Robust Video Transmission Over Wireless LANs

Pierre Ferré, Angela Doufexi, James Chung-How, Andrew R. Nix, and David R. Bull, *Senior Member, IEEE*

***Abstract*—Home wireless networks are mainly used for data transmission; however, they are now being used in video delivery applications, such as video on demand or wireless internet protocol (IP) television. Off-the-shelf technologies are inappropriate for the delivery of real-time video. In this paper, a packetization method is presented for robust H.264 video transmission over the IEEE 802.11 wireless local area network (WLAN) configured as a wireless home network. To overcome the poor throughput efficiency of the IEEE 802.11 Medium Access Control (MAC), an aggregation scheme with a recovery mechanism is deployed and evaluated via simulation. The scheme maps several IP packets (each containing a single H.264 video packet called a Network Abstraction Layer (NAL) unit) into a single larger MAC frame. Video robustness is enhanced by using small NAL units and by retrieving possible error-free IP packets from the received MAC frame. The required modifications to the legacy MAC are described. Results in terms of throughput efficiency and peak-signal-to-noise ratio (PSNR) are presented for the case of broadcast and real-time transmission applications. Compared to the legacy case, an 80% improvement in throughput efficiency is achieved for a similar PSNR video performance. For fixed physical layer resources, our system provides a 2.5-dB gain in video performance over the legacy case for a similar throughput efficiency. The proposed solution provides considerable robustness enhancement for video transmission over IEEE 802.11-based WLANs.**

***Index Terms*—Medium Access Control (MAC), packetization, peak-signal-to-noise ratio (PSNR), video transmission, wireless local area networks (WLANs).**

I. Introduction

Home entertainment, video on demand, and other multimedia communication products are now receiving considerable interest. New local wireless networks, such as the IEEE 802.11 a/g [1], [2], are emerging as possible solutions. Because of the high bit rates provided at the physical layer (PHY) (up to 54 Mb/s), video transport over wireless local area networks (WLANs) has become a reality. Applications of home video systems include Internet Protocol (IP) television (TV), video on demand, and compressed high-definition TV (HDTV) redistribution. For real-time and interactive applications, these home video systems are likely to be deployed with a limited number of users within a household, including a single TV with a set-top-box and several computers and/or media adapters, including personal digital assistants (PDAs), each accessing a video server connected to the core network via a wireless Access Point (AP). The architecture chosen by providers to deliver interactive time-bounded wireless video services around a home may consist of a low-cost IEEE 802.11 AP with a best-effort delivery system based on a User Datagram Protocol (UDP)/IP structure. This paper aims to provide good visual video quality for real-time interactive video delivery. This includes time-bounded delivery of video, improved visual quality, and an efficient means of accessing the medium, compared to existing standard solutions. This paper is organized as follows. Section II provides the necessary background information. Section III describes the framework and the weaknesses of typical video transmission over the legacy IEEE 802.11 Medium Access Control (MAC). Section IV describes the proposed packetization strategy, including the required modifications to the current IEEE 802.11 MAC. Section V provides statistical results, analysis, and a comparison with the legacy MAC. Finally, Section VI concludes this paper.

Manuscript received July 5, 2004; revised October 20, 2005, February 7, 2006, August 8, 2006, January 9, 2007, July 24, 2007, and August 23, 2007. The review of this paper was coordinated by Prof. D. O. Wu.

P. Ferré, A. Doufexi, A. R. Nix, and D. R. Bull are with the Centre for Communications Research, Department of Electrical and Electronic Engineering, University of Bristol, BS8 1UB Bristol, U.K. (e-mail: pierre.ferre@bristol.ac.uk).

J. Chung-How is with ProVision Communications Technologies Ltd., BS4 4EU Bristol, U.K.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2007.909258

II. Background

A. IEEE 802.11 a/g PHY and MAC

The physical layer technology used in this paper is coded orthogonal frequency division multiplexing (COFDM), conforming to the IEEE 802.11a and 11g PHY standards at 5 GHz [1] and 2.4 GHz [2], respectively. A detailed PHY layer description can be found in [3]–[5]. IEEE 802.11 a/g operates in one of eight different modes (or link speeds). Each mode offers different combinations of coding rate, modulation scheme, and nominal bit rate (from 6 to 54 Mb/s). Moreover, a large packet is more likely to result in a packet error.

The IEEE 802.11 MAC layer [6] provides shared access to a wireless channel, and the main (and mandatory) operating mode uses the distributed coordination function, which is based on carrier-sense multiple access with collision avoidance (CSMA/CA), to gain access to the medium. In-depth studies of the IEEE 802.11 MAC can be found in [7]–[9]. This MAC relies on a *Stop and Wait* ARQ retransmission scheme. Transmitted frames must be acknowledged, and if the transmitter does not receive an ACK within a short interframe spacing (*SIFS*), the frame is rescheduled. If the MAC frame has not been acknowledged after a given number of retransmissions, it is simply dropped. The interframe spacing (IFS) timings and the time duration to transmit the acknowledgment are fixed. Furthermore, an exponential random backoff is implemented prior to each transmission. The throughput efficiency of the MAC is analytically given by (derived from [7]–[9])

$$\text{Throughput}_{\text{eff}} = \frac{T_{\text{Data}}}{T_{\text{Success}}} = \frac{T_{\text{Data}}}{\text{DIFS}+\text{Backoff}+T_{\text{Data}}+\text{SIFS}+\text{ACK}} \tag{1}$$

with

$$T_{\text{Data}} = T_{\text{PLCP\_Preamble}} + T_{\text{PLCP\_Header}} + \left\lceil \frac{N_{\text{MAC Header}} + N_{\text{Data}} + N_{\text{FCS}}}{N_{\text{DBPS}}} \right\rceil \times T. \tag{2}$$

T_{Success} is the total duration of a successful transmission, including contention and transmission. T_{Data} is the duration of data transmission on the medium, $T_{\text{PLCP\_Preamble}}$ and $T_{\text{PLCP\_Header}}$ are the durations of the physical packet preamble and header, respectively, N_{Data} is the MAC frame payload size in bits, $N_{\text{MAC header}}$ is the MAC header size, N_{FCS} is the frame check sum (FCS) size, N_{DBPS} is the number of data bits per OFDM symbol, and T_s is the OFDM symbol duration. "Backoff" is the expected time duration of the backoff mechanism. "DIFS" and "SIFS" represent the time durations of the distributed and short IFS, respectively.

