STANDARDISATION OF THE ADAPTIVE MULTI-RATE CODEC

Kari Järvinen
Nokia Research Center
P.O.Box. 100, 33721 Tampere, FINLAND
Tel: +358 3 2725854; fax: +358 3 2725888
e-mail: kari.ju.jarvinen@nokia.com

ABSTRACT
European Telecommunication Standards Institute (ETSI) initiated a standardisation program in October 1997 to develop an Adaptive Multi-Rate (AMR) codec for GSM. After two competitive selection phases, ETSI chose in October 1998 a codec developed in collaboration between Ericsson, Nokia, and Siemens. The codec standard was finalised, characterised, and formally approved in ETSI during early 1999. The AMR codec provides the next step of speech quality improvement in GSM after the introduction of Enhanced Full-Rate (EFR) codec in 1996. AMR offers substantial improvement in error robustness by adapting speech and channel coding depending on channel conditions. By switching to operate in the GSM half-rate channel during good channel conditions, AMR provides also channel capacity gain over the EFR codec. In April 1999, the Third Generation Partnership Project (3GPP) adopted the AMR codec as the mandatory speech codec for the third generation WCDMA system.

1 INTRODUCTION
The GSM Enhanced Full-Rate (EFR) codec [1] was the first codec to provide digital cellular systems with quality equivalent to that of a wireline telephony reference (G.726 ADPCM standard at 32 kbit/s). The EFR codec gives substantial quality improvement compared to the previous GSM codecs, the full-rate (FR) and half-rate (HR) codecs. EFR provides wireline speech quality across all typical radio conditions down to carrier-to-interference ratio (C/I) of approximately 10 dB [2]. However, the codec still left some room for improvements. In particular, the performance in severe channel error conditions could be improved by employing a different bit-allocation between speech and channel coding. Also, the GSM half-rate channel was not yet able to provide high speech quality.

European Telecommunication Standards Institute (ETSI) launched a standardisation program in October 1997 to develop a new codec, Adaptive Multi-Rate (AMR) codec, for GSM. Already before that a one-year feasibility study had been carried out to validate the novel AMR concept. The AMR codec would operate both in the full-rate (22.8 kbit/s) and half-rate (11.4 kbit/s) channels of GSM. It would adapt to radio channel and traffic load conditions and select the optimum channel mode (full-rate or half-rate) and codec mode (bit-rate trade-off between speech and channel coding) to deliver the best combination of speech quality and system capacity.

This paper reviews the standardisation process of the GSM AMR codec. A brief review of the AMR codec and its performance is also given.

2 STANDARDISATION OF THE AMR CODEC
The AMR codec standardisation was carried out as a competitive selection process consisting of several phases.

2.1 Performance Requirements
After the launch of standardisation, detailed performance requirements were specified. In the FR channel mode, AMR was required to provide substantial improvement in error robustness over EFR. For clean speech, it would deliver at 4 dB C/I the same quality as EFR at 10 dB C/I. At 13 dB C/I and above, quality equal to EFR under error-free transmission was targeted. AMR was also expected to provide channel capacity gain over the EFR codec by switching to operate in the HR channel during good channel conditions. In the HR channel at low error conditions (C/I ≥ 16 dB), AMR would offer quality comparable to wireline. At higher error-rates in the HR channel, the quality would be equal to that of the GSM FR codec. Tables I and II show performance requirements for the best codec mode in each error condition.

