FINE-GRANULAR-SCALABILITY VIDEO STREAMING OVER WIRELESS LANS USING CROSS LAYER ERROR CONTROL

Xiaofeng Xu[+], *Mihaela van der Schaar*[*], *Santhana Krishnamachari*[†], *Sunghyun Choi*[‡], *Yao Wang*[+]

[+] Polytechnic University
Brooklyn, NY 11201, USA

*University of California, Davis
CA 95616, USA

†EG Technology

‡Seoul National University
Seoul 151-744, Korea

ABSTRACT

Real-time streaming of audiovisual content over the Wireless LANs (WLANs) is emerging as an important technology area in multimedia communications. Due to the error-prone, time-varying characteristics of wireless channels, there is a need for strong protection of video bitstreams. To cope with the variation of channel conditions, we introduce a novel cross-layer protection strategy that combines adaptive application-layer Forward Error Correction (FEC) and physical-layer modulation with Fine-Granular-Scalability (FGS) coding to improve the robustness of wireless transmission. Unlike data streams, different parts of a video stream have different priorities and hence merit the use of unequal error protection. Our schemes can dynamically select the optimal combination of FEC and modulation based on channel conditions and video data content. Experimental results show the stability of the scheme over a variety of channel conditions.

1. INTRODUCTION

Video delivery over wireless networks is very challenging. The presence of multi-path fading, co-channel interference, and noise disturbances makes the channel condition vary rapidly so that it is hard to guarantee for Quality of Service (QoS) parameters such as bandwidth and bit error rates. Many error control stratagies have been proposed to deal with such variability, e.g., [1-3]. These schemes can adaptively select the error control parameters such as FEC, ARQ and packet length, based on the network condition. However, these algorithms considered optimal selection in one network layer only.

In this paper, we propose a cross-layer optimization strategy to achieve robust end-to-end video quality under the IEEE 802.11a WLAN environment. More specifically, application layer, link layer and physical layer are all taken into consideration. The 802.11a standard is targeted for high rate multimedia applications and provides eight different physical layer (PHY) modes with different data transmission rates. The lower rate PHY modes are inherently more robust than the higher rate modes. Therefore, one can achieve desirable trade-offs between robustness and rate by choosing appropriate PHY modes. Furthermore, in application layer, adaptive forward error correction (FEC) has been shown to be effective to combat packet-loss rate variation. Although ARQ is more effective than FEC, retransmission of corrupted data frames introduces additional delay. For real-time traffic, due to the delay constraints, the number of retransmissions is limited and small. For this reason, ARQ is not directly addressed in this paper. In addition to traditional channel error control, recently, several scalable coding methods have been successfully proposed for video transmission through heterogeneous networks. One of these techniques is the MPEG-4 Fine-Granular Scalability (FGS) scheme, which can adapt in real time to the bandwidth variations over heterogeneous networks and the variation in terminal capabilities. Based on the above considerations, we exploit cross-layer error control combining application-layer FEC and physical layer modulations to achieve robust video streaming in the FGS framework. Although we focus on FGS video, the analysis performed in this paper can be easily extended, and will provide similar results when applied to alternative layered video coding schemes (e.g., data partitioning, spatial scalability, temporal scalability, and wavelet-coded video).

The rest of the paper is organized as follows. In section 2, we analyze the error performance of the IEEE 802.11 standard with application-layer FEC. In section 3, we briefly review the MPEG-4 FGS coding tool and propose a R-D model for FGS video. Based on this model, we derive the expected receiver-end distortion when a FGS stream is delivered through a packet lossy network, where the packet loss rate for each packet depends on the PHY mode and the FEC code rate applied. The cross-layer error control problem is formulated as to choose the optimal combination of PHY mode and FEC rate for each packet, so as to minimize the receiver distortion. In section 4, we compare the performance obtained with the proposed cross-layer adaptation with adapting PHY or FEC only. Section 5 concludes the paper.