It has been established that the throughput efficiency of the IEEE 802.11 MAC is sensitive to the MAC frame length [3], [7], [8]. Due to larger timing overheads, small frame lengths result in poor channel utilization, whereas larger frame lengths offer much better channel usage. Hence, for a given operating mode, since they provide a higher throughput, larger frames are preferable for high bit rate applications. A 32-bit cyclical redundancy check (CRC), calculated over the MAC header and payload, is appended at the tail of the *frame body* and is used for error detection. The IEEE 802.11 task group is currently proposing new MAC and PHY solutions to support higher bit rates (up to 600 Mb/s) for the coming IEEE 802.11n standard [10], [11]. The MAC is expected to follow the basis of the IEEE 802.11e MAC [12] with a *Block_ACK*-like scheme and the addition of an aggregation mechanism.

B. H.264 Video Standard

The new H.264/Advanced Video Coding standard [13] has recently been standardized. The main goal is to enhance the coding efficiency and to provide a "network-friendly" representation of the encoded video. The *Network Abstraction Layer* (NAL) is the interface between the Video Coding Layer (VCL) and the underlying layers [14]. It gives support for a packet-based approach with a real-time transport protocol (RTP)-based extension [15], [16] and for a byte-stream approach with the Annex B extension. In-depth studies of H.264 can be found in [17]–[20].

The output of the VCL is an NAL unit containing a single video slice. One slice consists of a fixed or variable number of encoded macroblocks. The slice structure is used for error resilience purposes. At the decoder, the concealment, or recovery, of a lost NAL unit (slice) is made easier if the NAL unit is small. Moreover, a smaller NAL unit contains less information, and its loss is less damaging. However, this results in a larger slice overhead. Slices provide points of temporal and spatial resynchronization, where the decoder can resume whenever the received video stream is corrupted. In adverse channel conditions, i.e., low received power, small NAL units are preferable [21], [22]. The *peak-signal-to-noise ratio* (PSNR) metric is used to measure the quality of a received video sequence.

C. Related Work

For video robustness purposes, large-sized NAL units are not desirable. However, a small NAL unit encapsulates into small MAC frames, and for MAC efficiency reasons, this is not desirable. To overcome this poor throughput performance, mechanisms have been designed using various techniques. In [23] and [24], the authors describe an optional No_ACK policy to enhance the IEEE 802.11e MAC. At the expense of greater degradation at high error rates, MAC frames are not acknowledged. In [25], the problem is tackled by proposing a simple concatenation mechanism where multiple MAC frames are concatenated and transmitted as a single (but longer) frame. Overheads are reduced, and the MAC throughput is improved. However, no recovery system is proposed, and whenever a longer frame is lost, the multiple concatenated frames are lost. In [26] and [27], the authors refer to MAC-level improvements to be added to a future version of the IEEE 802.11e MAC, where one IP packet is fragmented into several blocks and then aggregated into a single MAC frame. Each of the IP fragments is forward error correction (FEC) protected. The receiver keeps a copy of the correctly received IP fragments in the current frame after FEC decoding. The MAC frame is acknowledged by the receiver only if all the IP fragments of the MAC frame are reconstructed. If no acknowledgment is received, the transmitter reschedules the whole MAC frame, and the receiver combines the stored blocks with the error-free blocks retrieved from the retransmitted frame. This mechanism, therefore, reduces the number of transmissions. However, this scheme only considers fragmented IP packets, and no mechanism for partial retransmission is suggested.

The concept of aggregated frames is a core element in the enhanced IEEE 802.11n proposal for higher throughputs [11], [28]. Two types of aggregation are proposed 1) *A-MSDU*, where multiple IP packets are aggregated into a single MAC frame, and 2) *A-MPDU*, where multiple MAC frames are aggregated into a single PHY packet. No mechanism has yet been specified to retrieve error-free fragments, if any, in a corrupted MAC frame. The RTP format for H.264 [15] was designed so that the H.264-encoded output can be sent over packet-oriented networks. This paper describes several packetization schemes for encapsulating NAL units into RTP packets [19]. The simplest is the mapping of one NAL unit to one RTP packet, where the RTP payload is the NAL unit, and no special provisions are made for the carriage of H.264. The RTP format for H.264 offers an aggregation mechanism that allows several NAL units to be concatenated into larger RTP packets. However, at lower levels, the MAC discards any corrupted frames, and the content of the aggregated RTP packets is lost. Hence, the aggregation of RTP packets is not suitable.

The above techniques overcome the poor efficiency of the IEEE 802.11 MAC by using aggregation/concatenation. However, they do not allow the recovery of error-free segments, and they do not limit the retransmission process to errored segments (i.e., partial retransmission). More importantly, these schemes are not specially designed for video transmission. As far as the authors are aware, there is no development that incorporates both video and MAC issues to enhance the MAC throughput and aid the video error resilience. In this paper, we propose an aggregation scheme where multiple IP packets are aggregated into a single MAC frame and where error-free IP packets in the MAC frame are recovered. Our proposed scheme allows the video application to support and tolerate poor PHY layer performance. Our scheme 1) provides a good MAC throughput efficiency by ensuring good channel utilization (through the use of large MAC frames) with the aggregation of several small NAL units into a single larger MAC frame and 2) enhances the video quality by maintaining video robustness with the use of small NAL units and by retrieving possible error-free NAL units from an aggregated MAC frame.

