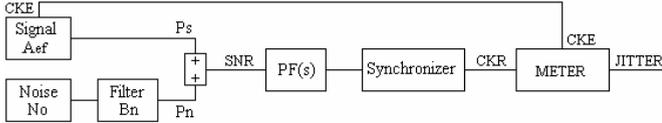


Fig.10 Block diagram of the test setup

The signal to noise ratio SNR is given by P_s/P_n , where P_s is the signal power and P_n is the noise power. They are defined as $P_s=A_{ef}^2$ and $P_n=No.B_n=2\sigma_n^2\Delta\tau B_n$. A_{ef} is the RMS amplitude, B_n is the external noise bandwidth, No is the noise power spectral density, σ_n is the noise standard deviation and $\Delta\tau$ is the sampling period (inverse of samples per unit time).

The prefilter is not used here, but can be useful in system with high noise quantities ($PF(s)=1$).

B. Jitter measurer

Fig.11 shows the jitter measurer (METER) that consists of a RS flip-flop which detects the variable phase of the recovered clock (VCO) relatively to the fixed phase of the emitter clock.

This relative phase variation is the recovered clock jitter.

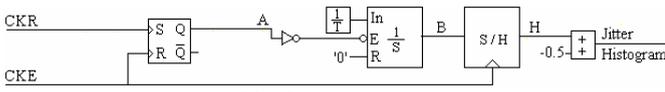


Fig11 The jitter measurer

The others blocks convert this phase variation into an amplitude variation, which is the jitter histogram.

Fig.12 shows the waveforms that illustrate the operation mode of the jitter measurer.

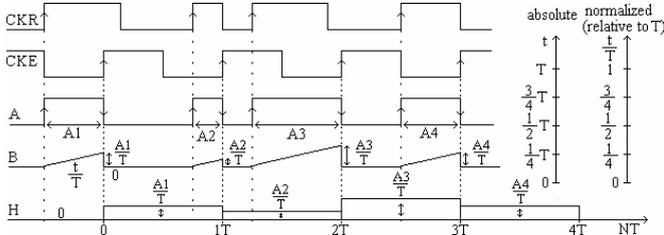


Fig.12 Waveforms at the jitter measurer

The jitter histogram is then sampled and processed by an appropriated program giving the average m , jitter variance in squared radians σ_n^2 , jitter standard deviation in unit intervals root mean squared UIRMS and jitter standard deviation in unit intervals peak to peak UIPP.

We have used also others jitter measurers with similar results.

C. Loop parameters design

To establish guaranteed comparisons it is necessary to test all the synchronizers in equal conditions.

We use a normalized transmission rate $t_x=1$ baud ($f_0=1$ Hz) what facilitates the analyses and allows one more easy extrapolation for other rhythms of transmission. We use an equivalent external noise bandwidth $B_n=5$ Hz for all SLL. For

the closed loop symbol synchronizers SLL, we use a loop noise bandwidth $B_l=0.02$ Hz.

For analog SLL, the relation between signal to noise ratio SNR and jitter variance σ_n is $SNR=A_{ef}^2/No.B_n=A_{ef}^2/(2\sigma_n^2.\Delta\tau.B_n)=(0.5)^2/(2\sigma_n^2*10^{-3}*5)=25/\sigma_n^2$. This relation is more complicated for the others symbol synchronizers.

We will now present the loop parameters design for the various PLLs considering the first (1st) and the second order loop (2nd).

- 1st order loop:

In the 1st order loop, the filter $F(s)=0.5$ Hz eliminates only the high frequency, but maintain the loop characteristics. This cutoff frequency $F(s)=0.5$ Hz is 25 times higher than $B_l=0.02$ Hz. Then the transfer function of the 1st order is

$$H(s)=\frac{G(s)}{1+G(s)}=\frac{KdKo}{s+KdKo} \quad (1)$$

the loop noise bandwidth for the SLLs is

$$B_l=\frac{KdKo}{4}=\frac{KfKo}{4}=0.02\text{Hz} \quad (2)$$

so for the analog SLL with $K_m=1$, $A=1/2$, $B=0.45$ we have

$$Ka\frac{K_mABKo}{4}=0.02\text{Hz} \rightarrow Ka=0.08\frac{2.2}{\pi} \quad (3)$$

for the hybrid SLL, with $K_m=1$, $A=1/2$ and $B=0.45$ we have

$$Ka\frac{K_mABKo}{4}=0.02\text{Hz} \rightarrow Ka=0.08\frac{2.2}{\pi} \quad (4)$$

for the combinational SLL ($K_f=1/\pi$) we have

$$Ka\frac{KfKo}{4}=Ka\frac{(1/\pi)2\pi}{4}=0.02\text{Hz} \rightarrow Ka=0.04 \quad (5)$$

and for the sequential SLL ($K_f=1/2\pi$) we have

$$Ka\frac{KfKo}{4}=Ka\frac{(1/2\pi)2\pi}{4}=0.02\text{Hz} \rightarrow Ka=0.08 \quad (6)$$

This formulas are useful in synchronizers design

- 2nd order loop:

The transfer function with $F(s)=\frac{1+sT_2}{sT_1}$ is

$$H(s)=\frac{sKdKo(T_2/T_1)+KdKo/T_1}{s+sKdKo(T_2/T_1)+KdKo/T_1} \quad (7)$$

$$=\frac{sA+B}{s^2+s2xWn+Wn^2} \quad (8)$$

and the loop noise bandwidth is

$$B_l=\frac{xWn}{2}\left(1+\frac{1}{4x^2}\right) \quad (9)$$

Taking ($\xi=1$ and $B_l=0.02$) and solving the above equations we obtain for $F(s)$

$$F(s)=\frac{1+s63}{s977} \quad (10)$$

so for the analog SLL we have

$$Kd=KaKf=Ka(1/2)(1/2) = \frac{1}{2p} \rightarrow Ka = \frac{2.2}{p} \quad (11)$$

for the hybrid SLL we have

$$Kd=KaKf=Ka(1/2)(0.45) = \frac{1}{2p} \rightarrow Ka = \frac{2.2}{p} \quad (12)$$

for the combinational SLL we have

$$Kd=KaKf=Ka \frac{1}{p} = \frac{1}{2p} \rightarrow Ka = 0.5 \quad (13)$$

and for the sequential SLL we have

$$Kd=KaKf=Ka \frac{1}{2p} = \frac{1}{2p} \rightarrow Ka = 1 \quad (14)$$

This formulas can be used in others synchronizers.

D. Results

We studied the jitter-noise behavior of four synchronizers namely the analog (ana), hybrid (hib), combinational (cmb) and the sequential (seq) synchronizers.

Then we select the sequential synchronizer developing versions operating at the rate (tx), at half rate (tx2), at a quarter rate (tx4) and at 1/N rate (txn).

Fig.13 shows the jitter-noise curves of the four synchronizers (ana, hib, cmb, seq).

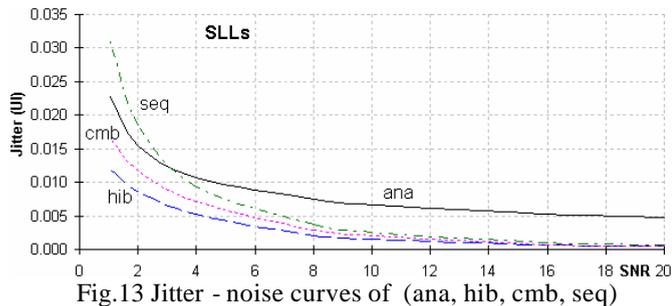


Fig.13 Jitter - noise curves of (ana, hib, cmb, seq)

We verify that generically the jitter UI diminishes when the signal to noise ratio SNR increases.

For low SNR, the synchronizer without input limiter (analog) is advantageous over the others with input limiter (hybrid, combinational, sequential).

However for high SNR the synchronizer without input limiter (analog) is disadvantageous over the others with input limiter (hybrid, combinational, sequential).

Fig.14 shows the jitter-noise curves of the synchronizers operating at sub data rates (tx, tx2, tx4, txn).

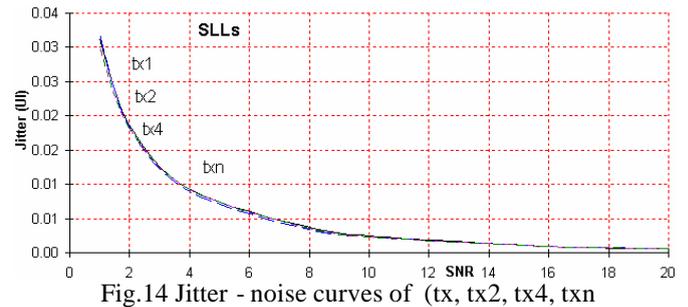


Fig.14 Jitter - noise curves of (tx, tx2, tx4, txn)

We note that, all the synchronizers at different sub data rates have the same jitter - noise curves.

V. CONCLUSIONS

We studied four synchronizers types namely analog (ana), hybrid (hib), combinational (cmb) and sequential (seq).

Then we take the sequential synchronizer and we developed versions operating at the rate (tx), half rate (tx2) a quarter rate (tx4) and 1/N rate (txn).

In the synchronizers (ana, hib, cmb, seq) we verify that for low SNR, the synchronizer without input limiter (analog) is advantageous over the others with input limiter (hybrid, combinational, sequential). This is comprehensible because the limiter provokes random gate commutations with noise spikes. However for high SNR the synchronizer without input limiter is disadvantageous over the others. This is comprehensible because the limiter margin noise ignores low noise spikes.

For the synchronizers operating at sub data rates, we verify that all them have similar jitter - noise curves. This is comprehensible because its signal comparator has the same operation mode and equal gain.

VI. ACKNOWLEDGMENTS

The authors are thankful to the program FCT (Foundation for sScience and Technology).

VII. REFERENCES

- [1] – Werner Rosenkranz, “Phase Locked Loops with limiter phase detectors in the presence of noise”, IEEE Transactions on Communications com-30. October 1982.
- [2] - Charles R. Hogge, “A Self Correcting Clock Recovery Circuit”, Journal of Lightwave Technology, Dec. 1985.
- [3] – A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho “Phase and Data Synchronizers in the Presence of Noise”, 5th International Conference on Telecommunications p.59, Zagreb 17-18 June 1999.
- [4] - A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho, “Manual and Automatic Data Synchronizers”, 7th International Conference on Telecommunications p.1076, Acapulco-MX 22-25 May 2000.
- [5] – A. D. Reis, J. F. Rocha, A. S. Gameiro, J. P. Carvalho “A New Technique to Measure the Jitter”, III Conference on Telecommunications p.64, F. Foz-PT 23-24 April 2001.

TECHNICAL WHITE PAPER ON IDM

(Intelligent Decision Maker)

Mr. Ajay Rana

Senior Lecturer
Amity School of Computer Sciences
Amity Campus Noida ,
Sector 44,
Noida (UP) 201303

E mail : ra_ajay@hotmail.com

Mr. Sandeep Chauhan

Technical Manager(IT)
ASD India Pvt. Ltd
NSIC – STP Complex
208 , Okhla Industrial Estate
New Delhi 110020

Email : chauhansandeep@hotmail.com

Abstract

IDM can be used to initialize decision-making system , this helps in analysis and designing the problems. Proficient decision making is the hallmark of professional managers and system analyst. Anyone can make bad decisions. All four of the basic managerial functions—planning, organizing, providing leadership, and ensuring system control—require skillful decision making.

This concept in implementation can help managers who prefer to rely on experience and intuition when dealing with strategic decisions.

Many tools and techniques have been developed for the systematic acquisition of domain knowledge, including knowledge elicitation (KE) methods to acquire knowledge from a human expert, machine learning (ML) algorithms that infer knowledge from data, and knowledge base refinement (KBR) tools that refine existing knowledge bases (KBs). As the number and sophistication of knowledge acquisition tools increases, it becomes progressively more difficult for users (notably domain experts) to choose between them for particular applications, especially when more than one is needed. We recognize the importance of driving this process by the epistemological requirements of the problem solver(s) that have been selected to solve a particular task.

Key Words : Knowledge Chaining , Fuzzy Logic , Data Mining , KDD

Introduction

In this paper, we present the IDM framework and illustrate it with a detailed description of the prototype currently being implemented

ASCS/ASD will develop a fully functional, generic Intelligent Decision Maker (IDM). IDM will be used

in many areas like troubleshooting guidelines; web based training, customer service improvements initiatives are few to list. The tool will be fully functional as application, or web-based service that can be used by its user.

IDM includes an advisory system coupled to several knowledge acquisition tools and problem solvers.

This advice-giving system compares the requirements of the selected problem solver with available sources of information (knowledge, data, human expert). As a result, it may recommend either the reuse of existing knowledge bases, or the application of one or more knowledge acquisition tools, based on their knowledge-level descriptions

This project is collaborative effort combining expertise in the various disciplines into one generic system. This will help its user to customize it as per the requirement for their use.

Technology

IDM has been developed based on a J2EE compliant n-tiered architecture. The architecture supports multiple-client access to the database server via the front-end portal, an Internet/Intranet connection, and any standard browser.

The functionality provided by the browser includes full drill down capabilities. The decision trees are implemented in XML for easy editing. The database is implemented on Oracle 8i.

APPROACH

Current IDM approach is based on the decision tree concept theory. Decision Trees are excellent tool for helping to choose between several courses of action. They provide a highly effective structure within which you can lay out options and investigate the possible outcomes of choosing those options. They also help you to form a balanced picture of the risks and rewards associated with each possible course of action.

The current solution is based on six different steps for simplifying the process. It uses knowledge chaining to implement the rules and actions based on the conditions. The different types of nodes in a decision tree are as follows:

Decision Pass

In this step IDM will execute an action at the backend and no user interaction will be there, as all of the processing will be done by certain programs/stored procedures. It make use of XML file from where it executes the decision tree and based on those information certain class file will be executed to perform the decisions.

Process/Execute Pass

In this step IDM will execute the action for intermediate processing and will pass the result to its next/specified step depending upon the decision.

Logic Pass

In this step IDM will display the messages for users and ask for the inputs like OK/CANCEL or YES/NO. Depending upon the user response a decision is made and control is passed to next/specified step. This pass uses the knowledge chaining.

Rule execution/Decision Pass:

In this step IDM will check inputs against a rule, which is there in underlying database or in property file and will make an action and pass the control accordingly to next step. The administrator depending on the different inputs sets the decisions. This part will be accomplished by the fuzzy logic implementation in future.

Display Decision Pass:

In this step IDM will display the messages for users that what kind of decision has been taken. Once the messages/decisions are displayed it will pass the control to next step. The pass is more of inforamatory pass that show the result of execution of a particular logic.

End/Finish Pass:

In this step IDM will display some messages and screen, which indicates the end of one decision process.

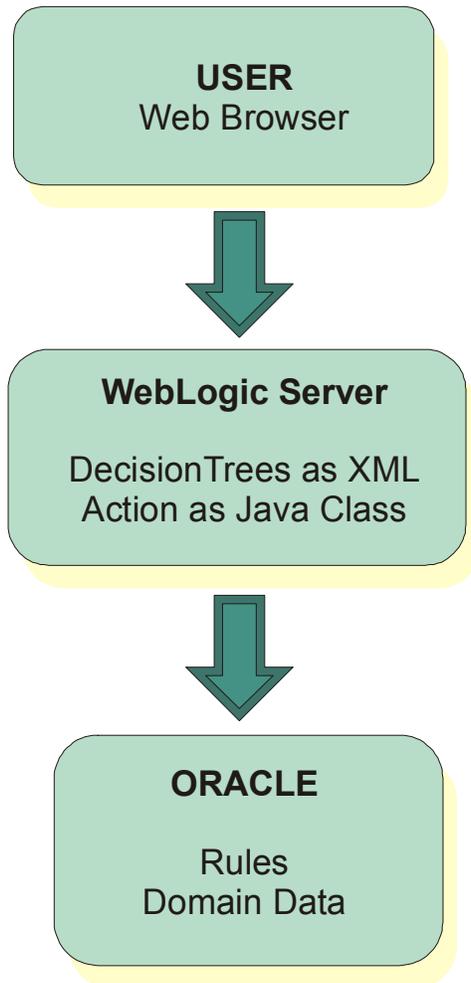
ARCHITECTURE

IDM uses a generic three-tier architecture. Layer 1 contains client in form of standalone browser or it may be a web-based client in form of applet. Layer 2 contains the J2EE compatible app server. Layer 3 contains the database server that can connect on legacy database for data mining incorporation. There are three architectural views defined for IDM.

Technical View:

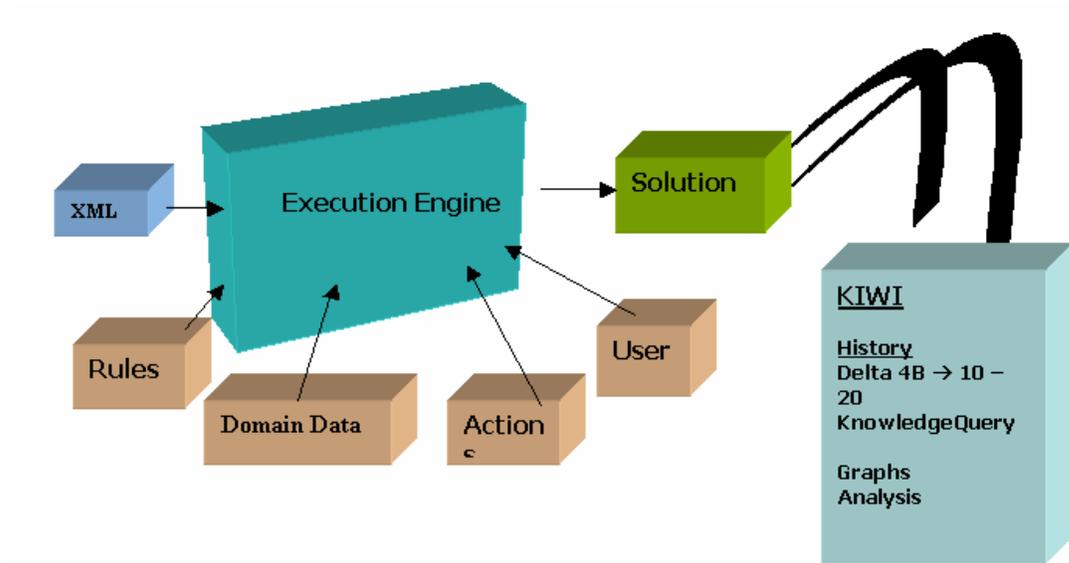
Technical view shows the three layers. Client layer involves user that may be a web browser using JSP/Servlets as its interface, may be an applet or it may be a java-swing application running on desktop. The client is interacting with the Application Server that come under the second layer. Application Server contains the decision tree in form of XML files and all the actions in form a java class files.

The third layer is database that stores the rules and the domain data for execution. Once we execute the decision tree, it interacts with the underlying database from where it executes the rules to get the results.



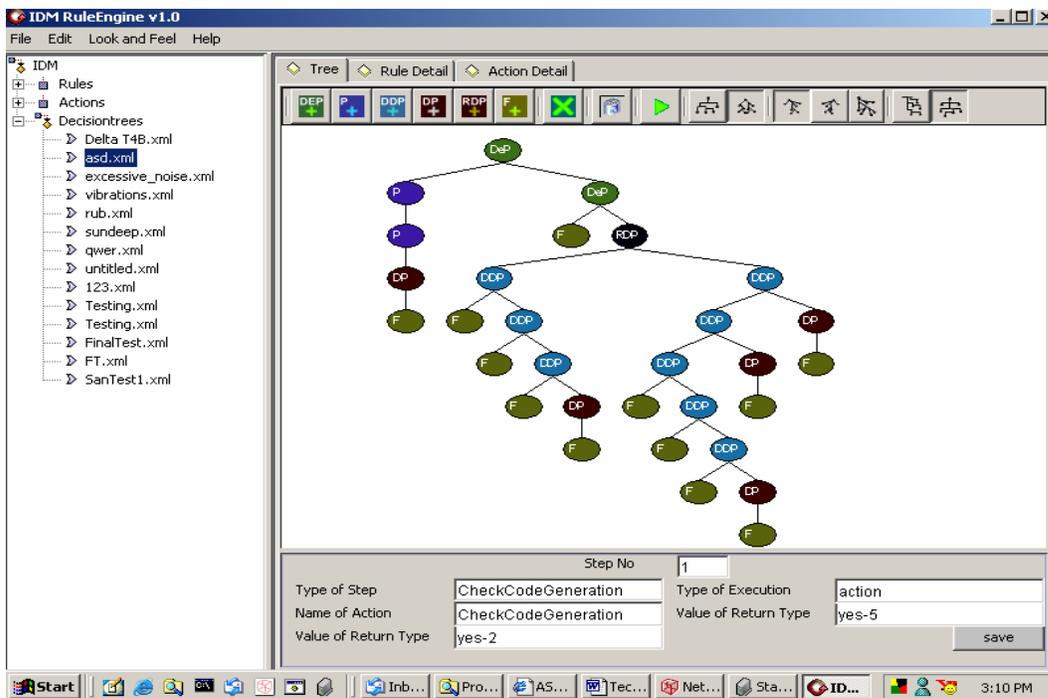
Functional View:

This view depicts the functional working of IDM tool. The Figure shows that execution engine involves XML, Rule, Domain Data, and Actions. When a requests to IDM tool , execution engine takes care of the request and process it in conjunction with the related items and throws the appropriate solutions .



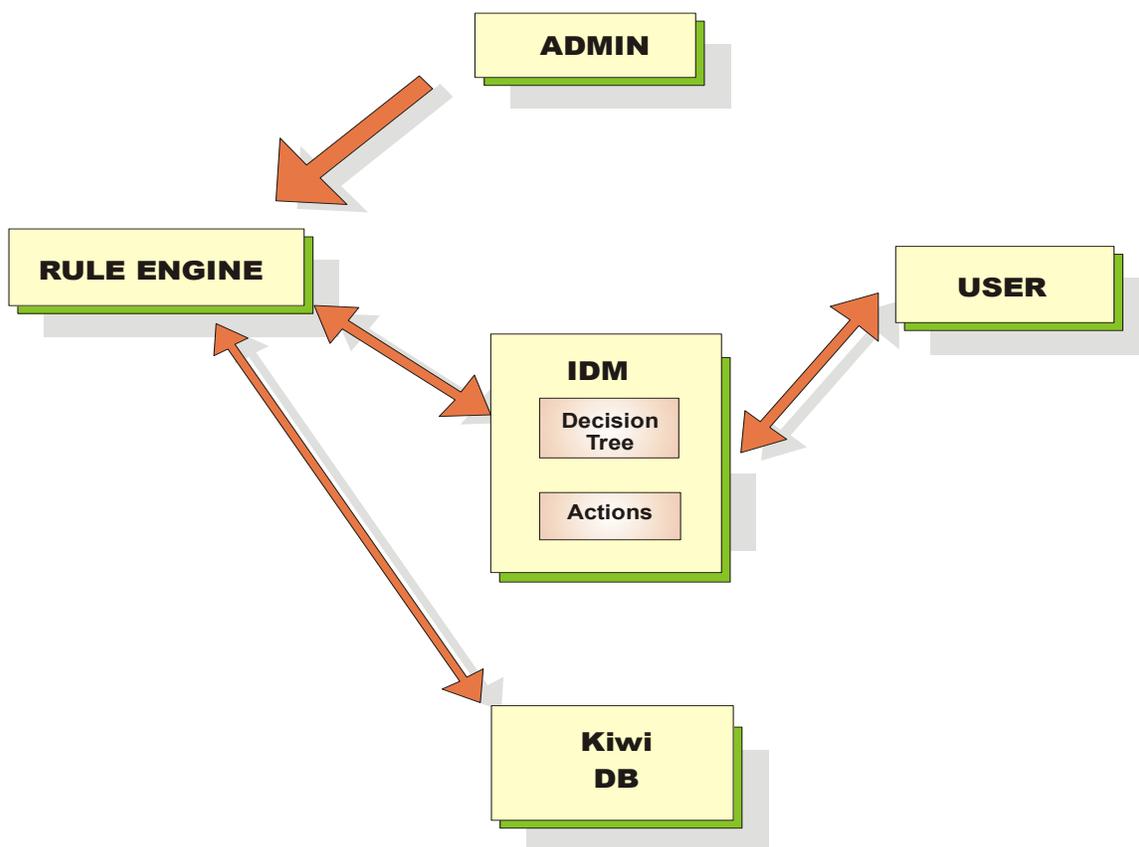
Visual Editor View:

This view depicts the actual interface, through which a admin/user will be interaction to IDM tool. A user can load the desired decision tree and he can execute it. Administrators can create/edit/modify the decision tree as per the problem domain.



WORK FLOW

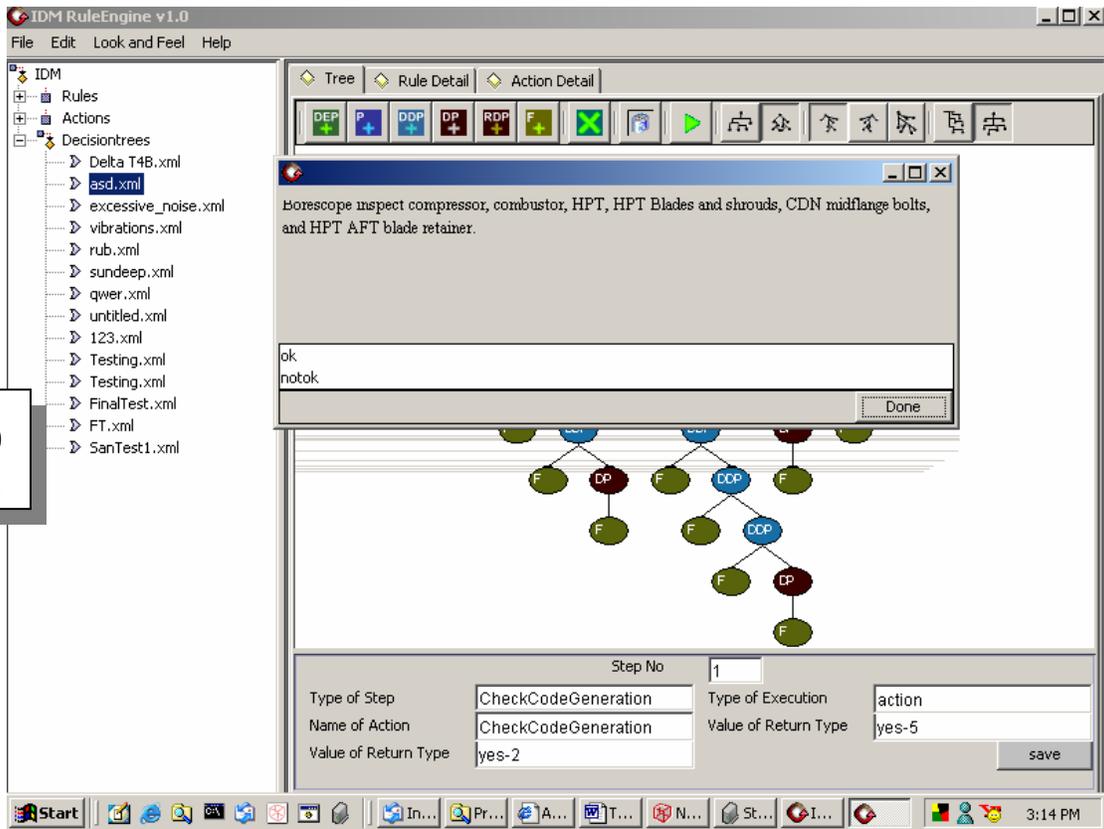
The current IDM implementation approach is based on rule based action system. The user first need to define the rules and based on these rules there will be an action for respective rule that will be fired by the decision tree depending on a condition. The approach is very effective and simple to understand and use.



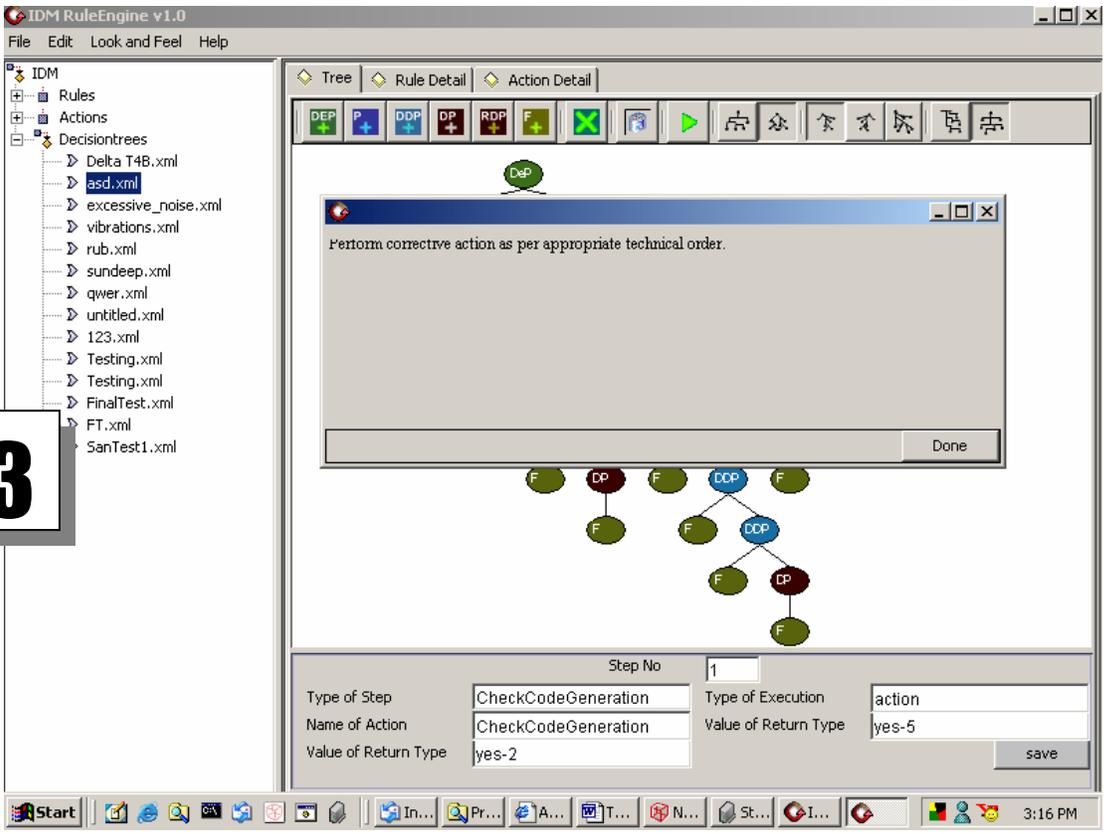
In the present workflow approach, administrator of the IDM system will set rules and actions by creating the decision tree. The information about the rules and actions are stored in a database and the decision tree will be stored in a XML file on server.