2. ERROR PERFORMANCE ANALYSIS FOR THE IEEE 802.11 NETWORK

In this section, we briefly review the IEEE802.11 standards. Also, we analyze the packet loss rate based on the frame structure of IEEE 802.11 MAC and application layer FEC.

2.1. PHY modes and Channel Model

The IEEE 802.11a PHY provides the interface between the MAC and the wireless medium. Based on Orthogonal Frequency Division Modulation (OFDM), the PHY provides eight PHY modes with different modulation schemes and different convolutional coding rates at the 5 GHz U-NII band. During transmission, each MAC Protocol Data Unit (MPDU) is provided with a Physical Layer Convergence Protocol (PLCP) preamble and a PLCP header as shown in Figure 1. The PLCP header, except the SERVICE field, is transmitted using PHY mode 1. The 16-bit SERVICE field of the PLCP header and the MAC frame are transmitted at the data rate specified in the PLCP header.

Mode	Modulation	Code Rate	Data Rate	BpS
1	BPSK	½	6	3
2	BPSK	¾	9	4.5
3	QPSK	½	12	6
4	QPSK	¾	18	9
5	16-QAM	½	24	12
6	16-QAM	¾	36	18
7	64-QAM	2/3	48	24
8	64-QAM	¾	54	27

Table 1: Eight different PHY modes of IEEE 802.11a

PLCP Preamble (12 Symbols)	PLCP header (20 bits)	SERVICE (16 bits)	MPDU	Tail bits (6 bits)

Figure 1: IEEE 802.11 PHY frame format

In this paper, we use the multi-path channel model of [4] to model the channel as a tapped delay line, where the distribution of path amplitude is chosen to be Rayleigh and the average power of different taps declines exponentially with delay. Using the above channel model and a typical receiver model, one can obtain Bit Error Rate (BER) for different Signal Noise Ratio (SNR) for different PHY modes of 802.11a. Using these bit error values, and assuming random errors, the probability that a frame of L bytes can have at least 1 bit error is:

$$P_e^m(L) = 1-(1-p_b^m)^{8L} \tag{1}$$

where $P_b{}^m$ is the BER of PHY mode m at a given channel SNR.

2.2. MAC PCF

The 802.11 WLANs standard defines the point coordination function (PCF) as medium access control mechanisms. In the PCF mode, the point coordinator (PC) centrally coordinates the access to the wireless medium. Based on a poll-and-response protocol, the central controller PC can control access to the shared wireless medium and eliminate contention among wireless stations. It makes use of the priority inter-frame space (PIFS) to gain control of the medium. After seizing control of medium, the PC starts a contention-free period (CFP). During CFP, the PC sends a data frame and then expects a CF-ACK frame from the corresponding station within a short inter-frame space (SIFS) time. After receiving the CF-ACK frame, the PC sends the data frame to the next station. If the data frame is received in error or if the frame is not received at all, the PC asks for the re-transmission of the data frame after a PIFS. If the data frame is received correctly, but the CF-ACK frame is failed, then the PC waits for a SIFS time and re-transmits the data frame.

Assuming that an L-byte frame body is transmitted using PHY mode m, the probability of a successful transmission is given by:

$$P^m_{good\_cycle}(L) = (1-P^m_{e,ack})(1-P^m_{e,data}(L)) \tag{2}$$

where $P^m_{e,ack}$ is the CF-ACK transmission error probability and $P^m_{e,data}$ is the data transmission error probability. These can be calculated as follows:

$$P^m_{e,data}(L) = (1-P_e^1(3))(1-P_e^m(38.75+L))$$

$$P^m_{e,ack}(L) = 1-(1-P_e^1(3))(1-P_e^m(38.75)) \tag{3}$$

where $P_e{}^1(3)$ is the error probability of the PLCP header, and 38.75 is MAC /PHY overheads.