III. Video Transmission With the Legacy IEEE 802.11 MAC

A. Scenarios

The IEEE 802.11a/g has potential application in a number of scenarios, ranging from compressed HDTV redistribution in the home to high-bit-rate (and low delay) outdoor wireless cameras. More traditionally, this technology provides a data link for PDAs and laptops. In the latter case, the bit rates per user are often low (< 1 Mb/s). Importantly, the IEEE 802.11a/g is likely to be used for real-time live multimedia distribution in a home network. In this case, a limited number of retransmissions at the MAC layer is appropriate. The relatively small number of contending users limits the probability of collision, and separate APs (i.e., with dedicated frequencies) can be allocated for video and data services to further reduce this risk. In such cases, packet error rates due to the channel condition dominate over packet errors due to multiple access collision. To illustrate these scenarios, we consider RTP/UDP/IP links with no MAC layer ARQ (Broadcast/Multicast) or with a limited number of MAC layer ARQ retries (Unicast). Moreover, the simple one-to-one RTP packetization case is applied.

B. Transmission Procedure

A simple data encapsulation scheme from the application layer to the MAC layer is applied, and the MAC frame then contends for

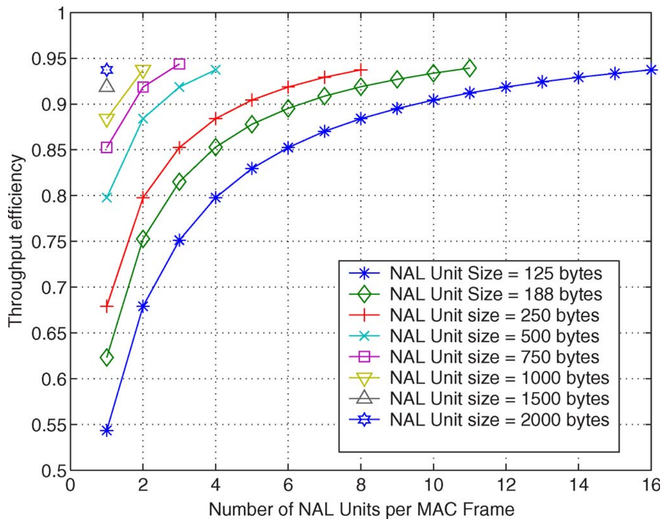

Fig. 1. Throughput efficiency. BPSK 1/2 rate. Different numbers of NAL units per MAC frame.

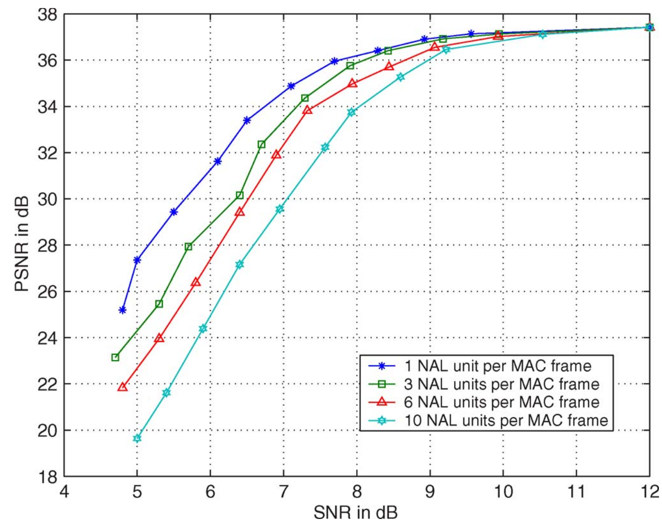

Fig. 2. PSNR with the IEEE 802.11 legacy MAC with one ARQ allowed.

transmission over the wireless channel. The reception scenario of the legacy MAC with a UDP link is as follows. The PHY layer decodes the received PHY packet, and the MAC checks the FCS field to detect errors. If the FCS is correct, the MAC sends back an ACK, and the MAC frame is deencapsulated and passed through the IP, UDP, RTP, NAL, and application layers. If the FCS is not correct, the MAC does not send an ACK. After one SIFS interval, the transmitter checks the retransmission status of the frame (MAC ARQ algorithm). If the number of retransmissions is less than or equal to the maximum retry count, the frame is rescheduled. If the number of retransmissions is greater than the maximum retry count, the frame is dropped, and the transmitter proceeds to the next frame.

The weaknesses of video transmission over the legacy IEEE 802.11 system results from the poor MAC throughput efficiency and a lack of video robustness. That is, 1) the MAC throughput and, therefore, the system throughput, greatly depend on the MAC frame length. At the MAC layer, large MAC frames are preferable. However, one large MAC frame corresponds to a large NAL unit at the application layer, and for error resilience reasons, large NAL units are not desirable and 2) from a video application point of view, small NAL units provide more robustness to the inherent packet losses of a wireless home network. However, these small NAL units lead to a very poor throughput at the MAC layer.

IV. Proposal for Improved Video Transmission

A. Motivation

Our proposal is based on the fact that a MAC frame can carry more than one IP packet, i.e., one NAL unit [21]. Due to our one-to-one mapping of NAL units to IP packets, mapping IP packets into one MAC frame and mapping NAL units into one MAC frame are conceptually equivalent. To illustrate the benefits in terms of the throughput of multiple NAL units (IP packets) transmitted over a single MAC frame, the throughput efficiency of the MAC for different numbers of NAL units per MAC frame is shown in Fig. 1 for the BPSK 1/2 rate mode of IEEE 802.11g and the legacy MAC (including RTP/UDP/IP overheads). When only one small NAL unit (125 B) is carried over a MAC frame, a very low throughput is observed. However, the throughput efficiency increases as the number of NAL units per MAC frame increases (because of the larger MAC frame): A 50% increase in throughput efficiency is achieved by mapping four NAL units, rather than a single NAL unit, per MAC frame.

