NETWORK-ADAPTIVE AND ENERGY-EFFICIENT MULTI-USER VIDEO COMMUNICATION OVER QOS ENABLED WLAN

Xin Ji(ab), Sofie Pollin(ab), Gauthier Lafruit(b),
Iole Moccagatta(b), Antoine Dejonghe(b) and Francky Catthoor(ab)

(a)ESAT, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001, Leuven, Belgium
(b)Interuniversity Microelectronics Centre (IMEC), Kapeldreef 75, 3001, Leuven, Belgium

ABSTRACT

This paper addresses energy efficient packet scheduling for the transmission of multiple video streams over a wireless LAN. By exploiting the temporal scalability present in the prediction dependencies among video frames, a centralized cross-layer optimization solution is proposed, which is built on the HCF Controlled Channel Access (HCCA) of the IEEE 802.11e standard. Our performance study shows that the proposed solution saves energy under good channel conditions, and achieves quality gains more than 4.5dB on bad channels or under heavily loaded network conditions.

1. INTRODUCTION

The increasingly fast integration of communication and computing on cost effective, low power, thumb-size devices, is quickly enabling a surge of new network applications. Transmission of multiple video streams over a wireless local area network (WLAN) is a typical example of this evolution. In this context, Quality of Service (QoS) provision for real-time applications is becoming more and more critical, as wireless networks are affected by highly error-prone and time-varying conditions, especially when a lot of users interact. Besides this QoS challenge, ensuring low-power consumption is imperative to enable the deployment of broadband wireless connectivity in battery-operated portable devices.

Performing high-quality video transmission over such wireless networks is a challenging task. The IEEE 802.11e standard was proposed to enhance QoS support over WLAN. In [1], a solution for scheduling transmission opportunities (TXOP) as function of the different traffic classes is proposed. Some recent studies improve the performance by also exploring the specificities of video traffic. For instance, different retransmission times are given according to data partition types provided by H.264/AVC encoding scheme[2].

Considering energy efficiency, a substantial body of prior work focuses on energy-efficient wireless transmission. Different and sometime conflicting approaches exist at Media Access Control (MAC) and physical (PHY) layers. In [5, 6], a cross-layer optimization methodology was proposed that combines the two approaches for delay bounded data transmission over non-congested network. The proposed solution consists of a two-phase systematic approach for optimally allocating the network resources and controlling the system configuration. There exist energy efficient cross-layer scheduling and resource allocation methods explicitly considering video transmission over wireless network [3, 4], but they only exploited the trade-off brought by application layer and PHY layer.

In [7], application awareness was introduced in the cross-layer framework proposed in [5,6] by integrating the video quality metric relying on SNR scalability of the embedded video bitstream as QoS constraint to obtain a protocol stack wide cross-layer optimization. In this paper, we study the temporal scalability of H.264/AVC video encoding and propose a cross-layer optimization for real-time video communication. Leveraging the MAC and PHY layer energy, delay and packet error rate (PER) calibration in [5, 6], a solution to solve network congestion under real-time delay constraints is integrated into the problem formulation and access control. The contribution of this work is to first develop a network adaptive reference frame choice scheme, where the temporal dependencies among frames have been built according to their playback deadline. Later on, the less important frames will be transmitted with a strategy towards a less reliable receiving probability which also results in less bandwidth cost. Following steps are involved in the proposed cross-layer optimization:

- Design-time phase: This phase consisting in the analysis of the Packet Error Rate(PER), transmission Delay and energy cost.
- Network-adaptive choice of the reference video frame.
- Expected quality-energy-delay trade-off calibration relying on temporal dependencies.
- Run-time phase: Based on different delay requirements of the data, the multi-constraints optimization problem is formulated as a linear programming problem and solved at run time.

The resulting solution enables better energy performance trade-off compared to SoA approach or link layer centric approach and achieves more than 8dB and 4.5dB quality gain respectively when transmitting video over heavily overload network.

The remainder of this paper is organized as follows. Section II briefly reviews the IEEE 802.11e WLAN standard and the deployed referenced frame choice scheme. Section III introduces the proposed cross-layer energy-efficient video scheduling strategy with rate-distortion awareness. In Section V, we examine the performance of our framework through simulations. Finally, concluding remarks are provided in Section VI.

