Adaptive Cross-Layer Protection Strategies for Robust Scalable Video Transmission Over 802.11 WLANs

Mihaela van der Schaar, *Member, IEEE*, Santhana Krishnamachari, Sunghyun Choi, *Member, IEEE*, and Xiaofeng Xu, *Student Member, IEEE*

***Abstract*—Robust streaming of video over 802.11 wireless local area networks poses many challenges, including coping with bandwidth variations, data losses, and heterogeneity of the receivers. Currently, each network layer (including physical layer, media access control (MAC), transport, and application layers) provides a separate solution to these challenges by providing its own optimized adaptation and protection mechanisms. However, this *layered* strategy does not always result in an optimal overall performance for the transmission of video. Moreover, certain protection strategies can be implemented simultaneously in several layers and, hence, the optimal choices from the application and complexity perspective need to be identified. In this paper, we evaluate different error control and adaptation mechanisms available in the different layers for robust transmission of video, namely MAC retransmission strategy, application-layer forward error correction, bandwidth-adaptive compression using scalable coding, and adaptive packetization strategies. Subsequently, we propose a novel adaptive cross-layer protection strategy for enhancing the robustness and efficiency of scalable video transmission by performing tradeoffs between throughput, reliability, and delay depending on the channel conditions and application requirements. The results obtained using the proposed adaptive cross-layer protection strategies show a significantly improved visual performance for the transmitted video over a variety of channel conditions.**

***Index Terms*—adaptive cross-layer error protocol, robust scalable video transmission, IEEE 802.11 WLAN.**

I. INTRODUCTION

THE USE of IEEE 802.11 wireless local area networks (WLANs) as an extension to the existing wired infrastructure, offering the convenience of mobility and portability in the enterprise environment, is growing at a rapid pace. The falling cost of WLAN products has also led to their increased use in consumer homes. Although currently WLANs are predominantly used for data transfer, the higher bandwidth provided by new WLAN technologies such as IEEE 802.11a and IEEE 802.11g will ultimately lead to their increasing use for multimedia transmissions. However, to achieve a high level of acceptability and proliferation of wireless multimedia, in particular wireless video, several key requirements need to be satisfied in order to provide a reliable and efficient transmission: 1) easy adaptability to bandwidth variations; 2) robustness to data loss; and 3) support for bandwidth, power, and device scalability.

In this paper, we investigate the robust and efficient transmission of video over WLANs. We specifically consider the recent WLAN standard, IEEE 802.11a [2], which offers high bit rates up to 54 Mb/s, enabling the transmission of delay sensitive audio/visual (AV) traffic. This paper proposes a novel vertical system integration that enables the joint optimization of the various protection strategies existing in the protocol stack. In the remainder of this paper, we refer to this vertical system integration strategy as *cross-layer* protection. The error control strategies that can be implemented at the various layers, namely, media access control (MAC) retransmission limit, application-layer forward error correction (FEC), and adaptive packet size selection will be investigated for the efficient transport of video over 802.11a. We will discuss cross-layer protection strategies, where for instance the FEC decoding is performed at the application layer and the retransmissions are handled at the link layer. Moreover, unlike previous papers that discussed optimal selection of physical layer modes and fragment lengths for the 802.11a distributed coordination function (DCF) [5], in this paper we specifically consider the point coordination function (PCF) mode for video transport, since it is the most effective scenario for video transmission over 802.11 WLANs. As will be explained further below, the PCF is based on the poll-and-response, and was developed in order to support real-time traffic like video.

To fulfill the requirements for wireless video identified at the beginning of the introduction, we employ MPEG-4 fine-grained scalability (FGS) for the compression of the video data (like in [11] and [6]), because it can provide easy adaptation to bandwidth variations and device characteristics [7].

In this paper, our adaptive cross-layer protection strategy is pursued as follows.

1) A multipath channel model is used to simulate the wireless indoor channel. This channel model provides the bit-error rate (BER) of the link for the eight different PHY modes of 802.11a under different channel signal-to-noise ratio (SNR) conditions.
2) Based on this channel model and the 802.11a PCF mode of operation, we analytically derive the packet loss ratios

Manuscript received October 15, 2002; revised May 1, 2003.

M. van der Schaar is with the the Department of Electrical and Computer Engineering, University of California, Davis, Davis, CA 95616-5294 USA.

S. Krishnamachari is with EG Technology, Atlanta, GA 30318 USA.

S. Choi is with the Seoul National University, College of Engineering, School of Electrical Engineering and Computer Science, Seoul 151-744, Korea (e-mail: schoi@snu.ac.kr).

X. Xu is with Polytechnic University, Brooklyn, NY 11201 USA.

Digital Object Identifier 10.1109/JSAC.2003.815231

and throughput efficiency at various channel conditions for a given packet size, a given number of retransmissions at the MAC, and an application layer FEC. The use of the application-layer FEC is advocated to improve the poor quality of the link under adverse channel conditions and to enable the use of unequal error protection for video traffic. We focus on the use of a small number of retransmissions to accommodate the strict delay constraints of real-time traffic.

3) An analytical model is developed to characterize the end-to-end distortion of FGS video based on the packet loss ratios obtained above. We show that the derived model matches the results obtained by the simulations.
4) Based on the end-to-end distortion model, cross-layer protection strategy is developed to dynamically adapt the following parameters in the video streaming system: a) the application layer FEC; b) the maximum MAC retransmission limit; and c) the packet sizes.

A. Related Work

In recent years, many papers proposed various solutions addressing one or several of the previously mentioned requirements. In [8], Girod and Färber give an excellent review of the existing solutions for combating wireless transmission errors. While their focus is on cellular networks, most presented protection strategies can also be applied to the transmission of video over WLANs. The focus is on channel-adaptive source coding schemes that are useful when real-time channel feedback is available to the encoder. Importantly, joint consideration of network and application layers is mentioned as an interesting area for further research.

In [9], Shan and Zakhor presented a novel integrated application-layer packetization, scheduling, and protection strategies for wireless transmission of nonscalable coded video. Cote *et al.* present in [10] a thorough survey of the different video-optimized error resilience techniques that are necessary to accommodate the compressed video bitstreams which are very sensitive to bit errors and packet losses. Various channel/network errors can result in a considerable damage to or loss of compressed video information during transmission, effective error concealment strategies become vital for ensuring a high quality of the video sequences in the presence of errors/losses. An excellent review of the existing error concealment mechanisms is given by Zhu and Wang in [3]. In [11], Majumdar *et al.* address the problem of resilient real-time video streaming over IEEE 802.11b WLANs for both unicast and multicast transmission. For the unicast scenario, a hybrid automatic repeat request (ARQ) algorithm that efficiently combines FEC and ARQ is proposed. For the multicast case, progressive video coding based on MPEG-4 FGS is combined with FEC. Similarly, in [6], Van der Schaar and Radha discussed the combination of MPEG-4 FGS with scalable FEC for unicast and multicast applications, and a new unequal error protection strategy referred to as fine-grained loss protection (FGLP) has been introduced. Hybrid ARQ schemes, where the rate of the associated FEC is adaptively changed based on the underlying channel conditions, have also been presented by Wang and Zhu in [3] and by Ma and Zarki in [4]. However, it should be pointed out that the protection strategies described in these papers are implemented at the application layer and do not exploit the mechanisms available in the lower layers of the protocol stack. In summary, the research efforts in the area of robust wireless transmission have mainly focussed on enabling adaptive error-control strategies at the application layer. However, in existing WLAN environments, different protection strategies exist at the various layers of the protocol stack and, hence, a joint cross-layer consideration is desirable in order to provide an optimal overall performance for the transmission of video.

In summary, the research efforts in the area of robust wireless transmission have mainly focussed on enabling adaptive error-control strategies at the application layer. However, in existing WLAN environments, different protection strategies exist at the various layers of the protocol stack and, hence, a joint consideration is desirable in order to provide an optimal overall performance for the transmission of video. Below, we briefly introduce the various protection strategies present in the different protocol layers, succinctly discuss their merits and benefits for cross-layer optimization.