It should be noted that the way the NAL units are handled at the MAC layer determines the overall quality of the received video sequence. More specifically, Fig. 2 depicts the effect on the video quality (PSNR) of varying the number of NAL units aggregated into a single MAC frame (using the legacy MAC) for different received signal-to-noise ratios (SNRs). In this figure, corrupted MAC frames are discarded after a single MAC ARQ to achieve low latency. As the number of mapped NAL units increases, the video quality reduces (decreasing PSNR). This occurs since, when a MAC frame is lost, all the content is discarded, i.e., all the NAL units. This illustrates the need for an error-free packet recovery mechanism if multiple NAL units are transmitted using a single MAC frame. The case of packet aggregation without error-free packet recovery is very similar to the mechanism developed in the upcoming IEEE 802.11n. In addition, as the frame length increases, a frame is more likely to be corrupted at the PHY layer. For a given NAL unit length, a MAC frame containing ten NAL units is more prone to error than a MAC frame containing three NAL units. Moreover, in a wireless system, errors are bursty. Error-free and corrupted segments could alternate in a PHY packet, depending on the channel conditions.

B. NAL Unit Aggregation With Packet Recovery

At the transmitter, the MAC layer gathers the incoming NAL unit/RTP/UDP/IP packets and appends a 4-B-long check sum field to their tail, as in [26] and [27]. The MAC layer then aggregates these "check-summed" IP packets together into a single MAC frame, as shown in Fig. 3. The MAC header is modified in a manner similar to [28] by adding an aggregation field, which contains the number of IP packets aggregated and their respective length, all encoded using 2 B. This allows the MAC layer at the receiver to separate and restore the IP packets. The CRC is applied only to the modified MAC header instead of applying it to the entire MAC frame as in the IEEE 802.11 legacy MAC. At the receiver, the MAC layer checks if the CRC is correct. If the MAC header is corrupted, the packet is dropped, and no ACK is transmitted. If the MAC header is correct, the MAC layer separates and restores the different IP packets from the MAC frame using the aggregation field in the modified MAC header and then checks their FCS. The MAC layer, therefore, has knowledge of which IP packets require retransmission and which IP packets can be passed to the upper layers. If the IP packet is corrupted, it is dropped. If the IP packet is

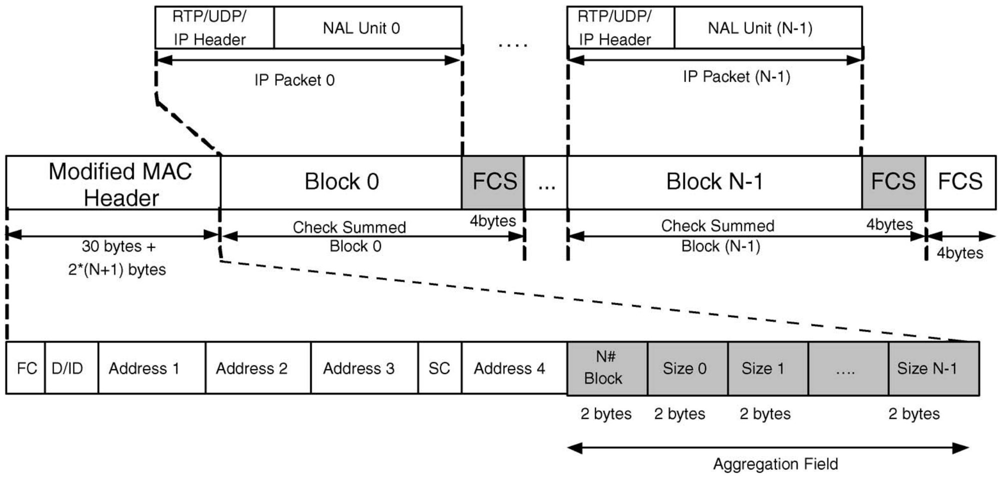

Fig. 3. MAC frame modification for multiple NAL unit MAC packetization.

not corrupted, an ACK frame is generated at the MAC layer, and the IP packet is passed to the upper layers.

This proposal uses an acknowledgment scheme similar to the *Block_ACK* in [12] and [23], combined with a retransmission mechanism for the aggregated IP packets at the MAC frame level. The ACK frame must be modified so that it incorporates a specific bit pattern (or *map*) describing the position of the IP packets to be retransmitted, which, again, is similar to the *Block_ACK* policy. We propose a bit pattern with "0" representing "*No Retransmission Required*" and "1" representing "*Retransmission Required*." The position of the bit corresponds to the position of the IP packet within the MAC frame (starting from 0). A 16-bit pattern allows up to 16 IP packets to be mapped. For example, the following bit pattern 0001 1110 0000 0000 would be used to indicate that IP packets 3, 4, 5, and 6 require retransmission. Note that the *Retry field* bit in the *Frame control* of the MAC does not need to be modified. In the legacy MAC, if set to 1, this bit indicates that the current MAC frame, with its NAL units, has already been transmitted. Originally designed to avoid duplicate MAC frames at the receiver, this remains valid in our proposal, with the *Retry field* bit relating to the whole MAC frame.

The transmitter keeps a copy of all the IP packets mapped into each MAC frame in a buffer. The received ACK contains information on the position of the IP packets that require retransmission. The transmitter then updates its copy buffer by discarding correctly by received IP packets, by keeping corrupted IP packets and by including new IP packets. The transmitter keeps the retransmission IP packet status (retry count) up to date and discards IP packets when the maximum retry count is reached. Fragmentation should be avoided since fragmenting may result in NAL units being split between several MAC frame fragments. This violates our video packetization scheme, which requires that a single NAL unit be sent in one, and only one, MAC frame. Furthermore, for video synchronization purposes, fragmentation is not desirable. These modifications are not standard compliant and add complexity to the MAC. The overheads introduced by the bytes in the modified MAC header, the 2 B in the ACK frame, and the 4-B FCS on each IP packet are small compared to the total MAC frame length.[1] The proposed mechanism allows us to 1) maintain and improve the throughput efficiency by using an aggregation mechanism at the MAC layer. By aggregating several NAL units into one single large MAC frame, a better usage of the channel resource is achieved. Note that possible throughput improvements due to the NAL unit retransmission mechanism within one MAC frame are negligible compared to the improvement due to the aggregation mechanism. It also allows us to 2) enhance the video quality and robustness by recovering error-free NAL units via the NAL unit retransmission mechanism within one MAC frame and by using small NAL units.