2.3. Application-layer FEC

We invoke application-layer FEC using the Reed-Solomon (RS) code. RS coding is applied across packets using an interleaver [7]. A *(n,k)* RS decoder can correct up to $n-k$ packet erasures. Therefore, the probability of error after RS decoding is given by:

$$P_r = 1-\sum_{i=0}^{n-k}\binom{n}{i}(P^m_{good\_cycle})^{n-i}(1-P^m_{good\_cycle})^i \tag{4}$$

3. CROSS-LAYER ERROR CONTROL: PROBLEM FORMULATION

3.1 R-D MODELING OF FGS CODED VIDEO

FGS has recently been adopted by MPEG-4 as a video-coding tool for streaming applications. The FGS framework consists of a non-scalable base-layer and a fine-granular enhancement-layer. The base layer can be compressed using a MPEG-4 compliant non-scalable encoder with motion-compensated encoding method. In addition to the base-layer, FGS consists of a single enhancement-layer coded in progressive (fine granular) manner. FGS can provide fine-granularity quality improvement for each additionally transmitted byte in enhancement layer. The base-layer is coded to a minimally acceptable quality of video with a bit-rate R_{bl}. Over the time-varying network, R_{bl} is almost certainly lower than the available bandwidth so that the base layer can provide a low but guaranteed level of quality. The enhancement-layer is progressively (bit-plane by bit-plane) coded by employing a low-complexity bit-plane embedded-DCT algorithm. The enhancement-layer improves upon the base-layer video, fully utilizing the available bandwidth at transmission-time. Hence, the enhancement layer can provide improvement in fine granularity. As long as the base layer video is reliably delivered, the packet or bit losses in the enhancement-layer do not propagate since the enhancement-layer frames are intra-coded in a progressive (fine-granular) manner.

Without any motion-compensation in the enhancement-layer, we can observe a linear dependency between the increase of enhancement data rate and the PSNR. Based on this observation, we propose the following rate-distortion (R-D) model for the FGS encoder:

$$Q_s = \theta\cdot(R-R_{bl})+Q_0 \tag{5}$$

where Q_s is the quality (in terms of PSNR) of the encoded sequence, R is the output data rate of the encoder, Q_0 is the quality of the base layer in the encoder, θ is the parameter of the RD model which depends on the characteristics of the encoded video sequence.

3.2 END-TO-END QUALITY ANALYSIS

In this section, we derive the expected receiver-end quality based on the R-D model of (6). For simplicity, we assume that the base layer is always received correctly. We assume that the enhancement layer is partitioned into N_q packets, with packet j containing R_e^j information bits. In general, each packet can have a different packet loss rate, depending on the FEC code used. We denotes the packet loss rate for packet j by P_{el}^{i} , which can be calculated according to Equations (1-4) based on the PHY mode and the FEC code rate used. A lost packet in an enhancement layer frame causes the remainder packets in that frame useless, so that the receiver-end quality depends on the number of packets received before the first lost packet. Therefore, the expected receiver-end quality is:

$$E(Q_s) = \theta \cdot \left(\sum_{i=1}^{N_q} \left(\prod_{j=1}^{i} (1-P_{el}^{j}) \right) P_{el}^{i+1} \left(\sum_{j=1}^{i} R_{e}^{j} \right) \right) + Q_0 \qquad (6)$$

The total bit rate of a FGS stream after FEC depends on the RS codes used for the packets. Let r_{bl} represent the redundancy rate of the RS code for the base layer packets and r_{el}^{i} the redundancy rate for the j-th enhancement layer packet.(For a (n,k) RS code, the redundancy rate is defined as $r=(n-k)/k$.) Then the total rate including the base layer packets and N enhancement layer packets is

$$R(N) = (1+r_{bl})R_{bl} + \sum_{i=1}^{N} (1+r_{el}^{i})R_{ep}^{i} \qquad (7)$$

The maximum number of packets that an enhancement layer frame can have, N_q, depends on the total bit-rate available at transmission time over the channel, R_t, which depends on the PHY mode used:

$$N_q = \arg\max_{R_t - R(N) \ge 0}(N) \qquad (8)$$

3.3 CROSS-LAYER OPTIMIZATION

At transmission time, we are interested in the following parameters: (i) the channel code rate r_{bl} , r_{el}^{i} ; (ii) the PHY mode m. Because enhancement layers cannot be decoded without base layers, any loss in base layers should be avoided. The packet loss rate in base layers should be as small as possible. Therefore, we reduce the search space with the constraint: $r_{bl} > r_{el}^{i}$. We define the set A as the available code rate set, and the set B as the available PHY mode set. The cross-layer optimization problem can be formulated as follows: given particular channel condition (e.g., in terms of SNR), encoding rate of base layers R_{bl} , and enhancement layer packet rates R_{e}^{i}, determine the parameter set $\{r_{bl}, r_{el}^{i}, m\}$ that minimizes the expected receiver-end distortion. We have

$$E(Q_s)^{*} = \min_{r_{bl}, r_{el}^{i}, m} E(Q_s) \qquad (9)$$

where r_{bl} , $r_{el}^{i} \in A$, m$\in B$.

4. EXPERIMENTAL RESULTS

4.1 VERIFICATION OF R-D MODEL FOR FGS

In our simulations, we use the MPEG-4 FGS coder [5]. For the experiments we used a CCIR601 resolution (720x576 Y pixles) video sequence coded at a frame rate of 25Hz. The base layer bit rate is 1Mbps. First, we verify the accuracy of the R-D model given in (5). In figure 1, the blue points are the real pairs of video quality and encoding rate at which the video is compressed by a FGS coder. The purple line is the linear model curve, which fits the real data very well. Note that the R-D curves for different video sequences may have different slopes. So for each video sequence, we need to find the slope based on selected R-D points using the least squares method.

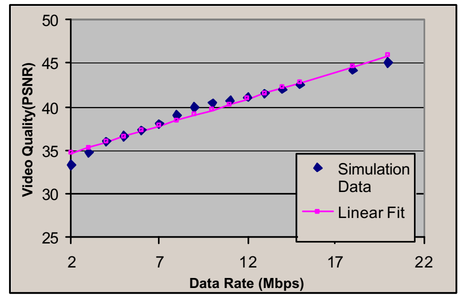

Figure 1 Quality-distortion curve

4.2 PERFORMANCE COMPARISON

We consider the cross-layer optimization within the following parameter space: the available channel code rate set A={0.476, 0.746, 0.936, 1}, the available PHY mode set B={1, 3, 5}. We assumed that the channel is congested with other traffic, including non-real-time traffic. The total available channel data rates for the available PHY mode are R_t={1.5Mbps, 4.5Mbps, 6Mbps}. We present the results when the video packet length is set to 2080 bytes. In our transmission scheme, the frames in a GOP are delivered based on their original encoded order. Base layers are first transmitted and then enhancement layers. If a packet exceeds its playback time, it is discarded.

Instead of allowing each packet having a different FEC code rate, we simplify the problem by dividing all the packets into three groups and assigning the same FEC rate for all the packets in the same group. The first group contains all base-layer packets, the second group include the first 1/3 of the enhancement-layer packets, and the third group the remaining enhancement-layer packets. This is based on the consideration that base-layer packets have higher priority than enhancement-layer packets and the lower part of an enhancement layer has higher priority than the higher part. The optimization problem is to search the best combination of the PHY mode and the FEC rates for the three groups that maximizes the expected receiver-end quality.

We compare the results obtained with this cross-layer optimization with non-adaptive approach (fixed PHY and FEC) and optimally adapting FEC or PHY only. Figure 2 shows the video quality using different PHY modes. From this figure, we observe that switching PHY mode can improve video quality. Figure 3 shows the video quality using our scheme versus adaptive application-layer FEC with a fixed PHY mode 5. From the figure, it can be seen that when the channel SNR is below 26dB, the video quality suffers severe performance degradation without using FEC. Although adaptive application-layer FEC can mitigate the degradation, the video quality still collapse when the channel SNR drops below 25dB. The cross-layer optimization scheme, on the other hand, can provide acceptable quality even when the channel SNR drops to as low as 20dB. Figure 4 compares our scheme with dynamically switching PHY mode without FEC. Our scheme always wins over the mode-switching algorithm. When the channel condition is changing from 21dB to 24dB or from 26dB to 28 dB, the video quality improves about 2dB by our scheme. This is because using a FEC code can counter the packet erasures that cannot be compensated by switching PHY mode. The cross-layer protection technique results in considerably better visual quality compared with adapting the application-layer FEC or the PHY mode alone. With the FEC or PHY adaptation only, the quality of the video data decreases abruptly when the channel SNR drops; with our cross-layer protection, the visual quality degrades more gracefully. The improvement introduced by our scheme is especially visible for highly time-varying channel.