1) At the physical layer, it is possible to dynamically reconfigure the modulation and channel coding techniques for each packet, based on the channel characteristics to allow for tradeoffs between throughput and reliability. Interleaving is also often applied at the physical layer to combat burst errors. Moreover, for multicast applications, where MAC or application-layer retransmission schemes cannot be successfully employed, this protection strategy could successfully be adopted. These mechanisms are very effective in combating losses due to interference, mobility, and fading, and enable on-the-fly tradeoffs between throughput, robustness, and delay. The complexity associated with this flexibility is limited. Nevertheless, the main disadvantage of the physical layer protection strategies is that they are solely based on the observed channel condition and, thus, the adaptation mechanisms do not consider the application requirements in terms of throughput, delay, etc. Consequently, cross-layer protection strategies would enable the adaptation at the physical layer to be done based on the application requirements. For instance, since video can tolerate a certain amount of losses, the best possible video quality given a particular channel condition can be obtained by performing tradeoffs between throughput, delay, and robustness provided by the various physical layer operation modes.
2) In the MAC layer, retransmissions are used for protection/error control. Note that the current MAC implementations of 802.11 WLANs do not employ FEC. The maximum number of retransmissions (i.e., retransmission limit) can be changed adaptively per packet to provide throughput, reliability, and delay tradeoffs. The flexibility in providing these tradeoffs is supported by most practical implementations of 802.11 WLANs and the complexity associated with this adaptation is relatively low. While adapting the retransmission mechanism is effective for unicast applications, multicast applications require employing alternative protection

mechanisms. Also, retransmission is not very effective in combating long burst of packet losses that occur due to interference or mobility. By using a cross-layer protection strategy employing application layer interleaved FEC besides the MAC retransmission, we can handle bursty errors, as well as adapt the protection strategy to the content characteristics.

3) At the application layer, FEC, ARQ, and hybrid ARQ schemes can be employed along with error resilient video coding schemes and error concealment strategies. Applying protection strategies at the application layer leads to a higher system complexity, but it has the advantage that it can be more specifically targeted toward the content characteristics, necessary levels of protection, etc.
4) Different packetization strategies can be adopted at the various layers (physical layer, MAC layer, transport layer, application layer), leading to various tradeoffs between throughput and reliability. However, for an optimal performance, the packetization strategies implemented at the various protocol layers should be designed jointly in order to maximize the throughput and robustness for particular channel conditions, while also taking into consideration the delay requirements imposed by the application.

Currently, the various protection strategies in 802.11 WLAN implementations are optimized separately at each network layer. However, for efficient real-time transmission of video over wireless networks, it is necessary to pursue a cross-layer optimization. Also, note that while the protection strategies implemented in the different protocol layers are often evaluated only from the perspective of maximizing the throughput and robustness for particular channel conditions, the actual evaluation of these strategies should be performed from the perspective of the application (i.e., by taking into consideration the quality, complexity, and delay requirements imposed by the application). For instance, for real-time video streaming applications, the various protection strategies should be evaluated in terms of their impact on the video quality perceived at the receiver side [expressed for instance in terms of peak-signal-to-noise ratio (PSNR) values].

B. Organization

The rest of the paper is organized as follows. In Section II, we present a general overview of the considered system including IEEE 802.11 MAC and 802.11a PHY, the application-layer FEC, and the adopted multipath model. We then analyze the throughput efficiency and delay of the employed system considering all the protocol overheads in Section III. In Section IV, we describe the mechanism to estimate the end-to-end distortion for the FGS transmission of video and we discuss different packetization strategies for FGS. In Section V, we present the experimental results obtained for the various error control strategies and the adaptive selection of these strategies based on the channel state. Section VI concludes the paper.

II. System Overview

The IEEE 802.11 WLAN standard [1] initially defined the specification for the MAC sublayer and three different low bit rate physical layers supporting 1 and 2 Mb/s. Due to the limited bit rate capabilities, the low rate systems have been predominantly used for data traffic. Two high-speed physical layers that were defined later [2], namely, 802.11b in the 2.4 GHz ISM band and 802.11a in the 5 GHz U-NII band, can offer bit rates up to 22 and 54 Mb/s, respectively. The 802.11a standard, which is considered in more detail here for video transmission, provides eight different physical layer (PHY) modes offering data transmission rates from 6 to 54 Mb/s. The lower rate PHY modes are inherently more robust than the higher rate modes. With the substantial increase in the bit rate available with 802.11a, real-time AV applications over WLANs become a reality.

The use of WLANs for the transport of video poses interesting and challenging problems due to the strict delay constraints of the video traffic and the inherent unpredictability of the wireless link characteristics. Unlike data traffic, video traffic is delay-sensitive and somewhat tolerant to packet loss through the use of error concealment techniques at the video decoder [3]. For nonreal-time data traffic, packet losses can always be countered by repeated retransmissions until the packet is received error free. However, for real-time traffic, due to the delay constraints, the number of retransmissions that can be used is limited and is usually small. Therefore, it may be necessary to use additional error control strategies, such as an application-layer FEC, to ensure reliable transport. The use of an application-layer FEC also has the additional advantage of offering the flexibility of unequal or selective error protection for video transport. Unlike data streams, different parts of video streams have different priorities and, hence, can benefit considerably from the use of unequal error protection as will be illustrated in this paper.

In this paper, we investigate the mechanisms necessary for robust video transmission from a video server to a wireless station. The video data is initially compressed using the FGS scalable coding format and stored on the PC/video server. Therefore, no real-time encoding is performed. At transmission time, the video data is streamed in real-time to the wireless station over an 802.11a network. In this section, we describe the various components of this wireless transmission system, including IEEE 802.11 MAC and 802.11a PHY, and the employed application-layer FEC.

A. 802.11 MAC Point Coordination Function (PCF)

For a wireless device transmitting AV content, periodic access to the shared wireless medium is very important. The 802.11 WLAN standard allows two different medium access control mechanisms, namely, the DCF and the PCF. The DCF is the basic access mechanism, which is based on carrier sense multiple-access with collision avoidance (CSMA/CA). In the DCF mode of operation, each station in the WLAN contends for the medium, and relinquishes control after transmitting a single frame. Clearly, this access mechanism is not very suitable for video streaming applications. In the PCF mode of operation, the access to the wireless medium is centrally controlled by the point coordinator (PC). The PC appropriately schedules the downlink traffic for delivery to different wireless stations and, for uplink traffic, grants the stations access to the medium through a polling mechanism. This mode is more appropriate for real-time applications. In this paper, we consider the PCF operation for downlink video traffic delivery from the video server (or equivalently wireline infrastructure) to the wireless station.

The PCF is based on a poll-and-response protocol to control access to the shared wireless medium and to eliminate contention among wireless stations. The PC is the central controller, which controls the access to the medium. The PC gains the control of the medium periodically. Once the PC gains control of the medium, it begins a contention-free period (CFP) during which the access to the medium is completely controlled by the PC; after a CFP is finished, a contention period (CP) during which the mandatory DCF is used starts. During the CFP, the PC can deliver downlink traffic to the individual stations without any contention. The PC can also send a contention-free poll (CF-Poll) that allows the stations to send uplink traffic to the PC. If the station that is being polled has uplink traffic to send, it can transmit one frame for each CF-Poll received. If the station does not have any pending frame, it responds with a data frame without any content, i.e., a Null data frame. During the CFP, a wireless station can only transmit after being polled by the PC.

B. PHY Modes of 802.11a

The 802.11a PHY transmits and receives data frames over the shared wireless medium. It is based on orthogonal frequency division multiplexing (OFDM) and provides eight different PHY modes (number from 1 through 8) with different modulation schemes and code rates offering data transmission rates ranging from 6 to 54 Mb/s (i.e., $\mathrm{DR}\,(\cdot) = \{6, 9, 12, 18, 24, 36, 48, 54\}$ Mb/s).

C. Employed Channel Model

The adopted channel model is presented very briefly here. We have used a multipath channel model. The channel is modeled 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. The angle of each arrival path is chosen as uniformly distributed in the range of 0 to 2π. The number of taps depends on the delay spread. In many of the measurements at 5 GHz, the delay spread varies between 50 and 300 ns. We have used a delay spread of 200 ns in our simulations. Using the above channel model and a typical receiver model, the performance curves of BER versus SNR are obtained for different PHY modes of 802.11a. For illustration, Fig. 1 shows the BER versus SNR performance of three different PHY modes of 802.11a. The BER values are obtained by selecting different statistical channel instances with a given SNR and averaging the resultant BER values over the different channel instantiations. Using these bit-error values and assuming random errors, the probability of error for a packet of length L bytes is obtained as

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

where p_b^m is the BER of PHY mode m at a given SNR. The analysis presented in this paper can be extended to include burst errors and multistate channel models.

