Performance Analysis of Video Transmission Over IEEE 802.11a/e WLANs

Sai Shankar N, *Senior Member, IEEE*, and Mihaela van der Schaar, *Senior Member, IEEE*

***Abstract*—This paper presents efficient mechanisms for delay-sensitive transmission of video over IEEE 802.11a/e Wireless Local Area Networks (WLANs). Transmitting video over WLANs in real time is very challenging due to the time-varying wireless channel and video content characteristics. This paper provides a comprehensive view of how to adapt the quality of service signaling, IEEE 802.11e parameters and cross-layer design to optimize the video quality at the receiver. We propose an integrated system view of admission control and scheduling for both contention and poll-based access of IEEE 802.11e Medium Access Control (MAC) protocol and outline the merits of each approach for video transmission. We also show the benefits of using a cross-layer optimization by sharing the Application, MAC, and Physical layer parameters of the Open Systems Interconnection stack to enhance the video quality. We will show through analysis and simulation that controlling the contention-based access in IEEE 802.11e is simple to realize in real products and how different cross-layer strategies used in poll-based access lead to a larger number of stations being simultaneously admitted and/or a higher video quality for the admitted stations. Finally, we introduce a new concept called time fairness, which is critical in enhancing the video performance when different transmitter–receiver pairs deploy different cross-layer strategies.**

***Index Terms*—Algorithm, cross-layer optimization, IEEE 802.11e Wireless Local Area Networks (WLANs), scheduler, video.**

I. Introduction

WIRELESS multimedia transmission across Wireless Local Area Networks (WLANs) has been gaining attention in the recent years because of the proliferation of technologies like Bluetooth, IEEE 802.11, 3G, and WiMAX. In particular, IEEE 802.11 WLAN [1] has emerged as a prevailing technology for (indoor) broadband wireless access because it supports real-time conversational multimedia applications like Voice over Internet Protocol and video conferencing. Since these applications are very delay sensitive (delays of around 200 ms are imposed), they require quality of service (QoS) support from the lower layers of the Open Systems Interconnection (OSI) stack to ensure timely delivery as well as to divide the available wireless resources in a fair manner among competing wireless stations (WSTAs). Several requirements have to be addressed to provide successful interactive multimedia applications over a network originally designed for generic data traffic and characterized by potentially high error rates. Multimedia data require a completely different network behavior: Throughput is no longer the only parameter in measuring network performance, and packet losses can be to some extent tolerated if counterbalanced by timely packet delivery. Multimedia communications have strictly bounded QoS requirements in terms of packet losses, end to end delays, and jitter. For example, video conferencing requires an end-to-end delay not greater than 200 ms, thus heavily restricting the possibility to retransmit lost or corrupted packets. Existing IEEE 802.11 standards, however, do not provide the necessary QoS to support such applications as they are aimed at solely replacing the wired Ethernet, which supports only a best-effort service (not guaranteeing any service level to users/applications). In order to address multimedia support over WLAN, the IEEE 802.11 Working Group defined a new supplement to support QoS known as 802.11e [2].

The new 802.11e Medium Access Control (MAC) enables applications like voice, video, and bandwidth intense data services. With the advent of IEEE 802.11e, the application (APP) layer has to pass on the parameters to the IEEE 802.11e MAC such that it can reserve adequate resources for the application to guarantee its bandwidth. Once the APP layer parameters are signaled to the MAC layer, the IEEE 802.11e standard uses the Traffic Specification (TSPEC) element to negotiate with the QoS access point (QAP). The new MAC protocol of IEEE 802.11e is called the Hybrid Coordination Function (HCF). The HCF is called "hybrid" since it combines a contention channel access mechanism, which is referred to as enhanced distributed channel access (EDCA), and a polling-based channel access mechanism, which is referred to as HCF controlled channel access (HCCA), each of which operates simultaneously and continuously within the QoS basic service set (QBSS). The list of abbreviations and symbols used in this paper are shown in Tables I and II, respectively.

A. Related Work in Multimedia Transmission and Contributions of This Paper

Previous works on multimedia transmission over WLANs considered different layers in isolation or made assumptions on different layers when considering the optimization from a one-layer perspective. An excellent review of channel-adaptive multimedia streaming research and error concealment strategies is provided in [8]. Potential solutions for robust wireless multimedia transmission over error-prone networks include APP-layer packetization, (rate distortion optimized) scheduling, joint

Manuscript received August 15, 2005; revised June 14, 2006 and July 28, 2006. The review of this paper was coordinated by Dr. W. Zhuang.