[1]Up to $2 + 16 \times 2 = 34$ extra bytes are used for the modified MAC header, and up to $16 \times 4 = 64$ extra bytes are used for the FCS on each IP packet. For example, five IP packets of 250 B each would generate a 7% overhead.

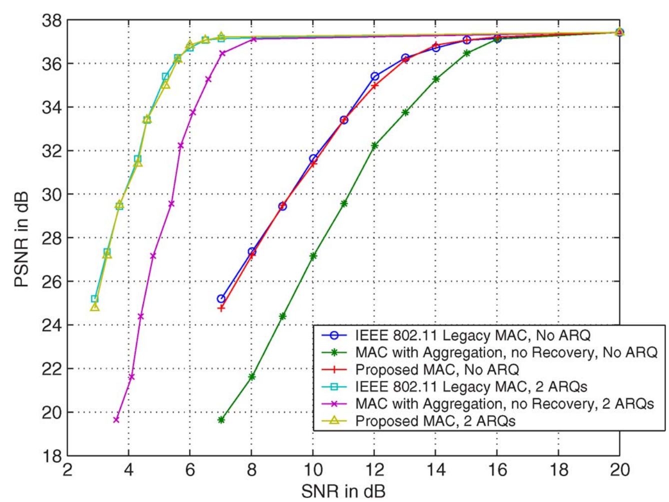

Fig. 4. PSNR comparison of the IEEE 802.11 legacy MAC/proposed MAC.

V. Results and Analysis

A. Simulation Conditions

Our own fully compliant 802.11 a/g PHY layer simulator, which meets the conformance requirements specified in [1] and [2] (Annex A), has been used to recreate accurate bit and packet error patterns [3], [4]. The simulator supports all the standardized operating modes and variable PHY layer packet lengths. Moreover, it implements all the components of the PHY layer with all the parameters configured in alignment to the standard and is capable of producing error patterns at any SNR level.

The channel model conforms to the European Telecommunications Standards Institute Broadband Radio Access Networks channel A specifications (nonline-of-sight office environment) with a

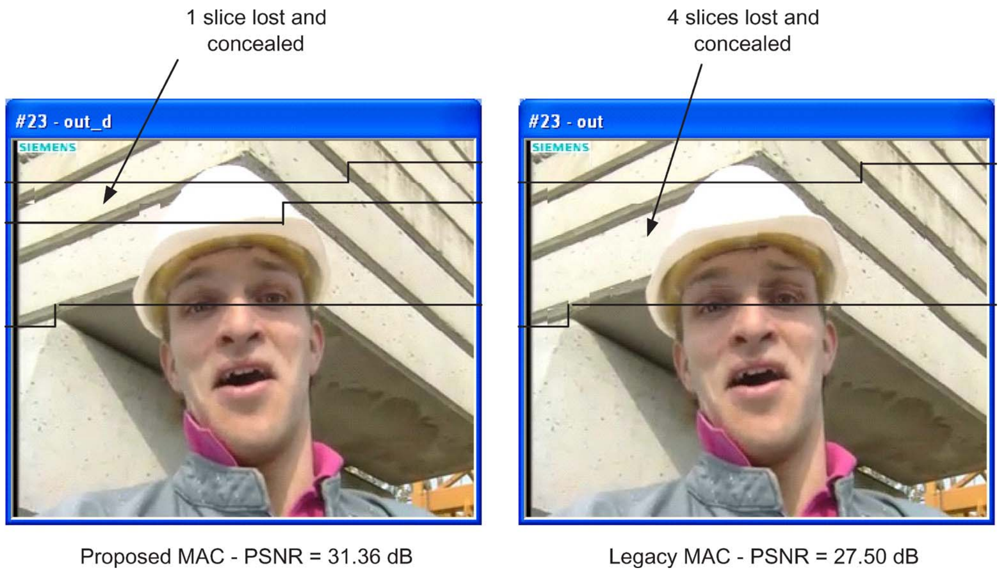

Fig. 5. Four NAL units per MAC frame at PER $= 9.5\ 10^{-3}$, with concealment.

root-mean-square delay spread of 50 ns. The key aim of the simulator is to accurately model the bit error distribution within a received PHY packet. Bit errors are bursty within a PHY packet, and in the case of heavily corrupted packets, the whole content of the PHY packet is lost. However, many PHY packets contain significant error-free segments (particularly those containing only a small number of bit errors).

We have encoded the *foreman* sequence with 300 video frames at CIF resolution (352 $\times$ 288 pixels) and at a frame rate of 30 Hz using the H.264 reference software [29]. The RTP format and a fixed maximum NAL unit size are chosen. Generated slices are encapsulated into UDP/IP packets. Missing slices are concealed using the advanced error concealment algorithm of the reference software [30]. For a range of received SNR levels, the video sequence is sent 100 times (using different initialization points in the error patterns) to create statistical results. The PSNR values (video quality) of the decoded sequences are then averaged. The transmission mode used is BPSK 1/2 rate. The results demonstrate the need for an aggregation scheme at the MAC layer combined with an error-free packet retrieval mechanism.