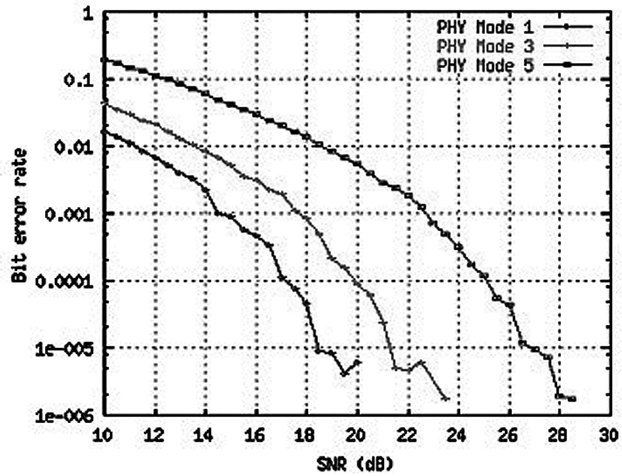


Fig. 1. BER versus SNR for various PHY modes.

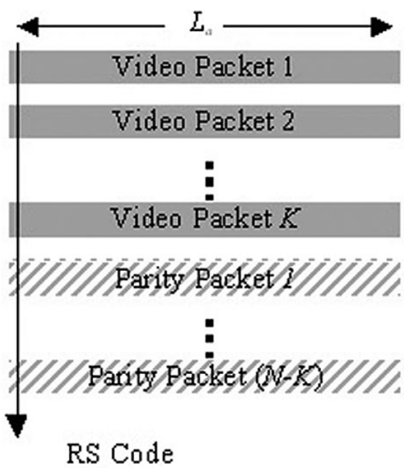


Fig. 2. Application-layer RS coding across packets.

D. Application Layer FEC Coding

We consider the use of Reed–Solomon (RS) codes for the application-layer FEC. When the FEC is used at the application layer, it is necessary to apply the RS coding across video packets. This is due to the fact that the nominal 802.11 MAC implementations discard the whole MAC frame in the event of an error. The erroneous frame at the receiving MAC is never passed on to the higher layer. Therefore, if RS coding is applied within a single packet at the application layer, the erroneous packet will not be available for error detection or correction at the application layer.

Therefore, RS coding at the application layer is applied across packets using an interleaver, i.e., K video packets each of length L_a bytes are buffered at the interleaver as shown in Fig. 2. The first symbol (byte) from each of the K video packets are sent through a (N, K) RS coder resulting in $(N - K)$ parity symbols each of which form the first byte of the $(N - K)$ parity packets. This is repeated for the L_a bytes resulting in $(N - K)$ parity packets each of length L_a generated by the RS encoder. Each video or parity packet is transmitted via an IEEE 802.11 MAC frame; if this frame is discarded at the receiving 802.11 MAC

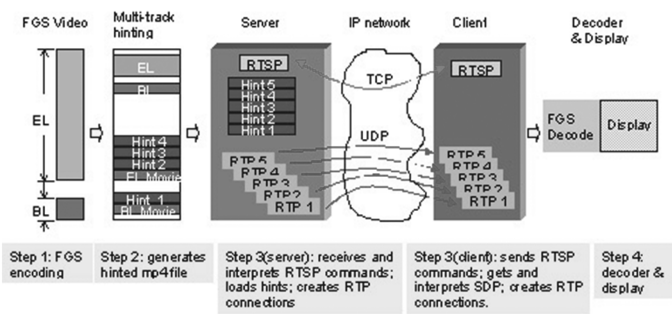


Fig. 3. FGS streaming system, in which BL and EL represent base layer and enhanced layer, respectively.

layer due to channel errors, this results in a symbol erasure at the RS decoder in the application layer. The RS decoder at the application layer can correct up to $(N - K)$ packet losses out of N packets over which the RS coding was applied.

E. FGS Streaming System

For the robust transmission of scalable coding over WLANs, we proposed in [13] a streaming system architecture different from that of conventional ones. This architecture can take advantage of the bit-rate scalability provided by scalable coding methods like FGS to accommodate the dynamic conditions that are common over wireless IP networks, the user requirements and/or receiving device's complexity/power requirements. In this context, we developed a generic scalable streaming architecture that includes client/server functions, control protocols, and related algorithms. In this section, we highlight only the "basic" features of the scalable streaming system that are important for the discussion of the cross-layer error protection strategies. The streaming system is illustrated in Fig. 3. (The interested readers are referred to [13] for more details.)

In our streaming system, the FGS enhancement layer is divided in several sublayers (see Section IV-A). To facilitate scheduling, rate adaptation, priority dropping, loss detection, and unequal error protection, different real-time transport protocol (RTP) payload types are associated to the base layer and to each of the enhancement sublayers. Then, by inspecting the RTP payload type, various application and MAC-layer unequal error protection strategies can be implemented for different layers. In addition, the various sublayers corresponding to one video frame are not scheduled sequentially (i.e., base layer of video frame Fn, followed by the most important FGS sublayer of frame Fn, and so on.) Alternatively, in our system, we first schedule the base-layer packets for a group of frames, followed by the next most important FGS enhancement sublayer for that group of frames, and so on. In this manner, the most important information is transmitted first. A simple "time-out" adaptation procedure is implemented that ensures that as soon as the deadline for transmitting the packets of that group of frames has passed, the less significant FGS enhancement sublayers are no longer sent and a new group of frames is transmitted. Note that if a MAC-layer ARQ is utilized, by employing this scheduling algorithm, a higher priority is inherently given to the transmission of the most important packets and their retransmission.

F. Estimation of Channel SNR

In order to dynamically change the packetization and error control strategies at transmission time, the SNR at the receiver side should be known. However, it is not possible to know it in reality and, hence, for a successful and practical implementation of such an adaptive scheme, the PC or the transmitter should be able to estimate the SNR at the receiver. This can be achieved by different means: one of the approaches is to assume that the channel is symmetric, i.e., the received signal SNR at the PC is the same as the SNR at the station. Under this assumption, the PC can estimate the SNR at any station by measuring the SNR of the signals it receives from a given station. Another possible approach is by means of periodic feedback from the receiver informing the transmitter about the current receiver SNR. The transmitter can then use this information, possibly with some prediction, in determining the error control strategies for the next set of frames. The interval of this feedback should depend on how fast the channel is expected to change over time. One such mechanism is currently being proposed for the upcoming improvements to 802.11a, called IEEE 802.11h. Another avenue for estimating the receiver SNR can be based on the number of unsuccessful MAC frame transmissions or based on the real-time control protocol (RTCP) reports from the receiver. When the PC transmits a set of frames to a given station, it can maintain a record of the number of transmissions for which the acknowledgment were not received, and then can use these statistics to estimate the SNR at the corresponding station. The RTCP reports from the receiver also contain information about the fraction of RTP packets that are lost within the RTP stream, but this channel estimation mechanism is slower than that based on the unsuccessful MAC frame transmissions.

III. Throughput Efficiency and Delay Analysis

Here, we analyze the throughput efficiency and delay performance of the 802.11a PCF mode with a given application-layer RS code and a MAC retransmission limit. The throughput efficiency is defined as the ratio of the useful data, say video data, which is received to the total amount of transmitted bits. Therefore, we take into account the overhead incurred due to use of the FEC, retransmissions, and headers associated with different protocol layers. For the analysis, we have made the following assumptions: 1) the video packets are of length L_a bytes and these packets are not fragmented in any of the lower layers and 2) the overhead of the higher layer protocols, like RTP, UDP, and IP is O bytes. The overhead of the MAC and PHY are not included in this. The average frame transmission duration computed in the following section accounts for the MAC and PHY overhead. The MAC layer retransmission limit is set to R.