S. Shankar N is with the Corporation R&D Systems Engineering Department, Qualcomm Inc., San Diego, CA 92121 USA (e-mail: sai.shankar@qualcomm.com).

M. van der Schaar is with the Electrical Engineering Department, University of California at Los Angeles, Los Angeles, CA 90095 USA (e-mail: mihaela@ee.ucla.edu).

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.2007.897646

TABLE I
LIST OF ABBREVIATIONS

Acronyms	Expansions
OSI	Open Systems Interconnection
WLAN	Wireless LAN
MAC	Medium Access Control
PHY	Physical Layer
APP	Application Layer
VoIP	Voice over Internet Protocol
QoS	Quality of Service
TSPEC	Traffic Specification
QAP	QoS Access Point
WSTA	Wireless Station
HCF	Hybrid Coordination Function
DCF	Distributed Coordination Function
EDCA	Enhanced DCF based Channel Access
HCCA	HCF Coordinated Channel Access
QBSS	QoS Basic Service Set
HC	Hybrid Coordinator
ACU	Admission Control Unit
CF-Poll	Contention Free Poll
MSDU	MAC Service Data Unit
EC	Effective Capacity
SI	Service Interval
SNR	Signal to Noise Ratio
SIFS	Short Inter Frame Space
PLR	Packet Loss Rate
PSNR	Picture Signal to Noise Ratio
SVC	Scalable Video Coder

TABLE II
LIST OF SYMBOLS

Symbols	Meaning
CW_{min}	Contention Window minimum
CW_{max}	Contention Window maximum
$TXOP$	Transmission Opportunity
P	Peak data rate
d	Delay
σ	Traffic burstiness
ρ	Mean data rate
δ	Channel burstiness
R	Minimum PHY rate
C	Channel capacity
L	Frame length
N	Number of frames
g	Guaranteed rate
$T^{overhead}$	Overhead in time
T_{CP}	Contention period time
r	Fraction of air time allocated to a WSTA
EA	Effective airtime
T_{ack}	Time for acknowledgement
Th	Threshold
$RSSI$	Received Signal Strength Indicator
ϕ	Weight

source-channel coding, error resilience, and error concealment mechanisms [9]. In [6] and [7], a QoS framework has been proposed that maps categorized video packets onto the relative differentiated service provided by the wireless channel using a predetermined pricing model. Video categorization is based on the relative priority index, which represents the relative preference per packet in terms of loss and delay. At the Physical (PHY) and MAC layers, significant gains have been reported by adopting a cross-layer optimization, such as link adaptation, channel aware scheduling, and optimal power control [10]. However, these contributions are aimed at improving the throughput or reducing power consumption without taking into consideration the multimedia content and traffic characteristics. Explicit consideration of multimedia characteristics and requirements can further enhance the important advances achieved in cross-layer design at the lower layers. Possible solutions and architectures for cross-layer optimized multimedia transmission have been proposed in [3]–[5]. Ansel *et al.* [11] proposed a scheduling scheme for IEEE 802.11e to improve the performance of multimedia in terms of delay and loss. The paper shows that the fixed transmission opportunity (TXOP) allocation is not efficient for video transmission as it does not consider the bursty characteristics of video traffic, and they propose improved scheduling schemes to alleviate this problem. In this paper, we shape the traffic that is arriving at the MAC buffer through a twin leaky bucket scheme that removes the burstiness of the video traffic and ensures that a simple scheduler can be deployed for efficient video transmission.

This paper proposes a new integrated system design for multimedia transfer over IEEE 802.11e WLAN using both EDCA and HCCA, where the admission control unit (ACU) and the scheduler are colocated with the QAP. We discuss all aspects of multimedia transmission, including admission control and scheduling at the MAC layer and an optimal cross-layer algorithm using APP/MAC and PHY layers to enhance the quality of video. The admission control and scheduling algorithm is based on the concept of effective bandwidth that uses the concept of statistical multiplexing in the wireless channel by admitting streams in a optimal way and ensures complete QoS for the admitted stream. This is different from the admission control proposed in the IEEE 802.11e standard [2] that is based on just mean or peak data rate allocation. The admission control based on mean data rate admits more streams, as compared to our scheme that is based on effective bandwidth, but it fails to satisfy the QoS demanded by the APP. The peak data rate scheme, however, admits less streams compared to our scheme, thus wasting wireless channel resources. Wireless channel characteristics change over time due to fading, and any design of admission control must take that into account. To incorporate channel fading, we include a new parameter called channel burstiness (δ) that captures the first-order characteristics of time-varying wireless channel capacity. This paper will also discuss a new concept of air or time fairness [12] and shows how this can enhance the video quality in the WLAN network. We also explain a cross-layer algorithm that is performed at the MAC layer considering different tradeoffs at the APP and PHY layers.