Table I: Performance requirements for clean speech (the best codec mode for each C/I)

<table>
<thead>
<tr>
<th>C/I</th>
<th>FR requirement</th>
<th>HR requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Errors</td>
<td>EFR No Errors</td>
<td>G.728 No Errors</td>
</tr>
<tr>
<td>16 dB</td>
<td>EFR No Errors</td>
<td>G.728 No Errors</td>
</tr>
<tr>
<td>13 dB</td>
<td>EFR No Errors</td>
<td>FR at 13 dB</td>
</tr>
<tr>
<td>10 dB</td>
<td>G.728 No Errors</td>
<td>FR at 10 dB</td>
</tr>
<tr>
<td>7 dB</td>
<td>G.728 No Errors</td>
<td>FR at 7 dB</td>
</tr>
<tr>
<td>4 dB</td>
<td>EFR at 10 dB</td>
<td>FR at 4 dB</td>
</tr>
</tbody>
</table>

Table II: Performance requirements for speech in background noise (the best codec mode for each C/I)

<table>
<thead>
<tr>
<th>C/I</th>
<th>FR requirement</th>
<th>HR requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Errors</td>
<td>EFR No Errors</td>
<td>EFR No Errors</td>
</tr>
<tr>
<td>16 dB</td>
<td>EFR No Errors</td>
<td>G.729 and FR No Errors</td>
</tr>
<tr>
<td>13 dB</td>
<td>EFR No Errors</td>
<td>G.729 and FR No Errors</td>
</tr>
<tr>
<td>10 dB</td>
<td>G.729 and FR No Errors</td>
<td>FR at 10 dB</td>
</tr>
<tr>
<td>7 dB</td>
<td>G.729 and FR No Errors</td>
<td>FR at 7 dB</td>
</tr>
<tr>
<td>4 dB</td>
<td>FR at 10 dB</td>
<td>FR at 4 dB</td>
</tr>
</tbody>
</table>

2.2 Pre-Selection (Qualification) Phase
The selection process started with a pre-selection during which five of the most promising codecs out of eleven candidates were chosen to enter selection tests. The qualification was based on in-house listening tests and on technical details of the proposed algorithms. All codecs were tested individually against the performance requirements according to a common test plan. The pre-selection took place during spring 1998 and was completed by June 1998.

2.3 Selection Phase
The selection phase was carried out from July until September 1998. Testing of the five qualified candidates was coordinated internationally with seven independent listening laboratories performing the testing in multiple languages: Chinese, English, French, German and Spanish. Each experiment in the tests was repeated for at least two languages to avoid any bias due to a particular language. Processing of speech samples was distributed...
among the candidate organisations themselves with cross checking performed by other candidates. The codecs were implemented in C-code with fixed-point arithmetics. The number of codec modes was limited to four in the listening tests during the pre-selection and selection phases.

The codec proponents delivered detailed documentation of their codec proposal including a justification of meeting all design constraints, e.g., limits for complexity and transmission delay. An independent analysis of codec complexity was carried out to verify the estimates given by the codec proponents.

The listening test results showed clear improvements over the existing GSM codecs where they were required vindicating the AMR codec. Based on the test results and technical details of the codec proposals SMG11 chose in September 1998 a codec developed in collaboration between Ericsson, Nokia and Siemens. The selection was formally approved in ETSI in October 1998.

2.4 Optimisation Phase

From October 1998 until January 1999, a short optimisation phase took place. The optimisation was essentially limited to bringing corrections to the codec C-code and making improvements for the channel coding part. During the optimisation, the complexity of channel coding was reduced while at the same time obtaining some performance improvements.

2.5 Verification Phase

After the selection phase, the codec was subjected to further analysis and testing to verify its suitability for GSM. A detailed analysis of implementation complexity and transmission delay was performed during this phase. The tests included assessment of the performance for a wide range of signals (music signals, various background noise types and talkers etc.) and for signalling tones and DTMF-tones. The results showed good performance for the chosen codec. The verification phase was essentially completed by January 1999, although some complementing results were provided later. ETSI approved the codec standard formally in February 1999 except for Voice Activity Detection (2 optional algorithms) and optimised channel coding. These remaining parts were included into the standard in June 1999.