2. SYSTEM OVERVIEW: VIDEO TRANSMISSION OVER QOS-ENABLED WLAN

We consider a setup consisting of multiple independent Mobile Terminals (MT) who want to upstream real-time video traffic to the Access Point (AP) of a Wireless Local Area Network (WLAN). These terminals transmit their data over a shared wireless channel, assumed to be slowly fading, which is typical of indoor propag protocols considered here are: the IEEE 802.11a standard for the physical layer (OFDM-based transmission in the 5GHz band), and the QoS-enabled IEEE 802.11e standard is considered for the MAC layer. The considered video standard is h.264/avc encoder, combined with a network adaptive choice of the reference frame. This will be detailed in the sequel.

2.1 MAC and PHY LAYERS: IEEE 802.11a and IEEE 802.11e (HCCA)

For the physical layer, the IEEE 802.11a standard is considered. IEEE 802.11a standard is based on Orthogonal Frequency Division Multiplexing (OFDM), and provides eight different PHY modes offering Data transmission Rates (DR), ranging from 6 Mbps to 54 Mbps. Our system modeling is based on an 802.11a direct conversion transceiver implementation with turbo coding [9]. Four control parameters have significant impacts on energy and performance for the OFDM transceivers: the modulation order N_{mod}, the code rate B_c, the power amplifier (PA) transmit power P_tx and the back-off $backoff$ characterizing the linearity of the amplification [10]. Let us represent a possible transmission configuration as a vector K (each specific transmission parameter corresponds to an entry in this vector). For reliably transmitting on a wireless network, a long application layer packet p is usually further fragmented into smaller data units. In this paper, we consider link layer fragmentation only. The energy and time needed to send a Mac Service Data Units (MSDU) is function of this configuration vector: $E_{MSDN}(K)$ and $TXOP_{MSDU}(K)$. The energy cost and time of transmitting an application layer packet is defined as $E_p(K)$ and $TXOP_p(K)$, and these values depend on the number of fragments that need to be transmitted or retransmitted for successful packet transmission.

For the MAC layer, the IEEE 802.11e standard is considered. IEEE 802.11e standard defines a set of enhancements for WLAN targeting QoS enhancement for multimedia applications [8]. Research of 802.11e HCF scheduling has recently received a lot of attention. Initial contributions were mainly concerned with the feasibility of the EDCA and HCCA mechanisms, which both difine the traffic classes. HCCA was used in this paper, as it provides significant benefits over EDCA for applications requiring strict QoS. In this section, HCCA-Based channel scheduling is briefly reviewed. With HCCA, there is a negotiation of QoS requirements between the QoS enhanced wireless station (QSTA) and the Hybrid Coordinator (HC). Once a stream for a QSTA is established, the HC allocates TXOPs via polling to the QSTA in order to guarantee its QoS requirements. The HC enjoys free access to the medium during the contention-free period and uses the highest EDCA priority during the contention period in order to 1) send polls to allocate TXOPs and 2) send downlink parameterized traffic. Once the HC has control of the medium, it starts to deliver parameterized downlink traffic to stations and issues QoS contention-free polls (QoS CF-Polls) to those stations that have requested parameterized services. The QoS CF-Polls include the TXOP duration granted to the QSTA. If the station being polled has traffic to send, it may transmit several packets for each QoS CF-poll received respecting the TXOP limit specified in the poll. HCCA operates during both the contention-free period and the contention period. The transmission strategy configuration can be piggybacked onto the CFPoll frame.