B. Motivation for Time Fairness

IEEE 802.11e can operate at multiple rates. The MAC chooses a particular physical transmission rate that is based on the experienced channel condition to satisfy the QoS needed by the real-time video streaming APP. This algorithm is called link adaptation [24], and it improves the station's throughput, but also it causes unfairness to the WSTAs that have high transmission rates. This unfairness arises due to the fact that a frame of size L takes more transmission time at lower physical transmission rate and, hence, reduces the overall network throughput and increases the delay of the higher physical transmission rate stations. To illustrate with a simple example, consider two

Octets 1	1	2	2	4	4	4	4	4	4
Element ID	Length	TS Info	Nominal MSDU Size	Mean Data Rate	Max. Burst Size	Peak Data Rate	Min. PHY. TX. rate	Delay Bound	Drop Rate

Fig. 1. TSPEC element format.

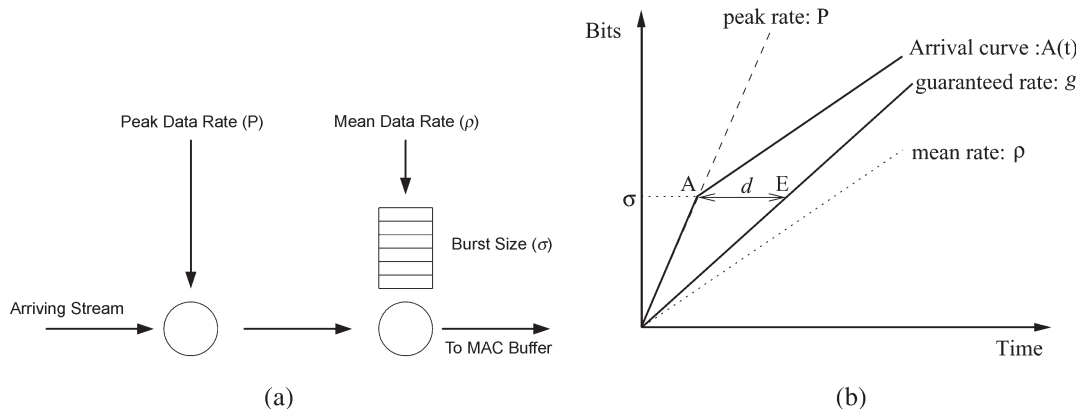

Fig. 2. Token bucket policer and the arrival curves. (a) Token bucket representation of admission control. (b) Arrival curve and calculation of guaranteed rate.

WSTAs transmitting a 1000-bit frame at 11 and 1 Mb/s, respectively. The WSTA with 11 Mb/s takes 111.11 μs to transmit the frame, whereas the WSTA with 1-Mb/s rate takes 1 ms to transmit the same frame. As a result, the TXOP of the WSTA with 11 Mb/s is deferred resulting in a degraded system throughput over a finite interval of measurement (t_1, t_2). As a remedy to this unfairness, we propose a new concept called air fairness.

C. Organization of This Paper

The rest of this paper is organized as follows. Section II explains the admission control based on effective (or guaranteed) bandwidth that also includes the channel burstiness and the simple scheduling algorithm that was implemented in our HCCA model of IEEE 802.11e. Then, we explain how to guarantee the airtime usage of a stream/WSTA using EDCA by controlling the parameters TXOP and $\mathrm{CW}_{\min}$ in Section III. Section IV outlines how the prioritized multimedia bitstreams (e.g., scalable video) can be efficiently mapped onto the 802.11e HCCA parameters to ensure an improved quality in terms of individual station performance as well as the overall system performance (i.e., more admitted station). Then, we outline a simple cross-layer strategy in Section V that could enhance the quality of multimedia by considering a small set of control parameters in APP, MAC, and PHY layers of the IEEE 802.11e system. IEEE 802.11 has multiple transmission rates depending on the location of an individual station. This has large effects on the video quality not only to this station but also for the entire network. Section VI explains the essence of time or air fairness and shows how it improves the quality of multimedia in these multirate or error environments. Section VII gives the conclusions and outlines necessary future research.