2.6 Characterisation Phase

The performance of the AMR codec was characterised in detail during spring 1999. Eight independent listening laboratories carried out the testing using six languages: Chinese, English, French, German, Italian, and Spanish. Each experiment in the tests was performed twice with different languages. Two host laboratories shared the responsibility of processing the speech samples through the AMR codec and a number of reference codecs. The tests included experiments for clean speech, background noise, transmission errors (constant C/I and dynamic error conditions), input level dependency, and tandeming. All experiments were carried out both for the FR and HR channels. The characterisation phase was finalised by June 1999. ETSI Technical Report on Performance Characterisation of the AMR codec summarises the results [3].

3 AMR BASIC OPERATION

The AMR codec contains a set of fixed rate speech and channel codecs, in-band signalling and link adaptation. Fig. 1 shows a basic block diagram of the AMR codec in GSM.

Each codec mode provides a different level of error protection through a different distribution of the available gross bit-rate between speech and channel coding. The link adaptation process bears responsibility for measuring the channel quality and selecting the optimal speech and channel codecs. In-band signalling transmits the measured channel quality and codec mode information over the air interface. The in-band signalling is transmitted along with the speech data.

The Mobile Station (MS) and the Base Transceiver Station (BTS) both perform channel quality estimation for the receive signal path. Based on the channel quality measurements, a Codec Mode Command (over downlink to the MS) or Codec Mode Request (over uplink to network) is sent in-band over the air interface. The receiving end uses this information to choose the
best codec mode for the prevailing channel conditions. A Codec Mode Indicator is also sent over the air interface to indicate the current mode of operation. The codec mode in the uplink may be different from the one used in downlink on the same air-interface, but the channel mode (FR or HR) must be the same.

The network controls the uplink and downlink codec modes and channel modes. The mobile station must obey the Codec Mode Command from the network, while the network may use any complementing information, in addition to Codec Mode Request, to determine the downlink codec mode. The mobile station must implement all the codec modes. However, the network can support any combination of them, based on the choice of the operator.

AMR contains also Voice Activity Detection and discontinuous transmission (VAD/DTX). These are used to switch off the encoding and transmission during periods of silence thereby reducing air-interference and extending battery life-time.

4 SPEECH CODING

The AMR speech codec utilises the ACELP (Algebraic Code Excitation Linear Prediction) algorithm employed also in GSM EFR [1] and D-AMPS EFR [4] codecs. The AMR codec is referred to as Multi-Rate ACELP (MR-ACELP) codec [5, 6]. The codec contains eight speech codecs with bit-rates of 12.2, 10.2, 7.95, 7.4, 6.7, 5.9, 5.15 and 4.75 kbit/s. As seen in Fig. 2, all the speech codecs are defined for the FR channel, while the six lowest ones are defined for the HR channel. All the codecs provide seamless switching between modes. The 12.2 and 7.4 kbit/s modes are identical to GSM EFR and D-AMPS EFR, respectively.

5 CHANNEL CODING

AMR contains eight channel codecs for the FR channel and six for the HR channel. Channel coding performs error correction and bad-frame detection. The error correction in all modes is based on Recursive Systematic Convolutional (RSC) coding with puncturing to obtain the required bit-rates. Each codec mode employs a 6 bit CRC (Cyclic Redundancy Check) for detecting bad frames. All channel codecs use convolution polynomials specified for the previous GSM traffic channels (either for speech or data traffic channels) to maximise commonality with the existing GSM system [7].

6 LINK ADAPTATION AND IN-BAND SIGNALLING

Link adaptation consists of channel quality measurement and codec/channel mode adaptation algorithms [8, 9]. Link adaptation in AMR is twofold: it adapts the bit-partitioning between speech and channel coding within a transmission channel (codec mode), and the operation in the full- and half-rate channels (channel mode). Depending on the channel quality and possible network constraints (e.g., network load), link adaptation selects the optimal codec and channel mode. Fig. 3 shows an example of how the codec mode adaptation operates in the FR channel under dynamic error conditions. Channel quality varies between about 22 and 2 dB in C/I. Based on estimated channel quality, one out of three codec modes (12.2, 7.95 or 5.9 kbit/s) is chosen. The in-band signalling supports adaptation between four active codec modes. The set of up to four active codec modes is selected at call set-up (and in handover). Codec Mode Command/Request and Codec Mode Indicator are transmitted in every other speech frame (alternating within consecutive frames). Therefore, the codec mode can be changed every 40 ms.