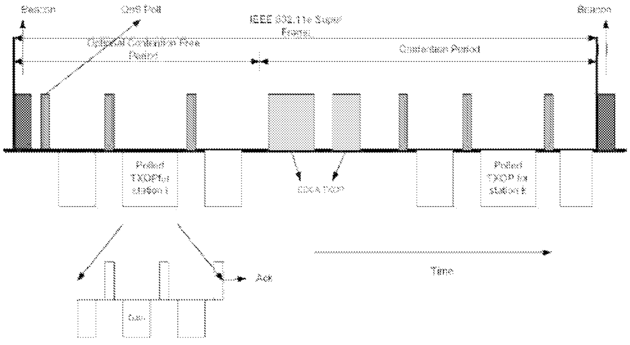

Figure 1: HCCA overview

We refer the interested reader to [5, 6] for delay bounded reliable data transmission over non-congested network. In [6], it is shown how to obtain a schedule to optimize the communication energy cost by leveraging MAC and PHY layer scaling and sleeping techniques, while working in the dimensions of Energy, TXOP and PER.

2.2 Application Layer: H.264/AVC and network adaptive reference frame choice

We consider a H.264/AVC encoder here. In a conventional encoding and transmission scheme without any awareness of network losses or scheduling strategy, an I-frame is typically followed by a series of P-frames, which are predicted from their immediate predecessors. This scheme is vulnerable to network errors since each P-frame depends on its predecessor and any packet loss will break the prediction chain and affect all subsequent P-frames.

We introduce the key frame concept here in manipulating the prediction dependency. Assume the QSTAs are aware of the Beacon period of the AP. When the real-time video is captured, the frame arrived closest to the next scheduling point will be encoded as a P-frame, which will also be denoted as key picture during this Beacon period. For the remaining frames that arrive within the same Beacon period, the closer it arrives to the scheduling point, the more it will be dependent on, as shown in Figure 2. After channel access is granted to a certain user, it will first transmit the least referred frame. Only when the referred frames are correctly received, the current frame will be transmitted, provided the TXOP limit has not been reached. The reason for this manipulation is that when network overload happens, we drop the least depended frames to avoid as much as possible to break the prediction dependencies.

Checking in Figure 2, frame B_1, B_2, P_1 arrives one after another during the first Beacon period. Assume at time tx_1, the AP grants the channel access to this user, to successfully display all these frames at receiver side, we need to transmit all these 3 frames within the period between tx_1 and the playback time B_1. If the channel capacity is not enough, we will have to discard some frames. By discarding B_1, both the data size decreased, and the delay allowed for the remainning 2 frames also increased. Upon WLAN, most of the delay comes from waiting the transmission opportunities when a lot of users interact. With this scheme, the discarded frames are those most delay stringent ones, which at the same time relaxes the bandwidth requ

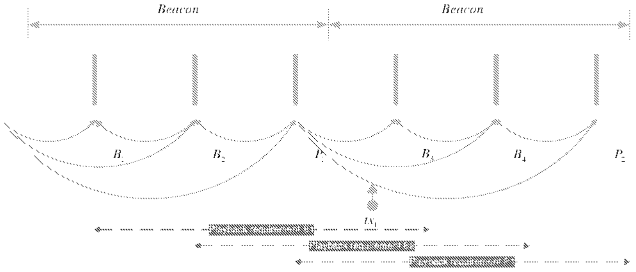

Figure 2: Deployed reference frame choice scheme

the most.

Assume whenever there is any frame loss happens, the former correctly received frame will be displayed. It is possible to calculate the expected distortion of each frame based on the prediction dependencies. A frame can be correctly decoded provide v referred frames have been correctly received. Let D_r denote the distortion corresponding to the reception of the frame and D_l denote the distortion associated with losing the current frame. Denoting the error probability of the referred frames v under configuration K_v as PER_{K_v}, the expected distortion of the current frame D_e can be written as:

$$D_e = \prod_{i=1}^{v}(1 - PER_{K_i})D_r + (1 - \prod_{i=1}^{v}(1 - PER_{K_i}))D_l \qquad (1)$$

During one beacon period, once the key picture P frame is decided for the current beacon period, the video encoding process begins and the calibration of expected distortion, energy cost, transmission delay for each possible configuration also starts at the meantime. In comparison with the video coding, this overhead is negligible when the GOP size is not big. A big GOP size is not suitable for real time communication considering the delay brought by waiting for the P-frames/key pictures to encode other frames. On the other hand, if the GOP size is big (e.g more than 5), we can definitely separate them into several GOPs by increase the key picture number.