II. Airtime Admission Control and Scheduling in IEEE 802.11e HCCA

Admission control is one of the most essential components in IEEE 802.11e. We will outline the admission control and the scheduling scheme that will provide per flow QoS for multimedia. To ensure user satisfaction, it is essential that a video stream once admitted by HC has guaranteed its QoS for the lifetime of that stream. Thus, there is a need to control how many streams are admitted into the system. The APP passes its request in form of the rate parameters to the MAC layer. The rate parameters that were demanded by the APP are then translated into time parameters by the MAC layer, and then, the HC (with the help of ACU) colocated at the QAP makes a decision to admit or reject a stream. Thus, the APP layer parameters are signaled to the MAC layer in form of a TSPEC element. The TSPEC element contains the set of parameters that characterize the traffic stream that the WSTA wishes to establish with the HC. Once the TSPEC request is received by the HC, it analyzes the TSPEC parameters and decides whether to admit the stream into the network using the admission control algorithm or not. The important TSPEC parameters of IEEE 802.11e are shown in Fig. 1. If the stream is admitted, the HC schedules the delivery of downlink traffic and/or QoS CF-Polls in order to satisfy the QoS requirements of the stream, as specified in the TSPEC. Among the parameters defined in the TSPEC element of IEEE 802.11e, we use a subset of the parameters to design the efficient admission control. These parameters are peak data rate (P), mean data rate (r), maximum burst size (σ), delay (d), nominal MAC service data unit (MSDU) size (L), minimum PHY TX rate (R), and maximum MSDU size ($M = 2304$ B). We also use the channel or link state to determine the additional percentage of channel resources (bandwidth or time) that need to be reserved to cover the losses that may occur in the wireless medium. The peak data rate, mean data rate, and the maximum burst size are part of the twin token bucket parameters that are supplied by the higher layer entities to the MAC. When the APP does not provide these parameters, the MAC needs to infer them over small intervals and pass them to the ACU colocated with the HC at QAP. Fig. 2 shows the twin token bucket policer and the calculation of effective or guaranteed rate.

To facilitate the efficient QoS support for multimedia applications over wireless networks, it is essential to model a wireless channel in terms of connection-level QoS metrics such as data rate, delay, and delay-violation probability. However, the existing wireless channel models, i.e., PHY layer channel models, do not explicitly characterize a wireless channel in terms of these QoS metrics. One of the earliest models to characterize wireless channels was reported in [17], wherein they try to fit the parameters of the two-state Markov chain using the parameters of Rayleigh or Riceian fading. In [18], a link-layer channel model that is termed as effective capacity (EC) has been developed. The key advantages of the EC link-layer modeling and estimation are ease of translation into QoS guarantees, such as delay bounds, simplicity of implementation, and accuracy, and, hence, efficiency in admission control and resource reservation. Apart from the above approximations found in the literature, we consider a simple approach to characterize the fading occurring in the wireless channel. This is termed as transmission burstiness δ. The transmission burstiness represents the shortage of bits that were not served by the wireless channel during a time interval $[0, t]$. As an example, if C is the original channel capacity of the wireless medium, then in any time interval $[0, t]$, the maximum amount of bits that can be transmitted is $C \times t$. However, due to interference and mobility, this channel may not be available for a short amount of time, and the amount of bits transmitted in time $[0, t]$ is less than or equal to $C \times t$. Let δ represent the total amount of bits wasted because of channel fading. This is obtained by determining the times when the channel is not usable in the time interval of $[0, t]$ and then multiplying this time by the channel capacity C. Hence, in any time t the lower bound on the channel capacity available to that stream or WSTA is $C \times t - \delta$. This will be taken into consideration when modeling admission control to admit multimedia. The value of δ can be easily calculated by obtaining the mean of the channel fading that follows a particular distribution such as Rayleigh or Riceian.