5. CONCLUSION

In this paper, we proposed a cross-layer error control scheme for delivering FGS video over wireless networks. We have specifically concentrated on 802.11a WLANs, which provides different PHY modes. We derive the relation between the residual packet loss rate and the FEC rate. Furthermore, we propose a R-D model for the FGS video. Together it enables us to derive the expected receiver-end quality achievable by a given combination of PHY modes and FEC rates. We show that by jointly adapting the application-layer FEC rate and the PHY mode to maximize the expected quality, one can obtain significantly better quality video than adapting the FEC or PHY mode alone, especially when the channel SNR is low.

6. REFERENCES

[1] D. Qiao and K.G. Shin, "A Two-step Adaptive Recovery Scheme for Video Transmission over Wireless Networeks," IEEE Infocom 2000, pp.1698-1704, 2000.

[2] K. Stuhlmuller, M. Link, and B. Girod, "Scalable Internet Video Streaming with Unequal Error Protection," in Proc. Packet Video Workshop'99, April 1999.

[3] D. Qiao and S. Choi, "Goodput enhancement of IEEE 802.11a Wireless LAN via Link Adaptation," Pro. IEEE ICC'01, Helsinki, June 2001.

[4] N. Chayat, "Tentative Criteria for Comparison of Modulation Methods," IEEE P802.11-97/96, September 1997.

[5] MoMusys code, "MPEG4 verfication model version 18.0" ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio, January 2001.

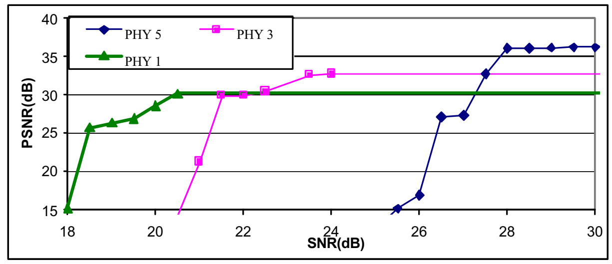

Figure 2 Different PHY modes without FEC

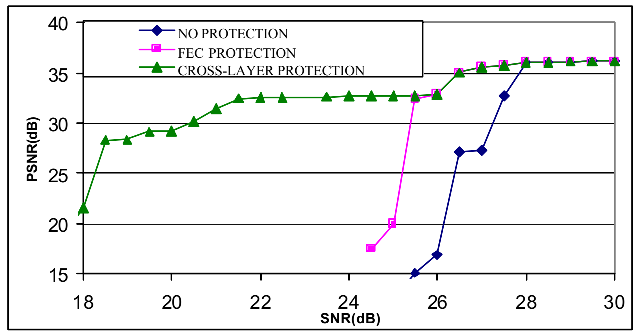

Figure 3 Cross-layer adaptation vs. adaptive application-layer FEC with a fixed PHY mode.

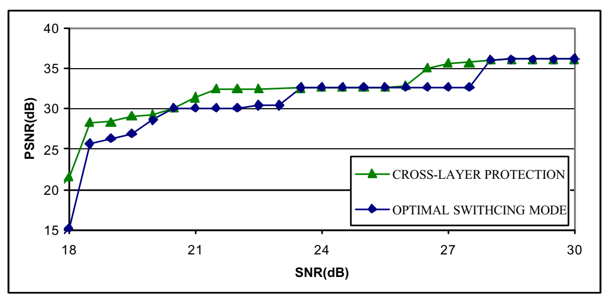

Figure 4 Cross-layer adaptation vs. optimal switching PHY modes without using FEC