When a user comes and he/she wants to execute the software to get the solution for his/her problem. The IDM tool executes for his/her problem, user needs to pass the attributes for his/her problem. Based on the user inputs IDM tool executes the decision tree that in turn gives the desired output to user

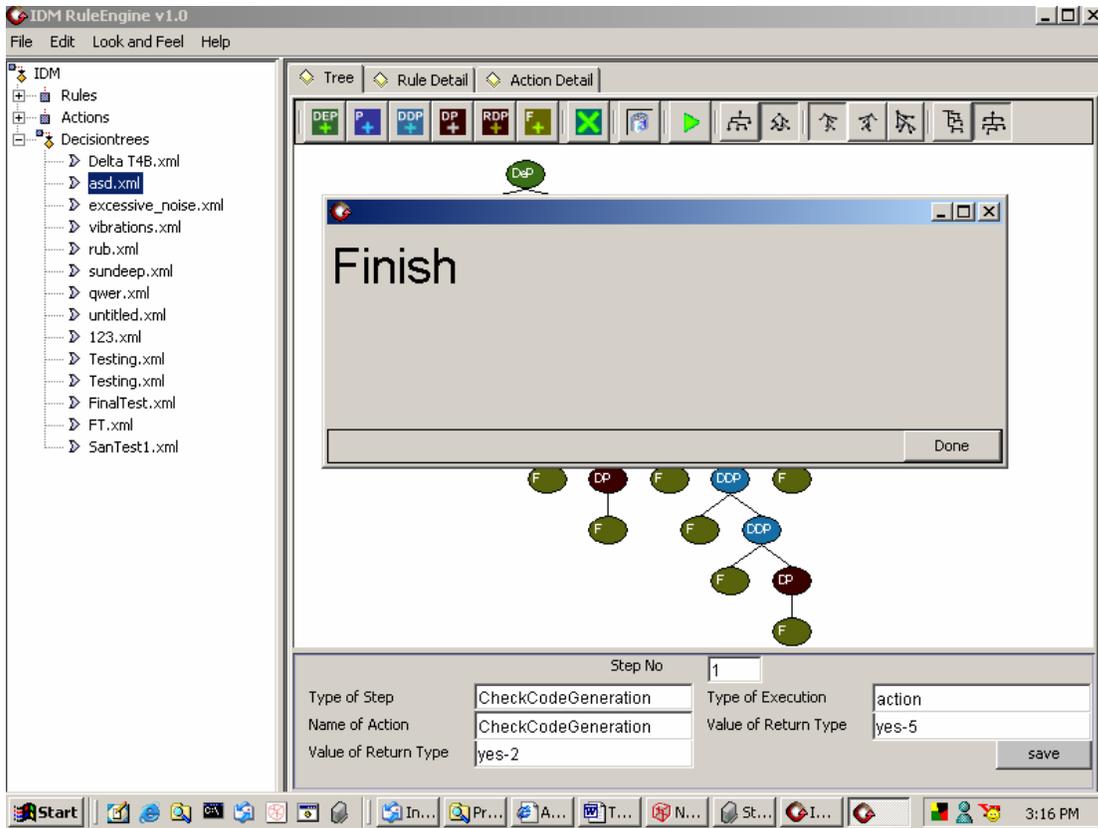
2



3



4



1. The first screen shows the dialog about the current condition of fan movement and other items depending on the use input 'ok/not ok' it proceed to next step
2. The second screen come when a user gave his input as ok and here it ask about the condition of HPT and HPT blades, when a user inputs 'not ok'
3. The third screen shows action to be taken on these conditions.
4. Once we are done and click on ok it shows us the last screen.

WHERE TO GO – NEXT STEP

The current approach is based on the crisp decision making approach. An intelligent system should take human like decisions. It should learn by its experience and it should be able to take some really intelligent decision. To achieve this we need to follow the Fuzzy Logic approach under Artificial Intelligence.

Moreover we need to put experience intelligent agents those will be responsible for learning and making the system more knowledgeable. Data mining and Knowledge Discovery (KDD) need to be explored for implementation. For summary these are the listed tasks that we have to take care

- 1) Fuzzy Logic implementation
- 2) Implementation of Intelligent Agents for learning.
- 3) Data Mining and Knowledge Discovery.
- 4) Incorporation of template decision trees to be used as base for creating new decision trees.
- 5) Database as data ware house for knowledge queries
- 6) History maintenance for tracking.

Conclusion:

IDM can be used to initialize decision-making system , this helps in analysis and designing the problems. Proficient decision making is the hallmark of professional managers and system analyst. Anyone can make bad decisions. All four of the basic managerial functions—planning, organizing, providing leadership, and ensuring system control—require skillful decision making.

This concept in implementation can help managers who prefer to rely on experience and intuition when dealing with strategic decisions.

Some decision tree software focuses on creating a decision tree from data, as for data mining and data classification. Others focus on risk management or another specialized vertical application. IDM concept focuses instead on having a smart user build a decision tree based on specialized knowledge, and then using that tree to gather the necessary inputs, process those inputs, and make decisions based on how the processed information fits into the logic of the tree. Furthermore, IDM suggests improvements to the tree based on actual use of the entire system

GLOSSARY

Knowledge Chaining:

An attribute in a decision tree can itself be represented by another decision tree. This is called knowledge "chaining".

Fuzzy Logic:

Fuzzy Logic is basically a multi-valued logic that allows intermediate values to be defined between conventional evaluations like yes/no, true/false, black/white, etc. In this way an attempt is made to apply a more human-like way of thinking in the programming of computers

Data Mining:

Data Mining is exploration & analysis, by automatic or semi automatic means of large quantities of data in order to discover meaningful patterns & rules.

KDD (Knowledge Discovery in Database):

Knowledge Discovery in Database is a non-trivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data.

References

1. Altman, R. B., Weiser, B., and Noller, H. F. (1994). Constraint satisfaction techniques for modeling large complexes: Application to central domain of the 16s ribosomal subunit. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, (pp. 10-18), Stanford, CA.
2. Angele, J., Fensel, D. and Studer, R. (1994). A model of expertise in KARL. Proceedings of the Second World Congress on Expert Systems, Lisbon, Portugal.
3. Breuker, J.A., and Van de Velde, W., Eds (1994). The CommonKADS Library for Expertise Modelling. Amsterdam: IOS Press.
4. Chandrasekaran, B. (1983). Toward a taxonomy of problem-solving types. AI Magazine, 4(4), 9-17.
5. Chandrasekaran, B. (1986). Generic tasks for knowledge-based reasoning: High-level building blocks for expert system design, ° IEEE Expert, 1(3), 23-30.
6. Eriksson, H., Shahar, Y., Tu, S.W., Puerta, A.R., and Musen, M.A. (1995). Task modeling with reusable problem-solving methods. Artificial Intelligence, 79, 293-326.
7. Finin, T., McKay, D., Fritzson, R., and McEntire, R., (1994). KQMLDA language and protocol for knowledge and information exchange. In Kazuhiro Fuchi and Toshio Yokoi (Eds.), Knowledge Building and Knowledge Sharing, Ohmsha and IOS Press.
8. Genesereth, M.R., and Fikes, R.E. (1992). Knowledge Interchange Format, Version 3.0 Reference Manual. Computer Science Department Technical Report Logic-92-1, Stanford University, CA.
9. Gennari, J. H., Altman, R. B., and Musen, M. A. (1995). Reuse with PROTÉGÉ-II: From elevators to ribosomes. Proceedings of the Symposium on Software Reuse, (pp. 72-80). Seattle, WA.
10. Gennari, J. H., Tu, S. W., Rothenfluh, T. E., and Musen, M. A. (1994). Mapping domains to methods in support of reuse. International Journal of Human-Computer Studies, 41, 399-424.
11. Goossens, L. (1995) From Commet to KrestT. Proceedings of the Ninth Banff Knowledge Acquisition for Knowledge-Bases Systems Workshop (pp. 41.1-41.20), Banff, CA.
12. Gruber, T.R. (1993). A translation approach to portable ontology specifications. Knowledge Acquisition, 5, 199-220.
13. Guarino, N., and Giaretta, P. (1995). Ontologies and knowledge bases: Toward a terminological clarification. In N.J.I. Mars (ed.), Towards Very Large Knowledge Bases, IOS Press, pp. 25-32.

14. Miller, R.A., Pople, H.E., and Myers, J.D. InternistI, an experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine* 307:468-476, 1982.
15. Musen, M.A., Gennari, J.H., and Wong, W.W. (1995). A rational reconstruction of Internist-1 using PROTÉGÉ-II. *Proceedings of the 19th Annual Symposium on Computer Applications in Medical Care* (pp. 289±293). New Orleans, LA.
16. Orfali, R., Harkey, D., and Edwards, J. (1996). *The Essential Distributed Objects Survival Guide*. John Wiley & Sons, New York.
17. Puerta, A.R., Egar, J. W., Tu, S.W., and Musen, M.A. (1992). A multiple-method knowledge-acquisition shell for the automatic generation of knowledge-acquisition tools. *Knowledge Acquisition*, 4, 171-196.
18. Rothenfluh, T.E., Gennari, J.H., Eriksson, H., Puerta, A.R., Tu, S.W., and Musen, M.A. (1996). Reusable ontologies, knowledge-acquisition tools, and performance systems: PROTÉGÉ-II solutions to sisyphus-2. *International Journal of Human-Computer Studies*, 44, 303-332.
19. Schreiber, A.T. and Birmingham, W.P., Eds. (1996). Special issue on the Sisyphus-VT initiative. *International Journal of Human-Computer Studies*, 44, 275-598.
20. Shadbolt, N., Motta, E. and Rouge, A. (1993). Constructing knowledge-based systems. *IEEE Software*, 10(6), 34-38.
21. Shwe, M.A., Middleton, B., Heckerman, D.E., Henrion, M., Horvitz, E.J., Lehmann, H.P., and Cooper, G.F. (1991). Probabilistic diagnosis using a reformulation of the Internist-1/QMR knowledge base: I. The probabilistic model and inference algorithms. *Methods of Information in Medicine*, 30, 241-255.
22. Singh, N. and Gisi, M. (1996). Coordinating distributed objects with declarative interfaces, *Proceedings of Coordination 1996*, pp.368-385, *Lecture Notes in Computer Science* 1061, Springer-Verlag.
23. Steels, L. (1990). Components of expertise. *AI Magazine*, 11(2), 28-49.
24. Studer, R., Eriksson, H., Gennari, J., Tu, S., Fensel, D. and Musen, M. (1996). Ontologies and the Configuration of Problem-Solving Methods. *Proceedings of the Tenth Banff Knowledge Acquisition for Knowledge-Bases Systems Workshop*, Banff, CA.
25. Tu, S. W., and Musen, M. A. (1996). Episodic refinement of ESPR. *Proceedings of the Tenth Banff Knowledge Acquisition for Knowledge-Bases Systems Workshop*, Banff, CA.

Three New Versions of Numerically Stable Fast Least Squares Algorithms (NS-FRLS) for Adaptive Filtering

M.Arezki^(*), A.Benallal^(*), A.Guessoum^(*), D.Berkani^(**)
 (*) : LATSI Laboratory - Department of Electronic , Faculty of Engineering
 Saad Dahlab University, BLIDA, ALGERIA. md.arezki@caramail.com
 (**) : Signal & Communication Laboratory - (ENP), ALGIERS, ALGERIA.

Key words : RLS, Propagation of Errors, Numerical Stability.

Abstract

In this paper, we present three new versions of numerically stable fast recursive least squares algorithms (NS-FRLS). These new versions are obtained by using some redundant formulas of the fast recursive least squares algorithms FRLS. Numerical stabilization is realized by using a propagation model of first order of the numerical errors. The simulation over very long duration for stationary signal did not reveal any tendency to divergence.

1. Introduction

The main identification bloc diagram of a linear system with finite impulse response (FIR), by adaptive filtering using an adaptation algorithm, is represented in figure 1.

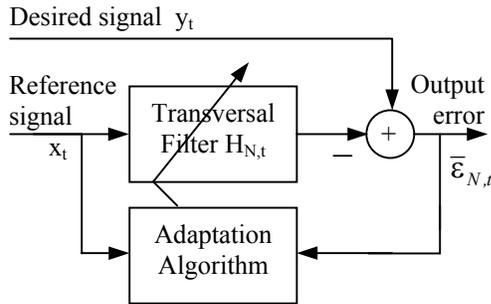


Figure 1 : Main bloc diagram of an adaptive filter.

The output a priori error $\bar{\epsilon}_{N,t}$ of this system is :

$$\bar{\epsilon}_{N,t} = y_t - H_{N,t}^T X_{N,t} \quad (1)$$

where t is the discrete-time index, $H_{N,t}$ is the vector column which gathers the N values of the system impulse response to be identified, $X_{N,t}$ is a vector which summarizes the past of the reference signal x_t on N points and y_t is the desired signal.

The exponent T symbolizes the transposition operator.

The filter $H_{N,t}$ is calculated by minimizing the least squares criterion according to :

$$J_{N,t} = \sum_{i=1}^t \lambda^{t-i} (y_i - H_{N,t}^T X_{N,i})^2 \quad (2)$$

where λ denotes the exponential forgetting factor ($0 < \lambda \leq 1$). The recursive solution is written as follow :

$$H_{N,t} = H_{N,t-1} + G_{N,t} \bar{\epsilon}_{N,t} \quad (3)$$

where $G_{N,t}$ represents the adaptation gain, given by :

$$G_{N,t} = \gamma_{N,t} \tilde{K}_{N,t}$$

where variables $\gamma_{N,t}$ and $\tilde{K}_{N,t}$ respectively indicate the likelihood variable and normalized Kalman gain vector. This latter is calculated, independently of the filtering part $H_{N,t}$, by a fast recursive least squares algorithm (FRLS) using a forward/backward linear prediction analysis over the signal x_t [1]. The calculations complexity of FRLS algorithm is of order N . This reduction of complexity, compared to that of recursive least squares algorithms (RLS, complexity of order N^2), made all algorithms FRLS numerically unstable. Several numerical solutions of stabilization, with stationary signals, are proposed in the literature [2],[3],[4]. Recently, other versions of numerically stable algorithms FRLS with stationary signals were proposed [5], [6]. In the following section, we propose three new versions of numerically stable algorithms FRLS.

2. New Versions of numerically stable algorithms FRLS

The general principle of the numerical stabilization method remains similar to that proposed in [2]; i.e. to modify the numerical properties of the algorithms without modifying the theoretical behavior of the algorithm. Then to make an analysis of the propagation of the numerical errors to the 1st order to choose the control parameters of numerical stability.

We define a variable " divergence indicator " $\xi_{N,t}$, theoretically null, given by [5]:

$$\xi_{N,t} = \bar{r}_{N,t}^c - [(1 - \mu_s) \bar{r}_{N,t}^{f_0} + \mu_s \bar{r}_{N,t}^{f_1}] \quad (4)$$

where $0 \leq \mu_s \leq 1$.

with

$$\bar{r}_{N,t}^c = x_{t-N} - B_{N,t-1}^T X_{N,t} \quad (5.a)$$

$$\bar{r}_{N,t}^{f_0} = \lambda \beta_{N,t-1} \tilde{k}_{N+1,t,N+1} \quad (5.b)$$

$$\bar{r}_{N,t}^{f_1} = \lambda^{-N+1} \gamma_{N,t-1} \alpha_{N,t-1} \tilde{k}_{N+1,t,N+1} \quad (5.c)$$

where the equations (5) represent the backward a priori prediction errors, theoretically equal, calculated differently: $\bar{r}_{N,t}^c$ is calculated by the relation of convolution using the backward prediction $B_{N,t}$ and the input signal, $\bar{r}_{N,t}^{f_0}$ is calculated by using in its formula the sum of backward prediction error squares $\beta_{N,t}$ and the component (N+1) of the normalized Kalman gain of order N+1, $\tilde{k}_{N+1,t,N+1}$ and $\bar{r}_{N,t}^{f_1}$ is calculated by using in its formula the likelihood variable $\gamma_{N,t}$, the sum of forward predictor squares $\alpha_{N,t}$ and $\tilde{k}_{N+1,t,N+1}$.

In practice, the variable divergence indicator $\xi_{N,t}$ is never null due to the precision of using machines. This variable which is theoretically null, does not modify the structure of the algorithm. On the other hand, its introduction in an unspecified point of the algorithm modifies its numerical properties.

To stabilize the algorithm, we define three backward a priori prediction errors, theoretically equivalents, which will be used to calculate the variables $B_{N,t}$, $\beta_{N,t}$ and $\gamma_{N,t}$:

$$\bar{r}_{N,t}^B = \bar{r}_{N,t}^c + \mu^B \xi_{N,t} \quad (6.a)$$

$$\bar{r}_{N,t}^\beta = \bar{r}_{N,t}^c + \mu^\beta \xi_{N,t} \quad (6.b)$$

$$\bar{r}_{N,t}^\gamma = \bar{r}_{N,t}^c + \mu^\gamma \xi_{N,t} \quad (6.c)$$

where the scalar parameters $(\mu^B, \mu^\beta, \mu^\gamma)$ allow to modify and control the propagation of the numerical errors in the algorithm.

These algorithms are differentiated by the various ways of calculating the likelihood variable $\gamma_{N,t}$. Stable Versions (SV1), (SV2) and (SV3) obtained are summarized in table 1.

Stable Version 1 (SV1) :

The algorithm respectively calculates the likelihood variable of order N+1 and order N by the following equations [1]:

$$\gamma_{N+1,t} = \frac{\lambda \alpha_{N,t-1}}{\alpha_{N,t}} \gamma_{N,t-1} \quad (7)$$

$$\gamma_{N,t} = \frac{\gamma_{N+1,t}}{1 - \gamma_{N+1,t} \bar{r}_{N,t}^\gamma \tilde{k}_{N+1,t,N+1}} \quad (8)$$

Let us note that $\mu_s = 0$ in equation (4) correspond to version (SV0) given in [2].

There are two other different manners to calculate the likelihood variable, they are given by the following versions :

Stable Version 2 (SV2) :

In this version, the algorithm calculates the likelihood variable of N+1 order by the formula of definition:

$$\gamma_{N+1,t} = \frac{1}{1 + \tilde{K}_{N+1,t}^T X_{N+1,t}} \quad (9)$$

and of order N given by equation (8).

Stable Version 3 (SV3) :

In this version , we avoid the recurrence in the calculation of the likelihood variable by using the definition directly:

$$\gamma_{N,t} = \frac{1}{1 + \tilde{K}_{N,t}^T X_{N,t}} \quad (10)$$

After a propagation analysis of the numerical errors of the 1st order and an asymptotic study of the equations of the propagation of the errors, and for appropriate choices we selected the following control parameters :

$$\mu^B = \mu^\beta = \mu^\gamma = 1 \text{ and } 0 \leq \mu_s \leq 1 \quad (11)$$

Stability conditions on the exponential forgetting factor λ obtained for a signal x_t white and Gaussian are [5]:

$$\lambda > \frac{4N+5}{4N+7} = 1 - \frac{1}{2N+3.5} \text{ (SV1) and (SV2) } \quad (12.a)$$

$$\lambda > \sqrt{\frac{N+1}{N+2}} = \left(1 - \frac{1}{N+2}\right)^{\frac{1}{2}} \cong 1 - \frac{1}{2N+4} \text{ (SV3) } \quad (12.b)$$

These conditions can be written in another simpler form :

$$\lambda = 1 - \frac{1}{pN} \quad (13)$$

where the parameter p is a real strictly greater than 2 to ensure numerical stability.

3. Results of simulation

We are interested here only in the correct operation of the forward/backward prediction part of the modified algorithms because divergences concerns only the prediction part ; the filtering part is robust compared to the numerical implementation, it requires only a forgetting factor lower than 1 [1] and the good working of the adaptation gain (Kalman gain) provided by the prediction part. It is stable if the latter is.

For that, we evaluate the good behavior of the divergence indicator variable $\xi_{N,t}$ and the likelihood variable $\gamma_{N,t}$. We will calculate thereafter the signal-to-noise ratio of the variable $\xi_{N,t}$ given by $\text{SNR}_{\text{dB}}(\xi)$ expressed in dB, calculated by the following equations :

$$\text{SNR}_{\text{dB}}(\xi) = 10 \log_{10} \left(\frac{\sigma_r^2}{\sigma_\xi^2} \right) \quad (14)$$

with

$$\sigma_r^2 = \frac{1}{L} \sum_{i=1}^L \bar{r}_{N,i}^2 ; \sigma_\xi^2 = \frac{1}{L} \sum_{i=1}^L \xi_{N,i}^2 \quad (15)$$

where L denotes the number of samples.

For a suitable choice of the parameters and a stationary input signal, we checked the validity of the numerical stability conditions on λ given by the equation (13). In addition, the no respect of these conditions make diverge the proposed algorithms.

For the selected choice (11) and a stationary input signal x_t and for a forgetting factor under good conditions [5], the stability of these versions was checked on very long sequences and for different orders of filter.

The signal used in our simulation is a white Gaussian noise, sampled at the frequency 16kHz, with mean zero and variance one.

For our simulation we took the choice of the control parameters : $\mu^B = \mu^\beta = \mu^\gamma = 1$.

We compare here the new versions (SV1), (SV2) and (SV3) with the version (SV0) which is version SV1 for $\mu_s = 0$ [2]. Tables 2 and 3 present the results of SNR_{dB} over the variable $N=32$, $N=256$ and $p = 3$ [5]. These results show well that the new proposed versions improve clearly the signal to noise ratio of $\xi_{N,t}$ better than (SV0) for various values of μ_s .

μ_s	0	0.25	0.5	0.75	1
SV1	198.22	198.86	199.79	201.09	200.72
SV2	199.25	199.88	201.10	201.27	198.73
SV3	202.52	203.26	203.85	202.88	198.61

Table 2 : Comparative table of various versions of $\text{SNR}_{\text{dB}}(\xi)$ compared to μ_s , ($N=32$, $p=3$)

μ_s	0	0.25	0.5	0.75	1
SV1	169.06	169.77	170.44	170.84	170.24
SV2	169.37	169.61	170.36	171.12	170.02
SV3	169.53	169.85	171.06	171.33	169.87

Table 3 : Comparative table of various versions of $\text{SNR}_{\text{dB}}(\xi)$ compared to μ_s , ($N=256$, $p=3$)

Figure 2 shows the temporal evolution of variables $\gamma_{N,t}$ and $\xi_{N,t}$ of version 3 (SV3) for a filter of order $N = 256$ and $p=3$. This figure illustrates the stability of this version, where the indicator of divergence remains very weak and the likelihood variable fluctuates around its optimal value.

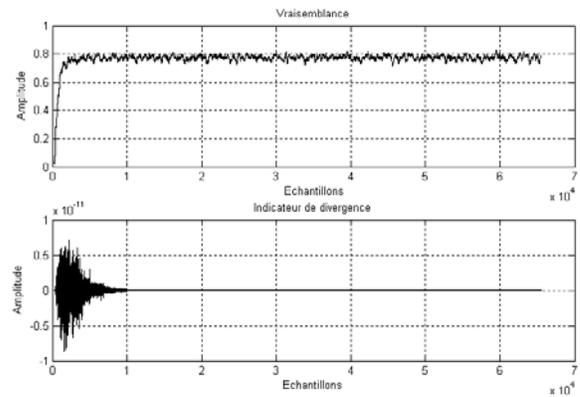


Figure 2 : Temporal evolution of the likelihood variable $\gamma_{N,t}$ and the divergence indicator $\xi_{N,t}$ $N = 256$, $p = 3$ and (SV3) for $\mu_s = 0.75$

4. CONCLUSION

We presented in this paper three new versions of numerically stable fast recursive least squares (NS-FRLS) algorithms. These algorithms thus modified are stable numerically for a suitable choice of the parameters $(\mu^B, \mu^\beta, \mu^\gamma)$ and the forgetting factor λ . These new versions increase the complexity of the initial algorithms slightly: of N additional multiplication for each method. Numerical stability was checked by simulation over very great duration of stationary signal x_t and for very orders of very low and very large filter.

BIBLIOGRAPHY

- [1] J.CIOFFI, T.KALATH "Fast RLS Transversal Filters for adaptive filtering" IEEE press. On ASSP 1984.
- [2] A.BENALLAL, A.GILLOIRE "A New method to stabilize fast RLS algorithms based on a first-order model of the propagation of numerical errors" Proc. ICASSP88, New York 1988.
- [3] D.T.M.Slock and T.Kailath, « Numerically stable fast transversal filters for recursive least squares adaptive filtering », IEEE transactions on signal processing, vol-39, N°.1, January 1991
- [4] A.P.Liavas and P.A.Regalia, « Numerical stability issues of the conventional recursive least squares algorithm, » in Proc. ICASSP, Seattle,WA, May 1998
- [5] M.AREZKI " Study of the fast recursive least squares least Algorithms in the presence of the speech signal " Magister. Blida's university 1997.
- [6]: M.Arezki, A.Guessoum, D.Berkani, «New version of numerically stable algorithm MCR for adaptive filtering», national conference on engineering of electronics, Oran, Algeria, pp.53-55, Decembre 2002.

- Prediction Part

Variables available at the moment t :

$$A_{N,t-1}; B_{N,t-1}; \tilde{K}_{N,t-1}; \gamma_{N,t-1}; \alpha_{N,t-1}; \beta_{N,t-1}; H_{N,t-1}$$

New information : x_t and x_{t-N}

Modeling of x_t and x_{t-N}

$$\bar{e}_{N,t} = x_t - A_{N,t-1}^T X_{N,t-1}$$

$$\alpha_{N,t} = \lambda \alpha_{N,t-1} + \gamma_{N,t-1} \bar{e}_{N,t}^2$$

$$\tilde{K}_{N+1,t} = \begin{bmatrix} 0 \\ \tilde{K}_{N,t-1} \end{bmatrix} + \frac{\bar{e}_{N,t}}{\lambda \alpha_{N,t-1}} \begin{bmatrix} 1 \\ -A_{N,t-1} \end{bmatrix}$$

$$\begin{bmatrix} \tilde{K}_{N,t} \\ 0 \end{bmatrix} = \tilde{K}_{N+1,t} - \tilde{k}_{N+1,t,N+1} \begin{bmatrix} -B_{N,t-1} \\ 1 \end{bmatrix}$$

$$A_{N,t} = A_{N,t-1} + (\gamma_{N,t-1} \tilde{K}_{N,t-1}) \bar{e}_{N,t}$$

$$\bar{r}_{N,t}^c = x_{t-N} - B_{N,t-1}^T X_{N,t}$$

$$\bar{r}_{N,t}^{f_0} = \lambda \beta_{N,t-1} \tilde{k}_{N+1,t,N+1}$$

$$\bar{r}_{N,t}^{f_1} = \lambda^{-N+1} \gamma_{N,t-1} \alpha_{N,t-1} \tilde{k}_{N+1,t,N+1}$$

$$\xi_{N,t} = \bar{r}_{N,t}^c - [(1 - \mu_s) \bar{r}_{N,t}^{f_0} + \mu_s \bar{r}_{N,t}^{f_1}]$$

$$\bar{r}_{N,t}^B = \bar{r}_{N,t}^c + \mu^B \xi_{N,t}$$

$$\bar{r}_{N,t}^\beta = \bar{r}_{N,t}^c + \mu^\beta \xi_{N,t}$$

$$\bar{r}_{N,t}^\gamma = \bar{r}_{N,t}^c + \mu^\gamma \xi_{N,t}$$

$$\gamma_{N+1,t} = \frac{\lambda \alpha_{N,t-1}}{\alpha_{N,t}} \gamma_{N,t-1} \quad (SV1)$$

$$\gamma_{N+1,t} = \frac{1}{1 + \tilde{K}_{N+1,t}^T X_{N+1,t}} \quad (SV2)$$

$$\gamma_{N,t} = \frac{\gamma_{N+1,t}}{1 - \gamma_{N+1,t} \bar{r}_{N,t}^\gamma \tilde{k}_{N+1,t,N+1}} \quad (SV1) \text{ and } (SV2)$$

$$\gamma_{N,t} = \frac{1}{1 + \tilde{K}_{N,t}^T X_{N,t}} \quad (SV3)$$

$$B_{N,t} = B_{N,t-1} + (\gamma_{N,t} \tilde{K}_{N,t}) \bar{r}_{N,t}^B$$

$$\beta_{N,t} = \lambda \beta_{N,t-1} + \gamma_{N,t} (\bar{r}_{N,t}^\beta)^2$$

- Filtering Part

$$\bar{e}_{N,t} = y_t - H_{N,t-1}^T X_{N,t}$$

$$H_{N,t} = H_{N,t-1} + (\gamma_{N,t} \tilde{K}_{N,t}) \bar{e}_{N,t}$$

Table 1 : New versions of numerically stable fast recursive least squares algorithms (NS-FRLS)

Three-pass JPEG-LS Image Compression

Mohammad Fathi

Faculty of Engineering, University of Kurdistan, Iran

Email: fathisam@yahoo.com

Hassan Taheri

M. A. Fasihi

Department of Electrical Engineering, Amirkabir University of Technology

Abstract:

JPEG-LS standard is a context-based lossless image compression scheme, which uses 180° type modeling contexts in one pass scanning an image. In this paper, for capturing high-order dependencies, statistical features and spatial configuration of an image, three-pass JPEG-LS scheme with 360° type modeling contexts is proposed. For context determination, local gradients of intensity which capturing the level of activity (smoothness, edginess) surrounding a pixel, is employed. For images with small dimensions, three-pass scheme and for large dimensions one-pass scheme result higher compression ratios.

1. Introduction

An image compression scheme consists of two components, an encoder and a decoder. The encoder takes an uncompressed image and encodes it into a more compact format. The decoder takes the encoded image and attempts to reconstruct the original uncompressed image. This process will either be lossless, near lossless or lossy. Lossless compression guarantees that the value of each pixel in the reconstructed image will match its corresponding original value. Recent lossless compression schemes have mostly been composed using a two-stage process involving prediction and coding [1], [2].

1.2 Prediction

It is assumed that an image is scanned using a raster-scan technique, that is to say the pixels are scanned from left to right, top to bottom. The common scanning method used by both the encoder and decoder allows the value of current pixel value to be predicted in the same way by the encoder and decoder.

1.2.1 Static Prediction

The first step is to use a fixed prediction function that predicts the value of the current pixel. An example of a prediction function is the Median Adaptive Predictor [3]. Its adaptive nature and low complexity give a good tradeoff between accuracy and speed.

1.2.2 Adaptive Correction

Once a predicted value is obtained from the static predictor(s), a scheme can perform adaptive correction on that prediction. The most straightforward method is to correct the predicted value by the current mean predicted value. This tends to involve "learning" the behavior of the prediction errors over time. Depending on the state of knowledge at that particular time, the encoder and decoder can correct the predicted value based on previous observed behavior. In principle, maximum-likelihood (ML) estimation of adaptive correction value would dictate a procedure based on the median of the prediction errors incurred so far in the context by the static predictor(s) [4].

2. Context Determinations and Modeling in JPEG-LS Standard

The context that conditions the encoding of the current prediction residual for pixel x in JPEG-LS is built out of the following differences.

c	b	d
a	x	

Figure 1: JPEG-LS context

$$g_1 = d - b; \quad g_2 = b - c; \quad g_3 = c - a; \quad (1)$$

These differences represent the local gradient, thus capturing the level of activity (smoothness, edginess) surrounding a sample, which governs the statistical behavior of prediction errors. By symmetry, g_1, g_2 and g_3 influence the model in the same way. Since further model size reduction is obviously needed, each difference $g_i, i=1,2,3$ is quantized into a small number of approximately equiprobable, connected regions by a quantizer $k(\cdot)$ independent of i [5]. In principle, the number of regions into which each context difference is quantized should be adaptively optimized. However, the low complexity requirement dictates a fixed number of "equiprobable" regions. To preserve symmetry, the regions are indexed $qi=k(gi)=-k(-gi) \quad i=1,2,3$.

For JPEG-LS, $T = 4$ was selected [6]. This number balances storage requirements (which are roughly proportional to the number of contexts) with high-order conditioning. To complete the definition of the contexts in JPEG-LS, it remains to specify the boundaries between quantization regions. For an 8-bit/sample alphabet, the default quantization regions are $\{0\}, \pm\{1,2\}, \pm\{3,4,5,6\}, \pm\{7,8 \dots 20\}, \pm\{e|e \geq 21\}$.

3. Three-pass Interlaced Predictive Coding Scheme

The majority of the current lossless image compression methods code the pixels in the raster scan order. As a result, the contexts available for image modeling cannot spatially enclose the modeled pixels. At any moment, only the pixels at the top and to the left of a pixel being coded are known to the both encoder and decoder so that they can be used in modeling and prediction which called 180° type context modeling, since the modeling pixels only form a semicircle around a modeled pixel. Many image features, such as the intensity gradient, edge orientation, and textures, can be better modeled in a completely enclosing contexts [7] called 360° type context modeling. The encoding, and accordingly the decoding, of an image is done in three passes. Each pass uses an interlaced sampling of the original image. Denote a continuous-tone image of width W and height H by $I(i, j), 0 \leq i \leq W, 0 \leq j \leq H$.