7 COMPLEXITY AND DELAY

The AMR codec is defined in fixed-point arithmetic using a set of basic operations defined by ETSI. This allows calculation of the complexity as a number of WMOPS (Weighted Million Operations Per Second) [3]. The theoretical worst case (TWC) computational complexity of the AMR speech codec is 16.75 WMOPS as shown in Table III below. This is only about 10% higher than for EFR.

Table III: Complexity of the AMR codec

<table>
<thead>
<tr>
<th>AMR complexity</th>
<th>WMOPS (TWC)</th>
<th>static RAM **</th>
<th>dynamic RAM**</th>
<th>Table ROM **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech codec</td>
<td>16.75</td>
<td>2241</td>
<td>3039</td>
<td>14571</td>
</tr>
<tr>
<td>FR channel codec*</td>
<td>5.21</td>
<td>658</td>
<td>1981</td>
<td>5236</td>
</tr>
<tr>
<td>HR channel codec*</td>
<td>2.85</td>
<td>886</td>
<td>1486</td>
<td></td>
</tr>
</tbody>
</table>
*) including link adaptation and in-band signalling  **) 16-bit words

All AMR speech codecs have a buffering delay of 20 ms with 5 ms lookahead. The MS-to-MS transmission delay in the GSM system, taking into account all processing and system delays, is about 200 ms in the FR channel and 177 ms in the HR channel.

8 AMR PERFORMANCE

In the GSM FR channel, the AMR codec extends the wireline quality operating region from about C/I ≥ 10 dB in EFR to about C/I ≥ 4-7 dB [3]. In poor channel conditions, AMR gives substantially improved robustness over EFR. At 4 dB C/I, AMR achieves about a 2 units of Mean Opinion Score (MOS) improvement over EFR. In typical dynamic error conditions in the FR channel, AMR provides up to over 1 MOS improvement compared to EFR. In the GSM HR channel at low error conditions (C/I ≥ 16 dB), AMR provides close to EFR quality (equivalent to G.728 for clean speech, and to G.729 and error-free GSM FR for
speech under background noise). Figs. 4 and 5 show the performance of the best codec mode for each C/I condition for clean speech. Fig. 6 shows the performance curves for each AMR FR codec mode (corresponding to Fig. 4). Figs. 7 and 8 show the performance of the best codec mode for each C/I condition under background noise. The AMR codec fulfills all the demanding requirements in Tables I and II except one single case (HR channel in background noise at 10 dB C/I). The AMR codec gives substantial improvement over previous GSM codecs where it was required meeting well the overall development target.

AMR provides good overall performance and high granularity of bit-rates making it suitable also for other systems and applications than GSM. In April 1999, the Third Generation Partnership Project (3GPP) adopted the AMR codec as a mandatory speech codec for third generation systems.

CONCLUSIONS

The GSM AMR codec standardisation was carried out as a competitive selection process involving several phases. AMR provides substantial performance improvement over the previous GSM codecs, in particular in error-robustness in the full-rate channel and providing high speech quality in the half-rate channel. The AMR codec provides high granularity of bit-rates between 12.2 and 4.75 kbit/s with seamless switching between modes.

ACKNOWLEDGEMENTS

The AMR codec was developed within ETSI SMG11. The new codec standard is a result of the efforts of many companies and persons contributing to its development within SMG11.

REFERENCES

[7] Digital cellular telecommunications system; Channel Coding (GSM 05.03), ETSI Technical Specification.