A. Average Frame Transmission Duration

In this section, we analyze the average transmission duration of a MAC frame under different conditions. This is used later in the computation of throughput efficiency with an application-layer RS coding. Assuming that a packet with L-byte payload is

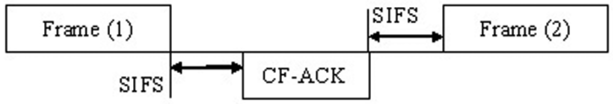


Fig. 4. Successful downlink frame transmission and associated timing.

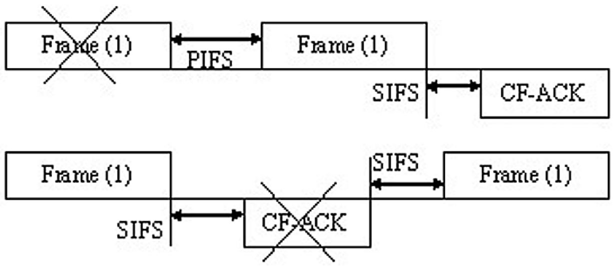


Fig. 5. Retransmission due to frame or CF-ACK transmission error.

transmitted using PHY mode m, the probability of a successful transmission is given by (see Fig. 4)

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

where $P_{e,\text{ack}}$ is the CF-ACK packet error probability, and P_e data is the data packet error probability. These can be calculated from the corresponding packet sizes (including the headers and the payload) and the BER. The average transmission duration for a good cycle T^m_{good}, where neither the data packet nor the CF-ACK packet is in error, can be obtained from the timing intervals given in Fig. 4. Similarly, the average transmission duration for a bad cycle T^m_{bad}, in a cycle where either the data packet or the CF-ACK packet is in error can be computed from the timing intervals given in Fig. 5. The average transmission duration for a packet with an L-byte payload, given that the transmission is successful with the retransmission limit of R, can be obtained as follows:

$$D^m_{\text{av,succ}}(L,R) = \sum_{i=0}^{R} \frac{P^m_{\text{succ}}(i\,|\,L)}{P^m_{\text{succ}}(L,R)}\left[iT^m_{\text{bad}}(L) + T^m_{\text{good}}(L)\right] \tag{3}$$

where the probability that the packet with L-byte data payload is successfully transmitted after i retransmissions under PHY mode m is given by

$$P^m_{\text{succ}}(i\,|\,L) = \left[1 - P^m_{\text{good\_cycle}}(L)\right]^i P^m_{\text{good\_cycle}}(L) \tag{4}$$

and the probability that the packet with L-byte data payload is successfully transmitted within the R retransmission limit under PHY mode m is given by

$$P^m_{\text{succ}}(L,R) = 1 - \left[1 - P^m_{\text{good\_cycle}}(L)\right]^{R+1}. \tag{5}$$

The average transmission duration for a packet with L-byte payload, given that the transmission is not successful with the retransmission limit R, can be obtained as follows:

$$D^m_{\text{av,unsucc}}(L,R) = (R+1)T^m_{\text{bad}}(L). \tag{6}$$

Now, the average transmission duration for a packet with L-byte payload and with a retransmit limit of R is given by

$$\begin{aligned} D^m_{\text{av}}(L,R) = {}& D^m_{\text{av,succ}}(L,R)P^m_{\text{succ}}(L,R) \\ &+ D^m_{\text{av,unsucc}}(L,R)\left(1 - P^m_{\text{succ}}(L,R)\right). \end{aligned} \tag{7}$$

B. Throughput Efficiency With Application Layer RS Code

The throughput efficiency of 802.11a with the use of the (N,K) RS erasure code at the application layer is computed based on the average frame transmission duration obtained earlier. The RS decoder can correct up to $N-K$ packet erasures. If there are more than $N-K$ packet erasures, then this results in a decoding failure. Therefore, the probability of error after RS decoding is given by

$$P^m_{\text{RS}} = 1 - \sum_{i=0}^{N-K} \binom{N}{i} \left(P^m_r\right)^i \left(1 - P^m_r\right)^{N-i} \tag{8}$$

where the resulting residual error probability P^m_r of the data frame after R retransmissions is given by

$$P^M_r = 1 - P_{\text{succ}}(L,R). \tag{9}$$

When a decoding failure happens, there are $N-i$ $(<K)$ correctly-received packets including both video and parity packets possibly. We utilize these video packets if there is any for the video decoding; on average, $(K/N)(N-i)$ packets out of $N-i$ correctly-received packets should be video packets. Therefore, the throughput efficiency, taking into the account the application-layer RS coding and the header overheads of the higher layer protocols, as shown in the equation, at the bottom of the page, where $D^m_{\text{av}}(\cdot)$ is from (7). $DR(m)$ is defined in Section II-B. The numerator here corresponds to the average number of useful video data bits that are received and the denominator corresponds to the total average number of bits that could have been transmitted over the medium in the time required to send those useful data bits successfully.

C. Delay Analysis

The overall delay in the system is a crucial factor for AV applications. In our system, the total delay comprises different components: the delay due to buffering for RS coding at the transmitter, the RS encoding delay, the delay incurred in the transmission and the retransmission of packets, the buffering delay at the receiver for RS decoding, and the RS decoding delay. It can be seen from Fig. 2 that there is no delay at the transmitter due to buffering. As each video packet is stored in

$$E^m_{\text{RS}}(L_a,R,N,K) = 8L_a \frac{\left(K(1-P^m_{\text{RS}}) + \sum_{i=N-K+1}^{N}(N-i)\frac{K}{N}\binom{N}{i}\left(P^m_r\right)^i\left(1-P^m_r\right)^{N-i}\right)}{ND^m_{\text{av}}(L_a+O,R)DR(m)} \tag{10}$$

the interleaver, it can be transmitted simultaneously, since the RS coding is applied across the video packets and the data transmission is along (not across) the video packets. We assume that the process delay due to RS encoding or decoding is small and can be neglected compared with the transmission delay. In certain applications, such as the transmission of a video stored in a residential media server, it may be even possible to perform the RS encoding before transmission. The maximum transmission delay depends on the length of a packet, the maximum number of retransmissions and the specific 802.11a PHY mode m that is used. The worst-case delay for R retransmissions is given by[1]

$$D_{\max}(m, L) = (R+1)\left(T^m_{\text{data}}(L) + T^m_{\text{ack}} + 2a\,\text{SIFSTime}\right). \tag{11}$$

If there are no packet erasures, then there is no buffering delay at the decoder. Each video packet can be delivered to the video decoder as soon it is received. In the presence of erasures out of the first k packets, the receiver needs to buffer up to n packets that belong to the current RS block before performing the erasure decoding. Therefore, the maximum buffering delay at the decoder is $N \cdot D_{\max}(m, L)$. Assuming a frame size of 2000 bytes, $R = 8$, and PHY mode 5, the value of $D_{\max}$ is 6.876 ms. Therefore, the total transmission delay for an RS code with $N = 63$ is 433.19 ms. This delay, on top of the video encoding and decoding delay, is well within the acceptable range for noninteractive video streaming applications. Note that for the most noninteractive video streaming applications, the acceptable delay ranges between 1–10 s.

IV. MPEG-4 FGS Video Coding for Wireless Transmission

A. FGS R-D Modeling

As mentioned in the introduction, we employed MPEG-4 FGS for the compression of the video data, because it can provide easy adaptation to bandwidth variations and device characteristics. The FGS framework consists of a nonscalable MPEG-4 compliant base layer using motion-compensation and a fine-granular (progressive) intracoded enhancement layer [7]. Under this framework, the scalable video content can be compressed in either real-time or off-line for later on-demand viewing. The base layer is coded at a minimally acceptable quality of video using a bit rate of B_{BL}. The enhancement layer improves upon the base-layer video, fully utilizing the available effective bandwidth B_{tot} for the video payload at transmission time. Since there is no motion compensation in the enhancement layer, we can model the overall video quality (PSNR) at the transmitter as a linear function of the transmission bit rate

$$Q_t = \theta \cdot (B_{\text{tot}} - B_{\text{BL}}) + Q_0 = \theta B_{\text{EL}} + Q_0 \tag{12}$$

where Q_t is the overall FGS video quality (PSNR) at R_{tot}, Q_0 is the video quality (PSNR) of the base layer, and θ is the parameter of the R-D model, which depends on the spatio-temporal characteristics of the video sequence (see [7]). B_{EL} is the bit-rate of the enhancement layer. The video quality at the receiver Q_d, equivalently, can be written as follows:

$$Q_d = \theta B_{\text{EL\_}d} + Q_0 \tag{13}$$

where $B_{\text{EL\_}d}$ is the effective received rate of the enhancement layer at the receiver after the channel losses.