The first pass encodes a sub sampled $W/2 \times H/2$ image denoted by η , with the following relation between I and η :

$$\mu(i, j) = \left\lfloor \frac{I(2i, 2j) + I(2i+1, 2j+1)}{2} \right\rfloor$$

$$0 \leq i \leq W/2, 0 \leq j \leq H/2 \quad (2)$$

$\mu(j, i)$ is average intensity of two diagonally adjacent pixels. This sub sampling is designed to benefit prediction and the context modeling in the subsequent passes. The $W/2 \times H/2$ image μ is encoded in raster scan sequence using a 180° type context as shown in Fig.2.

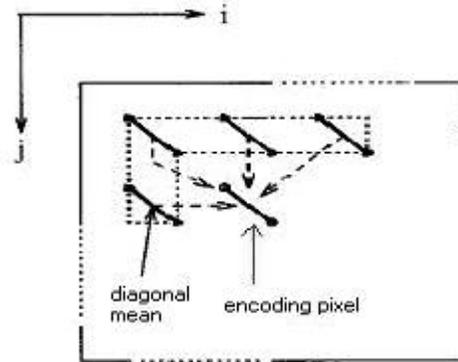


Figure 2: first pass

The prediction function is median estimation.

The second pass uses sub-sampled image as the prediction contexts to encode $NH/2$ pixels:

$$I(2i, 2j), I(2i+1, 2j+1)$$

$$, 0 \leq i \leq W/2, 0 \leq j \leq H/2 \quad (3)$$

Namely, the same pixels involved in the diagonal means are coded by the first pass. But in the second pass, individual pixel values will be resolved from the corresponding diagonal means. Again, the second pass proceeds in raster scan sequence. First, consider the encoding of $I(2i, 2j)$. Fig.3 is a snapshot of the second pass in which the X marks the pixel being currently coded, and the diagonals are the two-pixel means from the first pass.

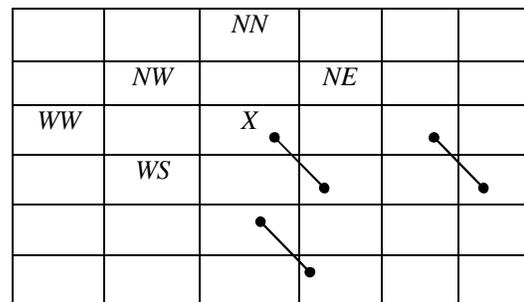


Figure 3: second pass

As shown by Fig.3, at this stage the sub sampled image η from the first pass is available to the right

and bottom of the current pixel, and the previously coded pixels in the second pass are available to the left and top. They provide 360° type contexts surrounding $I(2i, 2j)$. Once $I(2i, 2j)$ is reconstructed, the decoder can set

$$I(2i+1, 2j+1) = 2\mu(i, j) - I(2i, 2j)$$

without receiving any information on $I(2i+1, 2j+1)$.

The third pass encodes the remaining half of the original image. Namely, pixels interlaced in the checkerboard pattern, $I(2i+1, 2j+1)$ and $I(2i+1, 2j+1)$. The prediction contexts available to the third pass are spatially enclosing and adjacent to the modeled pixels. If the third pass of the image is also done in the raster scan order, then as illustrated by Fig.4 an unknown pixel x in the third pass can use a 360° type context consisting of all of its four-connected neighbors, and two of its eight-connected neighbors.

<i>NW</i>	<i>N</i>	<i>NE</i>
<i>W</i>	<i>X</i>	<i>E</i>
	<i>S</i>	

Figure 4: third pass

4. Results

Results for implemented schemes, which are one and three pass JPEG-LS schemes, are shown in following table. The comparison factor among implemented schemes is compression ratio (CR) that is calculated by dividing entropy of the original image by entropy of the compressed image. Images have been selected from Dr.S.Barre's web page [8].

Table 1: Comparison of compression schemes

Images	Pixels	1-pass JPEG-LS CR	3-pass JPEG-LS CR
MRI	208*256	1.78	2.86
MR-angio	256*256	1.16	2.34
MR-Knee	256*256	1.76	3.16
MR-Abdomen	256*256	1.72	2.9
CR-Chest	440*440	2.11	2.85
CT-Abdomen	512*512	2.82	2.54
CT-Ankle	512*512	2.84	2.56
Colon	512*512	2.86	2.82
Hip	512*512	2.9	2.88
CR-Abdomen	1976*1576	4.71	3.93
Average CR	-----	2.46	2.88

5. Conclusion

Analyzing the results is done by classifying images into two groups. One group consists of images with dimensions smaller than $512*512$ pixels and another with dimensions equal or larger than $512*512$ pixels. As shown at above table, compression ratios obtained from three-pass compression scheme are higher than one-pass scheme for images in first group. For second group, one-pass scheme have higher ratios.

Because of small data set (pixels) at first group, suitable training of context models and context dilution doesn't take place in one-pass scheme. But three-pass scheme that uses 360° type modeling contexts increases ability of isolation and resolution

between contexts and obtain higher compression ratios than one-pass scheme. Instead, because of higher image scanning in three-pass scheme, they have larger compression time than one-pass scheme, which is not suitable for real time applications. But for images with small dimensions, it is acceptable. By increasing image dimensions, enough data set (pixels) for training context models is provided. In this case, compression ratios of one-pass schemes are increased, so as for second image group, this scheme because of higher ratios and smaller compression time are suitable than three-pass scheme.

Introduction the image compression scheme to decoder, needs encoding excessive data to compressed image, which causes to increase entropy. So the optimal case in selecting image compression schemes, is independency to type and dimensions of images. For this selection, average of compression ratios for each scheme is calculated. As shown, three-pass JPEG-LS scheme has larger and is suitable for compression applications, independent from type and dimensions of images.

References

- [1] X. Wu, N. D. Memon, "Context-based, adaptive, lossless image coding", *IEEE Trans. on Communication*, Vol. 45, No. 4, pp. 437-444, Apr. 1997.
- [2] M. J. Weinberger, G. Seroussi, and G. Sapiro, "LOCO-I: A low complexity, context-based, lossless image compression algorithm", in *Proc. 1996 Data Compression Conference*, (Snowbird, Utah, USA), pp. 140-149, Mar. 1996.
- [3] M. J. Weinberger, J. Rissanen, "Applications of universal context modeling to lossless compression of gray-scale images," *IEEE Trans. on Image Processing*, Vol. 5, pp. 575-586, Apr. 1996.
- [4] M. J. Weinberger, "The LOCO-I Lossless Image Compression Algorithms: Principles and Standardization into JPEG-LS", *IEEE Trans. on Image Processing*, Vol. 9, No. 8, pp. 1309-1324, August 2000.
- [5] M. J. Weinberger, G. Seroussi, "Sequential prediction and ranking in universal context modeling and data compression", *IEEE Trans. on Information Theory*, Vol. 43, pp. 1697-1706, Sept. 1997.
- [6] Information Technology, "Lossless and Near-lossless compression of continuous-tone still images", 1999.ISO/IEC 14495-1, ITU Recommend. T.87.
- [7] X. Wu, "Lossless Compression of Continuous-Tone Images via Context Selection, Quantization, and Modeling", *IEEE Trans. on Image Processing*, Vol. 6, No. 5, pp. 656-664, May 1997.
- [8] S.Barre, "Medical Image Samples", available at <http://www.baree.nom.fr/medical/samples>.

Wavelet Footprints Representation of ECG Signals

Ho Tatt Wei and Varun Jeoti*

Electrical & Electronics Engineering Department

Universiti Teknologi PETRONAS, 31750 Tronoh, Perak, Malaysia

Email : (tattwei@hotmail.com, varun_jeoti@petronas.com.my*)

*For correspondence, please contact**

Abstract —Wavelet footprints is a new technique, which exploits the inter-dependence of wavelet coefficients around singularities in order to obtain parsimonious signal representation. In this paper, we report that wavelet footprints can be successfully applied to represent ECG signals more compactly. The ECG signal so reconstructed is accurate and sufficiently denoised. Compression ratios of 6 to 8X are achievable with PRD less than 1%. We also describe and justify the details of our implementation.

Keywords — wavelet footprints, boundary wavelets, boundary extension, ECG compression, ECG denoising

INTRODUCTION

The wavelet transform is a well-known and truly versatile signal processing tool, which has been successfully applied to myriad of problems including signal compression, denoising and classification. Typically in signal compression applications, the general signal trends (low pass features) are efficiently represented by only a few scaling functions while signal transients are represented by wavelets. Improvements in compression quality and performance may be achieved if a relationship between the wavelet coefficients of a signal can be formulated and exploited. Crouse, Nowak and Baraniuk suggested one such relationship, a statistical dependence based on Hidden Markov Models [3].

Wavelet footprints [1] is a recent attempt to exploit the relationship between wavelet coefficients at different resolutions to achieve a sparser signal representation. Wavelet footprints are a new type of over-complete basis functions, derived by exploiting the inter-dependence of wavelet coefficients (at different resolutions) at discontinuities in piecewise polynomial signals. In this paper, we explore the application of wavelet footprints to the representation of ECG signals. Early experimental results show excellent compression ratios while maintaining ECG diagnostic quality.

We present our work in the following structure: first we introduce wavelet footprints in the context of boundary wavelets. Our boundary wavelet construction of wavelet footprints [10][11][12] offers an advantage over the periodic wavelet construction suggested in [1]. The theory of wavelet transform is not reviewed since it has been well-researched for over 10 years within the signal-processing community and many excellent and authoritative texts exist e.g. [5][6].

In section 2, we present results on the performance of the wavelet footprint representation of ECG samples from the MIT-BIH arrhythmia database [7]. We conclude by highlighting some important advantages of ECG representation with wavelet footprints.

OVERVIEW OF WAVELET FOOTPRINTS

In [1], Dragotti and Vetterli showed that there was a dependency between the wavelet coefficients at discontinuity(s) in piecewise polynomial signals. Specifically, by selecting a wavelet with high enough vanishing moments, the polynomial portion of piecewise polynomials can be entirely represented by scaling functions, leaving only a few non-zero wavelet coefficients to represent the discontinuity(s). These non-zero wavelet coefficients will be clustered around the translational location(s) of the discontinuity(s) - called the “cone of influence” in [1].

A wavelet footprint, by definition, is the scale-space vector obtained by gathering together all the wavelet coefficients in one cone of influence and imposing unity norm. For a discontinuity in signal described by upto D^{th} order piecewise polynomials, the wavelet coefficients have only $(D+1)$ degrees of freedom; hence the discontinuity can be characterized by a linear combination of footprints spanning a subspace of dimension $(D+1)$. To characterize any arbitrary piecewise polynomial of finite-length, L and with maximum polynomial order n , we require the L translates of the scaling function and a dictionary of $(n+1)L$ wavelet footprints. This collection of scaling functions and wavelet footprints is called the footprint basis.

Wavelet Footprints Definition with boundary wavelets

Let the wavelet representation for some signal, $x(n)$ be defined as

$$x(n) = \sum_{l_b} c_{l_b} \Phi_b(n-l) + \sum_l c_l \Phi(n-l) + \sum_{l_b} \sum_{j_b} d_{j_b l_b} \Psi_b(2^j n-l) + \sum_l \sum_j d_{jl} \Psi(2^j n-l) \quad (1)$$

where Φ_b and Ψ_b are boundary scaling and wavelet functions and Φ and Ψ are the usual scaling and wavelet functions. Let $x(n)$ also be a piecewise polynomial with a single discontinuity at k defined mathematically [1] as

$$x(n) = \sum_{d=0}^D \alpha_0^{(d)} P^{(d)}(n) + \sum_{d=0}^D \alpha_k^{(d)} T_k^{(d)}(n) \quad (2)$$

where the basic polynomial pieces are defined as

$$P^{(d)}(n) = n^d, \quad (3a)$$

$$T_k^{(d)}(n) = 1_{[k, N-1]}(n-k+1)^d, \quad (3b)$$

$$d = 0, 1, \dots, D, \quad n \in [0, N-1];$$

If we know the maximum order polynomial component in $x(n)$, namely D , we must select a scaling function, Φ , whose vanishing moments equal or exceed $(D+1)$. This means that the corresponding wavelet, Ψ , must have at least $(D+1)$ vanishing moments. Then the wavelet domain representation of $x(n)$ becomes

$$x(n) = \sum_{l_b} c_{l_b} \Phi_b(n-l) + \sum_l c_l \Phi(n-l) + \sum_{l_k \in j, l} d_{jl} \Psi(2^j n-l) \quad (4)$$

where d_{jl} are the sparser non-zero wavelet coefficients within the ‘‘cone of influence’’ of discontinuity at k . The wavelet coefficients are zero elsewhere because the polynomial part is completely captured by the scaling functions, Φ and Φ_b with vanishing moments $(D+1)$ and there are no other discontinuities elsewhere. In [11], we show how Φ_b can be easily derived from Φ .

By definition,

$$\sum_{d=0}^D \alpha_k^{(d)} f_k^{(d)} = \sum_{l_k \in j, l} d_{jl} \Psi(2^j n-l) \quad (5)$$

where $f_k^{(d)}$ is the wavelet footprint to represent the discontinuity at k between the d^{th} order polynomial components in $x(n)$. Unlike [1], $k \neq 0$ because there are no boundary discontinuities. Readers may like to refer

to [1] for the definition of wavelet footprints in the context of periodic wavelets.

NUMERICAL EXPERIMENTS WITH ECG SIGNALS

In this section, we show the performance of wavelet footprints representation of selected ECG recordings from MIT-BIH arrhythmia database [7] as obtained by our numerical experiments with MATLAB. The ECG signals are each sampled at 360Hz with 11-bit resolution. For comparison purposes, three metrics are calculated to quantify the fidelity of the signal reconstructed from the wavelet footprint representation namely, the Correlation (CC), Percent RMS Difference (PRD) and Standardized RMS Difference (SRD). The following formula used are from Ishijima [4]:

$$\text{Correlation, (CC)} = \frac{\frac{1}{N} \sum_{i=1}^N (u(i) - \bar{u})(v(i) - \bar{v})}{\sigma_u \sigma_v} \quad (6a)$$

$$\text{Percent RMS difference (PRD)} = \frac{RMS_{Error}}{RMS_{Value}} \times 100\% \quad (6b)$$

$$\text{Standardized RMS difference (SRD)} = \frac{RMS_{Error}}{\sigma_u} \times 100\% \quad (6c)$$

where N is the length of the ECG signal, $u(i)$ are samples of original ECG signal, $v(i)$ are samples of the reconstructed signal (the time series constructed from the footprint representation), σ_u the standard deviation of signal $u(n)$ and the Root Mean Square (RMS) values are defined as follows:

$$RMS_{Value} = \sqrt{\frac{1}{N} \sum_{i=1}^N u(i)^2}, \quad (6d)$$

$$RMS_{Error} = \sqrt{\frac{1}{N} \sum_{i=1}^N (u(i) - v(i))^2} \quad (6e)$$

Matching Pursuit Algorithm to obtain the Footprint Representation

To obtain the wavelet footprint representation, we have used the Matching Pursuit algorithm suggested in [1]. One iteration of the Matching Pursuit algorithm detects the footprint components present in a signal by correlating the signal with every footprint in the footprint dictionary and extracting the footprints at the location, which gives the greatest correlation value. Several iterations of the algorithm are required to arrive at a refined approximation. However, the algorithm has no guarantee of convergence; after some iterations, the algorithm may diverge to give successively poorer approximations.

Implementation Details

ECG signals have a strong polynomial nature and are quasi-periodic with near-similar PQRST (see [8] for details on ECG signals) waves repeating approximately every second. We devised and used a simple strategy to separate the P and T waves (low order polynomial) from the QRS waves. From experiments, it is estimated that the P and T waves have maximum polynomial components of order 3 or 4. The ‘Daubechies Least Asymmetrical Wavelet’ of 5 vanishing moments [9] was employed in the construction of footprints.

In our implementation, the ECG signal is first segmented into short frames of approximately one cardiac cycle (one PQRST wave). For 360Hz sampled ECG signals, this corresponds to frame lengths of 512 points (frame lengths should ideally be powers of 2). For ECGs of different sampling rates, the frame length must be appropriately adjusted.

The frames are then individually processed to obtain the wavelet footprint representation by first performing a discrete wavelet transform (DWT) to obtain the scaling function components, removing the scaling function components from the original ECG and subsequently applying the Matching (Subspace) Pursuit algorithm [1] on the time series of the residual ECG signal. A DWT with boundary wavelets from zeroth order polynomial extrapolation was experimentally found to yield the best results.

The frame length greatly influences the compression power of wavelet footprints. If too short a length is chosen, too many redundant scaling coefficients are required and reconstruction quality possibly degrades. Conversely, too long a frame requires more footprints of higher orders for accurate representation; this increases computational effort without guaranteeing equal or better reconstruction quality.

Matching Pursuit (MP) algorithm was used because of its simplicity. Although the algorithm is not guaranteed to converge to an accurate footprint representation, we have found that 3 to 7 iterations are suitable for the ECG signals used.

Processing times

The transformation to footprint representation was performed with MATLAB on an Intel Pentium III 667MHz PC with 128MB RAM and Windows 2000 environment. Average processing time to obtain a footprint representation of a 10-minute ECG (216,000 samples) is 23 minutes. The footprint dictionary is separately constructed and typically requires 10 to 15 minutes to generate. Processing time will be shorter with code optimization.

Advantages

1) Improved Compression Ratios

For comparison of our results, we compare the footprint representation scheme to the performance of the SPIHT technique [2] of Z.Lu, D.Y. Kim and W.A. Pearlman. Lu et al provide a comprehensive comparison of SPIHT with existing ECG compression techniques and report that SPIHT outperforms these popular methods in terms of PRD for their chosen dataset from the MIT-BIH arrhythmia database [7].

Table 1: Performance of the SPIHT coding algorithm (reproduced from [2]).

	Dataset 1			Dataset 2		
Compression	4:1	5:1	8:1	4:1	6.6:1	8:1
PRD	1.19	1.56	2.46	1.11	2.04	2.50

We have used the same datasets to test the performance of our wavelet footprint representation i.e. first 10 minutes of lead 1 signals from records 101, 102, 103, 107, 118, 119, x109, x111, x115 and x117 for Dataset 1 and from 104, 201, 207, 208, 209, 212, 213, 214, x112, x116, x228, x231 and x232 for Dataset 2. The average performance results of our wavelet footprint representation are summarized in the following table.

Table 2: Performance of the Footprint Representation Scheme

	Dataset 1		Dataset 2
Compression ¹	6.73:1	8.26:1	8.26:1
PRD	0.6377	0.7966	0.8556
SRD	8.3176	10.3150	12.1124
CC	0.9963	0.9943	0.9914

Clearly, our wavelet footprint representation scheme outperforms SPIHT in both reconstruction quality and compression ratio by almost 2X! Note that this compacted representation is achieved by a simple transformation to the footprint representation. We achieve these compression rates by simple Pulse Coded Modulation (PCM) coding of the footprint coefficients, scaling coefficients and footprint locations (see Figure 1). Further compression is possible by application of other bit-manipulation techniques².

¹ Compression ratios of 6.73:1 was obtained by iterating the Matching Pursuit algorithm 5 times whereas ratio of 8.26:1 was obtained by 3 Matching Pursuit iterations.

² We have achieved greater compression ratios (upto 13:1) at negligible quality loss by introducing a quantization and entropy coding scheme [12]. This work is being compiled for future publication.

The compression ratio of our method is calculated as follows:

size (footprint coefficients)

= MP iterations * (D+1)*no. of frames

size (scaling coefficients)

= size (scaling coefficients per frame)* no. of frames

size (footprint locations) = MP iterations* no. of frames

C = size (footprint coefficients) + size (scaling)

Figure 1: Simplified Calculation of compression ratio

2) Denoising

The wavelet footprint representation was found to be an excellent denoising tool as well. We illustrate the denoising capability by appending below the 1st 2048 samples of the original ECG record x109 and the footprint reconstructed version (from 3 iterations of MP with 8.26X compression).

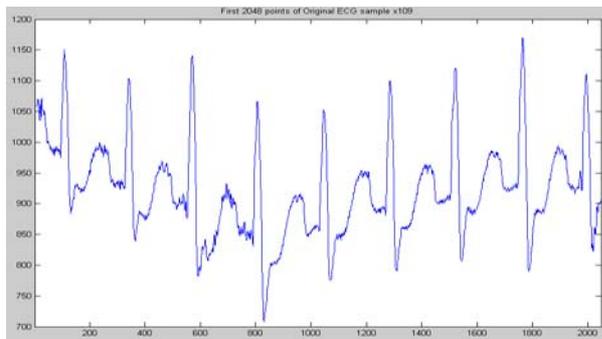


Figure 2: Original 2048 samples from MIT-BIH arrhythmia record x109

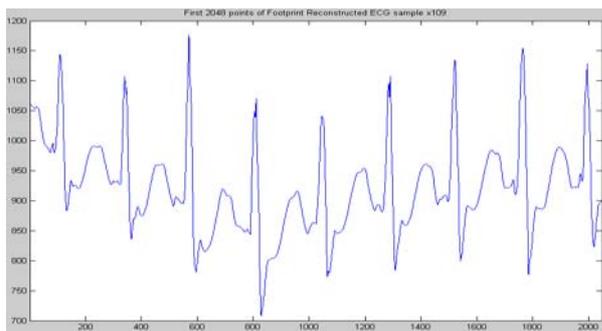


Figure 3: Footprint reconstructed version of first 2048 samples for MIT-BIH arrhythmia record x109

CONCLUSION

The wavelet footprint representation of ECGs is a promising compression technique that offers high compression rates, accurate signal reconstruction and effective signal denoising. In this paper, we have described an implementation method to obtain the wavelet footprint representation of 360Hz 11bit ECG samples with footprints from ‘Daubechies Least

Asymmetrical orthogonal wavelets’ with 5 vanishing moments [9]. To construct the footprints, we have alluded to our new method using boundary wavelets [11]. For other ECG signals sampled differently, minor modifications to the scheme based on guides presented in this paper should yield identical results. Experiments show that compression of 6 to 8 times are possible with PRD of less than 1%. However, our algorithm implementation is computationally intensive. We are confident of improving the compression rates and processing time in subsequent work [10][12].

Acknowledgements

The authors would like to thank Dr. Abdul Huq, Universiti Teknologi PETRONAS, Malaysia for useful discussions.

References

- [1] P. L. Dragotti, M. Vetterli, May 2003, “Wavelet footprints: theory, algorithms and applications,” *IEEE Trans. On Signal Processing*, 51(5):1306-1323
- [2] Z. Lu, D.Y. Kim, W.A. Pearlman, Jan 2000, Wavelet Compression of ECG Signals by Set Partitioning in Hierarchical Trees (SPIHT) Algorithm, *IEEE Trans. on Biomedical Engineering*.
- [3] Matthew S. Crouse, Richard G. Baraniuk, Robert D. Nowak, April 1998, “Wavelet Based Statistical Signal Processing Using Hidden Markov Models,” *IEEE Trans. on Signal Processing* 46 (4) : 886-902
- [4] M.Ishijima, Dec 1993, Fundamentals of the Decision of Optimal Factors in the ECG Data Compression, *IEICE Trans. in Information & Systems*, E76-D (12) : 1398- 1403
- [5] G. Strang & T. Nguyen, 1997, *Wavelets & Filter Banks*, Massachusetts, Wesley-Cambridge Press
- [6] S. Mallat 1999, A Wavelet Tour of Signal Processing 2nd edition, San Diego., Academic Press
- [7] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 101(23):e215-e220
- [8] J. Willis Hurst, 2001, *Interpreting Electrocardiograms Using Basic Principles and Vector Concepts*, Marcel Dekker, New York.
- [9] I. Daubechies, 1992, *Ten Lectures on Wavelets*, Philadelphia, Society for Industrial and Applied Mathematics (SIAM)
- [10] Ho Tatt Wei and Varun Jeoti, *ECG Signal representation using Wavelet Footprints*, to be communicated to *IEEE Trans. on Signal Processing*.
- [11] Ho Tatt Wei and Varun Jeoti, *Construction of Wavelet Footprints with Boundary Wavelets*, to be communicated.
- [12] Ho Tatt Wei and Varun Jeoti, *A Wavelet Footprint based ECG Compression Scheme*, to be communicated.

WEIERSTRASS APPROACH TO NONLINEAR BLIND SIGNAL SEPARATION

P. Gao, W.L. Woo and S.S. Dlay

School of Electrical, Electronic and Computer Engineering,

University of Newcastle upon Tyne,

NE1 7RU, Newcastle, United Kingdom

Email: pei.gao@ncl.ac.uk, w.l.woo@ncl.ac.uk and s.s.dlay@ncl.ac.uk

Abstract: *The fundamental objective of Independent Component Analysis (ICA) is to recover a set of statistically independent components from the output of the mixing system. However, most algorithms could achieve excellent performance only under linear assumption, which may not accord with practical situations. In this paper, a mathematical model of the nonlinear mixing system stemmed from the theory of functional analysis is investigated, achieving the formulation of a stable unique inverse solution. In addition, the optimum framework for the nonlinear demixer can be obtained directly from the derived mixing model. Furthermore, a novel method to acquire the desired solution is proposed using the Weierstrass approach, which displays high flexibility of nonlinearity. Finally, simulation results are presented to verify the proposed scheme.*

1. Introduction

Surging as the frontier of many signal restoration research avenues, independent Component Analysis (ICA) aims to extract the original source signals, which is assumed to be statistically independent, from only contaminated observations. It is closely relative to blind source separation and therefore attracts more and more attention for its various applications ([1] and references therein). However, most existing ICA algorithms have focused on linear distortion which may not accord with practical applications. For example, in biomedical cases, on account of the fact that many physiological signals suffer from the nonlinear distortion, e.g. the auditory nervous system is modelled as memoryless nonlinearity, the identification of nonlinear dynamics should be taken into consideration. Another instance is the recording of multiple speech sources by microphones, which incur nonlinearity by class-C amplifiers [2]. Therefore, for nonlinear mixing model, linear algorithms may fail to extract original signals and become infeasible, since there is no guarantee that the solution at the outputs of any nonlinear demixing system will correspond to the desired source signals even when the Kullback-Leibler Divergence is minimal, which is commonly used in linear ICA. In this paper, combined with a mono-nonlinear ICA model stemmed from the theory of the

functional analysis, we provide the nonlinear solution by using the Weierstrass Approach which can restore nonlinearly mixed signals uniquely in a blind mode.

2. Nonlinear Mixing and Demixing Model for ICA

Although linear model for ICA mixture is assumed previously in most of the literature by virtue of its simplicity to understand and analyse, a realistic mixture should be nonlinear and concurrently capable of treating the linear mixture as a special case for the sake of practical applications mentioned in Section 1. Starting from the following lemma in the theory of functional analysis [3], the nonlinear mixing model is set up by applying it into Blind Source Separation (BSS) situation.

Given two continuous group operations F and G , there exists a strictly monotonic continuous function f in which the equation $F[f(x), f(y)] = f[G(x, y)]$ holds. As a consequence, in nonlinear BSS case, the observed signals can be expressed as follows:

$$\begin{aligned} x_i &= F[s_1, s_2, \dots, s_q] \\ &= F[f(\mathbf{d}_1), f(\mathbf{d}_2), \dots, f(\mathbf{d}_q)] = f[G[\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_q]] \end{aligned}$$

where $s_j = f(\mathbf{d}_j)$ and $j=1, 2, \dots, q$. Assuming $G(\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_q) = m_{i1}\mathbf{d}_1 + m_{i2}\mathbf{d}_2 + m_{i3}\mathbf{d}_3 + \dots + m_{iq}\mathbf{d}_q$, the nonlinear mixture can be therefore modelled as

$$\begin{aligned} x_i &= F[s_1, s_2, \dots, s_q] = f[m_{i1}\mathbf{d}_1 + m_{i2}\mathbf{d}_2 + \dots + m_{iq}\mathbf{d}_q] \\ &= f[m_{i1}f^{-1}(s_1) + m_{i2}f^{-1}(s_2) + \dots + m_{iq}f^{-1}(s_q)] \\ &= f[\mathbf{m}_i^T f^{-1}(\mathbf{s})] \end{aligned}$$

Hence,

$$\mathbf{x} = \begin{bmatrix} f(\mathbf{m}_1^T f^{-1}(\mathbf{s})) \\ f(\mathbf{m}_2^T f^{-1}(\mathbf{s})) \\ \vdots \\ f(\mathbf{m}_p^T f^{-1}(\mathbf{s})) \end{bmatrix} = f(\mathbf{M} f^{-1}(\mathbf{s})) \quad (1)$$

where $\mathbf{M} = [\mathbf{m}_1 \ \mathbf{m}_2 \ \dots \ \mathbf{m}_p]^T$ with dimension $p \times q$ and $\mathbf{m}_i = [m_{i1} \ m_{i2} \ \dots \ m_{iq}]^T$. From (1), we recognise

that the nonlinear mixture is fundamentally a synthesis of two nonlinear functions, one of which is the inverse of the other and the model will reduce to a linear case when $f(u) = u$. Furthermore, according to the maximum likelihood estimation, the optimal solution for the demixer system is found to be

$$\hat{\mathbf{s}} = f[\mathbf{M}^{-1}f^{-1}(\mathbf{x})] = f[\mathbf{W}f^{-1}(\mathbf{x})] \quad (2)$$

Therefore, the above demixing system can be regarded as a simple implementation of the cascaded ‘linear-nonlinear-linear’ model given by the multilayer perceptron neural network [2], not only signifying its flexibility in nonlinear systems but also capable of handling the linear model as a special case. Moreover, due to the constraints on the nonlinear function f and mono-nonlinearity, the model tends to produce unique solution for nonlinear ICA.

3. Weierstrass Demixing System

Nonlinear demixers such as the Self-Organising Map (SOM), Radial Basis Function (RBF) [4] and FMLP with sigmoidal nonlinearity [2] are intrinsically nonlinear with the utilisation of fixed nonlinearity in the hidden neurons, tending to cause an oversized network, which inevitably subjects to huge computational complexity. Moreover, to obviate the need of generation of non-unique independent outputs the demixer is required to control the inherent capability from ‘overfitting’. To achieve this, the use of Weierstrass function expansion is required in order to generate flexible nonlinearity and to avoid non-unique independent outputs. In this section, the theoretical fundamental of using Weierstrass Network demixer is firstly described. The cost function is then formulated and the learning rules are derived for updating the parameters of the demixer. The obtained learning rules are then modified to incorporate the series reversion to mimic the desired structure as developed in Section 2.

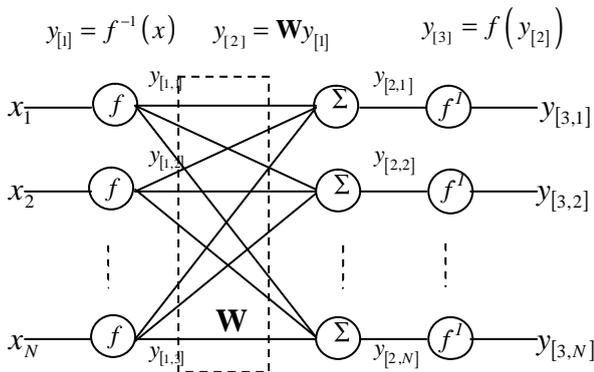


Figure 1: 3-layer Weierstrass Network as the nonlinear demixer

In the Weierstrass Approximation theorem [5], it has been pointed out that for every continuous function $f : [c, d] \rightarrow \mathbf{R}$, there always exists a

Weierstrass sequence $p(Np, a, u)$ which can uniformly approximate f with arbitrary accuracy.

$$p(Np, a, u) = a_0 + a_1 u + a_2 u^2 + \dots = \sum_{m=0}^{Np} a_m u^m$$

$$\forall \epsilon > 0, \exists Np > N(\epsilon), \forall u \in [c, d]: |f(u) - p(u)| < \epsilon$$

Hence, a general demixing system based on Weierstrass theory is proposed as shown in Figure 1 whose hidden layer neurons perform Weierstrass sequence to approximate the mixing mapping function $f(u)$ and $f^{-1}(u)$. According to the structure of Weierstrass Network, the output of the demixing system assumes the form of

$$y_{[3]} = f(y_{[2]}) = \sum_{m=0}^{Np_1} a_m y_{[2]}^m$$

$$y_{[2]} = \mathbf{W}y_{[1]}$$

$$y_{[1]} = f^{-1}(\mathbf{x}) = \sum_{n=1}^{Np_2} b_n (\mathbf{x} - a_0)^n \quad (3)$$

where a_m 's and b_n 's are the coefficients of the forward and reverse Weierstrass functions. For simplicity but without lost of generalisation, we assume that the number of the sources and the observations are equal to N as elucidated in the following sections.