We need to evaluate the effective bandwidth of the stream that is regulated by the first and second token bucket shown in Fig. 2 (see [15] for further details). This method of calculating the effective bandwidth is different from the previous works that were based on the mean and the peak data rate [2]. The mean-data-rate-based admission admits more streams at the expense of QoS, and the peak-data-rate-based admission provides QoS at the expense of number of streams admitted. Our effective bandwidth-based stream based on twin token bucket is optimal as it exploits the statistical multiplexing of all streams in the wireless channel without compromising the QoS requests of the admitted streams. Based on the twin leaky bucket analysis and including the frame error rate into the system, we get the effective bandwidth as

$$g_i = \frac{P_i}{\left[1 + d_i \frac{P_i - \rho_i}{\sigma_i + \delta_i}\right] [1 - p_e]}. \tag{1}$$

The previous effective bandwidth computation does not include overheads. Let us assume that for a stream i, the nominal MSDU size is L_i. For each frame, there is an overhead in time based on the acknowledgement (ACK) policy, interframe space (IFS) time, PHY layer convergence procedure preamble (PLCPPreamble), MAC and PHY layer headers, and polling overhead (only in case of upstream or sidestream). The scheduling policy also determines the polling overheads, as different scheduling policies determine how many times one has to poll a WSTA in the service interval denoted as SI. The value of the SI will be derived later. Assuming that the SI is known, the number of MSDUs per SI is given by

$$N_i = \left\lceil \frac{g_i * \mathrm{SI}}{L_i} \right\rceil \tag{2}$$

and the new guaranteed rate that is required for this connection, including overheads, is given by

$$g_i' = \frac{N_i(L_i + O_i)}{\mathrm{SI}} \tag{3}$$

where O_i in the above equation represents the overheads added for transmitting a frame. Now, the admission control policy, including the overheads, is given by the following equation:

$$g_{i+1}' + \sum_{k=1}^{i} g_k' \leq C \tag{4}$$

where g_{i+1}' is the guaranteed rate or effective bandwidth, including the overheads. Since the HC will poll (upstream/sidestream) or access the medium (downstream) for a specific time, the above rate parameters have to be converted to time parameter. We will convert the rate-based requirements of the stream given in (1) into the corresponding airtime requirement based on the minimum PHY TX rate parameter. During the admission control negotiation, the ACU and the QSTA/QAP negotiate this parameter as to what is the minimum PHY TX rate that a WSTA and QAP/WSTA shall communicate. The actual PHY rate can exceed the minimum PHY rate. If the rate of the source station (either QAP or WSTA) drops below the minimum PHY TX rate, the stream is considered to violate its initially negotiated QoS requirements and may be dropped by the ACU depending on the channel conditions. Thus, the sufficient condition for the minimum PHY TX rate is

$$R_i \geq g_i'. \tag{5}$$

Based on the above guaranteed rate, the ACU determines the number of frames that arrive in the SI for stream i, assuming that there are $i - 1$ streams already undergoing service in the QBSS

$$N_i^{\mathrm{SI}} = \left\lceil \frac{\mathrm{SI} * g_i'}{L_i} \right\rceil. \tag{6}$$

Then, the ACU calculates the TXOPs that are required to service all these MSDUs in an SI. This is given by

$$\mathrm{TXOP}_i = N_i^{\mathrm{SI}} * \left(\frac{L_i}{R_i} + T_i^{\mathrm{overhead}} \right) \tag{7}$$

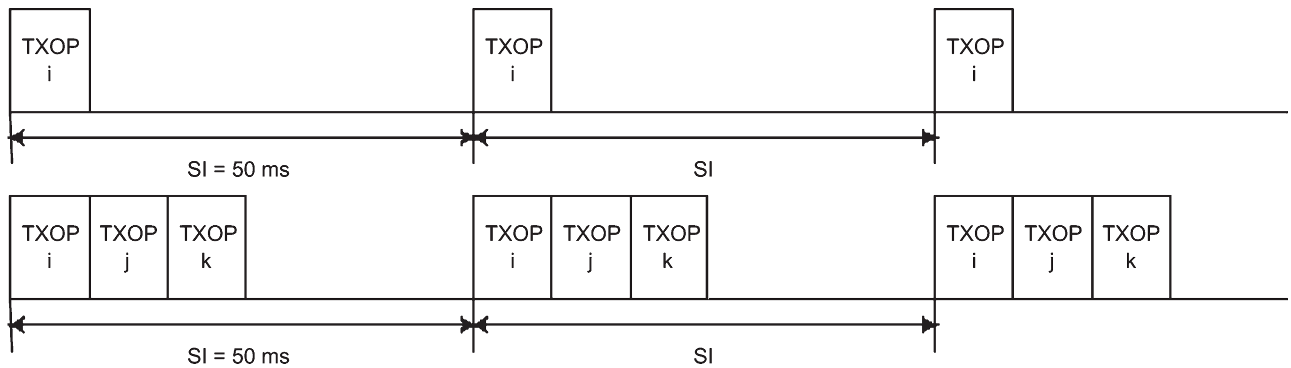

Fig. 3. Schedule allocation for stream from WSTA "I" and schedule allocations for streams from WSTAs "I" to "k."