B. Fixed NAL Unit Size

In the first simulation, the video is preencoded at 550 kb/s. The maximum NAL unit size is fixed and set to 188 B, i.e., to small NAL units that ensure video error resilience. Fig. 4 compares for different SNR levels the proposed MAC with the legacy MAC (which corresponds to one NAL unit mapped to one MAC frame) and with a MAC with aggregation but without error-free packet recovery. The case of ten aggregated NAL units is considered for the case of no ARQ (Broadcast) and for a maximum MAC layer retry limit of two (UDP Unicast). The figure shows that, in terms of PSNR video performance, our system is similar to the IEEE 802.11 MAC legacy case. Our system behaves as if the small NAL units aggregated in one large MAC frame were separately transmitted. The recovery mechanism allows us to preserve the error resilience features of the encoded video, i.e., small NAL units, that the IEEE 802.11 legacy MAC naturally conveys. Moreover, it can be seen that our schemes improve the PSNR by up to 3 dB for a received SNR of 11.5 dB (PHY PER $= 10^{-2}$) compared to when there is no recovery. This is explained in Fig. 5, which shows how one picture frame is affected when one MAC frame containing four NAL units is corrupted both with and without recovery. Without any recovery mechanism, four NAL units (slices) are dropped, whereas in our scheme, three error-free NAL units are retrieved, and only one NAL unit is dropped. The proposed scheme significantly benefits from the error concealment at the decoder, which provides an improvement in PSNR of around 4 dB. The concealment algorithm used to recover the missing slices at the video decoder uses adjacent available video blocks. The concealment when four slices are lost is, therefore, less accurate since it is already using concealed video blocks rather than correctly received ones.

With the use of small MAC frames on the medium, the channel usage of the IEEE 802.11 MAC has a poor throughput efficiency compared to our mechanism. It is shown in Fig. 1 that the legacy MAC offers a throughput efficiency of 52% (one NAL unit of 188 B), whereas our scheme provides a throughput efficiency of 91% (ten NAL units of 188 B each), i.e., an increase of 80%.

Fig. 6 shows the video PSNR for different SNR levels and compares our proposal to the MAC with aggregation but without recovery and with the legacy MAC for no ARQ (Broadcast/Multicast). We can see that for a given received SNR, our scheme allows constant quality, regardless of the number of NAL units mapped into each MAC frame. It also provides a better PSNR for a given SNR than the MAC with aggregation but without recovery. This is explained by the fact that for a given PER, the recovery mechanism allows a reduction of the NAL Unit Error (NER) to the same level as the legacy case and, therefore, allows us to maintain video quality but with improved throughput efficiency.

C. Fixed PHY Packet Size

In this second simulation set, our proposal is compared with the legacy MAC when the channel resources are fixed, with a PHY packet size of 752 B. With relatively large MAC frames, a good throughput efficiency is achieved, i.e., around 85% (see Fig. 1). The *foreman* sequence is encoded at 128 kb/s, and the influence of the number of NAL units per MAC frame is studied. As the number of NAL units per MAC frame increases, the number of slices per picture frame increases, and each slice contains a smaller part of the picture. Consequently, the slice overhead increases (up to 13% with eight NAL

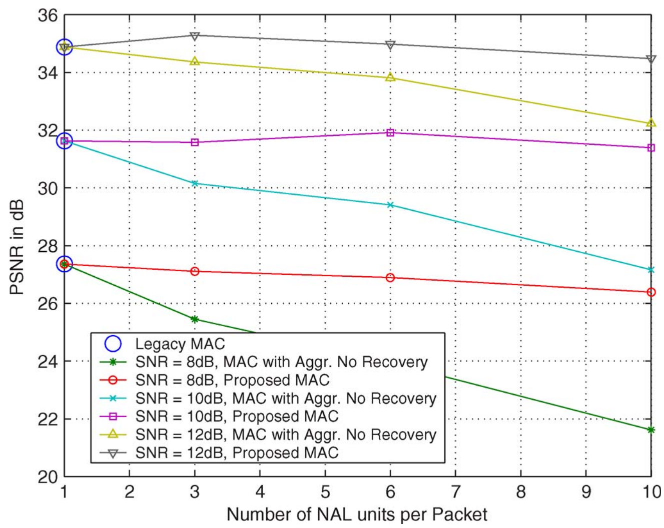

Fig. 6. PSNR comparison of the IEEE 802.11 legacy MAC/proposed MAC without ARQ for different SNR values (fixed NAL size).

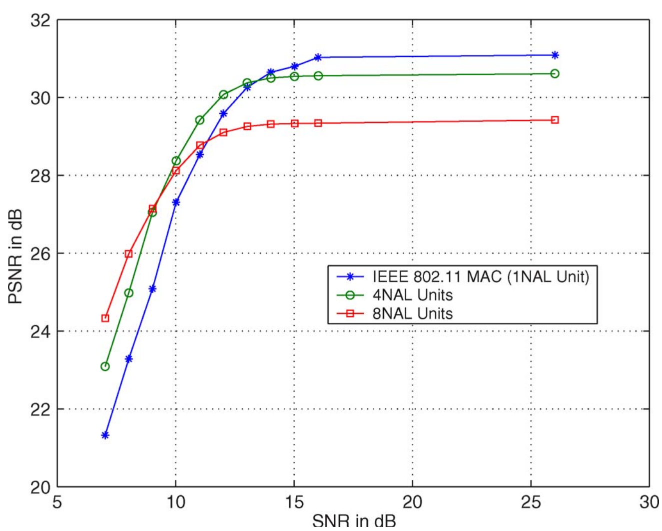

Fig. 7. PSNR comparison versus the SNR for the proposed MAC with no ARQ (*foreman* at 128 kb/s).

units per MAC frame), and the error-free PSNR is lower, since we encode at a fixed video rate. However, the resilience of the video to missing slices is enhanced with higher numbers of NAL units per MAC frame, since when a slice is lost, a smaller part of the video frame is lost, and this is then easier to conceal. This is illustrated in Fig. 7, which shows the PSNR performance with different numbers of NAL units per MAC frame. Up to 2.5-dB gain in PSNR is achievable over the legacy case at low SNRs, i.e., high PER, when eight NAL units are mapped into each MAC frame. This is explained by the fact that smaller NAL units provide more robustness to errors as well as resulting in a smaller NER. However, since there are fewer errors with increasing SNR, a very robust video sequence is not required at high received power levels. Ultimately, in an error-free environment, the legacy MAC offers the best PSNR.