3.1 Cost Function

Generally speaking, the goal of the demixer system is to obtain the estimated signals with independent entries. The original cost function rooted from the Kullback-Leibler Divergence (KLD) [1] is especially commonly used in linear source separation problem. In order to regulate the convergence of the network into a unique solution and reduce the effect of the nonlinear distortion, it is usually modified with additive set of constraints and expressed as follows:

$$J = -\log \left| \det \frac{dy_{[3]}}{d\mathbf{x}^T} \right| - \sum_{i=1}^N \log(c_i(y_{[3,i]})) + \sum_{i=1}^N \underbrace{\beta_i f_i^{(c)}(y_{[3,i]}, s_i)}_{\text{additive constraints}}$$

$$f_i^{(c)}(y_{[3,i]}, s_i) = \sum_{j=1}^{N_c} \left[\text{cum} \left(\overbrace{y_{[3,i]} \dots y_{[3,i]}}^{j \text{ times}} \right) - \text{cum} \left(\overbrace{s_i \dots s_i}^{j \text{ times}} \right) \right]^2 \quad (4)$$

where $c_i(y_{[3,i]})$ represents the marginal probability function (pdf) of the i^{th} estimated signal at the output of the Weierstrass demixer, and β_i 's are set of constants to control the importance of the additional constraints. It can be inferred that these constraints imply the *a priori* information about the source distributions in the form of cumulants which are intended to match the outputs with the original source signals.

3.2 Series Reversion

As shown in Figure 1, the implementation of the proposed demixer requires the reverse function of the Weierstrass sequence. According to the theory of

Series Reversion, if the function $g(\cdot)$ has a power series expression as $g(u) = a_1 u + a_2 u^2 + a_3 u^3 \dots$, then the reverse function will also have the similar form of $g^{-1}(u) = b_1 u + b_2 u^2 + b_3 u^3 \dots$ with the coefficients calculated by [6]

$$b_n = \frac{1}{a_1^n} \sum_{s,t,u,\dots} \left\{ (-1)^{s+t+u+\dots} \frac{(n-1+s+t+\dots)!}{(n-1)! s! t! u! \dots} \left(\frac{a_2}{a_1} \right)^s \left(\frac{a_3}{a_1} \right)^t \dots \right\} \quad (5)$$

where $s+2t+3u+\dots = n-1$; $s, t, u, \dots \geq 0$. The derivative of b_n with respect to a_m 's can be obtained as

$$db_n = \sum_{j=1}^{N_{p_1}} \left(\sum_{k_2, k_3, \dots}^{N_{p_1}} (-1)^{\sum_{i=2}^{N_{p_1}} k_i} \frac{(n-1+\sum_{i=2}^{N_{p_1}} k_i)!}{n! \prod_{i=2}^{N_{p_1}} (k_i!)} \left(\prod_{\substack{i=1 \\ i \neq j}}^{N_{p_1}} a_i^{k_i} \right) k_j a_j^{k_j-1} \right) da_j \quad (6)$$

where $k_2 + 2k_3 + 3k_4 + \dots = n-1$, $k_i \geq 0$, $i = 2, 3, 4, \dots$ and $k_1 = n - \left(n + \sum_{i=2}^{N_{p_1}} k_i \right)$. Therefore, the reverse function is

updated by using the relationship between $\{b_n\}_{n=1}^{N_{p_2}}$ and $\{a_m\}_{m=1}^{N_{p_1}}$. Furthermore, in the parameter learning algorithm, only $\{a_m\}_{m=1}^{N_{p_1}}$ and \mathbf{W} are the sets of variable parameters that need to be optimised.

3.3 Parameter Learning Algorithm

Assuming that zero-preserving functions are employed in the hidden layer, the generalised stochastic gradient algorithm expressed in (7) searches for the solution in a multidimensional space along the steepest descent direction.

$$\frac{dJ}{d\mathbf{W}} = -\mathbf{W}^{-T} - \left[\sum_{m=1}^{N_{p_1}} m(m-1) a_m \text{diag}(y_{[2]}^{m-2}) \right] F_v^{-1} y_{[1]}^T \quad (7)$$

$$+ \text{diag}(F_v) \mathbf{y} y_{[1]}^T$$

$$\frac{dJ}{da_m} = -F_v^{-T} \left\{ \begin{aligned} & m y_{[2]}^{m-1} \\ & + \left(\sum_{k=1}^{N_{p_1}} k(k-1) a_k \text{diag}(y_{[2]}^{k-2}) \right) \mathbf{W} \left(\sum_{n=1}^{N_{p_2}} \mathbf{x}_{[n,m]} (x-a_0)^n \right) \end{aligned} \right\}$$

$$- G_v^{-T} \left(\sum_{n=1}^{N_{p_2}} n \mathbf{x}_{[n,m]} (x-a_0)^{n-1} \right)$$

$$+ \mathbf{y}^T \left[\left(\sum_{k=1}^{N_{p_1}} k a_k \text{diag}(y_{[2]}^{k-1}) \right) \mathbf{W} \left(\sum_{n=1}^{N_{p_2}} \mathbf{x}_{[n,m]} (x-a_0)^n \right) + y_{[2]}^m \right]$$

where $F_v^{-1} = \begin{bmatrix} 1 & & & \\ & \dots & & \\ & & 1 & \\ & & & \dots \end{bmatrix}^T$,

$$F_v^{-1} = \begin{bmatrix} \frac{1}{\sum_{m=1}^{N_{p_1}} m a_m y_{[2,1]}^{m-1}} & \dots & \frac{1}{\sum_{m=1}^{N_{p_1}} m a_m y_{[2,N]}^{m-1}} \end{bmatrix}^T$$

$$G_v^{-1} = \begin{bmatrix} 1 & & & \\ & \dots & & \\ & & 1 & \\ & & & \dots \end{bmatrix}^T$$

$$G_v^{-1} = \begin{bmatrix} \frac{1}{\sum_{n=1}^{N_{p_2}} n b_n (x_1 - a_0)^{n-1}} & \dots & \frac{1}{\sum_{n=1}^{N_{p_2}} n b_n (x_N - a_0)^{n-1}} \end{bmatrix}^T$$

$$\mathbf{x}_{[n,m]} = \sum_{k_2, k_3, \dots} \left[(-1)^{\sum_{i=2}^{N_{p_1}} k_i} \frac{(n-1+\sum_{i=2}^{N_{p_1}} k_i)!}{n! \prod_{i=2}^{N_{p_1}} (k_i!)} \left(\prod_{\substack{i=1 \\ i \neq j}}^{N_{p_1}} a_i^{k_i} \right) k_m a_m^{k_m-1} \right]$$

$$\mathbf{y} = [\mathbf{y}_1(y_{[3,1]}) \dots \mathbf{y}_N(y_{[3,N]})]^T$$

and $\mathbf{y}_i(y_{[3,i]}) = -\frac{d \log \mathbf{c}_i(y_{[3,i]})}{dy_{[3,i]}}$ which can be estimated

using probability expansion series such as the Gram-Charlier or Edgeworth series. Taking the truncated 4th order Edgeworth series as an example $\mathbf{y}_i(y_{[3,i]})$ is approximated by the following formula:

$$\mathbf{y}_i(y_{[3,i]}) \approx \left(-\frac{1}{2} c_3 + \frac{7}{4} c_3^2 + \frac{3}{4} c_3 c_4 \right) (y_{[3,i]})^2 + \left(-\frac{1}{6} c_4 + \frac{1}{2} c_4^2 \right) (y_{[3,i]})^3 \quad (8)$$

where c_3 and c_4 are the third and fourth order cumulant of $y_{[3,i]}$, respectively.

4. Results

In order to evaluate the performance of the proposed approach, we assume that a general 3rd order Weierstrass function $f(u) = u + a_3 u^3$ is used in the nonlinear mixture model as in (1). The parameter $a_3 \in \mathbf{R}$ controls the amount of the nonlinearity in the function. In the experiment, a_3 is set to 2 and the output of the nonlinear mixing system for the case of 4 sources and 4 sensors can be derived and expressed as

$$x_i = f(m_{i1} f^{-1}(s_1) + m_{i2} f^{-1}(s_2) + m_{i3} f^{-1}(s_3) + m_{i4} f^{-1}(s_4)) = \mathbf{m}_i^T \mathbf{f} + a_3 (\mathbf{m}_i^T \mathbf{f})^3$$

where $\mathbf{f} = [\mathbf{f}_1(s_1) \dots \mathbf{f}_N(s_N)]^T$ and

$$\mathbf{f}_j(s_j) = \frac{1}{6a_3^{1/3}} \left(108s_j + 12 \sqrt{\frac{12 + 81a_3 s_j^2}{a_3}} \right)^{2/3} - 2a_3^{-2/3} \left(108s_j + 12 \sqrt{\frac{12 + 81a_3 s_j^2}{a_3}} \right)^{-2/3}$$

Four sub-gaussian sources are generated synthetically with 25dB white Gaussian noise perturbing each sensor and the parameters $\{m_{ij}\}$ are randomly selected from a Gaussian distribution. The original sources and the outputs of the mixture are displayed in Figure 2 and 3, respectively. The performance of the proposed algorithm will be studied using this mixture, compared with conventional linear methods [1]. The truncated 4th order Edgeworth series is used in the simulation to estimate the marginal entropy since all sources are sub-gaussian distributed.

The results of the restored signals based on linear ICA and gradient descent Weierstrass approach using Eqns. (7) are displayed in Figure 4 and 5, respectively. The step sizes used are set to 0.0005, which is the tradeoff between the stability of the system and the speed of convergence. From these plots, it is observed that the proposed Weierstrass algorithm has successfully recovered the original source signals while the linear method fails to do it under nonlinear mixture. From the theoretical point of view, we assume the Weierstrass function $f(u) = u$ for the linear case. Considering the Taylor series expansion, the linear model admits that the higher order ($=2$) of the expansion is equal to zero and therefore loses the capability of nonlinear analysis. On the other hand, the flexible Weierstrass approach we proposed could adaptively capture the nonlinearity of the mixture using high order nonlinear functions in the expansion, construct the inverse system of the mixture and recover the source signals.

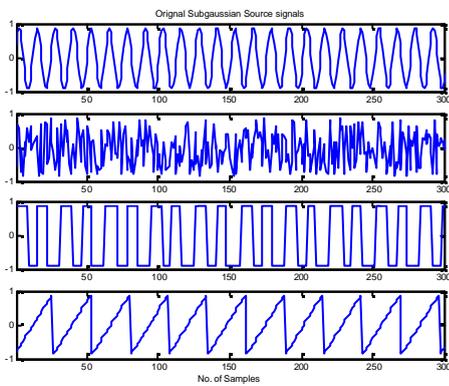


Figure 2: Original source signals.

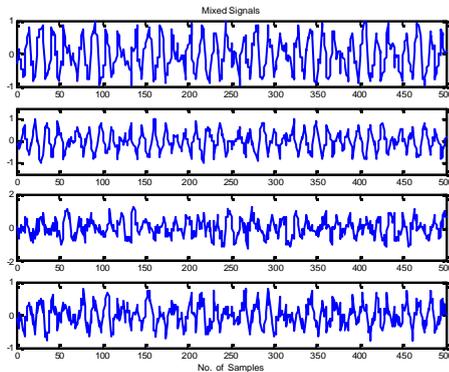


Figure 3: Outputs of the nonlinear mixture.

5. Conclusion

The new Weierstrass network for restoring nonlinearly mixed signals has been proposed with the derivation of the stochastic gradient descent algorithm for training the parameters. In addition, the generalised framework for mono-nonlinear mixing and demixing systems is also described, which reveals the form of a 3-layer multilayer neural network. Finally, the experiment has been conducted and analysed, demonstrating that the proposed algorithm outperforms the conventional

linear ICA method in reconstructing original sources under nonlinear mixture.

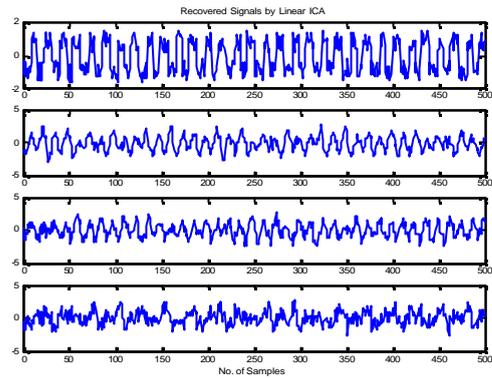


Figure 4: Recovered signals by linear ICA

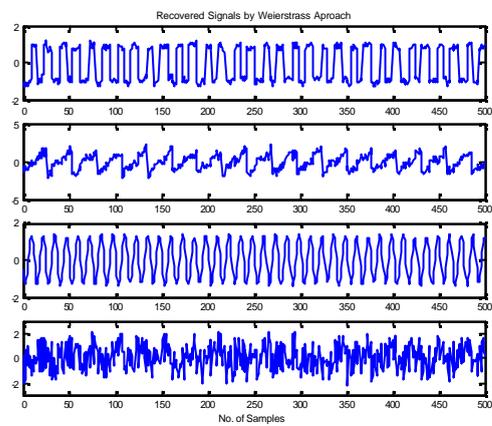


Figure 5: Recovered signals by Weierstrass approach.

6. References

- [1] S. I. Amari and A. Cichocki, "Adaptive blind signal processing - Neural network approaches," *Proceedings of the Ieee*, vol. 86, pp. 2026-2048, Oct. 1998.
- [2] W. L. Woo and S. Sali, "General Multilayer Perceptron Demixer Scheme for Nonlinear Blind Signal Separation," *IEE Proc. On Vision, Image and Signal Processing*, vol. 149, pp. 253-262, Oct. 2002.
- [3] M. Vidyasagar, *Nonlinear System Analysis*, Second Edition ed: Prentice Hall, 1993.
- [4] Y. Tan, J. Wang, and J. Zurada, "Nonlinear Blind Source Separation using a Radial Basis Function Network," *IEEE Transactions on Neural Network*, vol. 12, pp. 124-134, 2001.
- [5] C. D. Cantrell, *Modern Mathematical Methods for Physicists and Engineers*: Cambridge University Press, 2000.
- [6] W. H. Beyer, *CRC Standard Mathematical Tables and Formulae*, 30th ed. ed: Boca Raton, Florida : CRC Press, 1996.

Clock Jitter in Direct Sampling Systems

Theodoros Chalvatzis, Eric Gagnon[†], *Member, IEEE* and James S. Wight[‡], *Senior Member, IEEE*
 Dept of Electrical and Comp Engineering, University of Toronto, Canada (theo@eecg.toronto.edu)
[†]Nortel Networks, Ottawa, Canada

[‡]Dept of Electronics, Carleton University, Ottawa, Canada

Abstract—The effect of clock jitter on sampling systems is analyzed. Analytical expressions are derived for the Signal-to-Noise Ratio using the autocorrelation function and its properties. Special focus is given to direct sampling systems having raised cosine power spectral density.

Index Terms—Autocorrelation, analog-to-digital converter, clock jitter, clock phase noise, direct sampling, IS-95, raised cosine.

I. INTRODUCTION

THE progress in wireless radio systems in the recent years has changed the way circuits are designed and built. New ideas have been implemented and traditional transceiver structures have changed through time. The most widely used architecture, the superheterodyne receiver, is steadily being replaced by more sophisticated circuits that operate at higher data rates.

A typical radio receiver consists of two parts: the analog front-end and the digital back-end. Until recently, the focus of the design was on the analog part having many discrete components many of which were placed off chip (e.g. filters). Advances in VLSI technology enabled the use of fast and low cost digital circuits that can perform many of the functions previously done in the analog part of the circuit. For example we could employ digital downconversion and filtering, thus eliminating the need for expensive off-chip filters, mixers and VCOs. The ultimate goal is reduce the overall cost of wireless basestation by allowing flexibility to the receiver to support various different standards. This calls for higher performance receivers.

The evolution in RF receiver design is shown in Fig. 1. A superheterodyne receiver topology can be seen in Fig. 1(a) where the signal is downconverted from RF to IF. Further downconversion is necessary. On the other hand, only one stage is needed in the Direct Conversion of Zero-IF receiver of Fig. 1(b). However, the inherent DC leakage is difficult to filter out making the topology suitable for CDMA but not GSM where the channel bandwidth is small (30KHz). The receiver is also sensitive to I/Q demodulator mismatch. The Direct Sampling receiver of Fig. 1(c) requires minimum number of analog components since conversion from analog to digital is performed directly at RF. Most functions are carried out in the digital domain.

In a direct sampling receiver the ADC must have such performance so that the full receive bandwidth of the wireless system is supported. This calls for extra bandwidth, high dynamic range and resolution. Both noise and strong interferences limit the dynamic range in CDMA and GSM systems. The noise contribution of the ADC becomes significant at high frequencies where circuit and clock phase noise are larger than the quantization noise.

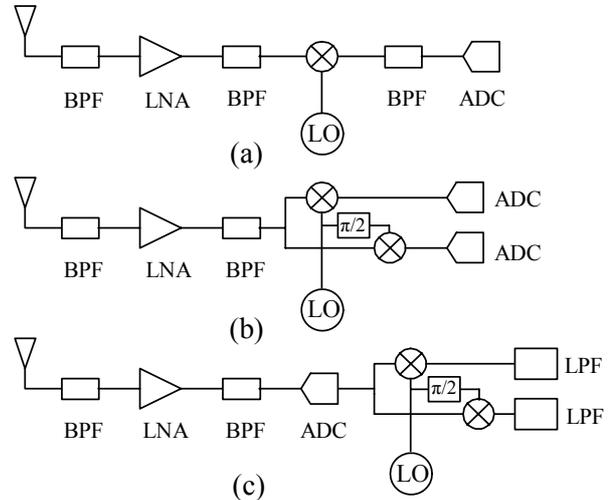


Fig. 1. Receiver architectures: (a) Superheterodyne, (b) Direct conversion (Quadrature sampling) and (c) Direct sampling (Digital IF).

II. JITTER IN DIRECT SAMPLING SYSTEMS

A. Definition of Clock Jitter

Clock jitter is an important parameter in radio systems as the sampling frequency f_s and the signal bandwidth increase. Noise in the time domain is described mostly as clock jitter, another way of expressing phase noise. This quantity is also known as aperture uncertainty, aperture jitter and clock uncertainty. Requirements regarding jitter performance in wideband radio can be extracted after the jitter effect is quantified.

The concept of jitter is better illustrated in Fig. 2, which shows an arbitrary signal being sampled at f_s . In an ideal sampling system, sampling instants are equally spaced in the time domain. Every T_s sec an analog sample is taken and converted to a digital word. In the presence of jitter, the intervals between sampling points are not equal. The uncertainty between each point represents the clock jitter. Jitter is a random process and treated in statistical terms such as the standard deviation σ .

B. Signal-to-Noise Ratio of a Sinewave Signal

The impact of jitter on a sampling system can be evaluated by a simple approximation found widely in literature [1], [2]. Let's assume a perfect sinewave with amplitude A and frequency f

$$x(t) = A \sin(2\pi ft). \quad (1)$$

If we assume that the signal is piecewise linear between two sampling instants, the SNR can be calculated. The expression relating the clock jitter with SNR is

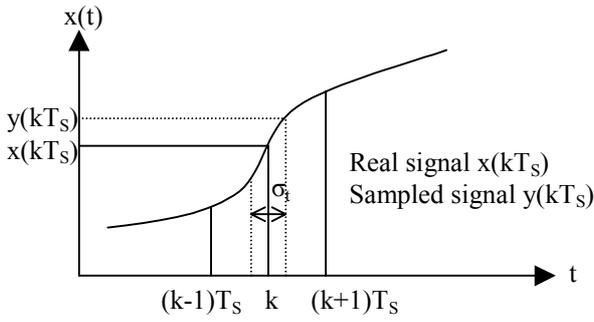


Fig. 2. Clock jitter in a sampling system.

$$SNR = -20 \log(2\pi f \sigma_t) \quad (2)$$

where σ_t is the rms value of jitter. The SNR degrades with increasing frequency for the same amount of jitter.

C. Generic SNR Expression of a Sampler

The assumption of linearity between two sampling instants that is used for the derivation of (2) can be extended for other types of Power Spectral Densities (PSD). As presented in [3] for the case of baseband rectangular PSD, the sampled signal can be approximated by a Taylor series expansion. If the input and output signals of the sampler are $x(kT_s)$ and $y(kT_s)$ respectively, the first order Taylor series becomes

$$y(kT_s) = x(kT_s) + x'(kT_s)t_j(kT_s) \quad (3)$$

where $t_j(kT_s)$ is jitter in the time domain.

It is known from signal theory that the average power of a stationary random signal is given by the autocorrelation function evaluated at zero shift in the time origin.

$$E[x^2(t)] = \sigma_x^2 = R_{xx}(0) = \int_{-\infty}^{+\infty} S_x(f) df \quad (4)$$

with

$$R_{xx}(\tau) = E[x(t)x(t+\tau)] \quad (5)$$

and $S_x(f)$ the PSD of the signal.

Using the autocorrelation function, the authors of [3] presented an expression for the SNR of a sampling system

$$SNR = 10 \log \left(\frac{R_x(0)}{-R_x''(0) \cdot R_x(0)} \right) \quad (6)$$

Equation (6) is a powerful expression showing that the SNR of a sampling system is inversely proportional to the second derivative of the autocorrelation function and the jitter total average power or variance ($R_x(0) = \sigma_x^2$ because $\mu_{x_j} = 0$). Thus, it is independent of the clock phase noise profile as long as its total average power remains constant. This makes the simulation of communication systems in the presence of ADCs less complicated since we can first select noise power levels that meet requirements and then translate them to phase noise profiles.

Another interesting property of the sampler is implied in (6). The autocorrelation function of low frequency signals is approximately flat around zero. Consequently, the second derivative takes small values and the SNR increases. On the other hand, high frequency signals are more susceptible to jitter and have lower SNR due to high autocorrelation function slope. These characteristics are particularly obvious in the following examples.

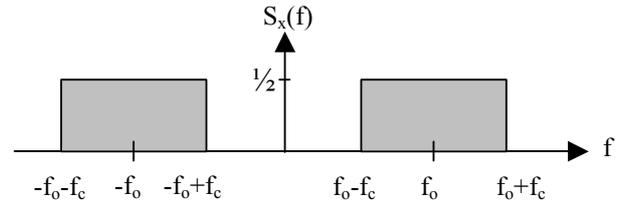


Fig. 3. Bandpass rectangular power spectral density.

D. Examples of SNR Calculation

Three examples below illustrate the efficiency of (6) in calculating SNR. The first example shows another way of finding the SNR of a signal with baseband rectangular PSD. In the second example, the input is a bandpass version of the previous random signal with flat power spectrum.

In a more practical scenario of a CDMA system, data undergo raised cosine pulse shaping (IS-95). Generally, a square root raised cosine filter is used in the transmitter and another one in the receiver. The third and fourth examples cover the case of a baseband and bandpass signal with raised cosine PSD.

1) *Baseband Rectangular PSD*: The mathematical expression of the flat PSD is

$$S_x(f) = \text{rect} \left(\frac{f}{2f_c} \right) \quad (7)$$

and the SNR resulting from (6) is [3]

$$SNR = 10 \log \left(\frac{3}{\omega_c^2 \sigma_t^2} \right) \quad (8)$$

2) *Bandpass Rectangular PSD*: When the flat PSD is modulated on a carrier with frequency f_0 (Fig. 3), we can derive an expression for the SNR. This case resembles the direct RF sampling scenario, when the ADC directly digitizes the RF signal. The PSD of the modulated signal becomes

$$S_x(f) = \frac{1}{2} \text{rect} \left(\frac{f - f_0}{2f_c} \right) + \frac{1}{2} \text{rect} \left(\frac{f + f_0}{2f_c} \right) \quad (9)$$

Using the modulation property of the Fourier Transform the autocorrelation is expressed

$$R_x(t) = 2f_c \sin(\omega_c t) / \omega_c t \cdot \cos(\omega_0 t) \quad (10)$$

The SNR can be found if the autocorrelation of (10) and its second time derivative are evaluated at zero. We can find the expressions using L'Hopital's rule where necessary.

$$R_x(0) = 2f_c \quad (11)$$

$$R_x''(t) = -\frac{\sin(\omega_c t) (\omega_c^2 t^2 - 2) + 2\omega_c t \cos(\omega_c t)}{\pi t^3} \cos(\omega_0 t) - 2\omega_0 \frac{\omega_c t \cos(\omega_c t) - \sin(\omega_c t)}{\pi t^2} \sin(\omega_0 t) - 2f_c \omega_0^2 \frac{\sin(\omega_c t)}{\omega_c t} \cos(\omega_0 t) \quad (12)$$

$$\lim_{t \rightarrow 0} R_x''(t) = -\frac{\omega_c^3}{3\pi} - 2f_c \omega_0^2 \quad (13)$$

From (6), (11) and (13) the SNR of a bandpass signal with rectangular PSD is

$$SNR = 10 \log \left(\frac{2f_c}{\left(\frac{\omega_c^3}{3\pi} + 2f_c \omega_o^2 \right) \sigma_i^2} \right) \text{ or} \quad (14)$$

$$SNR = 10 \log \left(\frac{1}{\sigma_i^2 \left(f_o^2 + \frac{BW^2}{12} \right)} \right). \quad (15)$$

3) *Bandpass raised cosine PSD*: Before being transmitted, data are pulse shaped to minimize the effect of InterSymbol Interference (ISI). Usually data are in a non-return to zero (NRZ) format and a pulse shaping filter reduces the spectrum sidelobes. The raised cosine filter is widely used for this purpose and particularly in CDMA. Pulse shaping is split between the receiver and transmitter in such a way that each one has a square root raised cosine filter. When a signal is filtered with a root raised cosine filter the output has a raised cosine PSD. The baseband PSD is thus

$$S_x(f) = \begin{cases} \frac{1}{2W}, & 0 \leq |f| < f_1 \\ \frac{1}{4W} \left[1 - \sin \left(\frac{\pi(|f| - W)}{2W - 2f_1} \right) \right], & f_1 \leq f < 2W - f_1 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

where a is the roll-off factor, $W=1/2T_b$ and $f_1=W-a$. Variable T_b is the chip rate and $\omega_b = 2\pi R_b$, where $R_b=1/T_b$ is the data rate after root raised cosine interpolation. If the signal is modulated on a carrier the PSD is

$$S_{x,BP}(f) = \frac{1}{2} S_x(f - f_o) + \frac{1}{2} S_x(f + f_o). \quad (17)$$

The autocorrelation function of the bandpass raised cosine PSD is the inverse Fourier Transform of (17)

$$R_{x,BP}(t) = \sin c(2Wt) \frac{\cos(2\pi a W t)}{1 - 4a^2 W^2 t^2} \cos(\omega_o t). \quad (18)$$

In order to simplify further calculations, we assumed $a=0.5$. Equation (18) evaluated at zero gives $R_x(0)=1$. The second derivative of (18) becomes

$$R_{x,BP}''(t) = \frac{-\frac{\omega_b^2 t^2}{2} \sin\left(\frac{\omega_b t}{2}\right) - 2\omega_b t \cos\left(\frac{\omega_b t}{2}\right) - 4 \sin\left(\frac{\omega_b t}{2}\right)}{\omega_b^3 t^3}.$$

$$\frac{\cos(\pi W t)}{1 - 4W^2 t^2} \cdot \cos(\omega_o t) + 2 \frac{\omega_b t \cos\left(\frac{\omega_b t}{2}\right) - 2 \sin\left(\frac{\omega_b t}{2}\right)}{\omega_b^2 t^2}.$$

$$\left[\frac{-\pi W \sin(\pi W t)(1 - 4W^2 t^2) + 8W^2 t \cos(\pi W t)}{(1 - 4W^2 t^2)^2} \cos(\omega_o t) - \right.$$

$$\left. \omega_o \frac{\cos(\pi W t)}{1 - 4W^2 t^2} \sin(\omega_o t) \right] +$$

$$\left[-\pi W \frac{\pi W \cos(\pi W t)(1 - 4W^2 t^2) - 8W^2 t \sin(\pi W t)}{(1 - 4W^2 t^2)^2} + \right.$$

$$\left. 8W^2 \frac{-\pi W t \sin(\pi W t) + [1 - 16W^2 t^2] \cos(\pi W t)}{(1 - 4W^2 t^2)^2} \right] \cos(\omega_o t) -$$

$$2\omega_o \sin(\omega_o t) \frac{-\pi W \sin(\pi W t)(1 - 4W^2 t^2) + 8W^2 t \cos(\pi W t)}{(1 - 4W^2 t^2)^2} -$$

$$\omega_o^2 \frac{\cos(\pi W t)}{1 - 4W^2 t^2} \cos(\omega_o t) \quad (19)$$

which yields

$$\lim_{t \rightarrow 0} R_{x,BP}''(t) = -\omega_b^2 / 12 - W^2 (\pi^2 - 8) - \omega_o^2. \quad (20)$$

The SNR of this signal is therefore

$$SNR = -10 \log \left(\left(\frac{\omega_b^2}{12} + W^2 (\pi^2 - 8) + \omega_o^2 \right) \sigma_i^2 \right). \quad (21)$$

The results calculated in this section can be supported with simulations.

III. SIMULATION RESULTS

Three systems were simulated in Matlab Simulink [4] and mathematical expressions were evaluated with satisfactory accuracy. The sampling system used is not a full ADC model but a Sample-and-Hold (S&H) block. The purpose of the simulations is to demonstrate the effect of jitter on baseband and bandpass signals under oversampling scenario but not to investigate the behavior of a specific type of ADC. In such a case, the SNR would depend on the architecture of the ADC and the interface between the clock and the device.

The structure of the Simulink models is similar in all cases. The system block diagram for all testbenches is depicted in Fig. 4. The generated signal is processed through the S&H block. The upconversion is used only in the bandpass cases. Two discrete S&H blocks exist in the design but are clocked from different sources: a noiseless clock for the upper path and a jittered clock for the lower path. Calculation of SNR is carried out by calculating the signal and noise powers in the time domain. If the noiseless signal (upper path) is subtracted from the jittered sampled signal (lower path), the difference equals the error due to jitter. The error power equals total noise power. The only source of noise in the system is jitter from the "Clock with

Jitter” block. Consequently, subtracting noise from signal power yields the SNR due to clock jitter.

A simulation testbench was first developed for the case of a sinusoid. Results are summarized in Table I. Using the same model structure, we performed simulations for a signal with flat PSD sampled by a high frequency clock. The signal passes through a 10th-order Butterworth lowpass filter. The cutoff (passband–edge) frequency is f_c and an approximate flat power spectrum is created. The SNR is calculated for 2^{15} samples. Results are shown in Table II. All frequencies are normalized to clock frequency f_{clk} in order to reduce simulation time. For $f_c=1/20$ Hz, the authors of [3] report SNR=54.31dB. The difference can be attributed to the numerical calculation of noise. The model used in Fig. 4 gives smaller discrepancy between expected and simulated values.

In the bandpass case, the signal was upconverted before being sampled. This was done by multiplying the baseband portion by a cosine signal at f_o . The remaining part of the model is identical to the baseband one. Results of a bandpass rectangular PSD can be found in Table III.