B. Fine Grained Loss Protection Strategies

The MPEG-4 nonscalable video data is divided into individually decodable packets using the definition of video packets. One or more video packets are encapsulated in RTP packets as defined in RFC 3016 [12]. (The video packets are not fragmented across several RTP packets.) The video packet size of the base layer is predetermined at packetization time and cannot be adjusted on-the-fly at transmission time based on the channel conditions. However, as will be shown in Section V, adapting the packet size at transmission time depending on the channel conditions leads to improved video quality performance, and this lends very well to the use of the FGS enhancement layer, where the sizes of the various enhancement-layer packets can be adjusted on-the-fly to obtain the best throughput versus resilience tradeoffs. In [6], the concept of FGLP has been introduced, where the enhancement layer is partitioned into an arbitrary number of n embedded sublayers that are differentially protected. Therefore, the total number of layers equals $n+1$ (one base plus the enhancement sublayers). Each sublayer k has an effective packet-loss ratio EP_k that is a function of the channel conditions and the error protection strategy adopted for that layer. Note that to achieve the proposed FGLP strategy, the application-layer RS codes are applied independently to each (priority) layer.

C. Concealment Strategy

Since the impact of packet-loss events on video quality is greatly influenced by the level of resynchronization supported by the video stream and the corresponding decoder, we provided three levels of resynchronization and concealment for the FGS base-layer and nonscalable streams at the decoder side. In addition to the sequence header (i.e., sequence-layer resilience), protection at the following three levels is provided.

- GOP (group of pictures)—the propagation effect of a packet-loss event is stopped by the periodically intracoded pictures in the stream.
- VOP (video object plane)—by allowing packets to contain data from only one VOP, the synchronization is always regained at the next VOP. (VOP is an MPEG-4 video picture.)
- Video packets—when the synchronization is lost within a VOP, the data between the synchronization point prior to the error and next marker is discarded. The data between two resynchronization markers is called "video-packet."

For concealment within the nonscalable streams, the lost image-area of the affected video packet is copied from the previous VOP. The VOP reconstructed in this manner is then used as a reference for the subsequent pictures in the GOP. For FGS, a packet loss within an enhancement-layer frame causes the remainder packets associated with that frame useless.

[1]Note that the worst-case delay happens when all of the first R transmissions fail due to the CF-ACK transmission failures, not the data frame transmissions failures.

D. Cross-Layer Bit Allocation Strategy

Based on this very simple error concealment method, we can determine statistically the bit rate of the video data that is received without errors. Since the wireless channel is time-varying, the effective video (i.e., source coding) bit rate correctly received by the decoder can be represented as a random variable, whose average value B_{avl} can be calculated by

$$B_{\mathrm{avl}} = B_{\mathrm{BL\_}d} + B_{\mathrm{EL\_}d}. \tag{14}$$

The effective base layer rate at the decoder is given by

$$B_{\mathrm{BL\_}d} = (1 - \mathrm{EP}_0)B_{\mathrm{BL}} \tag{15}$$

where EP_0 is the packet loss rate in base layer obtained from (8), B_{BL} is the encoding base-layer bit rate. For the enhancement layer, a single packet loss within an enhancement-layer frame causes the remainder packets associated with that frame useless. Assuming N_t enhancement-layer packets are sent for the current frame, the effective rate of the enhancement layer at the decoder given a packet loss ratio of EP_1 (assuming equal error protection among enhancement sublayers) is

$$B_{\mathrm{EL\_}d} = \left[\sum_{m=1}^{N_t}(1-\mathrm{EP}_1)^{m-1}\mathrm{EP}_1(m-1) + (1-\mathrm{EP}_1)^{N_t}N_t\right]FL_{\mathrm{EL}} \tag{16}$$

where L_{EL} is the enhancement-layer packet length and F is the coded sequence frame rate. N_t can be computed based on the effective FGS enhancement-layer bit rate (i.e., excluding the bit rate spent for its protection) B_{EL}

$$N_t = \frac{B_{\mathrm{EL}}}{FL_{\mathrm{EL}}}. \tag{17}$$

The B_{EL} can be computed at the server side from the total bit rate available at transmission time over the channel B_C, the effective base-layer bit rate B_{BL} and the throughput efficiencies (see (10))

$$B_C = \frac{B_{\mathrm{BL}}}{E_{\mathrm{RS}}^m(L_{\mathrm{BL}}, R_{\mathrm{BL}}, N_{\mathrm{BL}}, K_{\mathrm{BL}})} + \frac{B_{\mathrm{EL}}}{E_{\mathrm{RS}}^m(L_{\mathrm{EL}}, R_{\mathrm{EL}}, N_{\mathrm{EL}}, K_{\mathrm{EL}})} \tag{18}$$

where L_{BL} is the packet size, R_{BL} is the retransmission limit, $(N_{\mathrm{BL}}, K_{\mathrm{BL}})$ is the RS code used for the base-layer packets. The entities with the EL subscript correspond to the enhancement layer packet parameters. Hence

$$B_{\mathrm{EL}} = \left(B_C - \frac{B_{\mathrm{BL}}}{E_{\mathrm{RS}}^m(L_{\mathrm{BL}}, R_{\mathrm{BL}}, N_{\mathrm{BL}}, K_{\mathrm{BL}})}\right) \cdot E_{\mathrm{RS}}^m(L_{\mathrm{EL}}, R_{\mathrm{EL}}, N_{\mathrm{EL}}, K_{\mathrm{EL}}). \tag{19}$$

Based on (19), it can be established that we can trade the source coding bit rate of the enhancement layer (represented by B_{EL}) for the channel coding bit rate (represented by $R_{\mathrm{EL}}, N_{\mathrm{EL}}$, and K_{EL}). Based on the above equations, the effective enhancement-layer rate at the decoder is given by

$$B_{\mathrm{EL\_}d} = \left[\sum_{m=1}^{N_t}(1-\mathrm{EP}_1)^{m-1}\cdot\mathrm{EP}_1(m-1) + (1-\mathrm{EP}_1)^{\frac{B_{\mathrm{EL}}}{F\cdot L_{\mathrm{EL}}}}\cdot\frac{B_{\mathrm{EL}}}{F\cdot L_{\mathrm{EL}}}\right]\cdot F\cdot L_{\mathrm{EL}}. \tag{20}$$

Note that EP_1 is a function of $R_{\mathrm{EL}}, (N_{\mathrm{EL}}, K_{\mathrm{EL}})$ and can be obtained from (8). Similar equations can be derived for different FGLP strategies (i.e., where different error protection strategies are employed in the enhancement-layer). Consequently, the quality of the FGS video coded stream can be computed on-the-fly for various error protection strategies.

As mentioned previously, the employed concealment strategy assumes that if a higher priority packet is lost (i.e., a base-layer packet or a packet containing a more significant enhancement-layer bitplane), then the lower priority packets in that frame are discarded. Consequently, the packet loss rate of the base layer should be kept very small. In [6], the performance of nonscalable MPEG-4 base-layers has been determined for a variety of channel conditions, and it has been determined that for most sequences, if the base-layer packet-loss rate EP_0 is lower than 1%, the overall FGS performance remains unaffected. Hence, given a particular channel condition and the base-layer bitrate B_{BL}, we can determine the error protection strategy to keep EP_0 lower than 1%. Based on that and using (19) and (20), the overall quality at the receiver end Q_d (13) is optimized by maximizing $B_{\mathrm{EL\_}d}$ by appropriate selection of the enhancement layer error protection strategy. Based on the above analysis, the run-time optimal cross layer bit allocation algorithm can be summarized as follows:

- Estimate the channel condition (various methods exist to estimate the channel condition based on the receiver feedback; these are not discussed in this paper)
- For the estimated channel SNR and the chosen PHY mode, calculate the channel BER and the packet loss rates.
- Choose a judicious value of B_{BL} (for details, see the following section). The packet sizes for the base layer L_{BL} is fixed to 2256 bytes. Choose R_{BL} and $(N_{\mathrm{BL}}, K_{\mathrm{BL}})$ to obtain the packet loss rate in base layer EP_0 to lower than 1% (using (8)).
- For different combinations of R_{EL} and $(N_{\mathrm{EL}}, K_{\mathrm{EL}})$ obtain the optimal value of L_{EL} (using (10)). For this optimal value L_{EL}, using (19) and (20) find the combination that yields the maximum value $B_{\mathrm{EL\_}d}$. This combination of application layer FEC, MAC retransmission limit and the packet size will maximize the quality Q_d at the receiver.