3. PROBLEM FORMUATION OF CROSS-LAYER OPTIMIZATION FOR ENERGY EFFICIENT MULTI-USER VIDEO TRANSMISION

Denoting all the frames arriving during one Beacon period as a group of picture (GOP), the transmission opportunities of user i requires to transmit all the m frames within current GOP is: $TXOP_i = t_{i1} + t_{i2} + \cdots + t_{im}$. The distortion of that GOP is given as: $D_{e_i} = D_{e_{i1}} + D_{e_{i2}} + \cdots + D_{e_{im}}$ and the energy cost as $E_i = E_{p_{i1}} + E_{p_{i2}} + \cdots + E_{p_{im}}$.

From the prediction design of the former section, we see the least referred frames are also the most delay critical ones. Thus by discarding frames, the real-time playback delay requirement will also be relaxed. Assume that most of the delay comes from the waiting for channel access, for n GOPs waiting to be transmitted at the current scheduling point inside the network, the optimization problem is formulated as a linear programming problem to find for the GOPs of each user i, the optimal configuration K_i^* such that:

$$K_i^* = argmin(\sum_{i=1}^{n} D_{e_i}(K_i)) \qquad (2)$$

Subject to:

$$\begin{cases} TXOP_1(K_1) \leq T_{r1}, \\ TXOP_1(K_1) + TXOP_2(K_2) \leq T_{r2}, \\ \cdots \\ TXOP_1(K_1) + TXOP_2(K_2) + \cdots + TXOP_n(K_n) \leq T_{rm}. \end{cases} \qquad (3)$$

Where $T_{r1}, \cdots, T_{rm}$ are the delay requirement of every GOP from different users in ascending order[1]. The proposed solution consists in the solution of Linear programming problem, which is characterized by a bounded complexity. Although there is no simple analytical formula for the solution of a linear program, there are a variety of very effective methods for solving them [11]. A greedy backward algorithm is used at run-time for the feasible bandwidth allocation to reach the best quality.

To further decrease the complexity of the run time AP scheduling algorithm, we rely on a similar two-phase solution approach proposed in [5, 6]. At ***design time***, for each possible system state and packet size, the optimal operating points are determined according to their minimal PER, energy cost and network resource (TXOP) consumption. To that end, we introduce the Pareto concept for multi-objective optimization from microeconomics [12]. Compared to a convex hull approach, the Pareto Frontier retains more feasible settings at run-time.

After the system state of all the users is known at runtime and the key pictures are decided and encoded, the quality-energy Pareto Frontiers will be calculated by each user before the next scheduling period. According to the Pareto property, for one GOP g, if there exist two configurations g_1 and g_2, either they satisfy $D_e(g_1) > D_e(g_2)$ and $E(g_1) < E(g_2)$ or they satisfy $D_e(g_1) < D_e(g_2)$ and $E(g_1) > E(g_2)$. By finding the minimum distortion configuration settings in Eq. (3), we know if there is any transmission strategy that cost the less energy, it must also cause quality degradation.

A dual problem is that instead of allocating the bandwidth and transmission strategies to achieve best quality with minimum energy, the user can decrease the quality requirement for the sake of energy savings. In this approach, it is assumed that the different users can require different video quality levels. If the configurations achieve an expected quality better then required quality, only the minimum energy cost configuration will be retained in the Pareto configuration settings and sent to the AP. Each user experiences different channel and application dynamics, resulting in different system states over time. It is this important dynamic characteristic which makes it possible to exploit multi-user diversity for energy efficiency.

At start time of every Beacon period, the information of quality-energy Pareto configuration, and the corresponding required TXOP will be negotiated with AP as traffic specification (TSPEC) [8]. The overhead caused by this turns out to be very small compared to the amount of data.

---

[1] Notice that the delay requirement will be changed according to whether there is any frame to be discarded. And to discard a frame can be regarded as one of the configuration K which result $TXOP_p = 0$, $E_p = 0$ and $PER_p = 1$.