The last simulation testbench was developed for the raised cosine PSD. Random integers were generated and used as the input to a QPSK baseband modulation block. Data were in NRZ format. The QPSK pulses were shaped with a square root raised cosine filter. The filtered signal has raised cosine PSD. The ‘Raised Cosine Interpolator’ filter oversamples the input data by a factor of 2. Data were generated at rate f_d and upsampled at rate f_{ov} by the raised cosine filter-interpolator. The data rate at the S&H input is therefore f_{ov} . The baseband signal is upconverted at the carrier frequency f_o by a mixing process. The SNR is calculated for 2^{15} samples. Simulation results are presented in Table IV. In Tables III and IV, the SNR from theoretical

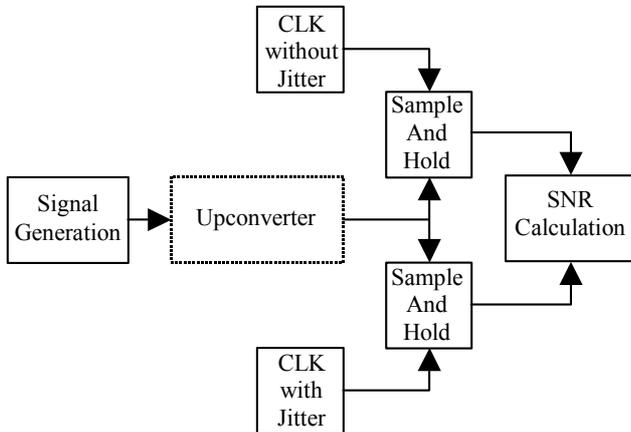


Fig. 4. Model for calculation of SNR in sampling systems with jitter.

expressions and simulation is compared to that given by (2).

Any discrepancy between analytical expressions and simulation results can be attributed mainly to the nature of the sampling block. The S&H operation does not fully capture the conversion since the oversampling ratio is finite compared to the assumption made in (2). In addition, we assumed a first order model for the noisy signal, which is not always true.

TABLE I
SAMPLING OF A SINEWAVE SIGNAL

f_o (Hz)	f_{clk} (Hz)	σ_t^2 (sec ²)	SNR Theory (dB)	SNR Simulation (dB)
1	10^{-4}	10^{-10}	84.036	84.08
10	10^{-4}	10^{-10}	64.036	63.34
100	10^{-4}	10^{-10}	44.036	43.01

TABLE II
SAMPLING OF A BASEBAND SIGNAL WITH RECTANGULAR PSD

f_c (Hz)	f_{clk} (Hz)	σ_t^2 (sec ²)	SNR Theory (dB)	SNR Simulation (dB)
0.05	1	1.78×10^{-4}	52.33	51.18
0.15	1	1.78×10^{-4}	42.79	41.66
0.20	1	1.78×10^{-4}	40.29	39.08

TABLE III
SAMPLING OF A BANDPASS SIGNAL WITH RECTANGULAR PSD

f_c (Hz)	f_o (Hz)	f_{clk} (Hz)	σ_t^2 (sec ²)	SNR (15) (dB)	SNR (2) (dB)	SNR _{sim} (dB)
0.05	0.5	10	1.78×10^{-6}	47.39	47.55	47.87
0.05	1	10	1.78×10^{-6}	41.49	41.53	41.77
0.40	2	15	9.68×10^{-6}	27.51	28.15	27.09

TABLE IV
SAMPLING OF A BANDPASS SIGNAL WITH RAISED COSINE PSD

f_d (Hz)	f_{ov} (Hz)	f_o (Hz)	f_{clk} (Hz)	a	σ_t^2 (sec ²)	SNR (21) (dB)	SNR (2) (dB)	SNR _{sim} (dB)
1	2	3	15	0.5	6×10^{-6}	26.55	26.71	26.17
1	2	4	15	0.5	6×10^{-6}	24.13	24.21	22.13
1	2	5	15	0.5	6×10^{-6}	22.23	22.28	20.40

IV. CONCLUSION

The effect of jitter in sampling systems was demonstrated through analytical expressions. The relationship between SNR, carrier frequency and jitter power is illustrated in (15) and (21). In direct sampling, the carrier frequency sets the SNR limit since the contribution of signal bandwidth is minimal for a given PSD. Consequently, the dynamic range of a direct sampling receiver is not only a function of strong interfering signals and quantization noise but it is also a function of the band location (carrier frequency) and clock spectral purity.

REFERENCES

- [1] B. Brannon, “Aperture Uncertainty and ADC System Performance”, Analog Devices Application Note, AN-501.
- [2] M. Shinagawa, Y. Akazawa, and T. Wakimoto, “Jitter Analysis of High-Speed Sampling Systems”, *IEEE J. Solid-State Circuits*, Vol. 25, February 1990.
- [3] N. Da Dalt, M. Harteneck, C. Sandner, and A. Wiesbauer, “On the Jitter Requirements of the Sampling Clock for Analog-to-Digital Converters”, *IEEE Trans. Circuits and Systems-I*, Vol. 49, September 2002.
- [4] Simulink, Version 5.0, The MathWorks Inc., Natick, MA, 2002.

Demonstration of 1.65 μm Optical Time-Domain Reflectometry in On-Line Monitoring 1.55 μm Distributed Fiber Raman Amplifier Systems

I-Yu Kuo *, Szu-Chi Tsai¹, Han-Ping Mi², and Yung-Kuang Chen²

* Optical Communication & Networking Technologies Dept., Computer & Communications Research Lab., Industrial Technology Research Institute, 195 Chung Hsing Rd., Sec. 4 Chu Tung, Hsin Chu, Taiwan.

¹ Department of Control-Engineering, National Taichung Industrial High School, Taichung, Taiwan.

² Institute of Electro-Optical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan

Abstract ? A 1.65 μm optical time-domain reflectometry (OTDR) on-line monitoring 1.55 μm 10-Gb/s dense wavelength division multiplexing (DWDM) distributed fiber Raman amplifier (FRA) transmission system was demonstrated. The link fault location is realized simply and such distorted traces give rise to an unaccurately measured fiber-loss coefficient and a reflection profile along the fiber link. This on-line monitoring technique can provide an in-serve monitoring to enhance the reliability of dense wavelength multiplexing (DWDM) system including distributed FRA's.

Index Terms ? Raman amplifier, supervisory technique, optical time-domain reflectometer (OTDR), fiber fault location.

I. INTRODUCTION

Silica fiber Raman amplifiers (FRAs) are particularly attractive for broadband application in dense wavelength division multiplexing (DWDM) systems [1]-[3] since they offer the advantages of greatly extended gain bandwidth, distributed amplification, and low noise performance with the installed fibers as the gain media. With the growing deployment of distributed FRA, the system should facilitate the on-line fault-location monitoring capability to enhance system reliability. In general, the 1.65 μm optical time domain reflectometer (OTDR) is a popular tool to provide on-line monitoring of optical fiber links carrying live signals in 1.55 μm DWDM systems [4], [5].

Since OTDR operates with high peak powers, the stimulated Raman scattering (SRS) effect in the transmission fiber gives rise to power depletion [6] of the data channel, and may degrade the bit-error-rate (BER) performance. In this work, we investigate and demonstrated the 1.65 μm OTDR on-line monitoring on the 1.55 μm distributed FRA transmission link

using forward- or backward- pumping schemes. The distorted OTDR trace, which resulted from the backscattered Raman amplified spontaneous emission light is observed. The system power penalty due to the OTDR monitoring and the residual pump lights in both 10 Gb/s forward- and backward-pumping systems are examined and compared.

II. EXPERIMENTAL SETUP

Fig. 1 shows the experimental setup of the 1.65 μm OTDR-monitored 1.55 μm distributed FRA system, in which the forward- and backward-pumping schemes are separately examined. At transmitter site, a 1550 nm DFB laser diode (DFB-LD) with a continuous-wave (CW) output power of +7dBm was externally modulated by a LiNbO₃ modulator with 232-1 NRZ pseudo-random bit-sequence (PRBS) data at 10 Gb/s. The modulated signal was amplified by a conventional-wavelength (C-band) erbium doped fiber amplifier (EDFA) with an output power of about +17 dBm and a noise figure of 5 dB. The power level of data channel was adjusted by a variable optical attenuator (VOA1) to 0 dBm at position A. The data signal combined with the 1656 nm Anritsu OTDR channel that operated with 10 μs pulse-width, peak power is about +5 dBm, and then launched into the fiber link through a 1.55/1.65 μm coupler (WDM1). The corresponding single-way dynamic range and the event resolution of 10 μs OTDR pulses is about 20 dB and 1 km, respectively.

Although it seems suitable to allocate the data channel at about 1585 nm in this experiment to catch the high gain of the distributed FRA, a single data channel with a central wavelength of 1550 nm was arranged for feasibility study and giving a separation of 106 nm relatively to the OTDR channel. Such separation matches the peak Raman gain coefficient of the used dispersion shifted fiber (DSF), and hence the data channel may encounter strong SRS interaction. The FRA consists of the pump module

and the 50.4 km DSF link as the gain medium. The DSF is used to offer higher Raman gain coefficient due to its small effective area of about $50 \mu\text{m}^2$. The pump module is composed of a pump power combiner and four fiber-grating-stabilized pump laser diodes (LDs) with the central wavelengths of 1460 nm, 1470 nm, 1480 nm and 1490 nm, respectively. The output power of each pump LD is about 160 mW, and the total pump power at the output port of the pump combiner is about 500 mW.

In forward-pumping scheme, the pump lights combined with both data and OTDR channels through a 1.48/1.55 μm coupler (WDM2), and then launched into the DSF link, in which the pump lights and the data and OTDR channels have the same propagating direction. In backward-pumping case, which is the operating scheme in most distributed FRA systems, the pump lights were launched into the DSF link through a 1.48/1.55 μm coupler (WDM3). In this case, the pump lights have opposite propagating direction with both data and OTDR channels.

All WDM couplers (WDM1, WDM2 and WDM3) are the thin-film-filter-based devices. The insertion

loss of WDM1 is about 1.8 dB in the region of 1400 to 1610 nm, and about 0.8 dB in the region of 1620 to 1670 nm. The channel isolation is about 59 dB at 1.55 μm port and 63 dB at 1.65 μm port. The insertion losses of both WDM2 and WDM3 are the same with about 0.8 dB in the region of 1420 to 1500 nm, and about 0.8 dB in the region of 1510 to 1670 nm. The channel isolations are about 65 dB at 1.55 μm port and 64 dB at 1.48 μm port. Since both WDM1 and WDM2 have low flat spectral characteristics and high channel isolations, the OTDR probe lights and also the Rayleigh backscattering lights can properly pass through both WDM1 and WDM2, and the on-line monitoring for this system can be achieved. At receiving site, a DWDM demultiplexer (DEMUX) with a channel spacing of 1.6 nm with a 3-dB bandwidth of 0.88 nm, an averaged channel insertion loss of 1.8 dB, and a channel isolation of 40 dB was used to demultiplex the data channel. The OTDR light is also separated from the data channel by this DEMUX. PINFET receiver (Rx) with sensitivity of -17.5 dBm is used for BER measurements.

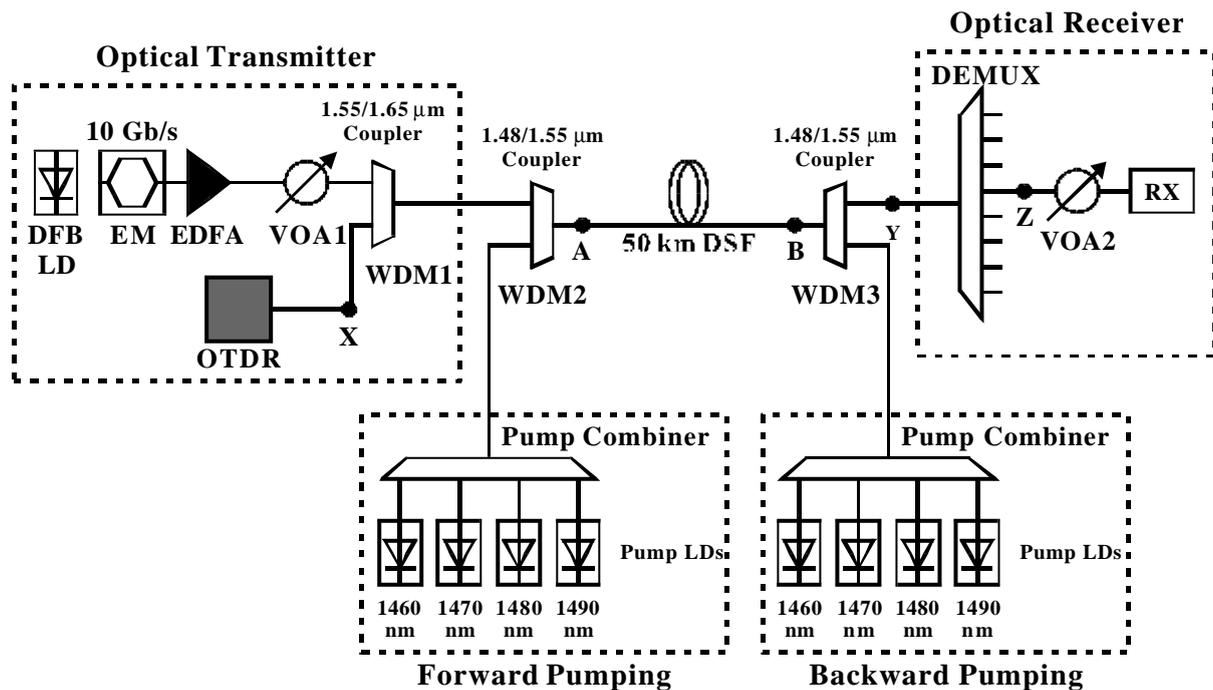


Fig.1 Experimental setup of the OTDR-monitored forward- or backward-pumping FRA system.

III. RESULTS AND DISCUSSION

Fig. 2 illustrates the measured optical spectra at position Y (as depicted in Fig. 1) while the OTDR was under (a) off and (b) on operation, and at positions (c) X and (d) Z in forward-pumping distributed FRA system. Fig. 3 shows the measured optical spectra at the corresponding positions in backward-pumping distributed FRA system. The upper and bottom

spectra, in Figs. 2 (a)-(b) and 3 (a)-(b), correspond to the cases with the pump light on and off, respectively. In forward-pumping case, the 1.55 μm optical Raman amplification gain was depleted from 11.1 dB to 9.4 dB while switching the OTDR on. In backward-pumping case, the 1.55 μm Raman gain was depleted from 16.0 dB to 15.9 dB while switching the OTDR on. The gain degradation, which resulted from the SRS-interaction-induced power depletion between the 1550 nm data and 1656 nm OTDR channels, is about

1.7 dB and 0.1 dB for forward- and backward-pumping cases, respectively. This is because the SRS interaction between the 1550 nm data channel and the 1656 nm OTDR channel in forward-pumping case is stronger than that the backward-pumping case.

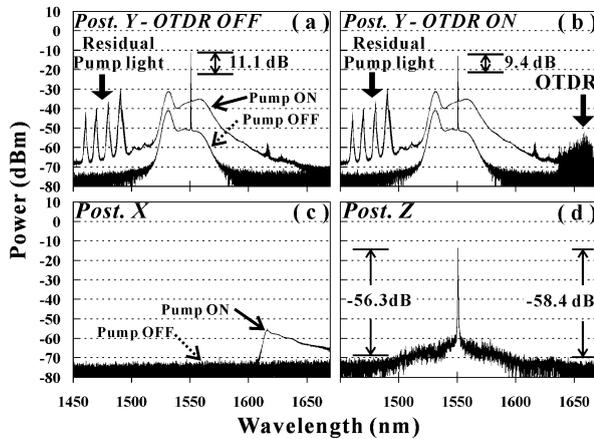


Fig.2 Evolution of optical spectra at position Y while OTDR was (a) off and (b) on operations, and at positions (c) X and (d) Z in forward-pumping distributed FRA system.

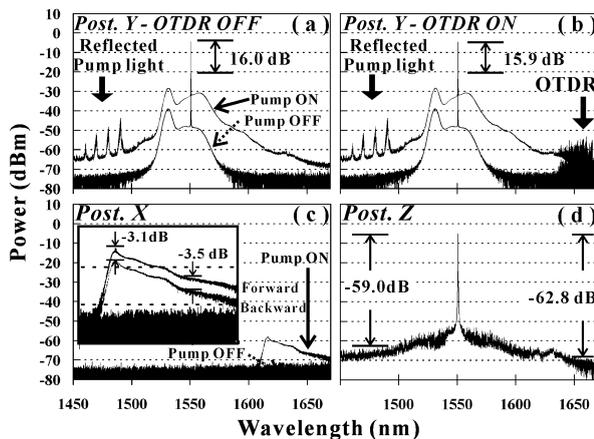


Fig. 3 Evolution of optical spectra at position Y while OTDR was (a) off and (b) on operations, and at positions (c) X and (d) Z in backward-pumping distributed FRA system.

While disconnecting the OTDR at position X, Fig. 2 (c) shows the measured optical spectra of the reflected and Rayleigh backscattered lights of both the pump lights and the data channel in forward pumping case. Fig. 3 (c) shows the corresponding measured optical spectra in backward-pumping case. Note that the spectral components of both Rayleigh backscattered lights and the reflected lights in the region of 1450 to 1610 nm were drastically attenuated and removed by the high channel isolations of WDM2 and WDM1. However, the Rayleigh backscattering CW lights, as shown in Fig. 2 (c) and 3

(c), of (1) the Raman amplifier amplified spontaneous emission (ASE) light and (2) the portion of amplified ASE light of EDFA in the region of 1620 to 1650 nm, contaminate the real OTDR Rayleigh backscattered pulsed light while the OTDR in on-line monitoring operation. Unfortunately, such backscattered ASE lights may give rise to distort the OTDR healthy traces (which to be described in next paragraph). For the demultiplexed data channel in Figs. 2 (d) and 3 (d), all residual or reflected pump lights, and the OTDR probe lights are drastically removed by the DEMUX. Note that the inter-channel crosstalk level of the demultiplexed data channel was about less than -56 dB at 1.48 μm and -58 dB at 1.65 μm in forward-pumping scheme, and less than -59 dB at 1.48 μm and -62 dB at 1.65 μm in backward-pumping case. Such low inter-channel crosstalk levels may give rise to negligible power penalty in BER measurements.

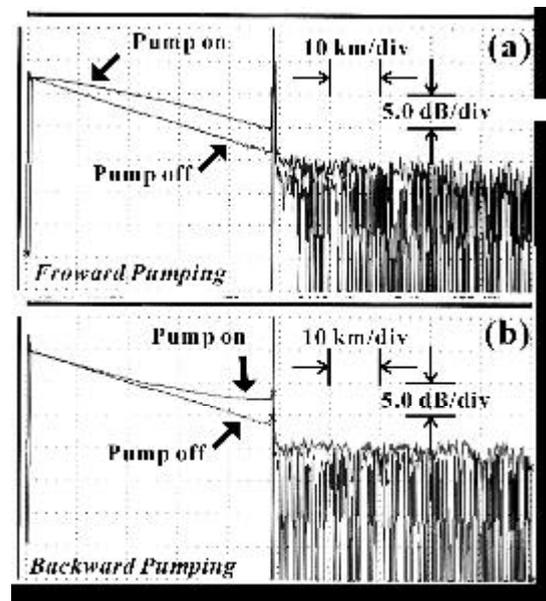


Fig. 4 Measured OTDR traces of the (a) forward-, and (b) backward- pumping distributed FRA systems with the pump light off and on operations.

Fig. 4 shows the OTDR traces of (a) forward- and (b) backward- pumping distributed RFA systems. The lower healthy trace corresponding to the case while switching all pump lights off, and the measured loss coefficient of the 50.4 km DSF is about 0.248 dB/km. In contrast, while switching the pump light on, the measured OTDR traces being distorted in both pumping cases. Such distorted OTDR trace is due to the unwanted Rayleigh backscattered light of the Raman-amplification-induced related ASE CW lights at 1.65 μm as shown in Fig. 2(c) and 3(c). Since such ASE CW light contaminated with the Rayleigh backscattered OTDR pulsed light, the OTDR detected signals were noisily added, and thus the OTDR trace distorted. The total power of backscattered ASE CW

lights detected by the OTDR for the forward-pumping case is at least 3 dB greater than the backward-pumping case as shown in the inset of Fig. 3(c). The larger ASE light as shown in Fig. 3(c) the worse distorted OTDR trace as shown in Fig. 4(b). The distortion of OTDR trace can be improved by (1) reducing such ASE lights into the OTDR by using a WDM1 with a sharp spectral filtering edge at about 1640 nm, or (2) by using a coherent-detection OTDR [4] with an inherent high background-noise-rejection capability to replace the used conventional OTDR. Although such distorted traces in Figs. 4(a) and (b) give rise to an un-accurately measured fiber-loss coefficient, the OTDR diagnosing capability still worked successfully. Note that the last reflecting spike at about 50.4 km coincides with the total link length of the system. When there is a fiber break occurred, the fiber fault location could be easily located and identified in such distorted OTDR trace.

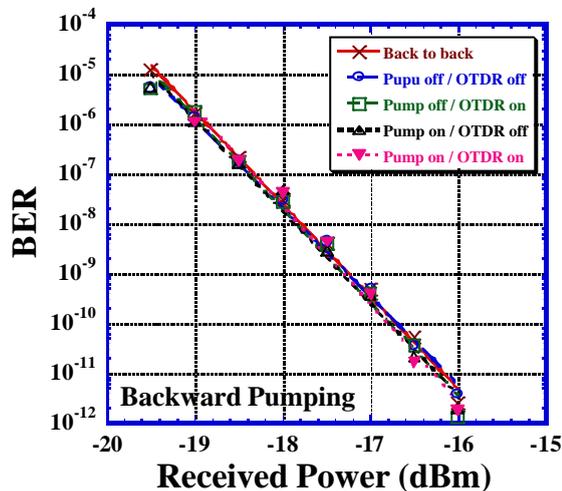


Fig. 5 10 Gb/s BER performance of the backward-pumped distributed 50.4 km FRA system.

Fig. 5 shows the 10 Gb/s BER performance of the backward-pumping 50.4 km FRA system. The BER floor was not observed below BER of 1×10^{-9} . Note that the OTDR-monitoring induced power penalty was negligible (< 0.1 dB) in spite of the pump light being on or off. This is due to the excellent low inter-channel crosstalk levels as shown in Fig. 2 (d). On the other hand, a similar BER performance has been obtained (not shown here) for the forward-pumping system. Either OTDR or pump-light induced BER power penalty is also less than about 0.1 dB in spite of the effect of SRS-induced power depletion while switching on the OTDR. The above experimental results confirm the feasibility of 1.65 μm OTDR on-line monitoring on 1.55 μm distributed FRA systems.

IV. CONCLUSION

We have successfully demonstrated the 1.65 μm OTDR on-line monitoring in forward- and backward-pumping distributed FRA systems. Negligible monitoring-induced power penalty has been achieved in both 10 Gb/s pumping systems. In addition, the reflective fiber faults occurred in the fiber link can be observed and identified. From the points of view of high Raman gain, low signal power depletion, and less distorted OTDR trace, the backward-pumping scheme is superior to the forward-pumping one. This supervisory technique can be easily to develop to an automated surveillance system to enhance the reliability of the distributed FRA systems.

REFERENCES

- [1] P. B. Hansen, L. Eskilden, S. G. Grubb, A. J. Stentz, T. A. Strasser, J. Judkins, J. J. DeMarco, R. Pedrazzani, and D. J. DiGiovanni, "Capacity upgrades of transmission systems by Raman amplification," *IEEE Photon. Technol. Lett.*, vol. 9, no. 2, pp. 262-264, 1997.
- [2] M. Nissov, C. R. Davidson, K. Rottwit, R. Menges, P. C. Corbett, D. Innis, and N. S. Bergano, "100 Gb/s (10 \times 10 Gb/s) WDM transmission over 7200 km using distributed Raman amplification," in *Proc. Eur. Conf. Optical Communication (ECOC'97)*, vol. 5, pp. 9-12, 1997.
- [3] S. Kado, Y. Emori, S. Namiki, N. Tsukiji, J. Yoshida, and T. Kimura, "Broadband flat-noise Raman amplifier using low noise bi-directionally pumping sources," in *Proc. Eur. Conf. Optical Communication (ECOC'01)*, vol. 6, pp. 38-39, 2001.
- [4] T. Otani, Y. Horiuchi, T. Kawazawa, K. Goto, S. Akiba, "Fault localization of optical WDM submarine cable networks using coherent-optical time-domain reflectometry," *IEEE Photon. Technol. Lett.*, vol. 10, no. 7, pp. 1000-1002, July 1998.
- [5] Y. K. Chen, W. Y. Guo, W. I. Way, and S. Chi, "In-service supervisory EDFA-repeated wavelength division multiplexing transmission system," *IEEE Photon. Technol. Lett.*, vol. 7, no. 8, pp. 923-925, Aug. 1995.
- [6] P. M. Kjeldsen, M. Obro, J. S. Madsen, S. K. Nielsen, "SRS induced depletion of 1540 nm signal co-propagating with 1630 nm OTDR pulses," *Electron. Lett.*, vol. 32, no. 20, pp. 1914-1916, Sept. 1996.

Efficient Memory Structures For Dynamic Multicast Routing Architecture

Lam S.K., *Srikanthan T. and Leow K.H.
Centre for High Performance Embedded Systems
Nanyang Technological University, Singapore
*Email: astsrikan@ntu.edu.sg

Abstract:

As real-time applications requiring dynamic group communication emerges, existing multicast routing algorithms must be improved to realize high-speed routing decisions. This paper describes an efficient technique for porting the widely used Dijkstra's shortest path algorithm to hardware to facilitate multicast routing. In particular, we increase the performance of the architecture by employing an efficient memory structure called the Bucket-Heap (BH) to store the network information. Software simulations on randomly generated networks show that significant performance gain in terms of reduction in the number of memory accesses can be achieved with the BH implementation. Finally, we demonstrate that further speed-up can be achieved by grouping independent clusters of strongly connected nodes as peer groups.

1. Introduction

With the emergence of new generation applications that require multipoint (or group) communication such as multimedia conferences, shared workspace, distributed games and distributed simulation, there became a pressing need for multipoint communication protocols. One of the most challenging issues researched in multipoint communication is multicast routing [1], which is defined as a routing process from a source node to multiple destination nodes in a particular multicast group. Members of dynamic multicast groups can leave and allow other members to participate in the session.

The Dijkstra's shortest path algorithm has been frequently employed for multicast routing [2]. However, in dynamic multicast sessions involving large multicast groups, the Dijkstra's algorithm may suffer from performance degradation when link faults arises and group membership changes. In this paper, a hardware routing engine based on the Dijkstra's shortest path algorithm that performs high-speed routing decisions is proposed for dynamic multicast route computations.

Efficient data structures such as heap implementations have been exploited to expedite the

computation of shortest paths in software. We will show that a variation of this data structure will lead to hardware-efficient realizations of the Dijkstra's shortest path computation by reducing the large number of memory accesses during run-time. As the number of data computations is proportional to the number of memory accesses, utilizing efficient memory structures will inherently increase the performance of the entire system. In this paper, we will present the Bucket-Heap (BH) memory structure, which is based on the heap data structure.

In the following section, the hardware implementation of the Naïve Dijkstra's shortest path algorithm is discussed. Next, the BH memory structure is presented to show the performance gain obtained by employing the efficient memory structure. Finally it will be demonstrated that the architecture's performance can be further increased, by grouping the independent clusters of strongly connected nodes into separate buckets of the BH structure.

2. Naïve Dijkstra's Implementation

Given a network $G=(V,E)$, with non-negative weights d_{ij} for all edges $(i,j) \in V$, a source node S , a path cost list D_k for all nodes $k \in V$, and a set of labeled nodes L , the Naïve Dijkstra's algorithm computes the shortest paths from S to all the other nodes. Tommiska et. al. [3] have briefly commented on the inefficiency of implementing priority queue in hardware, and we shall limit our discussion to the Naïve implementation as shown in Figure 1.

The Dijkstra's algorithm comprises of two main functions, Find-Min and Relax, which forms the basis of our discussion in the remainder of the paper. The Find-Min function identifies the node i that has the least path cost and insert it into the labeled list L . The Relax function computes the new path cost of the adjacent nodes of i , and updates the path cost list with the new value.

Find-Min: $D_{\min} = \infty$ For all $j \notin L$,

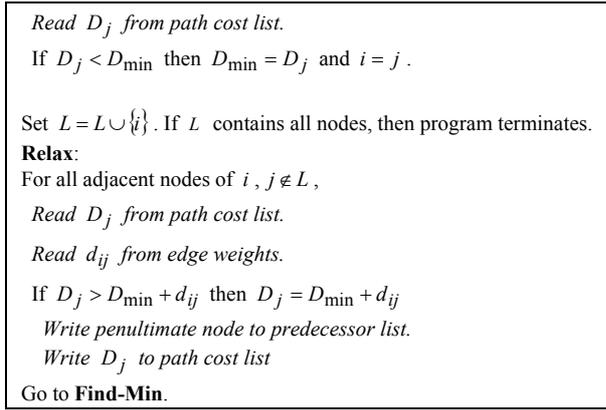


Figure 1: Naive Dijkstra’s implementation

The external memory access operations are denoted in italics and for a single bus architecture as shown in Figure 2, we can assume that the read and write operations are executed in distinct steps. A Memory Addressing Unit (MAU) is needed to compute the memory locations for accessing data during computations. The external memory comprises of the network representation, the labeled list, the path cost list, and a predecessor list that stores the penultimate nodes (the next-to-last nodes in each path). The network representation stores the adjacent nodes of each node in the network using the forward star data structure [4], which has been demonstrated to be the most efficient in past research.

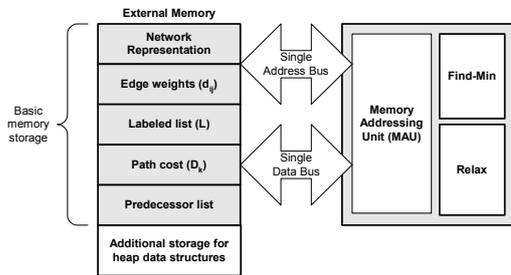


Figure 2: Single bus implementation

From Figure 1, the Find-Min operation reads in the elements from the path cost list D_j sequentially in the order of $O(n)$ to identify the minimum path cost value. In the Relax function, the adjacent weights d_{ij} and path costs D_j are read and checked if an update of a new path cost value is necessary. The operations within the function comprise of two memory-read accesses, an add/compare operation, and two possible write accesses to update the new path cost and penultimate nodes to external memory. This function has an order of $O(m)$ for a single iteration, where m is the maximum number of adjacent edges for each node.

3. Bucket-Heap Implementation

In this section, we propose a variation of the binary heap memory structure for hardware porting, which we call the Bucket-Heap (BH) memory structure. The binary heap data structure can be viewed as a complete binary tree, which is implemented with an array H , where the children of the node in $H[i]$ are found in array positions $H[2i]$ and $H[2i + 1]$ [5]. The Dijkstra’s shortest path algorithm utilizes the heap to maintain the cost magnitude order, where each parent in the heap stores the path cost of an unlabelled node in which its children has higher or the same path cost. The binary heap data structure simplifies the Find-Min process as the root of the tree represents the node with the minimum path cost. However, to maintain the order of the tree throughout the computation, the heap must be readjusted after each Find-Min and Relax operation.

The readjustment process is called Heapify, which incurs a computational complexity of $O(\log n)$, where n is the number of nodes in the network. It is noteworthy that the Heapify function lends well to efficient hardware realization as the determination of parent and child locations in a heap array comprises of simple shift and add operations. Different variation of heaps for shortest path computation in software have been evaluated and concluded that the binary heap provides an implementation with the least memory requirement [5]. In addition, the binary heap is only occasionally out-performed by more advanced data structures. Hence, in this paper we will limit the discussion to the implementation of binary heap.