V. Experimental Results

In this section, we present the performance of the MPEG-4 FGS coder in IEEE 802.11a WLAN under various adaptive cross-layer protection strategies. We start by highlighting the various choices employed in different protocol layers. At the application-layer, we used a standard-definition resolution video sequence in progressive (i.e., noninterlaced) format coded using a frame rate of 25 Hz that we compressed using the MPEG-4 Momusys FGS software. The base-layer bit rate $B_{\rm BL}$ equals 1 Mb/s and the enhancement layer is adapted based on the channel conditions and applied protections strategies. Furthermore, at the application layer, we considered the use of four different RS codes for the adaptive protection of the base and enhancement layer—(63, 63), (63, 59), (63, 47), (63, 30). The case of (63, 63) actually corresponds to no RS coding. The FEC is applied across packets with an interleaving of $N = 63$ packets.

- At the MAC layer, the retransmission limit is not allowed to be larger than 8 to ensure the short delay necessary for real-time wireless transmission of video.
- The video packets are packetized using RTP, user datagram protocol (UDP), and Internet protocol (IP), having typical header sizes of 12, 8, and 20 bytes, respectively. Between the IP and the 802.11 MAC sublayer, the 802.2 logical link control (LLC) sublayer, along with subnetwork access protocol (SNAP) is typically used with a combined overhead of 8 bytes. Therefore, the total overhead of the higher layer protocols amounts to $O = 48$ bytes. With the maximum size of the frame body of the 802.11 MAC restricted to 2304 bytes and $O = 48$, the maximum value of L_a is restricted to 2256 bytes. Considering packetization overhead, we believe that using a payload size smaller than 16 bytes is meaningless and, hence, the minimum payload size is assumed to be 16 bytes. Accordingly, we consider enhancement-layer packet sizes $L_{\rm EL}$ ranging from 64 to 2256 bytes. The base-layer packet size was fixed and equal to 2256 bytes.
- At the physical layer, we assume that the video data is transmitted using a fixed PHY mode equal to 5. Adaptation of the physical layer modulation mode could also have been included in the overall cross-layer protection strategy, but has been left out for the purpose of simplicity.

Note that due to the random nature of packet losses over the wireless network and their impact on compressed video, for each tested channel condition (SNR) and protection strategy, 50 different runs of the experiments were conducted. The PSNR value computed from each experiment was used to determine the average PSNR plotted in the following results figures of this section. Furthermore, the objective measurements were verified by subjective evaluations performed on an extensive set of sequences containing various motion characteristics and textures.

A. Run-Time Adaptation Procedure

Before presenting the experimental results, we summarize our run-time adaptation procedure based on what we described thus far.

Step 1) Channel estimation using the mechanisms described in Section II-F (based on MAC ARQ, not RTCP).

Step 2) Optimal adaptive packet size selection based on estimated channel SNR for the FGS enhancement-layer packet sizes L_a ranging from 64 to 2256 bytes. The base-layer packet size is fixed at 2256 bytes.

Step 3) Adaptive application-layer FEC with different protections for base and enhancement layers based on channel condition to maximize the PSNR (quality performance) using the analytical model.

Step 4) Adaptive MAC ARQ performance and MAC ARQ with different protections for base/enhancement.

Step 5) Cross-layer adaptive application-layer FEC + MAC ARQ.

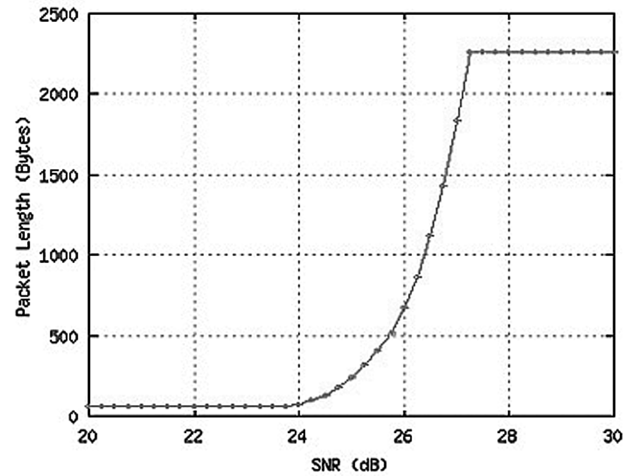


Fig. 6. Optimal packet size selection for different SNRs.

B. Adaptive Packet Size Selection

In this section, we present the results that specify the optimal FGS enhancement-layer packet size that should be selected to maximize the throughput efficiency for a given RS code and retransmission limit. The optimal packet size L_a^* is obtained from (10) as follows:

$$L_a^* = \arg_{L_a} \max E_{\rm RS}^m(L_a, R, N, K) \tag{21}$$

where $E_{\rm RS}^m$ was the throughput efficiency that takes into the account the application-layer RS coding and the header overheads of the higher layer protocols. The above equation can be solved by evaluating the $E_{\rm RS}$ function for all possible values of L_a. In a practical implementation, we can use a lookup table by precomputing the values. Using PHY mode 5, Fig. 6 shows the optimal packet size for different SNRs using a (63, 47) RS code with no retransmissions. As expected, the optimal packet size increases with increasing SNR, with a corresponding increase in efficiency.

Fig. 7 shows the optimal packet sizes for RS code (63, 49) with three different numbers of maximum retransmissions, $R = 0, 1$, and 2. It can be seen that for low SNRs, the optimal packet size is the largest for the case corresponding to $R = 2$. For low SNRs, using the maximum allowed number of link layer retransmissions makes the link more reliable and, hence, allows the use of larger packet sizes. As the SNR improves, the optimal packet size corresponding to the case of $R = 0$ increases rapidly. This is due to the fact that as the underlying link becomes more reliable at higher SNRs, the resultant packet erasures can be handled by the application-layer FEC even in the absence of any retransmissions.

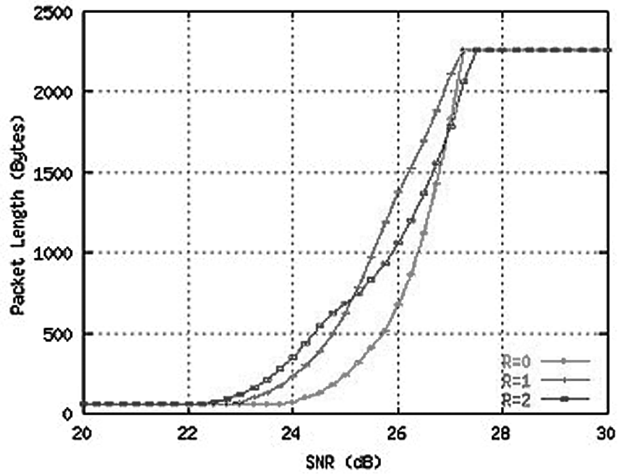


Fig. 7. Comparison of optimal packet sizes for different retransmission limits.

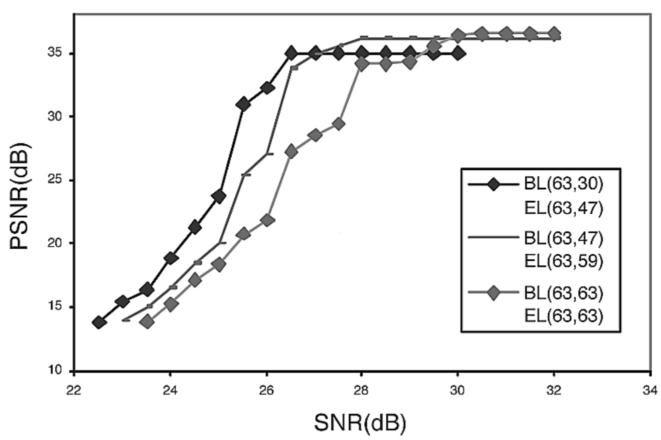


Fig. 8. Decoded picture quality (PSNR) of MPEG-4 FGS using various application-layer FEC protection schemes for different SNRs.