4. NUMERICAL RESULTS

4.1 Simulation Setup

In the experiments, Foreman sequence at CIF (352x288, 4:2:0) resolution and with 30 frames per second is considered here. The sequence was encoded using JSVM software under MVC mode. We encoded the sequence around 350 kbps with an average encoded quality of about 35 dB without transmission errors. The user number is increased until 30 to represent the network condition from no congestion to heavily overload. MSDU size is set to 1000 Bytes. The Beacon period is set to 100ms and the end-to-end delay requirement of each frame is set to 200 ms according to video conference requirement [13]. The mobile devices are uniformly distributed around the AP with a radius of 10m. And the transceiver energy consumption unit is shown in Joule.

To determine the channel impact on the loss probability, the fading channel was discretized into 8 classes, corresponding to a 2dB difference in received SNIR (Signal-to-Noise-and- Interference-Ratio) for reaching a given turbo code block error rate target. In order to derive a time-varying link-layer error model, we associate every channel class to a Markov state, each with a probability of occurrence based on the channel fading statistics [10]. The increasing channel numbers are corresponding to worse channel conditions.

In order to evaluate the relative performance of the proposed approach, we provide results for the three following transmission strategies:

- **SoA reference point:**The transceiver uses the highest feasible modulation and code rate that will successfully deliver the packets. After that, it switches to sleep mode. This approach is proposed by commercial 802.11 interfaces [14], which aims to maximize the sleep duration.
- **Link centric XL:**in this transmission strategy, we use our approach to transmit every video packet with a configuration resulting in a PER smaller than 1e-3 (by doing so, we assume an error free transmission).
- **Application - aware XL, proposed method**: in this transmission strategy, we introduce the expected visual distortion into the optimization framework. In this strategy, we show not only the approach targeting for best quality (Optimal Quality approach) but also the quality compromise approach for the sake of bandwidth and energy efficiency (Quality Con = * dB).

In the aforementioned strategies, when congestion happens, the most delay stringent packets will be discarded until the bandwidth is enough for the remaining data.

4.2 Result analysis

In this part, we not only show the best quality approach, but also analyze the influence of lowering the quality requirement. From the results, we see that without congestion, both the energy gains and quality gains can be achieved under good channel conditions. When simulating on bad channels or under heavy overloaded network conditions, significant quality gains can be observed.

Figure 3 shows the average quality observed at the AP side and the energy cost with five users upstreaming the video bitstream on different channel types, where hardly any congestion happens.

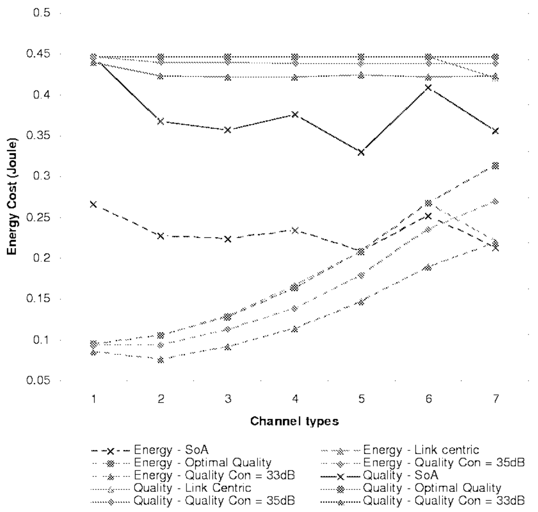

Figure 3: Average Quality and Energy cost of 5users on different channels

Under most of the channel states, we observe the end-to-end quality of our Optimal Quality approach is slightly better or equal to Link Centric, while the energy cost is the similar or less. Except under the worst channel (channel 7), with 30% more energy cost, the Optimal Quality approach increases the average user visual quality with 2.3dB. The reason for this is that: to achieve a small enough PER, the Link Centric approach requires more bandwidth for each packet. When bandwidth is not sufficient, the user has to discard packets. Since they discard too many packets, the energy cost also decreases. But this is accompanied with the severe degradation of video quality. Even though they discard the less important packets, it still results in a much lower quality. The Optimal Quality approach, does not simply discard the less important packets, but also configures the less important packets with a less reliable setting which results in less bandwidth cost. By doing so, an optimal average quality can be achieved and the energy cost is the minimum one to achieve this quality.