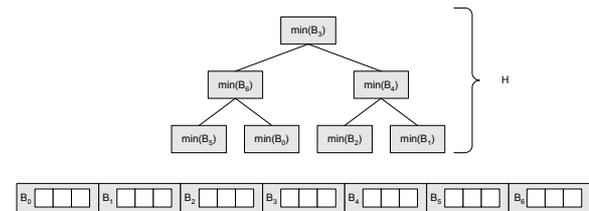


Figure 3: Bucket Heap data structure

In the BH data structure, the path cost list D_k is grouped into b buckets, where each bucket contains c number of path costs and $c = \lceil n/b \rceil$. Let B_k denote the buckets for $k = 0, 1, \dots, b-1$, where $B_k = \{D_{kc}, D_{kc+1}, \dots, D_{kc+c-1}\}$. The elements of the heap contain the minimum path cost of each bucket (i.e. $\min(B_k)$) and the number of levels in the heap is $H = \log b = \log n - \log c$. The heap property such that the child values are always less or equal to the parent value is also maintained in the BH data structure. In the case where $c = 1$ and $b = n$, the data structure is equivalent to the conventional binary heap. Figure 3

shows the conceptual view of the BH data structure with $n = 21$, $b = 7$ and $c = 3$.

Figure 4 describes the algorithm for the BH implementation in hardware. During the Find-Min function, the minimum path cost can be determined by reading the root of the heap from the external memory in the order of $O(1)$ memory access. When the minimum path cost is obtained, the heap must be readjusted to maintain the heap order. First a new path cost must be computed from the bucket in which the minimum path cost is found. This incurs an order of $O(c)$ memory accesses and computations. Next, the Heapify process ($O(\log b)$) is performed to maintain the heap order. Hence, the order of Find-Min in the BH architecture is of $O(1) + O(c) + O(\log b) = O(c + \log b)$. This is an improvement over $O(n)$ for the Naive implementation when $c < n$. In the case of conventional binary heap where $c = 1$ and $b = n$, the order of Find-Min becomes $O(\log n)$, which is less than $O(n)$ for the Naive implementation. However, when $c = n$, there is no performance advantage of the BH implementation over the Naive implementation.

<p>Find-Min: Read minimum path cost D_i from root of heap $H[\text{root}]$. $H[\text{root}] = \infty$. For all $j \in L$ and $D_i, D_j \in B_k$, Read minimum path cost D_j. If $H[\text{root}] > D_j$, then $H[\text{root}] = D_j$</p> <p>HEAPIFY Set $L = L \cup \{i\}$. If L contains all nodes, then program terminates.</p> <p>Relax: For all adjacent nodes of $i, j \in L$, Read D_j from path cost list. Read d_{ij} from edge weights. If $D_j > D_i + d_{ij}$ then $D_j = D_i + d_{ij}$ Write penultimate node to predecessor list. Write D_j to path cost list If $D_j \in B_k$ and $H[B_k]$ exist then If $H[B_k] > D_j$ then $H[B_k] = D_j$ Else $H[B_k] = D_j$</p> <p>HEAPIFY. Go to Find-Min.</p>

Figure 4: Hardware implementation for BH

The adoption of the BH data structure incurs additional memory accesses and computations to the Relax function. The Heapify process is performed when the path cost list D_j is updated with a new value. The heap element corresponding to the bucket containing D_j is updated with the value of D_j if the

value of the element is larger in the order of $O(1)$. Hence the worst-case computation of the Relax function is of the order $O(m(\log n - \log c))$, which is higher than $O(m)$ in the Naive implementation. In the conventional heap data structure where $c = 1$, the order of Relax becomes $O(m \log n)$, which is worst than $O(n)$ for the Naive implementation.

4. Simulation Results

Software simulations were written in C to compile the total number of computations and accesses required for computing the shortest paths using the implementations discussed in the paper. Networks with various node sizes and random density were generated and the same set of networks is used for a particular node size to maintain a standard test case for each implementation.

	Number of nodes				
	64	128	256	512	1024
Naive	8698	33690	133971	525921	2102398
	11329	34080	108396	356641	1266387
BH-4	9810	30324	99926	337535	1225046
BH-8	8667	27708	93783	324488	1196228
BH-16	8192	26328	90363	316726	1178316
BH-32	8427	26496	90123	315249	1172906
BH-64	9738	29164	95107	323495	1187431

Table 1: Comparison of memory accesses

Table 1 shows the number of memory accesses (Find-Min + Relax) of the Naive and BH implementation using various node and bucket sizes in randomly generated networks. The term BH- x denotes the implementation of BH with $c = x$. In general, the BH implementation has an overall better performance as compared to the Naive implementation. For small network size (i.e. 64), the BH implementation using small bucket sizes (i.e. BH-2 and BH-4) has a lower performance than the Naive implementation. This can be explained from the complexity analysis whereby for small number of nodes and small bucket sizes, the performance gain of the Find-Min operation in the BH implementation ($O(n) - O(\log n)$) is less than the gain in the Relax function of the Naive implementation ($O(m \log n) - O(m)$). However, as the number of nodes increases, the performance gain of the BH implementation over the Naive method becomes more significant. From Table 1, we can also observe that the performance of the BH implementation degrades when the bucket size increases beyond $c = 16$ for each set of node size. This can also be explained from the complexity analyses, whereby for large bucket sizes, the order of the Find-Min and Relax functions in the BH implementation converges to that of the Naive implementation.

Table 2 describes the number of memory accesses for the Find-Min and Relax functions in the Naïve and BH implementations with bucket size of 16. We can observe that for the Naïve implementation, the number of accesses for the Relax function is significantly higher than that of the Find-Min function for the Naïve implementation. The difference between these accesses grows exponentially with increasing number of nodes and hence, the computation bottleneck lies in the Relax function. It can be observed from Table 2 that the Find-Min operation of the BH implementation has lower number of accesses than that of the Naïve implementation. This demonstrates the effectiveness of using the tree data structure to obtain the minimum path cost.

	Number of nodes				
	64	128	256	512	1024
Naive (Find-Min)	4162	16516	65800	262672	1049632
Naive (Relax)	4535	17173	68170	263248	1052765
BH-16 (Find-Min)	2565	6460	15823	38700	91274
BH-16 (Relax)	5627	19868	74539	278026	1087042

Table 2: Memory accesses for the Find-Min and Relax function

5. Improved Implementation for Bucket Heap

In this section, we propose a technique to improve the performance of the BH implementation by grouping clusters of independent strongly connected nodes of a network into individual peer groups as suggested by Van Mieghem in [6]. A strongly connected group of nodes in a network comprises of a maximal set of nodes that are reachable from each other. By allocating each set of strongly connected nodes into individual buckets of the BH structure, the Heapify process in the Relax function shown in Figure 3, needs only to be executed only once for each Relax operation. This is possible as all the adjacent nodes of the minimum cost node are guaranteed to reside in the same bucket. Table 3 compares the complexity analyses between the Naïve, BH and improved BH implementation. It can be observed that the complexity of the Relax function of the improved BH implementation is significantly lower than the BH implementation.

	Find-Min	Relax
Naive	$O(n)$	$O(m)$
BH	$O(c + \log n - \log c)$	$O(m(\log n - \log c))$
Improved BH	$O(c + \log n - \log c)$	$O(m) + O(\log n - \log c)$

Table 3: Comparison of complexity analysis

Although we have shown that grouping the clusters of strongly connected nodes as peer groups will increase the performance of the multicast route computations, the benefits of this improvement will vary based on the topological characteristics of the network. Future work in this area includes identifying other criteria for selecting the peer groups to be allocated into the buckets of the BH structure for high-speed multicast route computations.

6. Conclusion

In this paper, we have proposed a Bucket-Heap technique for the efficient implementation of multicast routing in hardware. Although the BH based implementation leads to superior performance, employing large bucket sizes does not lead to further improvements. This can be evidenced by the fact that the Naive implementation has the same order of complexity and less memory storage. However, when the number of bucket sizes is kept below 64, the BH implementation performs significantly better than the Naive implementation. Finally, we show that the performance of the BH implementation could be further improved by allocating strongly connected nodes into the buckets of the BH structure.

7. References

- 1 Laxman H. Sahasrabudde and Biswanath Mukherjee, "Multicast Routing Algorithms and Protocols: A Tutorial", IEEE Network, Vol. 14, No. 1, January-February 2000, pp. 90-102.
- 2 Salama H.F., Reeves D.S. and Viniotis Y., "Evaluation of Multicast Routing Algorithms for Real-Time Communication on High-Speed Networks", IEEE Journal On Selected Areas In Communications, Vol. 15, No. 3, April 1997, pp. 332-345.
- 3 Tommiska M. and Skytta J., "Dijkstra's Shortest Path Routing Algorithm in Re-configurable Hardware", Lecture Notes in Computer Science, Springer-Verlag, 2001, pp. 653-657.
- 4 Ahuja R.K., Magnanti T.L. and Orlin J.B., "Network Flows: Theory, Algorithms and Applications", Englewood Cliffs, NJ: Prentice Hall, 1993.
- 5 Oberhauser G. and Simha R., "Fast Data Structures for Shortest Path Routing: A Comparative Evaluation", IEEE International Conference on Communication, Vol. 3, June 1995, pp. 1597-1601.
- 6 Van Mieghem P., "Dividing a Network into PNNI peer groups", Proceedings of 1st International Workshop on the Design of Reliable Communications Networks, May 1998.

Implementation of an RLS Equalizer for OFDM-based WLAN

Gabriel Caffarena, Angel Fernández, Carlos Carreras and Octavio Nieto-Taladriz

Dept. Ingeniería Electrónica (IEL), Universidad Politécnica de Madrid (UPM)
E.T.S.I. Telecomunicación, Ciudad Universitaria s/n, 28040, Madrid, Spain
{gabriel|angelfh|carreras|nieto}@die.upm.es

Abstract

In this paper we address the implementation of a time-varying channel equalizer for high-speed WLAN systems. An RLS equalizer that works in conjunction with ABL pre-equalization is designed for a 200 Mbps OFDM system. Implementation results on Xilinx Virtex II devices are presented.

1. Introduction

This paper addresses the implementation on FPGAs of an RLS equalizer for a 200 Mbps OFDM-based WLAN. The equalizer works in conjunction with ABL pre-equalization to achieve 4G transmission rates. OFDM [1] has become a popular choice for WLAN systems because, amongst other features, it provides mitigation of the multipath channel distortion thanks to the cyclic prefix (CP), which eliminates inter symbol interference (ISI). At the same time, adaptive bit-loading (ABL) [3] combats the capacity loss due to the dips in the frequency spectrum of the indoor channel by means of adapting the transmitted power and the constellation size for each subcarrier. RLS equalization [2] allows to estimate the channel and to track its changes in time. The improvement is significant: an RLS equalizer on its own requires a

signal-to-noise ratio (SNR) higher than 40 dB to achieve a symbol error rate (SER) of 10^{-3} transmitting 200 Mbps, while the mixture of RLS with ABL requires around 20 dB.

Among the different ABL approaches, we selected the Fischer and Huber's algorithm [3], since it enables definition of a fixed bit-rate with a low computational load. On the other hand, the RLS equalizer presents some advantages: fast convergence and accuracy on the results. Moreover, when applied to multicarrier applications, its implementation turns to be very simple. The fact that there is no ISI for OFDM makes possible to implement a scalar version of the algorithm, avoiding the use of systolic arrays [4].

The design process was divided into three main stages: i) algorithm analysis, ii) fixed-point refinement and iii) hardware implementation. During the first stage, the algorithm was analysed and some implementation decisions were taken. The second one is essential to obtain fast and low-cost hardware implementation. One of the goals of this stage, normally excessively time-consuming, was to limit its duration to reasonable time bounds. Finally, the system was implemented using information from the previous stages and knowledge of the target architecture: Xilinx Virtex II devices.

The paper is structured as follows. In section 2 the system is described and the design constraints are listed. Section 3 deals with the algorithm analysis. In section 4 the fixed-point refinement process is revealed and the final quantization results are shown. Section 5 presents the FPGA implementation results. Finally, the conclusions are drawn in the last section.

Table I.- System specifications

Bandwidth	50 Mhz
Bit rate	225 Mbps
Number of subcarriers	512
M-QAM	4-16-64-256-1024
Cyclic prefix duration	1 μ s
OFDM symbol duration	9 μ s
OFDM frame size	100 OFDM symbols
SER _T	10^{-3}
SNR _T	21.4 dB
SNR loss	0.2 dB
RLS memory factor λ	0.9
Channel delay spread	150 ns
Channel Doppler spread	14 Hz

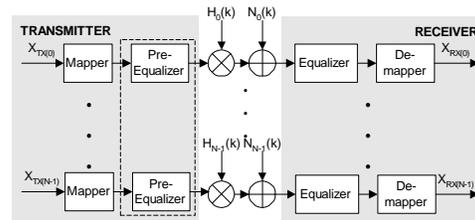


Figure 1.- Frequency model of OFDM modem

2. System description

Figure 1 depicts a frequency model of an OFDM-based communication system where perfect time and frequency synchronization and no ISI were supposed. The bitstream to be transmitted is split into small groups of bits and assigned to different subcarriers. The ABL pre-equalizer determines the constellation size and the transmitted power for each individual subcarrier depending on the current state of the channel. Upon reception, each subcarrier signal has been degraded by the channel ($H_i(k)$) and the noise added by the analogue front-end ($N_i(k)$). The RLS equalizer recovers the original information sent. It estimates the channel by means of an OFDM symbol that contains known information (pilot), and then tracks its changes by a decision-directed strategy.

The frequency model simplifies the simulations by substituting the convolution between the transmitted signal and the impulse response of the channel by a multiplication in the frequency domain. Thus, the model of the channel must be in the frequency domain and only phasors at the frequencies of the different subcarriers are needed. The channel was modelled by a stochastic radio channel model [5].

The system specifications are shown in Table I. The timing constraints are determined by the symbol period (9 μ s). The equalizer must be able to process $56 \cdot 10^6$ subcarriers per second (512/9 μ s).

Noise constraints are defined by the target symbol error rate (SER_T). A SER_T of 10^{-3} is achieved at a SNR equal to 21.4 dB. The degradation produced by the quantization process is measured in terms of SNR loss (ΔSNR):

$$\Delta SNR_Q = SNR_Q - SNR_T$$

where SNR_T and SNR_Q are the SNRs which produce error rates equal to SER_T in the floating-point (FIP) and fix-point (Fxp) systems, respectively. ΔSNR_Q must be smaller than 0.2 dB, thus the quantized system will achieve a SER of 10^{-3} at an SNR smaller than 21.6 dB.

The ABL block tries to keep a transmission rate of 225 Mbps, corresponding to 16-QAM constellations. Regarding the RLS channel tracking, the pilots are sent every 100 OFDM symbols and the memory factor λ is around 0.9. Only even QAM constellations are used to simplify the design of the slicers.

3. Algorithm analysis

The equations of the RLS algorithm for a single subcarrier are shown in Fig. 2. The variables in bold are complex numbers and n is the time step. The OFDM frames are composed of N OFDM symbols ($N=100$). The first OFDM symbol contains known pilots (\mathbf{Y}_{pilot}) required to estimate the inverse of the channel phasor (\mathbf{W}) for each subcarrier. For the remaining $N-1$ OFDM symbols, the initial estimation is used to track the channel variations.

Table II shows the arithmetic operations involved in one iteration of the RLS algorithm. The complex multiplications were assigned 4 real multiplications and 2 real additions.

$$\begin{aligned} \mathbf{w}[n] &= \mathbf{P}[n-1] / \mathbf{R}[n-1] \\ \mathbf{Y}_S[n] &= \mathbf{Y}[n] \times \mathbf{w}[n] & \mathbf{R}[n] \Big|_{n \bmod N = N-1} &= 0 \\ \mathbf{Y}_H[n] &= \text{ slicer}(\mathbf{Y}_S[n]) & \mathbf{P}[n] \Big|_{n \bmod N = N-1} &= 0 \\ \mathbf{R}[n] &= \lambda \cdot \mathbf{R}[n-1] + \mathbf{Y}[n] \times \mathbf{Y}^*[n] & \mathbf{Y}_H[n] \Big|_{n \bmod N = 0} &= \mathbf{Y}_{pilot} \\ \mathbf{P}[n] &= \lambda \cdot \mathbf{P}[n-1] + \mathbf{Y}_H[n] \times \mathbf{Y}^*[n] \end{aligned}$$

Figure 2.- RLS equations

The main speed limitation of the algorithm in Fig. 2 is due to the feedback loops ($\mathbf{R}[n-1]$, $\mathbf{P}[n-1]$). These produce an iteration bound equal to the critical path delay ($path(\mathbf{R}, \mathbf{P} \rightarrow \mathbf{W} \rightarrow \mathbf{Y}_S \rightarrow \text{ slicer} \rightarrow \mathbf{P})$) that is likely to be bigger than the minimum delay allowed ($T=1/56 \cdot 10^6=17$ ns). A way to alleviate this situation is to apply an C-slow transformation that multiples any delay by a factor C , in fact replacing $\mathbf{R}[n-1]$ and $\mathbf{P}[n-1]$ by $\mathbf{R}[n-C]$ and $\mathbf{P}[n-C]$. This transformation reduces the iteration bound C times and forces an interleaved computation [6]. The first feature allows cut-set retiming with the result of notably reducing the critical path (theoretically by a factor of C). The latter is in fact desirable, since the subcarriers are received sequentially from the FFT block allowing reuse of the same equalizer block for all subcarriers. In our case study there are 512 subcarriers.

The architecture of Virtex II devices is suitable for DSP implementations since it has high integration density, fast embedded signed-multiplier blocks (18x18 bit) and RAM blocks (18 Kbit), and admits clock frequencies up to 400 MHz. Both multipliers and dividers have high area costs, so it may be interesting to reuse them to reduce implementation cost. On the other hand, multipliers with input word-lengths smaller than 18 bits can be implemented by using the embedded ones to further reduce costs. The three 512-tap delays necessary to store \mathbf{R} and \mathbf{P} can be implemented using block-RAMs.

Table II.- Number of operations involved

Additions/Subtractions	13
Multiplications	12
Divisions	2
Slice	2
Memory access (read)	3
Memory access (write)	3

4. Fixed-point refinement

Fixed-point refinement, also called quantization, performs the translation of the algorithm specification from floating point (FIP) to fixed-point (Fxp), meeting the design constraints, for the purpose of achieving low-cost hardware implementations. Traditionally, the quantization of systems like our case study has been performed using a simulation-based approach due to the complexity of the system itself and to the presence of non-linear blocks [7-9].

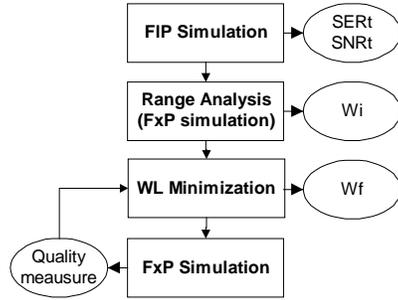


Fig. 3.- Simulation-based quantization methodology

In [10] we stated that simulation-based quantization of OFDM systems is a very time-consuming task due to: a) the iterative nature of this process, that requires multiple long Monte-Carlo (MC) simulations to assess the effect of finite word-lengths; b) the existence of multiple subcarriers, which increases the amount of simulation input data; and c) the presence of a time-varying channel, which makes necessary a high number of samples to model a WLAN link. We address this long time required for the quantization of OFDM adaptive equalizers twofold: i) selecting a word-length minimization procedure tuned to the system under study, to reduce the number of quantization iterations; and ii) reducing MC simulation time and providing a method to handle the inaccuracy introduced.

4.1. Simulation-based quantization

Fig. 3 shows a typical simulation-based quantization methodology. Prior to quantization, the algorithm is evaluated by running FIP simulations that assess its suitability to meet a specific SER_T .

Next, the FxP format must be applied, which is composed of the following parameters:

- S Sign (0 if unsigned, 1 if signed)
- W_i Integer word-length
- W_f Fractionary word-length

Thus, the word-length of an FxP number is:

$$W_L = S + W_i + W_f$$

The first step carried out is determining the integer word-lengths W_i , and this is usually done through a simple FIP simulation that collects statistics for all variables. Hence, W_i of variable V can be determined by using the expression:

$$W_i \geq \lceil \log_2 (\max(|MIN_V|, |MAX_V|)) \rceil$$

where the interval $[MIN_V, MAX_V]$ comprises the possible values of V obtained during the simulation.

The next step is to find W_f through a word-length minimization procedure that states the ordering of variables for FxP format application, and the strategy

to follow if the quality measure exceeds a certain threshold. The quality measurement is performed using FxP simulation and corresponds to the SNR loss (ΔSNR) in communications systems.

4.2. Word-length minimization procedure

Assuming that W_i determination has already been performed, the word-length minimization that we chose is the following:

1. Estimate W_f for the output of operators that require infinite precision (i.e. dividers).
2. Estimate suitable W_f for feedback signals (e.g. in IIR filters), as they increase their fractionary part with each iteration and may end up having too long word-lengths.
3. Estimate W_f for input signals if their FxP format is not specified in the design specifications.
4. Propagate W_f requirements through the algorithm to obtain FxP formats of the remaining signals.

In step 1 signal W is quantized. In 2, both R and P variables are formatted (to simplify the process they are assigned uniform word-lengths). Finally, in 3 signal Y is quantized. From 1 to 3, the W_f values are selected using a quality criterion that is computed through simulation, further explained in the next section. After step 3, the whole set of variables in the filter is implicitly quantized. The format propagation in 4 is then carried out by one final FxP simulation that collects information about the maximum number of fractionary bits necessary for each remaining variable.

4.3. Quality measurement

MC simulations necessary to compute SER values require a considerable large number of input vectors to estimate low SERs (e.g. $N > 1/(SER \cdot e^2)$, where e is the relative precision). This makes unfeasible to use an accurate value of the SER to compute the SNR loss produced by the quantization of a particular signal. We therefore decided to trade-off accuracy for simulation time. The procedure to decide if a FxP format applied to a particular signal was acceptable was the following:

1. Set the signal word-length to a value that produces a negligible increase in the SER.
2. Reduce this value until the relative error from the last SER calculated and the current one is bigger than a certain threshold S . The maximum word-length that complies with S is selected as the word-length of the signal.

The FxP simulation must be performed using the same input vectors (data, noise and channel) and setting the number of them to a relatively small figure to reduce the duration of the process. The more restrictive S the smaller the SNR loss produced.

4.4. Simulation framework

Matlab [11] is commonly selected to model DSP systems. While it works fine at the algorithmic level, it hides information regarding the internal variables of many of its operations and does not fully support

Table III.- Quantization results

W (S, W_i , W_f)	R (S, W_i , W_f)	P (S, W_i , W_f)	Y (S, W_i , W_f)	ΔSNR_0 (dB)
1, 2, 8	0, 8, 7	1, 6, 9	1, 3, 7	0.17

quantization. For this reason, a hardware-wise language is required to implement bit-true operations. Examples are [7] and [8], all of them based on the casting of floating-point (FIP) into fixed-point (Fxp) using C/C++. We have developed a C++ framework called Abaco [12] that includes a library of data types and some analysis and presentation tools that enable the mixed simulation of FIP and Fxp code to obtain signal statistics and traces. The library supports several overflow modes, including wrap-around and saturation, and several underflow modes, including rounding and truncation.

4.5. Quantization results

We only used 128 subcarriers assuming that the results will closely resemble those for 512 subcarriers. The number of input vectors were

128,000,000 QAM symbols for FIP evaluation

12,800,000 QAM symbols for Fxp simulations

The word-lengths were selected by choosing a SER relative error smaller than a 19 %. Table III shows the Fxp format of the main algorithm signals. The SNR loss is approximately of 0.2 dB, which complies with the specifications. The total time of quantization was around 11 hours and only 30 Fxp simulations were required.

5. Implementation results

The equalizer was implemented applying both pipeline and resource sharing techniques. The timing specifications allowed us to apply a folding factor of two in order to share half of the multipliers and dividers, thus reducing hardware cost. Adders were not reused since a folding factor of two does not provide any gain in area: the area of two adders is equal to the area of one adder and a multiplexer in Virtex II devices. A digital clock manager (DCM) was used to generate two clocks signals: a 58 MHz clock was employed to control registers and adders, while a 116 MHz clock controlled the multiplexing of multipliers and dividers. The divider was implemented using a core generated by Coregen from Xilinx. The quantization obtained enabled the use of the embedded multipliers for all multiplications. The FIFOs were implemented using block-RAMs. In fact one 18Kbit block-RAM can store around 1024 16-bit words, and that suffices to implement the 512-tap delays necessary to store R and P .

The equalizer was developed as a parameterizable core generator written in C++, taking as input the word-lengths and generating synthesizable VHDL code. Synthesis and place and route stages were performed using Xilinx ISE 6.1i.

Table IV displays the implementation results. The area cost due to the divider core is about half of the

Table IV.- Implementation results on XC2V250-6

Slices	18x18 Mult.	18 Kb Block-RAM	DCM	T _{CK}
1090	7	3	1	7.143 ns

total area (593 slices) and it can be seen how the use of embedded resources significantly reduces FPGA area cost. Note that the timing specifications are met loosely since the clock period is much smaller than the value necessary (17 ns).

6. Conclusion

In this paper we have addressed the implementation of an RLS channel equalizer for high-speed WLAN systems. An FPGA-based pipeline architecture is presented that achieves a bit-rate of 225 Mbps with an error rate of 10^{-3} . Implementation results on Xilinx Virtex II devices have been presented.

7. Acknowledgements

This work has been partially supported by the Spanish Ministry of Science and Technology under project TIC2003-09061-C03-02.

References

- [1] J. Bingham, "Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come", *IEEE Communications Magazine*, May 1990.
- [2] J. G. Proakis, *Digital Communications*, McGraw-Hill, 2001.
- [3] R. F. H. Fischer, J. B. Huber, "A New Loading Algorithm for Discrete Multitone Transmission", in *Proceedings GLOBECOM '96*.
- [4] B. Yang, J. F. Bohme, "Rotation-based RLS algorithms; unified derivations, numerical properties, and parallel implementations", *IEEE Transactions on ASSP*, May 1992.
- [5] R. Heddergott, U. P. Bernhard, B. H. Fleury, "Stochastic radio channel model for advanced indoor mobile communication systems", in *Proc. PIMRC '97*.
- [6] K. Parhi, *VLSI Digital Signal Processing Systems: Design and Implementation*, Wiley, 1999.
- [7] S. Kim, K. Kum and W. Sung, "Fixed-Point Optimization Utility for C and C++ Based Digital Signal Processing Programs", *IEEE Trans. on Circuits and Systems*, 1998.
- [8] R. Cmar, P. Schaumont, S. Vernalde, I. Bolsens, "A Methodology and Design Environment for DSP ASIC Fixed Point Refinement", in *Proc. DATE '99*.
- [9] G. A. Constantinides, "Perturbation Analysis for Word-length Optimization", in *Proc. FFCM '03*.
- [10] __, "Fixed-point Refinement of OFDM-based Adaptive Equalizers: An Heuristic Approach", *accepted at EUSIPCO '04*.
- [11] *Using Matlab*, The MathWorks Inc.
- [12] J. A. Lopez, C. Carreras, G. Caffarena, O. Nieto-Taladriz, "Fast characterization of the noise bounds derived from coefficient and signal quantization", in *Proceedings ISCAS '03*.

Iris Filter: The VLSI Perspective

H. Tian, T. Srikanthan and C. S. Lim

Centre for High Performance Embedded Systems, Nanyang Technological University,
Nanyang Avenue, Singapore

Abstract

Iris filter is an edge detection filter that demonstrates superior capabilities in enhancing indistinct boundaries of regions in an image. This paper investigates hardware implementation issues of the Iris filter. A CORDIC-based hardware architecture design for Iris filter is proposed. Efforts for the design are targeted towards simplifying the need for complex computation of trigonometric functions required by the Iris filter operation. The proposed design is implemented in FPGA, the speed of operation and the resource utilisation of the implemented design is presented.

1. Introduction

Iris filter is an edge detection filter that demonstrates superior capabilities in enhancing indistinct boundaries with weak contrast to the background [1]. It is because of its strength in enhancing low contrast boundaries, that the Iris filter has been adopted for use as the main process for edge enhancement in the segmentation algorithms proposed in [2][3] for accurate lumen region segmentation in complex endoscopic images. In practical application of lumen region detection, real-time operation is necessary and therefore hardware implementation of the Iris filter is required. The inherent Iris filter operations are suitable for parallel processing architectures, however it involves much complex arithmetic operations such as trigonometric computation, which does not translate well to a high-speed and low cost implementation in VLSI.

This paper investigates hardware implementation issues of the Iris filter. CORDIC modules are implemented as alternative means for the computation of trigonometric function required by the Iris filter. CORDIC is used as it requires only shift and add operations that lends well for hardware implementation. The hardware architecture designed for Iris filter employs two CORDIC modules, a vector CORDIC is used to compute the arctangent function and another rotation CORDIC is used to compute the cosine function, both of them use ten iterations to achieve the required accuracy. From this implementation, it is observed that the rotation CORDIC computational process contributes to the critical path of the hardware implementation of Iris

filter and limit the speed of operation. An improvement was therefore devised for the implementation, the vector CORDIC that is used to compute the arctangent function is retained and the rotation CORDIC that is used to compute the cosine function is replaced with a look-up table of the same accuracy to improve the speed efficiency of the design. Results from implementation show that the architecture with look-up table is superior in terms of speed of operation.

2. CORDIC Theory

The CORDIC rotator is normally operated in one of two modes. The first, called rotation by Volder[4], rotates the input vector by a specified angle (given as an argument). The second mode, called vectoring, rotates the input vector to the x axis while recording the angle required making that rotation.

In rotation mode, the angle accumulator is initialised with the desired rotation angle. The rotation decision at each iteration is made to diminish the magnitude of the residual angle in the angle accumulator. The decision at each iteration is therefore based on the sign of the residual angle after each step. Naturally, if the input angle is already expressed in the binary arctangent base, the angle accumulator may be eliminated. For rotation mode, the CORDIC equations are:

$$x_{i+1} = x_i - y_i \cdot d_i \cdot 2^{-i} \quad (1)$$

$$y_{i+1} = y_i + x_i \cdot d_i \cdot 2^{-i} \quad (2)$$

$$z_{i+1} = z_i - d_i \cdot \tan^{-1}(2^{-i}) \quad (3)$$

where

$$d_i = -1 \text{ if } z_i < 0, +1 \text{ otherwise}$$

which provides the following result:

$$x_n = A_n [x_0 \cos z_0 - y_0 \sin z_0] \quad (4)$$

$$y_n = A_n [y_0 \cos z_0 + x_0 \sin z_0] \quad (5)$$

$$z_n = 0 \quad (6)$$

$$A_n = \prod_n \sqrt{1 + 2^{-2i}} \quad (7)$$

In the vectoring mode, the CORDIC rotator rotates the input vector through whatever angle is necessary to align the result vector with the x axis. The result of the vectoring operation is a rotation angle and the scaled magnitude of the original vector (the x component of the result). The vectoring function works by seeking to minimize the y component of the

residual vector at each rotation. The sign of the residual y component is used to determine which direction to rotate next. If the angle accumulator is initialised with zero, it will contain the traversed angle at the end of the iterations. In vectoring mode, the CORDIC equations are:

$$x_{i+1} = x_i - y_i \cdot d_i \cdot 2^{-i} \quad (8)$$

$$y_{i+1} = y_i + x_i \cdot d_i \cdot 2^{-i} \quad (9)$$

$$z_{i+1} = z_i - d_i \cdot \tan^{-1}(2^{-i}) \quad (10)$$

where

$$d_i = +1 \text{ if } y_i < 0, -1 \text{ otherwise}$$

then

$$x_n = A_n \sqrt{x_0^2 + y_0^2} \quad (11)$$

$$y_n = 0 \quad (12)$$

$$z_n = z_0 + \tan^{-1}(y_0 / x_0) \quad (13)$$

$$A_n = \prod_n \sqrt{1 + 2^{-2i}} \quad (14)$$

The CORDIC rotation and vectoring algorithms as stated are limited to rotation angles between $-p/2$ and $p/2$. This limitation is due to the use of 2^0 for the tangent in the first iteration. For composite rotation angles larger than $p/2$, an additional rotation is required. Volder[4] describes an initial rotation $\pm p/2$. This gives the correction iteration:

$$x' = -d \cdot y \quad (15)$$

$$y' = d \cdot x \quad (16)$$

$$z' = z + d \cdot p/2 \quad (17)$$

where

$$d = +1 \text{ if } y < 0, -1 \text{ otherwise}$$

There is no growth for this initial rotation. Alternatively, an initial rotation of either p or 0 can be made, avoiding the reassignment of the x and y components to the rotator elements. Again, there is no growth due to the initial rotation:

$$x' = d \cdot x \quad (18)$$

$$y' = d \cdot y \quad (19)$$

$$z' = z \text{ if } d = 1, \text{ or } z - p \text{ if } d = -1 \quad (20)$$

$$d = -1 \text{ if } x < 0, +1 \text{ otherwise} \quad (21)$$

Both reduction forms assume a modulo $2p$ representation of the input angle. The style of first reduction is more consistent with the succeeding rotations, while the second reduction may be more convenient when wiring is restricted, as is often the case with FPGAs.