C. Adaptive Application-Layer FEC

Here, we present the results of the adaptive use of different RS codes and different retransmission limits, as well as their combinations based on the SNR. Throughout the following discussion, it is assumed that whenever a particular RS code and retransmission limit are selected, the packet size used corresponds to the value of optimal value obtained in (21). First, we compare the dynamic selection among the (63, 30), (63, 47), (63, 59), and (63, 63) RS codes with a fixed retransmission limit of $R = 0$. We assume that the enhancement sublayers are equally protected. The results are shown in Fig. 8. From the figure, it can be seen that when the channel SNR is under 26 dB, the RS code for the base-layer $(\mathrm{RS}^{\mathrm{BL}})$ of (63, 30) and the RS code of the enhancement-layer $(\mathrm{RS}^{\mathrm{EL}})$ of (63, 47) should be employed for the best visual quality performance. In the range of 26.5 to 30 dB, the use of $\mathrm{RS}^{\mathrm{BL}}$ equal to (63, 47) code and $\mathrm{RS}^{\mathrm{EL}}$ equal to (63, 59) is optimal, while for higher SNRs, no FEC for both the base and enhancement layers gives the best performance. This is expected because when the channel condition is poor, using a low rate RS code is beneficial to counter the packet erasures caused by the unreliable link. However, as the channel SNR improves, the probability of packet erasures goes down and, hence using a high rate code is sufficient to correct the erasures, as well as to reduce the overhead due to the RS coding.

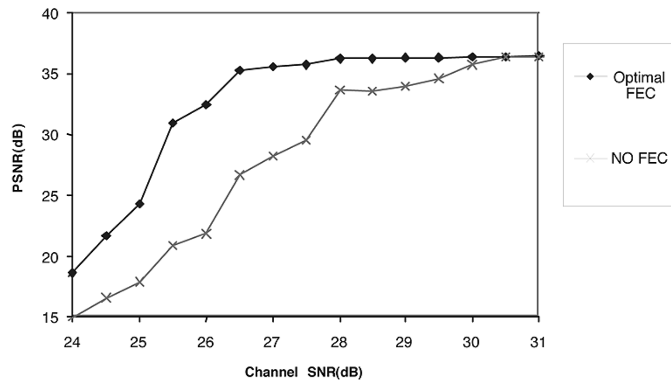


Fig. 9. Decoded picture quality (PSNR) of MPEG-4 FGS for optimal selection FEC and no FEC for different SNRs.

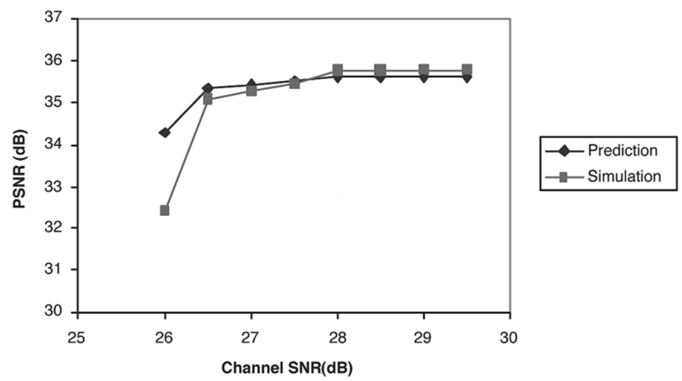


Fig. 10. Modeling and simulation of PSNR performance for various channel conditions.

To further illustrate the utility of the adaptive FEC strategies at the application layer, the best PSNR performance given the optimal selection out of four different RS codes is plotted as a function of the channel characteristics. For comparison, the PSNR performance obtained when no FEC is used is also plotted. As can be seen from Fig. 9, the usage of application-layer FEC is especially important for improving the video quality performance for poor channel conditions. Furthermore, the adaptive application-layer FEC combined with scalable/prioritized coding of the video can ensure graceful degradation across a large range of channel conditions.

The previous result shows the utility of applying adaptive application-layer FEC based on channel conditions. However, to be able to select an optimal FEC on-the-fly, we need to employ the model introduced in Section IV-D to determine the resulting quality for a specific FEC strategy at a particular channel condition. To establish the validity of the model derived in Section IV-D, we computed the sequence PSNR performance under different channel conditions assuming that the base-layer bit rate has been protected using the (63, 30) RS code, and the enhancement layer is equally protected using the (63, 47) RS code. The results using both simulations and the modeling are portrayed in Fig. 10, and show the validity of the model. For channel conditions under $\mathrm{SNR} = 26.5$ dB, the performance of the computed quality using simulations and

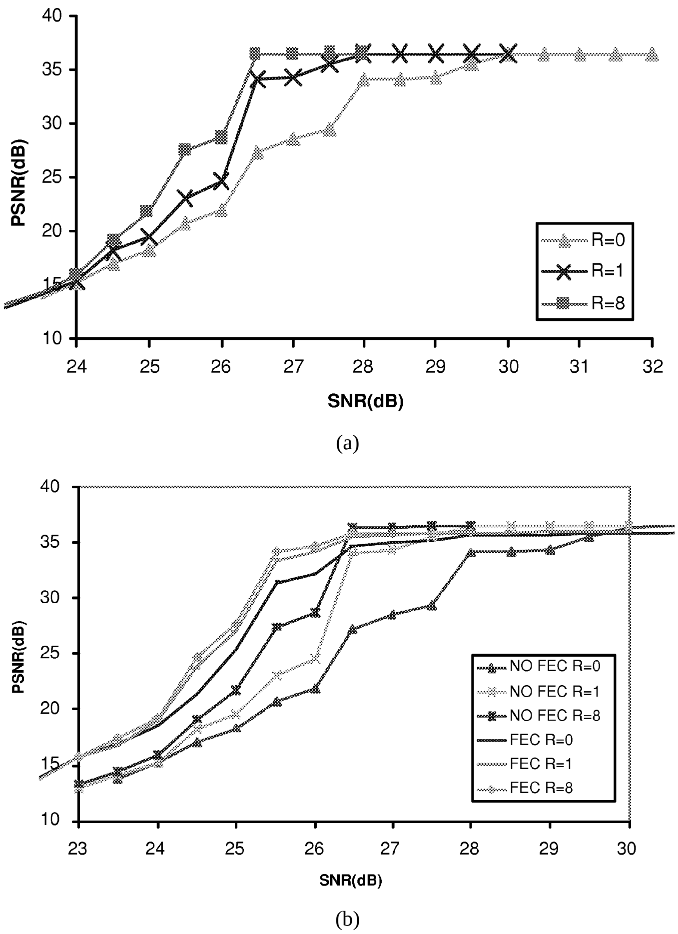


Fig. 11. Decoded picture quality (PSNR) of MPEG-4 FGS using different retry limits (a) without and (b) with application-layer FEC.

predicted quality start to diverge due to the incurred loss in the base layer that exceeds 1%. For base-layer packet losses higher than 1%, the model derived in Section IV-D does not hold. This is because the model was derived under the assumption that the base layer is received without losses and this assumption does no longer hold. The performance of the nonscalable base layer under losses does not decrease linearly like in the FGS enhancement-layer case [6].

Note that the PSNR results presented in the sequel are obtained through simulations over various channel conditions rather than by modeling. The modeling is only used for determining on-the-fly the appropriate RS codes to be used at the application layer based on the instantaneous channel conditions.

D. Adaptive Application-Layer FEC and MAC-Layer ARQ

We now investigate a cross-layer protection strategy combining the application-layer FEC and MAC-layer ARQ. For a better understanding of the interaction between these various protection strategies, the results have been generated with and without application-layer FEC. The retransmission limit is changed adaptively in the range of $R = 0$ to 8 to maximize the video quality performance. Fig. 11(a) shows the PSNR performance for different retransmission limits, assuming that no application-layer FEC is used. From the figure, it can be concluded that MAC-layer ARQ is essential for achieving a graceful degradation over a large range of SNR channel conditions. For low SNRs, the channel BER is large and, hence using the maximum allowed number of retransmissions improves the link reliability. When the channel SNR is higher than 26.5 dB, note that there is only limited performance gain compared with $R = 1$. Note that these results are expected, since an increasing number of retransmissions always results in the better video performance. Unlike in the FEC case that leads to a constant overhead irrespective of loss, ARQ overhead happens only when errors occur. However, it should be noted, that in practical implementations, the MAC retransmission limit should be selected based on the delay constraints of the application.