where T_i^{overhead} is the overhead in time incurred per frame transmission. The admission control algorithm can be rewritten as

$$\frac{\text{TXOP}_i}{\text{SI}} + \sum_{k=1}^{i-1} \frac{\text{TXOP}_k}{\text{SI}} \leq \frac{T - T_{\text{CP}}}{T} \tag{8}$$

where T is the beacon interval, and T_{CP} is the time reserved for EDCA traffic. The ACU might also consider additional time to allow for retransmissions. We will explain the construction of a simple modified round robin scheduler in IEEE 802.11e MAC. If the delay bound is specified by the flow in the TSPEC, the scheduler uses the delay bound value for the calculation of the schedule. The schedule generated by any scheduler should meet the normative behavior set by the standard [2]. The normative behavior states that the HC shall grant, for every flow, the negotiated TXOP in an SI. The schedule for an admitted stream is calculated in two steps:

1) calculation of the scheduled SI;
2) calculation of the TXOP duration for the stream for a given SI.

The TXOP is already calculated as specified in (7). The calculation of the SI is done as follows. First, the scheduler calculates the minimum of all delay bounds for all admitted streams. Then, it takes half of the minimum delay so that any arbitrary schedule to any flow within the SI by the scheduler residing at the HC satisfies the delay and jitter constraints demanded by the APP. This is represented by

$$\text{SI} \leq \frac{1}{2} * \min\{d_1, d_2, \ldots, d_n\}. \tag{9}$$

The TXOP was obtained in the previous section assuming the SI. An example is shown in Fig. 3. A stream from WSTA "i" is admitted in the network. Therefore, the HC allocates a TXOP for the WSTA/stream. The beacon interval is fixed at 100 ms, and the SI is calculated from the above equation.

The same process is repeated continuously if more than one WSTA/stream is in the network. Each WSTA is polled in a round robin manner and granted a specified TXOP duration according to the requirements of the stream. An example is shown in Fig. 3. The TXOP allocation by the HC ensures that all the flows will get the above time, irrespective of their physical transmission rate, thereby ensuring fair airtime allocation in the wireless medium. This type of allocation has been well considered in literature in the context of multiprogramming by Liu and Layland [19], who also prove the optimality of such schedules.

This scheduler is different from the one proposed in [27]. In the calculation of the effective bandwidth, which is represented by g_i in (1), we account for the variability of the arrival rate of the traffic through the burstiness parameter σ and the channel variability through δ. σ is determined at the APP layer based on the video traffic. The PHY layer determines the δ parameter. Thus, the TXOP is calculated based on the effective bandwidth, which accounts for both the video traffic variations as well as the changes in channel conditions. When the arrival rate exceeds the effective bandwidth, the excess traffic is buffered at the MAC and is serviced during subsequent SIs. Similarly, if the signal-to-noise ratio (SNR) experienced by the video flow is poor, the video packets are buffered and serviced at a later time. In summary, the calculation of effective bandwidth takes into these variations, and thus, the scheduling policy can successfully cope with these situations.

III. Controlled Airtime Usage in IEEE 802.11e EDCA

In this section, we show how to realize the airtime requirement of the stream using EDCA. There are two methods to control each station's airtime usage in the EDCA outlined in [20]: 1) controlling the TXOP limit of each station and 2) controlling the frequency of a station's access to the wireless medium. By using the first method, all stations choose the same EDCA parameters, but each station occupies the wireless medium for a different amount of time during each access depending on their requirement. By using the second method, each station occupies the medium for the same amount of time during each access but has a different medium "accessing frequency" or probability of transmission.