The CORDIC rotator described is usable to compute several trigonometric functions directly and others indirectly. Judicious choice of initial values and modes permits direct computation of sine, cosine, arctangent, vector magnitude and transformations between polar and Cartesian coordinates.

2.1. Arctangent computation

The arctangent, $f = \text{Atan}(y/x)$, is directly computed using the vectoring mode CORDIC rotator if the angle accumulator is initialised with zero. The argument must be provided as a ratio expressed as a vector (x,y) . Presenting the argument as a ratio has the advantage of being able to represent infinity (by setting $x=0$). Since the arctangent result is taken from the angle accumulator, the CORDIC rotator growth does not affect result.

$$z_n = z_0 + \tan^{-1}(y_0 / x_0) \quad (22)$$

2.2. Sine and cosine computation

The rotational mode CORDIC operation can simultaneously compute the sine and cosine of the input angle. Setting the y component of the input vector to zero reduces the rotation mode result to:

$$x_n = A_n \cdot x_0 \cos z_0 \quad (23)$$

$$y_n = A_n \cdot x_0 \sin z_0 \quad (24)$$

By setting x_0 equal to $1/A_n$, the rotation produces the un-scaled sine and cosine of the angle argument, z_0 . Very often, the sine and cosine values modulate a magnitude value. Using other techniques (e.g., a look-up table) requires a pair of multipliers to obtain the modulation. The CORDIC technique performs the multiply as part of the rotation operation, and therefore eliminates the need for a pair of explicit multipliers. The output of the CORDIC rotator is scaled by the rotator gain. If the gain is not acceptable, a single multiply by the reciprocal of the gain constant placed before the CORDIC rotator will yield un-scaled results. It is worth noting that the hardware complexity of the CORDIC rotator is approximately equivalent to that of a single multiplier with the same word size.

3. Iris Filter Implementation

3.1. Iris filter operations

Let us denote the intensity of an image and its gradient vector at (x,y) as $I(x,y)$ and $g(x,y)$ respectively. The Iris filter is applied to the gradient vector field of the image. The region of support of the Iris filter R_p is a union of N half-lines radiating from the pixel of interest P , as shown in Figure 1.

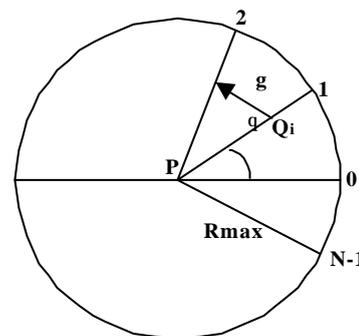


Figure 1: Region of support R_p

The convergence index of the gradient vector g at a point Q_i towards the pixel of interest is defined:

$$f(g | x', y' = f(Q_i)) = \begin{cases} \cos \mathbf{q}, & |g| \neq 0 \\ 0, & |g| = 0 \end{cases} \quad (25)$$

where Q_i is an arbitrary pixel on the i th half-line whose relative coordinates from the pixel of interest are denoted by (x', y') and \mathbf{q} is the orientation of the gradient vector g with respect to the i th half line.

The output of the Iris filter at the pixel of interest (x, y) is defined as:

$$C(x, y) = \frac{1}{N} \sum_{i=0}^{N-1} \max_{Q_i} \frac{\int_P^{Q_i} \cos \mathbf{q}(x', y') dl}{PQ_i}, \quad 0 < \overline{PQ_i} < R_{\max} \quad (26)$$

3.2. Hardware implementation

The block diagram of hardware implementation is shown in Figure 2.

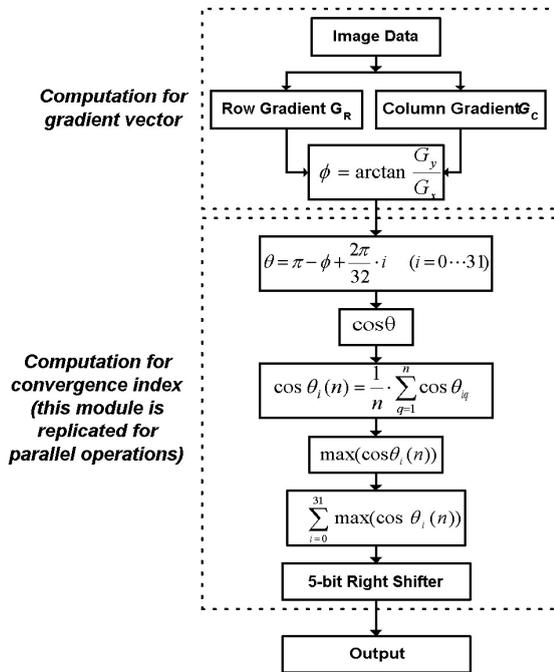


Figure 2 Block diagram of Iris filter implementation

In the computation for gradient vector module, the input intensity image data is represented with 8 bits data. Both row and column gradients, G_R and G_C are obtained by a 3×3 Prewitt-type operator. ϕ is the orientation of gradient which is computed by an arctangent function. This process converts the intensity image to its gradient image for further processing. For each pixel of the gradient image, the computation for convergence index module can be performed in parallel. Cosine calculation is performed for each pixel of each half-line; therefore there are quite a number of trigonometric calculations required for the Iris filter process. The main concern would therefore be the efficient implementation of trigonometric computations.

CORDIC, described earlier, is used for the computation of trigonometric function required by the Iris filter as it requires only shift and add operations that lends well for hardware implementation.

3.3. Iris filter implementation using CORDIC modules

As shown in Figure 2, the Iris filter hardware implementation contains two parts. First part is the computation for gradient vector and the other is the computation for convergence index. In computation for gradient vector, a digital intensity image with 8-bit resolution is provided as input. Prewitt operators for obtaining row and column gradients, G_R and G_C are implemented by adders with signed data, the eventual representation for G_R and G_C are in 9bit. For the implementation of gradient vector module, arctangent computation is required. A CORDIC vector architecture is implemented for the arctangent computation, which require 10 iterations to obtain the angle between 0 to $\pi/4$. This range is used for reducing the number of iterations required and obtained results can then be mapped onto 0 to 2π according to the corresponding symmetric property of tangent and arctangent. 10 iterations of CORDIC corresponds to 10 bits for the range from 0 to $\pi/4$, achieving an accuracy of three floating points. The data width of 14 bits is used to cover the whole 2π range.

In the computation of convergence index, the implementation becomes more complex. Inputs to the computation of convergence index module are based on the results obtained from the first module. Each pixel in the gradient image will be considered as a pixel of interest. Iris filter, which is the union of 32 half-lines radiating from the pixel of interest, will be applied in parallel. For each pixel on each half-line, the convergence index $\cos \mathbf{q}$ of the gradient vector at this pixel is computed. In full CORDIC operation implementation, this can be performed by CORDIC rotator architecture, in which, 10 iterations are used and 10-bit data width is necessary to represent -1 to 1 . Similarly, to reduce the iteration cycles required by CORDIC rotator, 0 to 2π range is mapped from 0 to $\pi/4$ according to the corresponding symmetric property of cosine. Thereafter 10-bit divider and comparator are implemented to obtain the maximum convergence index of gradient vectors in each half-line. Accumulator is used for summing the maximum of the cosine values of 32 half-lines. A 5-bit right shifter is eventually used to obtain the output of Iris filter at the pixel of interest, which is in 14-bit representation.

3.4. Improvement to the Iris filter implementation

In the above implementation of Iris filter with CORDIC, the $\cos \mathbf{q}$ computation in the second module

for obtaining the convergence index requires 32 rotation CORDIC computations for all of the 32 half-lines of each pixel in the region of interest. Also each of this rotation CORDIC computation requires 10 iterations. This method of implementation will affect the processing speed since each iteration need one-clock cycle. As expected, the rotation CORDIC is the processing bottleneck in the eventual hardware implementation of Iris filter. A look-up table with multi-port RAM structure is proposed to replace the rotation CORDIC to improve the speed performance. However a similar method could not be applied to the first CORDIC module in computation of gradient vector, to replace it with a look-up table for arctangent computation of ϕ , a two-dimension table would be necessary. The eventual table required will be of considerable size since the range of gray level is from 0 to 255. Therefore the look-up table option for arctangent implementation would be tedious and not improve the overall performance.

4. Implementation Analyses and Simulation Results

In order to verify the feasibility of the proposed architecture, both architectures using all CORDIC modules and CORDIC with look-up table (LUT) had been modelled using VHDL and synthesized with Synplicity Synplify Pro 7.0.3 targeted at the Xilinx Virtex x2v2000 FPGA device to obtain the area-time analysis of an experimental application. The area-time comparison results of both architectures are listed in the Table 1.

Table 1 Complexity comparison by FPGA

	Full CORDIC	CORDIC with LUT
Time	691 ns	245.38 ns
Area	7513 out of 10752 (69%)	9266 out of 10752 (86%)
External IOBs	13 out of 408 (3%)	228 out of 408 (55%)

From the above table, it is observed that CORDIC with look-up table architecture could achieve 3 times better speed performance when compared to full CORDIC architecture.

Figure 3(a) shows a 256x256 test image with a background intensity of zero. Its intensity patterns is given by

$$f(i, j) = \begin{cases} 190 \exp\left(-\frac{i^2 + j^2}{2s^2}\right), & \sqrt{i^2 + j^2} \leq 2.5s \\ 0, & \sqrt{i^2 + j^2} > 2.5s \end{cases}$$

The standard deviations of the object are 25 pixels. The contrast between the rounded object and its background at its boundary is very weak. After applying the proposed Iris filter implementation

architecture, the boundary of the object is enhanced and exhibits a striking contrast to its background, which is shown in Figure 3(b).

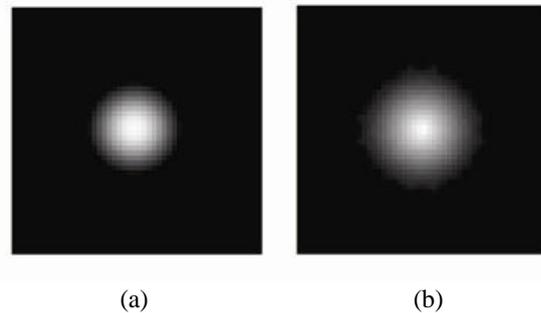


Figure 3 Gaussian function with (a) vague boundary before entering the Iris filter architecture process (b) enhanced boundary after Iris filter architecture process

5. Conclusion

Hardware architectures for implementation of the Iris filter have been proposed. Iris filter operation involves much computational intensive trigonometric computations, direct implementation of such computations are tedious and therefore efforts are on using alternative computational methods to achieve similar computation functions. The design employs two CORDIC (vector and rotation CORDIC, respectively) modules in the implementation. However this implementation have speed concerns due to the high number of trigonometric computations required. The improved design incorporates a look-up table to replace the rotation CORDIC, which contributes to the bottleneck for the first implementation. FPGA based investigations show that the architecture that incorporates look-up table achieves a computation speed-up of about 3 times when compared to the first implementation that uses all CORDIC modules.

References

- [1] H. Kobatake and S. Hashimoto, "Convergence index filter for vector fields", *IEEE Trans. Image Processing*, vol. 8, pp. 1029-1038, August, 1999.
- [2] H. Tian, T. Srikanthan and K. Vijayan Asari, "Automatic segmentation algorithm for the extraction of lumen region and boundary from endoscopic images", *IEE Journal of Medical & Biological Engineering & Computing*, vol. 39, no. 1, pp. 8-14, 2001.
- [3] H. Tian, T. Srikanthan, K. Vijayan Asari and C. S. Lim, "A hardware efficient technique for rapid lumen segmentation from endoscopic images", in *Proceedings of Third International Symposium on Communication System, Networks, and Digital Signal Processing*, pp. 260-263, July 2002.

- [4] J. E. Volder, "The CORDIC trigonometric computing technique", IRE Trans. Electronic Computers, vol. 8, pp. 330-334, September 1959.

Local Optimal Algorithm for High Performance Solution of Reconfigurable VLSI Arrays

Wu Jigang, Srikanthan Thambipillai (Senior Member, IEEE)
Centre for High Performance Embedded Systems
Nanyang Technological University
Singapore, 639798
{ asjgwu, astsrikan }@ntu.edu.sg

Abstract

Given an $m \times n$ mesh-connected VLSI array with some faulty elements, the reconfiguration problem is to find an maximum-sized fault-free sub-array under the row and column rerouting scheme. This problem has already been shown to be NP-complete. The power awareness problem of the reconfigurable array is first proposed in this paper. A local optimal algorithm based on dynamic programming has been presented for the same. The algorithm is simple but very efficient. The performance of the proposed algorithm is more powerful than that of the older algorithm, without loss of harvest.

Key words: Degradable VLSI array, reconfiguration, heuristic algorithm, fault-tolerance, NP-completeness.

1. Introduction

Speed, area and cost have always been the major concerns of VLSI (Very Large Scale Integration) designers. The designs were made such that they were high in speed, small in area and low cost. In recent years, the growth of personal computing devices (portable computers and real time audio and video-based multimedia products) and wireless communication systems (personal digital assistants and mobile phones) has forced designers to make high performance systems that can work on very low power. Also, with the development of deep sub-micron technologies like 90nm, the power dissipation due to the switching capacitance would be lesser if the interconnect wire lengths in the VLSI array are shorter. It is clear that in the absence of lower-power design techniques, reliability is strongly affected by power consumption. Usually, high power dissipation implies higher operating temperatures, which, in turn, may induce a failure in the system. Also, low power dissipation implies that the system could be clocked at higher

speeds without increasing the power penalty. This leading to much higher performance at very low power.

Mesh-connected topology is of importance due to its simple structure and its good performance in practice and is becoming popular for reliable and high-speed communication switching. Reconfiguration in a mesh structure has been investigated extensively, e.g., [1-3] for redundancy approach and [4-9] for degradation approach. However, no previous work has been carried out in order to provide high performance sub-mesh for degradable VLSI arrays with faults. This paper proposes the problem of finding a fault-free sub-array which demonstrate higher performance without increase in power dissipation in a two-dimensional mesh-connected VLSI arrays, and then we propose a simple but efficient heuristic algorithm that has higher performance requirements, while maintaining the same harvest.

2. Preliminaries

Let host array H be the original array obtained after manufacturing. Some of the elements in this array may be defective. A target array T is a fault-free subarray of H after reconfiguration. The rows (columns) in the host array and target array are called the *physical row (columns)* and *logical rows (columns)*, respectively.

In this paper all the assumptions in architecture are the same as that in [4-7]. Neighboring elements are connected to each other by a four-port switch. A target array T is said to contain $\{R_1, R_2, \dots, R_k\}$ if each logical column in T contains exactly one fault-free element from each of the rows. The previous problems and the related algorithms are as follows.

Problem \mathcal{RP} : Given an $m \times n$ mesh-connected host array, find a maximal sized target array under the row and column rerouting scheme that contains the selected rows.

The problem \mathcal{RP} is optimally solvable in linear time, and an algorithm, called *Greedy Column Rerouting (GCR)*,

was proposed for solving \mathcal{RP} in [5, 6]. GCR constructs the target array in a left-to-right manner. In each iteration, it produces the current leftmost logical column. A detailed description of GCR can be found in [6].

As shown in Fig. 1, there are 6 possible types of link-ways for a target array. They can be categorized into two classes based on the number of the switches used. In this paper, a link-way that uses only one switch to connect neighboring elements in a target array is called a *high performance interconnect*, while a link-way using two switches is called a *low performance interconnect*. In Fig. 2, (a) and (d) are high performance interconnects, while the others are low performance one. (a), (b) and (c) are used for row rerouting, while (d), (e) and (f) are used for column rerouting. Obviously, the smaller the number of the low performance interconnects the target array has, the lesser is the system delay, and therefore, the higher the performance.

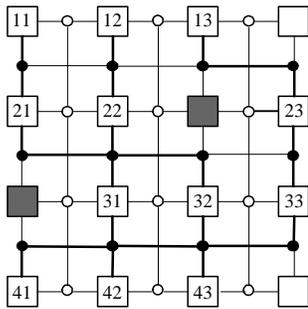


Figure 1. An example of target array.

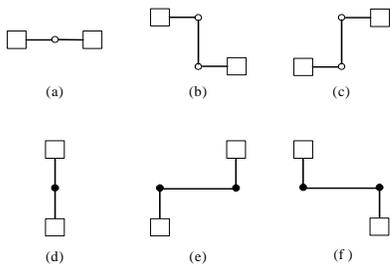


Figure 2. All possible link-ways in the target array.

In this paper, the maximal sized target array with the minimal number of low performance interconnects is called the high performance target array or high performance solution. The problem that we proposed is described as follows.

Problem \mathcal{HP} : Given an $m \times n$ mesh-connected host array, find a high performance solution under the row and column rerouting scheme that contains the selected rows.

We aim to propose a heuristic algorithm to solve this optimization problem in this paper.

Given a host array H of size $m \times n$, let U be the set of logical columns that pass through each of the rows. Thus, each logical column in U contains exactly one fault-free element from each of the rows. For any two logical columns, C_p and C_q in U , and $p \neq q$, We say that $C_p \leq C_q$ if the i th element in C_p lies to the left of, or is identical to, the i th element in C_q , for $1 \leq i \leq m$.

Assume $\mathcal{B}_l, \mathcal{B}_r \in U$ and $\mathcal{B}_l \leq \mathcal{B}_r$. We use $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r)$ to indicate the area consisted of the fault-free elements bounded by \mathcal{B}_l and \mathcal{B}_r , including \mathcal{B}_l but not including \mathcal{B}_r . \mathcal{B}_l and \mathcal{B}_r are called the left boundary of $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r)$ and the right boundary of $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r)$, respectively.

3. Local Optimal Algorithm

Assume the solution of GCR is the target array T with k logical columns: C_1, C_2, \dots, C_k . We revise each C_i , for $i = 1, 2, \dots, k$. The proposed algorithm, denoted as $LGCR$ revises the solution of GCR to obtain an approximate high performance solution. It starts from the logical column C_k , revises it and then revises C_{k-1}, \dots, C_1 one by one. In each iteration, $LGCR$ works in $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r)$. In the first iteration, \mathcal{B}_l is set to be C_k , and \mathcal{B}_r is set to be a virtual column that lies to the right of the n th physical column of the host array. Let $col(u)$ and $col(v)$ denote the physical column index of the element u and v , respectively.

Definition 1 For each fault-free element u in R_i ,

1. $Adj^-(u) = \{v : v \in R_{i-1}, v \text{ is fault-free, and } |col(u) - col(v)| \leq 1\}$ for $2 \leq i \leq k$.
2. $Adj^+(u) = \{v : v \in R_{i+1}, v \text{ is fault-free, and } |col(u) - col(v)| \leq 1\}$ for $1 \leq i \leq k - 1$.
3. For arbitrary $v \in Adj^-(u)$ ($Adj^+(u)$), v is called the left adjacent, the middle adjacent or the right adjacent of u if $col(v) - col(u) = -1$, $col(v) - col(u) = 0$ or $col(v) - col(u) = 1$, respectively.

Definition 2 A logical column is called the local optimal column related to \mathcal{B}_l and \mathcal{B}_r if and only if it is the logical column of the minimal number of the low performance interconnects in $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r)$.

In this section, we propose an algorithm based on dynamic programming to find the local optimal column, and use it to replace the original logical column obtained by GCR .

Given a left boundary \mathcal{B}_l and a right boundary \mathcal{B}_r , we model the local optimal column in $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r)$ as the shortest path of the directed network $G = (V, E)$. Here, V is the set of the vertices which represent the fault-free elements between the two boundaries. E is the subset of $V \times V$.

There exists a directed edge, say $e(u, v)$, from u to v if and only if $v \in Adj^+(u)$. Each edge in E has associated with it a non-negative cost. The cost of $e(u, v)$ is defined as

$$c(u, v) = \begin{cases} 0, & \text{if } col(u) - col(v) = 0, \\ 1, & \text{if } |col(u) - col(v)| = 1. \end{cases}$$

The cost of a path is defined as the sum of the costs of the all edges in the path. Hence, E is the set of these edges with each edge having a cost of 0 or 1. According to the definitions of Adj^- and Adj^+ , each node in V has at most 3 input edges and at most 3 output edges.

For the simplification of denotation, let $c(x, y)$ be the cost of the shortest path from x to y , where x and y are vertices in V or the rows. Then, for arbitrary $u \in \mathcal{A}[\mathcal{B}_l, \mathcal{B}_r]$, $c(u, R_i)$ satisfies the following condition

$$c(u, R_i) = \begin{cases} 0, & \text{if } row(u) = i, \\ \min_{v \in \mathcal{L}} \{c(u, v) + c(v, R_i)\}, & \text{if } row(u) < i \leq m. \end{cases} \quad (1)$$

where $\mathcal{L} = Adj^+(u) \cap \mathcal{A}[\mathcal{B}_l, \mathcal{B}_r]$. For the special case when \mathcal{L} is empty, we define $c(u, R_i) = \infty$. The cost of the shortest path in $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r]$ from R_1 to R_m is defined as

$$c(R_1, R_m) = \min_{u \in R_1} \{c(u, R_m)\}. \quad (2)$$

The local optimal column related to \mathcal{B}_l and \mathcal{B}_r is the shortest path from the row R_1 to the row R_m in $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r]$ in the network. The calculation of the shortest path is based on bottom-up technique. The informal overview of the algorithm, denoted as *LGCR*, is that:

1. Initialize $c(v, v)$ to be 0 for each $v \in R_m \cap \mathcal{A}[\mathcal{B}_l, \mathcal{B}_r]$.
2. Then, calculate $c(u, R_i)$ by using the formula (1) for each $u \in R_{i-1} \cap \mathcal{A}[\mathcal{B}_l, \mathcal{B}_r]$, for $i = m - 1, m - 2, \dots, 1$.
3. Select the minimal one according to the formula (2) as the overhead cost of the shortest path.
4. Output the calculated shortest path in the trace-back process.

In the trace-back process, we select the rightmost shortest path as the local optimal column related to the current left boundary and the right boundary in order to save as many fault-free elements in $\mathcal{A}[\mathcal{B}_l, \mathcal{B}_r]$ as possible for constructing the next local optimal column.

It is not hard to show that the worst time complexity of *LGCR* is bounded by $O((n - k) \cdot N)$, and the average time complexity of *LGCR* is bounded by $O(N)$, where N is the number of the fault-free elements.

4. Experimental Results

The proposed algorithms and *GCR* are implemented in C and run on a Pentium IV 2.2 MHZ computer with one GB RAM. The whole experiment is divided into two parts - one is for uniform fault distribution in the whole mesh, which is corresponding to random fault model. And, the other is for the uniform fault distribution in different localized portions of the mesh, which is corresponding to the clustered fault model. In Table 1 and Table 2, the attribute *nois* denotes the **number of the low performance interconnects**. *imp* stands for the improvement in *nois* over *GCR*. It is calculated by

$$\left(1 - \frac{nois_of_LGCR}{nois_of_GCR}\right) \times 100\%.$$

In Table 1, data are collected for host arrays of different sizes from 32×32 to 256×256 . For each algorithm, for smaller or medium sized host arrays, *nois* increases with the increase in the number of faulty elements in the host array. This is because that there will be a percentage faulty density beyond which *nois* will fall, as *nois* is equal to 0 both for 0% fault and for 100% fault. While for larger sized host arrays, *nois* first increases and then decreases. This decrease is due to heavy decrease in target size as the fault density increases beyond a certain point. The improvement of *LGCR* over *GCR* is significant, especially for the host arrays of small size or small fault density, which is occurred more frequently in applications.

In Table 2, data are collected for 256×256 sized host array, averaged over 10 random instances, each for a localized fault in the center 1/8, 1/4, 1/2 and 3/4 of the host array. Here, center 1/8, 1/4, 1/2 and 3/4 in Table 2, is the situation when faulty elements are located in center area consisting of $32 * 32$ elements, $64 * 64$ elements, $128 * 128$ elements, $192 * 192$ elements and the full host array, respectively. For center 1/8 and 1/4, considering 10% fault density was not possible as 10% of 256×256 comes out to be 6554 elements, while center 1/8 comes out to be a $32 * 32$ mesh, i.e., 1024 elements, and center 1/4 comes out to be 64×64 mesh, i.e., 4096 elements. So it is clear that these two cannot accommodate 10% faulty elements. Hence, the row for these cases is marked as NA (Not Applicable).

From table 2, we observe that, for 1% faulty density and localized fault in center 1/8 portion of the host array, *nois* comes out to be 0. This is because the center gets so much concentrated with the faulty elements that no logical column passes through that portion in which the faulty elements are located. *GCR* is not better for performance because it uses the leftmost strategy. The improvement over *GCR* is more significant for localized fault distribution than for random fault distribution. It increases as the spread of faulty elements in the host array decreases.

As can be seen from above tables, in all the cases, the

Table 1. The performance comparison of the algorithms *GCR* and *LGCR* for uniform fault distribution, average of 10 random instances of different size.

Host array		Target array	Performance		
Size $m \times n$	Fault (%)	Size $m \times k$	<i>GCR</i> <i>nois</i>	<i>LGCR</i> <i>nois imp</i>	
32 × 32	0.1	32 × 31	33	0	100%
32 × 32	1	32 × 30	265	76	71%
32 × 32	10	32 × 23	452	179	60%
64 × 64	0.1	64 × 63	292	59	80%
64 × 64	1	64 × 60	1471	394	73%
64 × 64	10	64 × 47	1992	835	58%
128 × 128	0.1	128 × 126	2012	486	76%
128 × 128	1	128 × 122	8168	2651	68%
128 × 128	10	128 × 95	8294	3885	53%
256 × 256	0.1	256 × 253	13827	3678	73%
256 × 256	1	256 × 246	38155	14533	62%
256 × 256	10	256 × 193	34167	17218	50%

Table 2. The performance comparison of the algorithms *GCR* and *LGCR* for clustered faulty distribution, average of 10 random instances of different size.

Host array		Target array	Performance		
Size $m \times n$	Fault (%)	Size $m \times k$	<i>GCR</i> <i>nois</i>	<i>LGCR</i> <i>nois imp</i>	
1/8	0.1	256 × 249	3290	154	95%
	1	256 × 224	8446	0	100%
	10	NA	NA	NA	NA
1/4	0.1	256 × 251	5911	524	91%
	1	256 × 231	10179	814	92%
	10	NA	NA	NA	NA
1/2	0.1	256 × 252	10175	1449	86%
	1	256 × 241	16918	4081	76%
	10	256 × 145	15430	881	94%
3/4	0.1	256 × 253	12553	2707	78%
	1	256 × 244	27090	8889	67%
	10	256 × 176	26208	7640	71%

revised algorithm gives much better results than the older *GCR* algorithm. Even for arrays of a specific size, as the percentage of faulty elements in the array increases, our new algorithm performs much better than the old algorithm. We can therefore say that for all cases our algorithms gives us a high performance solution that consumes much low power than the solutions of the older algorithms like *GCR*.

5. Conclusions

High performance VLSI design has emerged as a major theme in the electronics industry today. It can be achieved at all levels of the VLSI system. In this paper, we have proposed a new problem based on a high performance solution for reconfigurable VLSI arrays and have presented an algorithm for the same. The proposed algorithm is based on a dynamic programming that is easy to implement. Its performance is very high and the harvest remains unchanged. Experimental results reflect the underlying characteristics of the proposed algorithm.

References

[1] Y. Y. Chen, S. J. Upadhyaya and C. H. Cheng, "A Comprehensive Reconfiguration Scheme for Fault-tolerant VLSI/WSI Array Processors", *IEEE Trans. on Computers*, vol. 46, no. 12, pp. 1363-1371, Dec. 1997.

[2] T. Horita and I. Takanami, "Fault-tolerant Processor Arrays Based on the 1.5-track Switches with Flexible

Spare Distributions", *IEEE Trans. on Computers*, vol. 49, no. 6, pp. 542-552, June 2000.

[3] Li Zhang, "Fault-Tolerant meshes with small degree", *IEEE Trans. on Computers*, col. 51, No.5, pp.553-560, May, 2002.

[4] S. Y. Kuo and I. Y. Chen, "Efficient reconfiguration algorithms for degradable VLSI/WSI arrays," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 10, pp. 1289-1300, Oct. 1992.

[5] C. P. Low and H. W. Leong, "On the reconfiguration of degradable VLSI/WSI arrays," *IEEE Trans. Computer-Aided Design of integrated circuits and systems*, vol. 16, no. 10, pp. 1213-1221, Oct. 1997.

[6] C. P. Low, "An efficient reconfiguration algorithm for degradable VLSI/WSI arrays," *IEEE Trans. on Computers*, vol. 49, no. 6, pp.553-559, June 2000.

[7] W. Jigang, T. Srikanthan, "A Run-time Reconfiguration Algorithm for VLSI Arrays", *Journal of Systems Architecture*, vol. 49, pp. 23-31, 2003.

[8] Wu Jigang, Schroder Heiko & Srikanthan Thambipillai, "New architecture and algorithms for degradable VLSI/WSI arrays", in Proc. Of 8th International Computing and Combinatorics Conference, 2002, Singapore (COCOON'02), *Lecture Notes in Computer Science*, vol. 2387, pp.181-190, Aug, 2002.

[9] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, "The design and analysis of computer algorithms", Addison-Wesley, Reading, Mass., 1974.

Multiple Carries Prediction to Enhance the Performance of a 32-bit Estimated Carry Adder

E. M. ASHMILA, S. S. DLAY, O. R. HINTON

School of Electrical, Electronic and Computer Engineering

University of Newcastle upon Tyne, Merz Court, NE1 7RU, UK

E.M.Ashmila@ncl.ac.uk, S.S.Dlay@ncl.ac.uk, oliver.hinton@ncl.ac.uk

Abstract

This paper presents a novel methodology to significantly improve the speed performance of a 32 bit estimated carry (ESTC) adder and extends the theory of the ESTC adder, and shows that further performance benefits can be achieved by using the statistical approach together with multiple carries. The simulation results for multiple carries, shows that the 32-bit adder can achieve dramatic speed advantages over other adders. Furthermore, the comparison in term of delay-area product shows a saving of 44.2% over a ripple adder, 41% over carry select adder; with ripple adder elements, and over 26% on carry select look-ahead adders.

Introduction

The performance of any digital systems or processors is very influenced by the speed of their adders. Many fast adders for high speed have been recently are introduced [1, 2, 3]. However, always high-speed adders are necessary for high performance in digital applications. Thus, research into improving the performance of digital adders is an important field.

Adders can be implemented with both synchronous and asynchronous circuit design. However, in the recent years, asynchronous circuits design have attracted much more attention than synchronous circuits [4, 5, 6]. Because they offer several benefits compared with synchronous circuit, including early completion (performance), low power, improved electromagnetic emissions, superior adaptability and the elimination of a global clock signal [7, 8, 9]. The addition speed is limited by the speed of the propagation carry across the word because the carry propagation delay has a significant effect on system performance. Thus to achieve high performance adder, the main issues to be addressed are the length of the carry propagation path in the addition process and its reduction. Resolving these issues helps in the design of faster adders, and in turn, produces a faster CPU data path. Thus, many different adder architectures have been proposed over the last few years with various trade offs; each one has its advantages and disadvantages [10, 11].

Recently, our research team introduced a new adder, estimated carry adder (ESTCA) [12]. It uses statistical probability of a carry being in a particular state to

design a 32-bit adder, which will operate at the speed of a 16-bit adder for majority of additions. This adder is appropriate for asynchronous designs and has speed advantages compared with other adder designs.