Lowering the expected quality requirement can be used to save the energy in expected quality approach. From the result we see, by lowering the expected quality to 35dB, with 18.6% more energy, we can achieve 1.6 dB in video quality increase. If lowering the quality expectation to 33 dB, with actually slightly less energy, we still achieve 0.3 dB quality increase compared to Link Centric approach.

We check the results for a moderate congestion in Figure 4, where 10 users want to upstream the video to AP. Graceful degradation of quality can always be achieved by expected quality approaches. While considering the same channel states, to achieve the similar video quality, the energy cost is actually lowered for expected quality approaches. E.g. by lowering the quality requirement to 33dB, on channel 5, the energy saves 6.6% with same end-to-end video quality.

The results of heavily congested environments are shown in Figure 5. In this setup, 20 users tried to upstream the video to AP. From the results, we see the AP already begins to refuse users on worst channels for Link Centric approach. By lowering the reliability of less important packets, hence providing more bandwidth for all the users, the expected quality approaches still provide acc

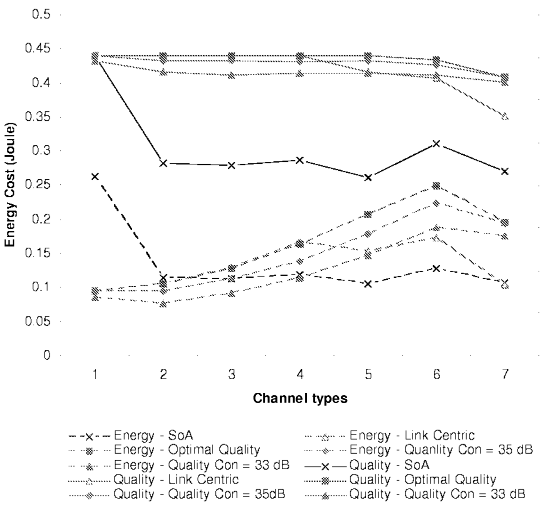

Figure 4: Average Quality and Energy cost of 10users on different channels

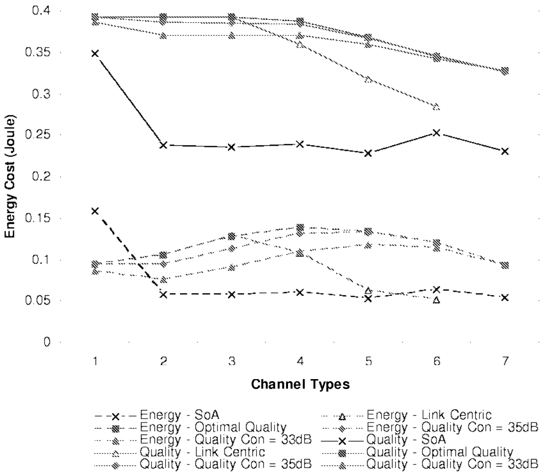

Figure 5: Average Quality and Energy cost of 20users on different channels

nels, while at meantime supporting more users.

The impacts of the time variant channel's influence to different scheduling schemes are shown in Figure 6 and Figure 7 respectively. To simulate a time varying channel, we assume the channel will change to another state every Beacon period with a probability obtained in [10]. The user number increased until 30. If a user was refused to access the channel because of the congestion on a bad channel and the delay requirement is expired, the key pictures will be kept in the waiting queue until better channel states.

From the results, we observe, in the environment of little congestion (less than 5 users), the Optimal Quality approach gains both energy and quality. The quality constraint approach meets the quality requirement well and achieves more energy gains. Compared to the SoA approach, energy gains of 200% to 70% are achieved. And compare to Link Centric approach, the energy gains can reach up to 45%. When congestion becomes more and more severe, the expected quality