Fig. 8 depicts the PSNR for different SNR levels versus the number of NAL units per MAC frame. The PSNR of the error-free video decreases as more NAL units are mapped into a single MAC frame and is maximized in the legacy case (one NAL per MAC frame), i.e., where the slice overheads are minimal. However, for a particular SNR,

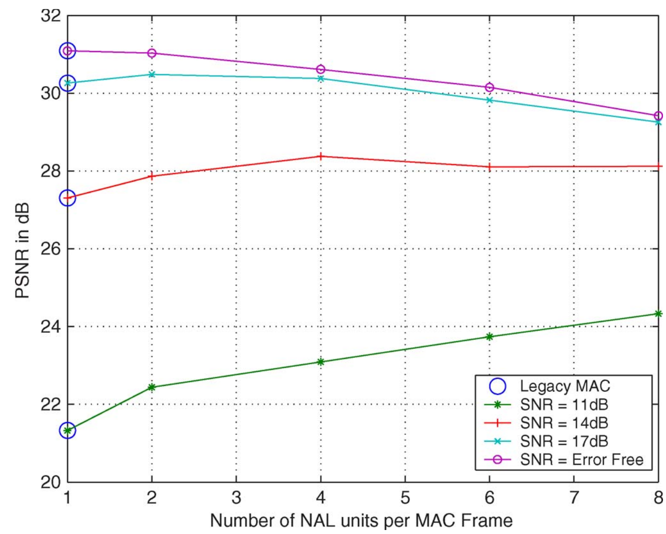

Fig. 8. PSNR comparison versus the number of NAL units for the proposed MAC with no ARQ (*foreman* at 128 kb/s).

i.e., PER, the PSNR is now optimized when there is more than one NAL per MAC frame. With an SNR of 11 dB, the best mapping uses four NAL units per MAC frame. To the right of this optimal point, a greater number of NAL units per MAC frame offers more robustness than required (and lowers video quality). To the left of the optimum point, the video is not robust enough compared to the level of error. The proposed scheme, therefore, improves the video robustness and video quality when compared to the legacy case for equal throughput efficiency performance.

VI. Conclusion

In this paper, we have proposed and characterized one packetization strategy for enhanced H.264 video transmission over an IEEE-802.11-based MAC using a UDP/IP protocol. The proposed scheme provides good throughput for video transmission and enhances video robustness. This is realized via an aggregation scheme and a recovery mechanism at the MAC layer, where several NAL unit/RTP/UDP/IP video packets are aggregated into a single larger MAC frame. Video error resilience and robustness are enhanced by using small NAL units and by performing a recovery mechanism to retrieve error-free NAL units in the MAC frame. Some modifications were necessary to support this retrieval process. These included the following: 1) an *aggregation field* in the MAC header; 2) the addition of a *Block_ACK*-type acknowledgment scheme; and 3) the use of a retransmission mechanism for aggregated NAL units at the MAC frame level. Enhancements to the MAC throughput efficiency, as well as improvements in terms of video quality PSNR, were presented.

References

[1] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band*, IEEE Std. 802.11a/d7.0, 1999.

[2] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Further High-Speed Physical Layer in the 2.4 GHz Band*, IEEE Std. 802.11g/d1.1, 2001.

[3] A. Doufexi, S. Armour, P. Karlsonn, M. Butler, A. Nix, and D. Bull, "A comparison of the Hiperlan/2 and IEEE802.11a wireless LAN standards," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 172–180, May 2002.

[4] A. Doufexi, S. Armour, M. Butler, A. Nix, and D. Bull, "A study of the performance of Hiperlan/2 and IEEE 802.11a physical layers," in *Proc. IEEE VTC—Spring*, Rhodes, Greece, 2001, pp. 668–672.

[5] A. Hettich and M. Schrother, "IEEE 802.11a or ETSI-BRAN HIPERLAN 2: Who will win the race for a high speed wireless LAN standard?" in *Proc. Eur. Wireless*, Munich, Germany, Oct. 1999, pp. 169–174.

[6] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. 802.11, 1999.

[7] P. Ferré, A. Doufexi, A. Nix, and D. Bull, "Throughput analysis of the IEEE 802.11 and IEEE 802.11e MAC," in *Proc. IEEE WCNC*, Atlanta, GA, Mar. 2004, pp. 783–788.

[8] D. Qiao and S. Choi, "Goodput analysis and link adaptation for IEEE 802.11a wireless LANs," *IEEE Trans. Mobile Comput.*, vol. 1, no. 4, pp. 278–292, Oct. 2002.

[9] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[10] K. Assaf, *EWC: HT PHY Specification, v1.27*, Dec. 2005.

[11] S. Trainin, *EWC: HT MAC Specification, v1.24*, Jan. 2006.

[12] *Draft Supplement to Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specification: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)*, IEEE Std. 802.11e, Nov. 2002.

[13] J. V. T. J. of ISO/IEC MPEG and I.-T. VCEG, *ITU-T H.264—Series H: Audiovisual and Multimedia Systems—Advanced Video Coding for Generic Audiovisual Services*, Mar. 2005.

[14] T. Stockhammer, M. M. Hannuksela, and S. Wenger, "H.26L/JVT coding network abstraction layer and IP-based transport," in *Proc. IEEE ICIP*, Rochester, NY, Sep. 2002, pp. II-485–II-488.

[15] S. Wenger, M. M. Hannuksela, T. Stockammer, and M. Westerlund, *Request for Comments—RFC 3984—RTP Payload Format for H.264 Video Format*, Feb. 2005.