A. Controlling the TXOP Limit

Let r_i' be the fraction of airtime that station i should obtain for stream i and TXOP_i be the value of station i's TXOP limit. Also, let r_i denote the amount of airtime that a WSTA requires to serve all its streams in a time interval of 1 s. Let T_i be the amount of time required to transmit a frame with size of L_i

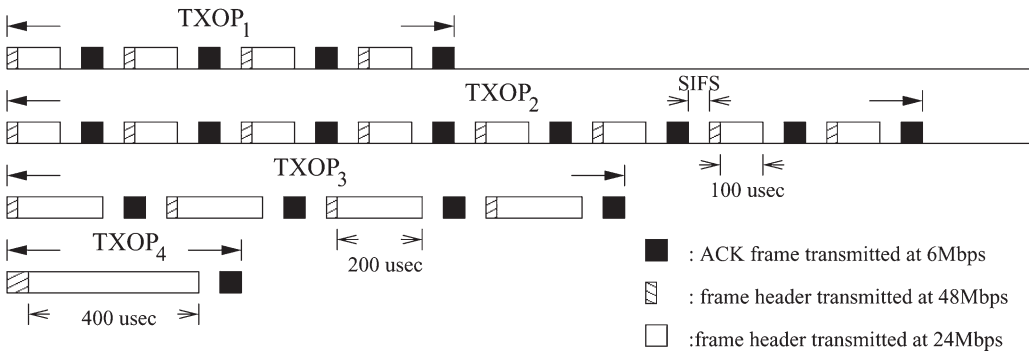

Fig. 4. Example 1—Selection of TXOP limits. Given that SIFS $= 16\ \mu$s, frame header size $= 34$ B, and ACK frame size $= 14$ B in the IEEE802.11a standard, we have $\text{TXOP}_1 = 619.6\ \mu$s, $\text{TXOP}_2 = 1255.2\ \mu$s, $\text{TXOP}_3 = 1019.6\ \mu$s, and $\text{TXOP}_4 = 512.5\ \mu$s.[1]

(excluding the frame header) from stream i at the negotiated minimum PHY rate R_i. T_i is obtained by

$$T_i = \frac{L_i}{R_i}. \tag{10}$$

Let M be the index of the stream such that $T_M = \max_i T_i$. Then, the TXOP_i can be chosen as

$$\text{TXOP}_i = \frac{r_i T_M}{r_M T_i}\frac{L_i + H}{R_i} + \left(2\left\lceil\frac{r_i T_M}{r_M T_i}\right\rceil - 1\right)\text{SIFS} + \left\lceil\frac{r_i T_M}{r_M T_i}\right\rceil T_{\text{ack}} \tag{11}$$

where T_{ack} is the amount of time to transmit an ACK frame. Here, the term $r_i T_M / r_M T_i$ represents the number of frames that will be served for a stream i when it accesses the medium. The second and third terms include the overheads for the turnaround time accounted by SIFS and the ACK transmission time, respectively. As an example, let us consider four streams with $L_i =$ 600, 600, 1200, and 1200 B, respectively. We assume that these four streams are required to transmit at least at the PHY rates of 48, 48, 48, and 24 Mb/s, respectively. Based on (10), we have $T_M = 1200{*}8/24{*}10^{-6} = 400\ \mu$s. If we assume r_i for each stream to be 0.1, 0.2, 0.2, and 0.1, respectively, we have $N_i = r_i T_M / r_M T_i = 4$, 8, 4, and 1, and N_i is actually the number of data frames that stream i should transmit during each access to the wireless medium. The values of TXOP_i are illustrated in Fig. 4. In the case when N_i is not an integer number, frame fragmentation is required for precise airtime control.

With the values of TXOP_i chosen by (11) and the fact that each station has a statistically equal probability to access the medium (because of using the same EDCA parameters), each station will obtain the amount of airtime proportional to its r_i' value. The maximum amount of airtime that station i can get within a 1-s period $r_{\max,i}$ is

$$r_{\max,i} = \frac{r_i}{\sum_i r_i}\text{EA} \geq \frac{r_i}{\sum_i r_i}\sum_i r_i \geq r_i \tag{12}$$

given that (8) is held true by the ACU. Here, EA is the effective airtime, which represents the percentage of airtime available in the medium for video data transmission. This percentage is less than 100% due to the incurred transmission overheads such as beacons, polls, preambles, protocol layers' headers, etc.