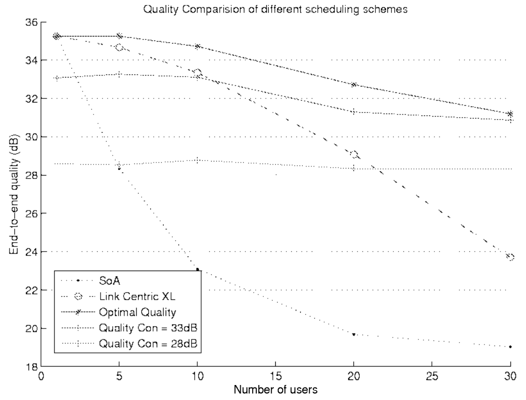

Figure 6: Quality Comparison of scheduling schemes

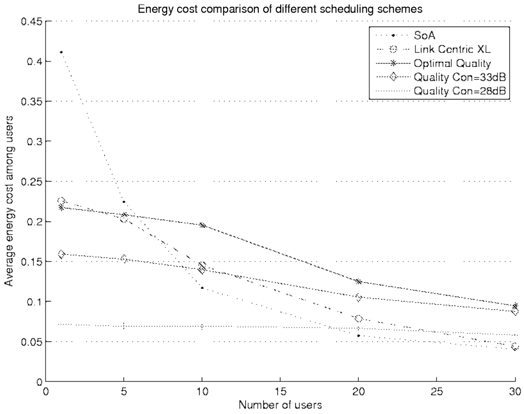

Figure 7: Energy cost comparison of scheduling schemes

approaches have been observed with stable quality. Compared to SoA and Link Centric approaches, quality gains more than 10 dB and 4.5 dB are respectively observed when 30 users interact.

5. CONCLUSIONS

We have introduced an energy efficient packet scheduling mechanism for the transmission of multiple H264/AVC video streams over a wireless LAN. By exploiting the temporal scalability brought by the prediction dependencies among video frames, a centralized cross-layer optimization solution is proposed, which is built upon the HCCA in the IEEE 802.11e standard. Our performance study shows that the proposed solution saves energy on good channel conditions, and achieves significant quality gains on bad channels or under heavily overloaded network conditions.

REFERENCES

[1] Skyrianoqlou. D., Passas. N. and Salkintzis.Ar, "Traffic scheduling for multimedia QoS over wireless LANs," in *Proc. ICC 2005*, May 16-20. 2005, pp. 1266–1270.

[2] A. Ksentini, A. Gueroui and M. Naimi, "Toward an improvement of H.26

802.11e through a cross layer architecture," *IEEE Communications Magazine*, pp. 107–114, Jan. 2006.

[3] A. K. Katsaggelos et al., "Energy-Efficient Wireless Video Coding And Delivery," *IEEE Wireless Communications*, pp. 24–30, Aug. 2005.

[4] C. Luna et al, "Joint Source Coding and Data Rate Adaptation for Energy Efficient Wireless Video Streaming," *IEEE JSAC.*, Vol. 21, No. 10, pp. 1710-1720, Dec. 2003.

[5] R. Mangharam, et al., "Optimal fixed and scalable energy management for wireless LANs," in *INFOCOM'05*, Miami, Mar, 2005.

[6] S.Pollin, et al, "MEERA: cross-layer methodology for energy-efficient resource allocation for wireless networks," *IEEE Trans. on Wireless Communications*, to appear.

[7] X. Ji, et al., "Energy Efficient Bandwidth Allocation for Multi-user Video Streaming over WLAN," in *ICASSP*, Apr. 2007.

[8] IEEE Std 802.11e, *Medium Access Control (MAC) Enhancements for Quality of Service (QoS)*. IEEE 2005.

[9] B. Bougard, et al., "A scalable 8.7nJ/bit 75.6 Mb/s Parallel Concatenated Convolutional Turbo Codec," in *IEEE ISSC*, Feb. 2003.

[10] B. Bougard, et al., "Energy-Scalability Enhancement of Wireless Local Area Network Transceivers," in *IEEE SPAWC*, Lisboa, Jul. 11-14 2004.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2003.

[12] Giguan G. Lin, "Multiple-objective Problems: Pareto-Optimal Solutions by Method of Proper Equality Constraints," *IEEE Trans. on Automatic Control*, Vol. AC-21, No. 5, Oct 1976.

[13] Stephan Wenger, "H.264/AVC Over IP," *IEEE Trans. on CSVT*, Vol. 13, No. 7, Jul. 2003.

[14] Atheros White Paper, *Power Consumption and Energy Efficiency Comparisons of WLAN Products*, 2003.