Fig. 11(b) presents the video performance when both application-layer FEC and MAC ARQ are simultaneously employed. The application-layer FEC utilizes different RS codes: (63, 30) for the base-layer, (63, 47) for the first enhancement-layer queue and (63, 63) for the second enhancement-layer queue. Note that at poor channel conditions (SNR below 26 dB), using application layer FEC and $R = 8$ offers the best performance. At channel SNRs above 26 dB, using no FEC and $R = 8$ offers the best performance. These results can be explained, because when the channel is poor, using only retransmission is not very useful, since every retransmission fails, using a strong FEC can improve the performance. However, as the channel improves, retransmission becomes the best protection strategy, because there is no overhead associated with retransmissions if there are no losses unlike FEC.

VI. Conclusions and Further Research

In this paper, we focus on the robust and efficient transmission of video over WLANs. We specifically address the recent WLAN standard, IEEE 802.11a, which offers high bit rates, enabling the transmission of delay sensitive AV traffic. This paper proposes a novel vertical system integration, referred to as "cross-layer protection," that enables the joint optimization of the various protection strategies existing in the protocol stack. We have specifically concentrated on 802.11a WLANs in the PCF mode. We have presented an analysis of the throughput efficiency for the downlink traffic based on a realistic channel model, 802.11a MAC operation, and various header overheads associated with different layers of protocols. We modeled the end-to-end distortion of MPEG-4 FGS video for various channel conditions using different unequal error protection strategies, and showed that the derived model matches the results obtained by simulations. Based on this model, a strategy for the adaptive selection of application-layer FEC, maximum MAC retransmission limit, and packet sizes depending on the channel condition to maximize the video quality under different multipath channel conditions is developed.

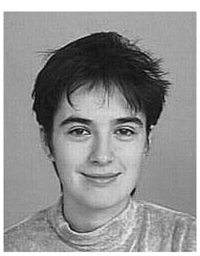
While in this paper a simple set-up was investigated in which one video stream in transmitted to a single receiver, an interesting topic for further research is to consider different network topologies, and the transmission of multiple AV streams to one or multiple receivers.

ACKNOWLEDGMENT

The authors would like to thank the three anonymous reviewers who provided excellent and thorough feedback that improved the quality of this paper.

REFERENCES

[1] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Aug. 1999. IEEE 802.11-1999 Standard.
[2] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications High-Speed Physical Layer in the 5 GHz Band*, Sept. 1999. IEEE 802.11a, Supplement to IEEE 802.11 Standard.
[3] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communications: A review," in *Proc. IEEE*, vol. 86, 1998, pp. 974–997.
[4] H. Ma and M. El Zarki, "MPEG-2 video transmission over wireless access networks using type-I hybrid ARQ schemes for VoD services," presented at the Packet Video Workshop'99, Apr. 1999.
[5] D. Qiao and S. Choi, "Goodput enhancement of IEEE 802.11a wireless LAN via link adaptation," *IEEE ICC'01*, June 2001.
[6] M. van der Schaar and H. Radha, "Unequal packet loss resilience for fine-granular-scalability video," *IEEE Trans. Multimedia*, vol. 3, pp. 381–394, Dec. 2001.
[7] H. Radha, M. van der Schaar, and Y. Chen, "The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP," *IEEE Trans. Multimedia*, vol. 3, pp. 53–68, Mar. 2001.
[8] B. Girod and N. Farber, "Wireless video," in *A Chapter in Compressed Video Over Networks*. New York: Marcel Dekker, 2001.
[9] Y. Shan and A. Zakhor, "Cross layer techniques for adaptive video streaming over wireless networks," *Proc. IEEE Int. Conf. Multimedia and Expo (ICME)*, pp. 277–280, Aug. 2002.
[10] G. Cote, F. Kossentini, and S. Wenger, "Error resilience coding," in *A Chapter in Compressed Video Over Networks*. New York: Marcel Dekker, 2001.
[11] A. Majumdar, D. Sachs, I. Kozintsev, K. Ramchandran, and M. Yeung, "Multicast and unicast real-time video streaming over wireless LANs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 524–534, June 2002.
[12] *RTP Payload Format for the MPEG-4 Visual*. IETF RFC 3016.
[13] Q. Li and M. van der Schaar, "A flexible streaming architecture for efficient scalable coded video transmission over IP networks," presented at the Contribution to the MPEG-Meeting, M8944, Oct. 2002.



Mihaela van der Schaar (M'98) received the M.Sc. and Ph.D. degrees in electrical engineering from Eindhoven University of Technology, Eindhoven, The Netherlands.

In 1996, she joined Philips Research Laboratories, Eindhoven, The Netherlands, where she became a Research Scientist in the TV Systems Department. In 1998, she became a Senior Member of Research Staff in the Wireless Communications and Networking Department, Philips Research, Briarcliff Manor, NY, where she led a team of researchers working on scalable video coding, networking, and streaming algorithms. She is an Assistant Professor in the Department of Electrical and Computer Engineering, University of California, Davis, since July 2003. From January to September 2003, she was also an Adjunct Assistant Professor at Columbia University, NY. Since 1999, she is an active participant to the MPEG-4 standard, contributing to the scalable video coding activities. She was coeditor of the MPEG-4 *Fine Granularity Scalability* standard. She is currently Chairing the MPEG Ad-Hoc Group on scalable video coding and is also Co-chair of the Ad-Hoc Group on multimedia test bed. Her research interests include multimedia coding, processing, networking, and architectures. She has coauthored more than 70 conference papers, journal papers, and book chapters in this field and holds seven patents.

Dr. van der Schaar is a Member of the Technical Committee on multimedia signal processing of the IEEE Signal Processing Society and is an Associate Editor of IEEE TRANSACTIONS ON MULTIMEDIA.

Santhana Krishnamachari received the Ph.D. degree in electrical engineering from University of Maryland, College Park, in 1995.

He is the Director of Technology at EG Technology (EGT), Inc., Atlanta, GA, where he leads the algorithm development and DSP software teams for video compression and transmission systems. Prior to EGT, he was with Philips Research, Briarcliff Manor, NY, as a Project Leader and a Senior Member Research Staff since 1997. At Philips Research, he was actively involved in MPEG-7 standardization activities and lead projects on video streaming over wireless LANs. He has published several papers, book chapters and holds patents in these areas. His research interests are in the areas of video processing, video compression and video transport over IP networks.

Sunghyun Choi (S'96–M'00) received the B.S. (*summa cum laude*) and M.S. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, in 1992 and 1994, respectively, and the Ph.D. in electrical engineering and computer science from the University of Michigan, Ann Arbor, in September 1999.

He is an Assistant Professor at the School of Electrical Engineering, Seoul National University (SNU), Seoul, Korea. Before joining SNU in September 2002, he was with Philips Research, Briarcliff Manor, NY, as a Senior Member Research Staff and a Project Leader for three years. His current research interests are in the area of wireless/mobile networks with emphasis on the QoS guarantee and adaptation, in-home multimedia networks, wireless LAN and PAN, MAC protocols, and mobility management. He coauthored about 40 technical papers and book chapters in the areas of wireless/mobile networks and communications.

Dr. Choi is an active participant and contributor of the IEEE 802.11 WLAN Standardization Committee. He was a recipient of the Korea Foundation for Advanced Studies Scholarship and the Korean Government Overseas Scholarship during 1997–1999 and 1994–1997, respectively.

Xiaofeng Xu (S'00) received the B.S. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1999 and 2001, respectively. He is currently working toward the Ph.D. degree at Polytechnic University, Brooklyn, NY.

In the summer of 2002, he was an Internship Student with Philips Research, Briarcliff Manor, NY. His research interests are in the field of video coding for unreliable networks.