This paper extends further the theory for the ESTC methodology and shows a novel methodology to achieve further performance benefits by using the statistical approach together with multiple carries to predict a correct carry from the least significant (LS) to the most significant (MS) half of a 32-bit ESTC.

Multiple Carry Prediction for ESTC Adder

A 32-bit ESTCA, Fig.1, derived from the idea of carry select adder but it eliminates one of the MS adders and the multiplexer and instead uses a control circuits to generate the carry out from LS adder [12] based on that the bits of the operands of LS adder can be used to predict carry out from LS to MS adder. Thus, it reduces the overall average time delay, and the amount of hardware overhead for the adder. It is divided into LS (A_{0-15} and B_{0-15}) and MS (A_{16-31} and B_{16-31}) adders. The carry out from LS to MS adders can be estimated by using MS bit of the first half, A_{15} and B_{15} , and it will be correct for 75% of additions, and in this case a 32-bit adder will operate as 16-bit adder in parallel. For incorrect estimation, the true carry out C_{15}^T is applied to the MS adder, and the addition time will be delay time of a full 32-bit adder.

The performance improvements achieved by using single carry shows that the speed of this adder could be improve further by investigating how the previous carries (C_{14} , C_{13} , ... etc) affect the correct prediction. By including A_{14} , B_{14} bits in our analysis we found, by inspecting the full truth table; Table 1, that the probability of the C_{15}^E being equal with C_{15}^T is 0.75 when logic $A_{15} \cdot B_{15}$ is used for prediction. For incorrect estimates addition logic is required to provide the results needed when the first logic fails.

The new logic could modify in the light of C_{14}^E carry and takes the following form:

$$A_{14} \cdot B_{14} (A_{15} + B_{15})$$

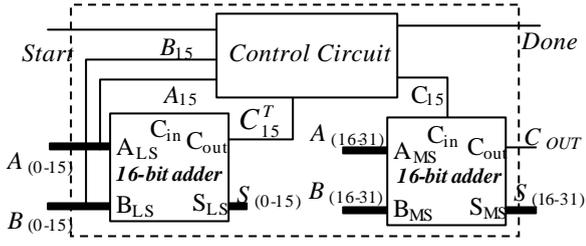


Fig. 1: A 32-bit Estimated Carry Adder

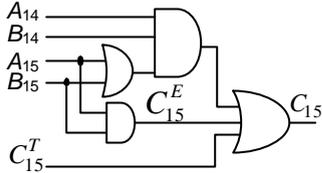


Fig. 2: Carry Control Circuit Using Two Carries for Prediction

Using this logic the C_{15}^T can be estimated for a further 50% of the incorrect cases (Table 1). As result of this, the probability of prediction being correct has increases from 75% to 87.5% of the additions, and the full estimation can be represented as the following:

$$P(C_{15}^E = C_{15}^T) = P((A_{15} = B_{15}) + (A_{15} \neq B_{15}) \cdot (C_{14}^E = C_{14}^T)) = 0.875 \quad (1)$$

$$P(C_{15}^E \neq C_{15}^T) = P((A_{15} \neq B_{15}) \cdot (C_{14}^E \neq C_{14}^T)) = 0.125 \quad (2)$$

Where $P(C_{15}^E = C_{15}^T)$ is the probability of the estimated carry (C_n^E) and true carry (C_n^T) being equal, and $P(C_{15}^E \neq C_{15}^T)$ is the probability of the estimated and true carry being unequal.

Using these logics a new carry control circuit has been designed, Fig. 2. The modification amounts to two extra gates and the enlargement of the OR gate in the original control circuit [16] from two inputs to three inputs, and the control circuit remains compact with this configuration and is improved to give 87.5% correct estimates. Thus, the average time delay for the adder will be reduced since for 87.5% of addition time, the 32-bit adder will operate as two 16-bit adders in parallel. For the incorrect estimates, the C_{15}^T is applied to the MS adder and the addition will be performed in the full 32-bit adder time delay plus the time needed to detect the state and apply the correct value. The carry control circuit, Fig.2, has been used to design the control Circuit for ESTCA that generate the carry out from LS adder to MS adder, Fig.3. The carry estimation is performed using these logics:

$$A_{15} \cdot B_{15} \text{ and } A_{14} \cdot B_{14} (A_{15} + B_{15})$$

In most of the cases the first logic could manage to generate the correct estimated carry, however, when it fails the second logic generates the correct carry.

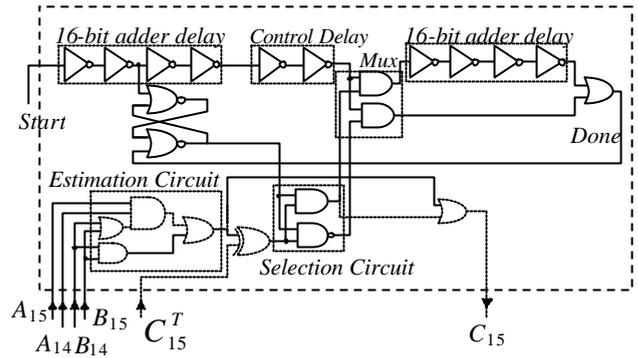


Fig. 3: The Schematic of the Control Circuit Using Two Carries for Prediction

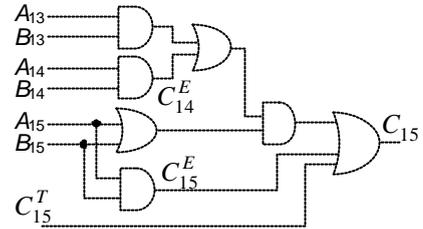


Fig. 4: Carry Control Circuit Using Three Carries for Prediction

For four specific cases of inputs neither of the logics could estimate the correct carry. Here the C_{15}^T must be applied to the MS adder, i.e. the circuit waits for the LS adder to finish and propagates the C_{15}^T , then the MS adder continues with the addition (a 32-bit delay). Further improvement can be achieved by looking further back to other bits and by using the same idea; the C_{15}^E carry can be modified in the light of previous carries. By using three carries for prediction, the results of C_{15}^T being correct are increases to 93.75%, and the probabilities of the estimated and true carry obtained as following:

$$P(C_{15}^E = C_{15}^T) = P((A_{15} = B_{15}) + (A_{15} \neq B_{15}) \cdot (C_{13}^E = 1 + C_{14}^E = 1)) = 0.9375 \quad (3)$$

$$P(C_{15}^E \neq C_{15}^T) = P((A_{15} \neq B_{15}) \cdot (C_{13}^E = 0 + C_{14}^E = 0)) = 0.0625 \quad (4)$$

Thus by this design the correct estimates improved to 93.75% and the control circuit is given in Fig. 4.

The theory can be further extended and it suggests that when a new carry is used for prediction the speed performance increases by half of the time taken by the previous 32-bit addition, see Table 2. The graph shown in Fig.5 indicates an increase in the probability of a correct prediction as a new carry is used for prediction. In addition, the performance will increase dramatically by looking back to say carry C_{11} . However, by going further back the increase in performance versus area is less spectacular.

Table 2: Probabilities for Multiple Carry

Multiple Carry	Probability of a correct prediction	Probability of incorrect prediction
C ₁₅	0.75000	0.25000
C ₁₄	0.87500	0.12500
C ₁₃	0.93750	0.06250
C ₁₂	0.96875	0.03125
C ₁₁	0.98438	0.01562
C ₁₀	0.99219	0.00781
C ₉	0.99609	0.00391
C ₈	0.99805	0.00195
C ₇	0.99902	0.00098
C ₆	0.99951	0.00049
C ₅	0.99976	0.00024
C ₄	0.99988	0.00012
C ₃	0.99994	0.00006
C ₂	0.99997	0.00003
C ₁	0.99998	0.00002

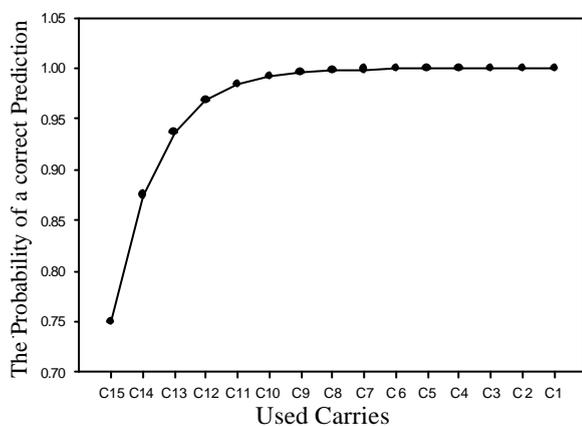


Fig. 5: Performance Improvement of ESTC Adder

The results for Multiple ESTC adder

In most implementations of the adders, the area and speed trade off is the most important measure that characterizes the adders. Thus the delay-area product has been used as a measure. Simulations have been performed using PSPICE with parameters from a 0.125 μm CMOS technology. The delay, transistor number, and delay-area product for all configurations have been computed for ESTCA and are listed in Table 3. The results are organized into two groups for the ESTCA/RA and ESTCA/CLA adders. It clearly can be seen that the area requirements grow as more carries are used for prediction, but the delay is reduced. Since the speed advantage gained after C_{11} is not practical because it is limited by the complexity, the comparison is carried out until five carries. It clearly can be seen that the area requirements grow as more carries are used for prediction, but the delay is reduced.

Since the speed advantage gained after C_{11} is not practical because it is limited by the complexity, the comparison is carried out until five carries.

The results of performance comparisons for the multiple carry ESTC adder with other designs show that

Table 3: Comparison for 32-bit ESTC adder using multiple carry for prediction

Adder Structure	Carry for Prediction	Delay Ps	Trans. No.	Delay-Area Product
32-bit ESTCA/RA	Single Carry	627	1530	959310
	Two Carries	571	1546	882550
	Three Carries	544	1556	845842
	Four Carries	533	1568	835509
	Five Carries	531	1582	839647
32-bit ESTCA/CLA	Single Carry	166	1756	291496
	Two Carries	151	1772	267413
	Three Carries	144	1782	255824
	Four Carries	141	1794	252147
	Five Carries	140	1808	252758

there is considerable performance achieved to be made when a new carry is used for prediction. E.g., using five carries for prediction, the ESTC/RA adder offers a speed advantage of 50.1% over the RA/RA for only 12.4% of increasing in hardware area, and 9.1% over the CSA/RA and it uses 34.7% less area. Comparing the adders that are based on CLA elements, the results show that our ESTCA/CLA is 31.1% faster than the RA/CLA for only 7.6% more area, and CSA/CLA is 13.4% faster but uses 56.6% more area. Also in terms of delay-area product, the ESTCA multiple carry for prediction, using four carries for example, shows that ESTCA/RA offers a saving of 44.2% over RA/RA and 41% over CSA/RA, and the ESTCA/CLA is a saving of over 26% on CSA/CLA and RA/CLA designs.

The comparison of ESTC/RA in terms of delay-area product; the multiple carry prediction for ESTCA, using four carries for example, shows that ESTCA/RA offers a saving of 44.2% over RA/RA and 41% over CSA/RA, and the ESTCA/CLA is a saving of over 26% on CSA/CLA and RA/CLA designs. The results of the ESTC demonstrate that the ESTC adder is relatively fast and area efficient for VLSI applications.

Conclusion

This paper presents a novel methodology to improve the speed performance of a 32-bit ESTC adder, and reduces the amount of hardware overhead for the adder. The methodology uses multiple carries to predict a correct carry from the LS to the MS half of a 32-bit adder. Theoretically when a new carry is used for prediction the performance in time increases by half of the time taken by the previous 32-bit addition, and this could increase to nearly 100% but the complexity and area overhead of the control circuits start to outweigh any speed advantages.

The simulation results show that the 32-bit ESTC adder using multiple carries can achieve dramatic speed advantages over other adders. The overall results demonstrate the ESTC adder is simple, fast and is suitable for implementation using VLSI technology.

Since the methodology has shown substantial improvement, the research will be continuing and

concentrates on: implementing the completion control circuits for 64-bit adders, and investigate and extend the methodology for other common digital functions such as multipliers.

References

- [1] Sun Xu-guang, Mao Zhi-gang, Lai Feng-chang: 'A 64 bit parallel CMOS adder for high performance processors', IEEE Proc. of Asia-Pacific Conference on ASIC, 2002, pp.205-8.
- [2] Sakamoto M, Hamano D, Morisue M.: 'Design of a multiple-operand redundant binary adder', Systems&Computers, Japan, 33(10), 2002,pp.1-9.
- [3] Chih-Jen Fang, Chung-Hsun Huang, Jinn-Shyan Wang, Ching-Wei Yeh: 'Fast and compact dynamic ripple carry adder design', IEEE Asia-Pacific Conference on ASIC, 2002, pp.25-8.
- [4] Sklavos N, Koufopavlou O.: 'Asynchronous low power VLSI implementation of the International Data Encryption Algorithm', 8th IEEE International Conference on Electronics, Circuits and Systems, IEEE Part vol.3, 2001, pp.1425-8.
- [5] Singh M, Nowick SM.: 'Fine-grain pipelined asynchronous adders for high-speed DSP applications', Proceedings IEEE Computer Society Workshop on VLSI, System Design for a System-on-Chip Era. IEEE Compu. Soc., 2000, pp.111-18.
- [6] Beerel PA.: 'Asynchronous circuits: an increasingly practical design solution', Proc. Inter. Symposium on Quality Electronic Design. IEEE Comput.Soc. 2002, pp.367-72.
- [7] Scott Hauck: 'Asynchronous design methodologies an overview', IEEE Proc., 83 (1), 1995, pp 69-93.
- [8] D. Johnson, V. Akella: 'Design and analysis of asynchronous adders', IEE Proc. Comput. Digit. Tech., 145 (1), 1998, pp 1-7.
- [9] A. Davis and S. M. Nowick: 'Introduction to asynchronous circuit design', University of Utah technical report, Departme nt of Computer Science, UUCS-97-013, Sept. 1997.
- [10] Mark A. Franklin and Tienyo Pan: 'Performance comparison of asynchronous adders', IEEE Proc. of Inter. Symposium on Advanced Research in Asynchronous Circuits and Systems, ASYNC-94, 1994, pp. 117-125.
- [11] C. Nagendra, M. Irwin and R. Owens: 'Area-time-power trade-offs in parallel adders', IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Proc., 43 (10), 1996, pp 689-702.
- [12] W.F.Wallace, S.Dlay,and O. Hinton: 'Probabilistic Carry Estimate for Improved Asynchronous Adder Performance', IEE Proc.-Comput. Digit. Tech., vol. 148 (6), 2001, pp 221-226.

Table 1: Carry Logic Truth Table for Multiple Carry Prediction Using Bits A_{14}, B_{14} .

C_{13}	A_{14}	B_{14}	C_{14}^T	C_{14}^E	A_{15}	B_{15}	C_{15}^T	C_{15}^E	C_{15}^{EM*}	Logic gate	Performing mode
0	0	0	0	0	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	0	0	0	0	0	1	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	0	0	0	0	1	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	0	0	0	0	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel
0	0	1	0	0	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	0	1	0	0	0	1	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	0	1	0	0	1	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	0	1	0	0	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel
0	1	0	0	0	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	1	0	0	0	0	1	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	1	0	0	0	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel
0	1	1	1	1	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
0	1	1	1	1	0	1	1	0	1	$A_{14}B_{14}(A_{15}.B_{15})$	Two 16-bit in parallel
0	1	1	1	1	1	0	1	0	1	$A_{14}B_{14}(A_{15}.B_{15})$	Two 16-bit in parallel
0	1	1	1	1	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel
1	0	0	0	0	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
1	0	0	0	0	0	1	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
1	0	0	0	0	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel
1	0	1	1	0	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
1	0	1	1	0	0	1	1	0	0	True carry is applied	Full 32-bit CPA
1	0	1	1	0	1	0	1	0	0	True carry is applied	Full 32-bit CPA
1	0	1	1	0	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel
1	1	0	1	0	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
1	1	0	1	0	0	1	1	0	0	True carry is applied	Full 32-bit CPA
1	1	0	1	0	1	0	1	0	0	True carry is applied	Full 32-bit CPA
1	1	0	1	0	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel
1	1	1	1	1	0	0	0	0	0	$A_{15}.B_{15}$	Two 16-bit in parallel
1	1	1	1	1	0	1	1	0	1	$A_{14}B_{14}(A_{15}.B_{15})$	Two 16-bit in parallel
1	1	1	1	1	1	0	1	0	1	$A_{14}B_{14}(A_{15}.B_{15})$	Two 16-bit in parallel
1	1	1	1	1	1	1	1	1	1	$A_{15}.B_{15}$	Two 16-bit in parallel

C_{15}^{EM} : Carry Estimated Modified

CSNDSP 2004

Papers by Author

Author	Title of Paper	
Abbas, Mr Fawad	Analysis and Evaluation of Traceback Schemes for Distributed Denial of	314
Adamovic, Miss Ljiljana	A New Traffic Engineering Approach for IP Networks	287
Aguni, Mr Julius, Nkenchior	Mobility of Wireless Networks for M-Business	367
Ahamed, Mr Faruque	Design of a digital BPSK demodulator using Field Programmable Gate A	103
Al-Dweik, Dr A.	A New Non Data-Aided Frequency Offset Recovery Technique for OFD	27
Al-Dweik, Dr A.	The Effect of Timing and Fequency Offsets on OFDM Systems Using O	229
Al-Kabi, Mr Amin Hamdalla	Wireless Communication System Performance Enhancement Using Adap	252
Als, Mr Adrian	Simulation of All-Optical Recirculating Fibre Loop Buffers employing S	212
Ameer Ali, Mr Mohammad	FUZZY IMAGE SEGMENTATION CONSIDERING OBJECT SURF.	507
An, Mr Daewon	Pseudo PCA for Cluster Analysis of Time-Series Data	544
Apostolopoulos, Prof Theod	OverTester: A Peer-to-Peer, Modular Toolkit for Network Performance I	392
Arezki, Mr MD	New Method of Comparison NS-FRLS and NLMS Algorithms for the A	528
Arezki, Mr MD	Three New Versions of Numerically Stable Fast Least Squares Algorithm	576
Armstrong, Dr Jean	Optimum noise thresholds in decision directed impulse noise mitigation f	168
Ashmila, Mr Esmail Milad	Multiple Carries Prediction to Enhance the Performance of a 32-bit Estim	617
Attar, Mr Ali Reza	MODULATION AND PROTOCOL RECOGNITION IN COMMUNI	136
Ball, Dr Peter	Billing solutions for 3G data services	75
Balster, Mr Eric	Combined Spatial and Temporal Domain Wavelet Shrinkage Algorithm f	499
Bizinelli, Mr Fabiano	Analysis of HARQ Error Control Scheme for TCP Applications based on	63
Björnemo, Mr Erik	Bluetooth Interference Suppression in IEEE 802.11b	318
Burian, Mrs Adina	The use of clustering methods for combining equalization and decoding i	237

Author	Title of Paper	
Busawon, Dr Krishna	On nonuniform sampling of one-dimensional signals	536
Caffarena, Mr Gabriel	Implementation of an RLS Equalizer for OFDM-based WLAN	604
Carvalho, Prof Paulo	A Multi-constrained QoS Aware Scheduler for Class-based IP Networks	271
Chai, Mr Colin	Simulation Study of WFQ Packet Scheduling Mechanism for different co	217
Chalvatzis, Mr Theodoros	Clock Jitter in Direct Sampling Systems	592
Chan, Mr Erwin	Coherence-free Photonic Notch Filter	495
Chang, Prof Jin-Fu	An Adaptive HARQ-II Protocol using RCPC Codes	55
Chatprechakul, Mr Nattapo	Using Mobile Agents for Data Collection, Data Correlation and Respons	429
Chatzimisios, Mr Periklis	Improving performance through optimization of the RTS/CTS mechanis	363
Chatzimisios, Mr Periklis	IEEE 802.11 WLANs: Performance analysis in presence of bit errors	355
Chee, Mr Chin Yi	Transmission of M-ary Digital Signals Using Controlled Projective Sync	241
Chen, Mr Char-Long	A Novel MAC Protocol and Traffic Scheduling for HFC Network with C	294
Chen, Prof Joy Iong-Zong	Performance Evaluation of MRC for MC-CDMA Communication Syste	176
Cheng, Prof Xiao-rong	Optimization of OFDM modulation for Broadband Power Line Commun	387
Chiang, Mr Po-Chun	Gigabit Ethernet over OC-48 Transponder for Metro-Optical Access	342
Chrisanthopoulos, Mr Panag	A minutiae extraction fingerprint algorithm	445
Das, Mr Suvra Sekhar	Low Complexity Residual Phase Tracking Algorithm for OFDM-based	128
Davarpanah, Ms Hashem	A Fast and Accurate Intrusion Detection System	261
Dennett, Mr Christopher	A Comparison of Complexity and Performance of Turbo Codes over Ray	5
Driberg, Mr Micheal	A Comparative Study of 1x1, 2x1 And 2x2 MIMO-OFDM Systems Bas	1
Dutta, Mrs Kamlesh	Graphical Simulation Tool for Mobile Networks Routing	346
Eissa, Dr Hussein	A New Proposed Burstness Factor For Real-time Traffic Pricing across t	283
Erwu, Dr Liu	A Simple Efficient Adaptive Algorithm For Jitter Reduction in TDM Ov	39
Fathi, Dr Mohammad	Three-pass JPEG-LS Image Compression	580
Fernandes, Mr Alberto	Offering Broadband services via the Telephone system in a Volatile Econ	383

Author	Title of Paper	
Flores Lucio, Mr Gilberto	Improving a Simulated Annealing Approach for Network Optimisation u	359
Forsyth, Dr David Ian	Dual semiconductor optical amplifier polarisation independent wavelengt	107
Franklin, Mr Daniel	A New Robust Frequency-Domain Decision Feedback Equaliser for DM	31
Fricker, Mr Martin	Failure of Adaptive RED in Dynamic Networks	334
Gao, Ms Pei	Weierstrass Approach to Nonlinear Blind Signal Separation	588
Gazi, Mr Boran	Sigmoid Function Based Dynamic Threshold Scheme for Shared-Buffer	412
Giunta, Prof Gaetano	Pre-Post-Rake Schemes for DS/SS Mobile Communications	195
Govinda, Mr Sachidananda	A Modified Niederreiter Public Key Cryptosystem for Combined Coding	18
Grey, Mr Stuart	ARTIFICIAL NEURAL NETWORK ASSESSMENT OF THE PROGNO	483
Griffin, Mr Barry	Demodulation Schemes for Interferometric Optical Fiber Sensors	503
Guyen, Miss Neslihan	A NEW APPROACH TO PREDICTIVE POWER CONTROL OF WC	23
Hegazy, Mr Islam Mohamed	A design of softbots for intelligent, reactive intrusion detection systems	437
Hoang, Mr Vinh Dien	The effect of using multiple ARs in MANETs for global connectivity	421
Holte, Prof Nils	A new model for the probability distributions of crosstalk in twisted pair	279
Huang, Mr Pi	Modelling IrDA Transport Layer TinyTP	371
Huang, Mr Pi	Analysis of the IrFM Digital Payment System	474
Huang, Dr Yung-Fa	Performance of Adaptive Fuzzy-Based Partial Parallel Interference Cance	180
Jean-Pierre, Prof CANCES	Super-Orthogonal Space-Time Trellis Codes together with Blind Channel	221
Jeoti, Dr Varun	Wavelet Footprint Representation of ECG Signals	584
Jimaa, Dr Shihab	lower order based adaptive receiver for ds-cdma communication system	132
Jirón, Mr I	The combination of hyperelliptic curves and block codes for data encryp	225
KALANTZIS, Mr EMMAN	Performance Investigation of a Broadband OFDM System Incorporating	400
Kastamonitis, Mr Kostas	A Novel Detection Technique for Uplink Power Control	35
Kazakov, Dr Vladimir	Sampling-Reconstruction Procedure of Gaussian Processes with a Finite	548
Khor, Ms LC	SIGNAL SEPARATION OF NONLINEAR MIXTURES USING FUZZY	552

Author	Title of Paper	
Knudsen, Mr Thomas Phillij	A Method for Upper Bounding on Growth of Network Access Speed	267
Konstantakos, Mr Dikaio	Adaptive MMSE Receiver Structures, Employing Spread Spectrum and I	47
Koutsouvelis, Mr Kostas	Turbo Trellis Coded Modulation with modified transition metric for Ray	245
Krile, Dr Sreko	New SLA Creation and Optimal Resource Management	152
Kuo, Mr I-Yu	Demonstration of 1.65 um Optical Time-Domain Reflectometry in On-Li	596
Lam, Mr Siew Kei	Efficient Memory Structures for Dynamic Multicast Routing Architectur	600
Le Bihan, Prof Jean	Simple expressions for the locations and amplitudes of the extrema of the	556
Le Goff, Dr Stephane Yves	Constellation Shaping for Turbo-Coded Modulation	87
li, Mr shumu	Arbitrarily Shaped Video Object Compression with Weighted Finite Aut	479
Lim, Dr Chai Soon	Iris Filter: The VLSI Perspective	608
logothetis, Dr Michael	Generalized Max-Min Fairness Policy for Elastic Traffic based on Linear	338
Lopez, Dr M J	e-learning promotes students' deep learning:	330
Luna, Dr Amjad	A Low-Complexity Hybrid FIR/IIR Algorithm for Multi-Channel Line E	441
Luo, Mr Guang-Yeu	Simulating Heterogeneously Timed Networks in Network Simulator NS	207
Luo, Mr Qinglin	Multilevel HARQ on Wireless Channels	144
Mahmoudi, Dr Fariborz	QoS-based Call Admission Control using Multi-code for Voice/Data Traf	199
Manioudakis, Mr Stylianos	ROBUST BLIND CHANNEL ESTIMATION FOR SPACE-TIME CO	203
Marghani, Mr Khaled	MORPHOLOGICAL ANALYSES OF HISTOLOGICAL IMAGES USI	520
Meghdadi, Dr Vahid	New Accurate Upperbounds for Maximum Likelihood Decoder of STBC	148
Menon, Mr Raghav	Biometric User Authentication	487
Mohamed-pour, Dr k	A New Method for Linear-Phase FIR Filter Designing with Desired Shap	452
Nahapetyan, Mr Vahan	BER Improvement in the Zero-IF Receivers	71
Ngah, Mr Razali	Noise and Crosstalk Analysis of SMZ Switches	156
Ngo, Mr Vu Duc	Expurgated Tangential Bound of Low Density Parity Check	120
Noussi, Miss Eleni	Adaptive MF-TDMA: Burst Length Control as a Rain Fade Countermea	43

Author	Title of Paper	
Nowak, Mr Dawid	Bandwidth Scheduling Techniques for Differentiated Services Support in	67
Nowakowski, Mr Marcin	UNIVERSAL QUADRATURE BROADCAST MODULATOR WITH	425
Ntagkounakis, Mr Konstanti	Novel Rotated Channel Assignment Schemes for Non-Line-of-Sight Fixe	379
Nwoye, Mr Ephraim	Classification of Cancerous Cells Images using Clustered Fuzzy-Neural	491
Obiebi, Mr Job Akpofure	Adaptive PN Code Acquisition in DS-CDMA Systems	51
omar, Dr hasan	Performance of Quantized Turbo Decoding for Decoders	188
Ontiveros, Mr B	Construction of a Convolutional ElGamal Elliptic Curve Cryptosystem	95
Othman, Mr Al-Khalid	Node Discovery and Self Initialization Protocol for Distributed Underwat	375
Pathak, Mr Bhumini	Analysis of Boundary Conditions for Real-time Facsimile Transmission c	59
Penttilä, Mrs Katariina Maa	The Evolution of a Wireless RFID and a Cell Phone	233
Qazi, Mr Saad Ahmed	NESTED WIGNER DISTRIBUTIONS AND THEIR APPLICATION 1	524
Qian, Mr Xin	Optical Fibre Refractive Index Reconstruction from Mode Near Field	164
Raad, Mr Muhammad Wasi	PARAMETER ESTIMATION& DIGITAL PEAK LOCALIZATION	541
Radvánszki, Mr Tamás	Virtual Transmission Based MAC Protocol for Wireless Access	433
Rana, Dr Ajay	A survey paper on regression teting techniques	466
Rana, Dr Ajay	technical White paper on IDM	564
Reis, Mr Antonio	Symbol Synchronizers Operating at Sub Data Rate	560
Reis, Mr Antonio	Effects of the Prefilter Bandwidth on Symbol Synchronizers	111
Sabagh, Dr Mohamad Reza	Estimating The Position and velocity of Mobiles with Neural Network	115
Shaikh, Mr Siraj	Performance of Authentication and Encryption algorithms in IPSec Tunn	184
Sigüenza-Tortosa, Mr David	Packet Scheduling Configuration in Proteo Network-on-Chip	396
Sizeland, Mr Robert	A New Architecture and MAC Layer for Quality of Service Support in V	275
Smith, Dr Dave	Using Invisible Region Wave Vectors for Determining the Properties of	248
Sodha, Dr Janak	Channel SNR Estimate for Turbo Decoding	83
Soongsathitanon, Mr Somph	A Novel Fast Search Algorithm For MPEG Standard	462

Author	Title of Paper	
Soyjaudah, Dr K M Sunjiv	A low complexity joint Huffman and Convolutional decoder employing s	14
Soyjaudah, Dr K M Sunjiv	A comparative study of combinations of concealment techniques with err	257
Soyjaudah, Dr Sanjeev	Multifunctional unit for Rayleigh Fading Channels combining Huffman,C	140
Speidel, Dr Ulrich	Constructing Finite Pseudo-Random Strings using a Codeset-Based Entr	91
SRINIVASAN, Mr M	INTERACTIVE SPEECH RECOGNITION SYSTEMS USING COMM	515
Srinonchat, Mr Jakkree	An efficient codebook design for Speaker Dependent Coding System	471
Suraweera, Mr Himal Asang	OFDM Peak-to-Average Power Reduction Scheme with Spectral Maskin	160
Tabatabaian, Mr Seyed	A New Search Algorithm for Vector Quantization	458
Tekiner, Mr Firat	The Antnet Routing Algorithm – A Modified Version	416
Toossian Shandiz, Dr Heyda	New method to count vehicles in highway by using some pixels and thres	533
Vafi, Mr Sina	Iterative turbo decoder design with convolutional interleavers.	124
Venkataraman, Miss Gayath	A Size-limited Cluster Maintenance Technique for Mobile Ad-hoc Netwo	302
Verma, Dr Seema	Performance Evaluation of Medium Dependence of Turbo Codes	172
Vial, Mr Peter	Graphical User Interface to access RMON probe in Java	351
Walker, Dr Jacqueline	An Improved Jitter Time-stamp Method for Source Timing Recovery in	310
Wang, Dr Dajin	An Energy-Ecient Clusterhead Assignment Scheme for Hierarchical Wire	306
Wang, Dr Yongji	Congestion control of AQM routers by an improved PID controller	322
Wei, Mr Chen	Information Theoretic Approach to Blind Deconvolution under Impulsive	511
Wen, Miss Hong	A Family of Irregular LDPC Codes Constructed Based on Optical Orthog	10
William, Prof Peter E.	A New Image Compression Technique Based on Combining Artificial Ne	448
Wu, Dr Jigang	Local Optimal Algorithm for High Performance Solution of Reconfigura	613
Xiaohu, Dr Ge	A Prediction Method of Self-Similar Network Traffic Based on the Covar	298
Xu, Mr Dengyuan	Provision of Virtual Line-switched Service over DWDM with PFTS.	408
Yasukawa, Mr Kenta	Proposal and Evaluation of Dynamic Class Assignment Method: On Sim	404
Yu, Dr Shaokai	Dynamic Resource Control for QoS-Aware Multimedia Wireless Networ	326

Author	Title of Paper	
Zhou, Mr Hongwei	Blind Adaptive Multiuser Detection Based on Subspace Tracking and $K\alpha$	79
Zou, Dr Chuanyun	Decoding with Arbitrary Pulse-Shape in Optical PPM Communications	99