[16] H. Shchulzrinnel, *Request For Comments—RFC 3550—RTP: A Transport Protocol for Real-Time Applications*, Jul. 2003.

[17] T. Wiegand, G. J. Sullivan, B. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[18] T. Stockhammer, M. M. Hannuskela, and T. Weigand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 6–13, Jul. 2003.

[19] S. Wenger, "H.264/AVC over IP," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 645–656, Jul. 2003.

[20] R. Schafer, T. Wiegand, and H. Schwarz, "The emerging H.264/AVC standard," in EBU Tecn. Rev., Jan. 2003. Tech. Rep.

[21] P. Ferré, A. Doufexi, J. Chung-How, A. Nix, and D. Bull, "Packetisation strategies for enhanced video transmission over wireless LANs," in *Proc. IEEE PV Workshop*, Irvine, CA, Dec. 2004.

[22] J. Chung-How and D. Bull, "Loss resilient H.263+ video over the Internet," *J. Signal Process.: Image Commun.*, vol. 16, no. 9, pp. 827–908, Jun. 2001.

[23] S. Mangold, S. Choi, G. R. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11e for QoS support in wireless LANs," *Wireless Commun.*, vol. 10, no. 6, pp. 40–50, Dec. 2003.

[24] J. del Prado Pav'on and S. Shankar, "Impact of frame size, number of stations and mobility on the throughput performance of the IEEE 802.11e," in *Proc. IEEE WCNC*, Atlanta, GA, Mar. 2004, pp. 789–795.

[25] Y. Xiao, "Concatenation and piggyback mechanisms for the IEEE 802.11 MAC," in *Proc. IEEE WCNC*, Atlanta, GA, Mar. 2004, pp. 1642–1647.

[26] S. Choi, "IEEE802.11e MAC-level FEC performance evaluation and enhancement," in *Proc. IEEE GLOBECOM*, Taipei, Taiwan, R.O.C., Nov. 2002, pp. 773–777.

[27] X. Xu, M. van der Schaar, S. Krishnamachari, S. Choi, and Y. Wang, "Adaptive error control for fine-granularity-scalability video coding over IEEE 802.11 wireless LANS," in *Proc. IEEE ICME*, Baltimore, MD, Jul. 2003, pp. I-669–I-672.

[28] Y. Xiao, "IEEE 802.11n: Enhancements for higher throughput in wireless LAN," *Wireless Commun.*, vol. 12, no. 6, pp. 82–91, Dec. 2005.

[29] *H.264/AVC Software Coordination.* [Online]. Available: http://iphome.hhi.de/suehring/tml/

[30] Y.-K. Wang, M. M. Hannuksela, V. Varsa, A. Hourunranta, and M. Gabbouj, "The error concealment feature in the H.26L test model," in *Proc. IEEE ICIP*, Rochester, NY, Sep. 2002, pp. II-729–II-732.

A Study on Gossiping in Transportation Networks

Sarit Kraus, Roni Parshani, and Yuval Shavitt

***Abstract*—To alleviate road congestion, suggestions have been made to equip cars with wireless communication to allow drivers to exchange information. This information is used to bypass congested areas. We study the dynamics of this solution using a hybrid microsimulation tool that we have developed and show that gossiping is an efficient method of information propagation. An increase in the number of gossiping agents leads to a faster and wider distribution of information. On the other hand, as in other information models, when the number of agents obtaining information about road conditions increases, their routing performance may decrease (unless smarter algorithms are deployed) since they will all attempt to use the same uncongested roads. Nevertheless, when the number of gossiping agents is balanced (20%–30% in our simulations), the average traveling time of gossiping agents is similar to the average traveling time of those who obtain information from a centralized information center.**

***Index Terms*—*Ad hoc* networks, agents, routing.**

I. Introduction

Road congestion is a known and acute urban menace with no signs of disappearing. There are apparently many suggested approaches to tackle this problem; one of them is to supply vehicles and drivers with up-to-date information about road conditions. There are two kinds of approaches to supply drivers with information that can aid them avoid congestion: One approach is based on fixed-structure communication networks, for example, cellular networks or frequency modulation/amplitude modulation radio [1]–[3], and the other approach is based on *ad hoc* communication networks. The latter approach saves the need to deploy expensive infrastructure. Several innovative projects propose using *ad hoc* networks as the communication infrastructure, for example, FleetNet [4] and CarNet [5].

The advance in technology in recent years has helped bring forth sophisticated onboard navigation systems for vehicles at a reasonable price. Such systems contain a computing device with a detailed road map, Global Positioning System for locating the vehicle on the map, and means of communication. One can use *ad hoc* communication networks (such as IEEE 802.11p or, in the future, dedicated short-range communications) to exchange information between neighboring vehicles. When two vehicles are within communication range, they can exchange their information regarding road conditions. Road condition information is thus propagated in the network without the need for an external or central infrastructure. Each time new information is obtained by a vehicle, the onboard navigation systems recalculate the optimal route from its current location to the destination. For example, if the navigation system receives information that one of the streets in its planned path is blocked, it will plan a new path that avoids the blocked road; the new path will be the shortest path from the vehicle's current position to the destination, taking into account the blockage.

Manuscript received July 13, 2005; revised September 19, 2006, January 6, 2007, September 19, 2007, and October 3, 2007. This work was supported in part by the Israel Science Foundation Center of Excellence Program under Grant 8008/03. This work was part of the Master's thesis of R. Parshani. The review of this paper was coordinated by Dr. L. Cai.

S. Kraus is with the Department of Computer Science, Bar Ilan University, Ramat-Gan 52900, Israel (e-mail: sarit@cs.biu.ac.il).

R. Parshani is with the Department of Physics, Bar Ilan University, Ramat-Gan 52900, Israel (e-mail: parshani.roni@gmail.com).

Y. Shavitt is with the School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel (e-mail: shavitt@eng.tau.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2007.912339