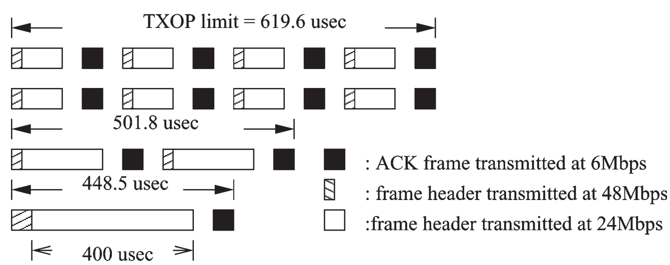

Fig. 5. Example 2—Selection of the network-wide unified TXOP limit. In this example, the TXOP limit for all stations is 619.6 μs.

Equation (12) shows that each station can always obtain the required amount of airtime (determined by the ACU) by using this simple control method. In fact, one of the greatest advantages of using the EDCA is that the amount of airtime a station can get is determined by the ratio of stations' r_i values and not by the absolute value of r_i. For example, assume that station 1 needs 0.1 s out of every 1-s period (i.e., $r_1 = 0.1$) for a stream and station 2 needs 0.2 s (i.e., $r_1 = 0.2$) for another stream. Based on (12) and given that EA $= 0.6$, the actual amount of airtime that station 1 can obtain is 0.2 s and that for station 2 is 0.4. When more streams join the wireless LAN, the amount of airtime that station 1 can get decreases [automatically adjusted by the EDCA via (12)], but it will not get less than 0.1, according to (8).

B. Controlling the Accessing Frequency

Instead of controlling the duration of a TXOP, we can use a fixed TXOP duration for all stations but control their access frequencies AF_i, so that stations can still acquire the desired amount of airtime. This TXOP has to be chosen so that each station uses the same amount of airtime—during each access to the wireless medium—to transmit data frame at the negotiated minimum PHY rate. Therefore, the TXOP limit is chosen as

$$\text{TXOP limit} = \max_i \left\lceil\frac{T_M}{T_i}\right\rceil \frac{L_i + H}{R_i} + \left(2\left\lceil\frac{T_M}{T_i}\right\rceil - 1\right)\text{SIFS} + \left\lceil\frac{T_M}{T_i}\right\rceil T_{\text{ack}}. \tag{13}$$

As shown in Fig. 5, the TXOP limit of the above example is 619.6 μs, and all four streams will transmit 400-μs worth data

[1] Physical layer overhead is not included in the computation.

frames given this TXOP limit (i.e., streams 1 and 2 send four frames, stream 3 sends two frames, and stream 4 sends one frame).

Several EDCA parameters can be used for controlling AF_i, including minimum/maximum contention window size ($\mathrm{CW}_{\min,i}/\mathrm{CW}_{\max,i}$) and arbitration IFS ($\mathrm{AIFS}_i$). The relation between these parameters and the access frequency can be found as [21]

$$\sum_{i=1}^{n_1} \mathrm{BT}_i^{(1)} = \sum_{j=1}^{n_2} \mathrm{BT}_j^{(2)} + \sum_{h=1}^{n_1+n_2-1} D_h \tag{14}$$

where $\mathrm{BT}_i^{(j)}$ is the ith backoff time chosen by STA j and is mainly determined by $\mathrm{CW}_{\min,j}$ and $\mathrm{CW}_{\max,j}$, D_h is referred to as the "decrementing lag" in [21] and is mainly decided by AIFS_i value, and n_i represents the total number of times STA i has backed off during the observing time interval and is proportional to AF_i. Based on (14) and by setting

$$\frac{\mathrm{AF}_i}{\mathrm{AF}_j} = \frac{r_i}{r_j} = \frac{n_i}{n_j} \tag{15}$$

we can determine the adequate EDCA parameters using the algorithms given in [21]. By using the approximation in [21], one simple solution is to choose $\mathrm{CW}_{\min}$ as

$$\frac{\mathrm{CW}_{\min,i}}{\mathrm{CW}_{\min,j}} = \frac{r_j}{r_i} \tag{16}$$

which will give a very good control on AF_i. One can easily reach the same conclusion drawn from (12) that stations can always acquire, at least, the required amount of airtime in a distributed manner.

C. Comparison of HCCA and EDCA

In this section, we compare the polling-based HCCA and the contention-based EDCA for their QoS support via simulations. We will emphasize the advantages of using the enhanced EDCA for QoS support and to verify the effectiveness of the integrated airtime-based admission control and enhanced EDCA. The simulations are carried out in Optimum Network (OPNET) simulator. In this scenario, we compare the system efficiency in terms of the number of streams being admitted into a wireless LAN under the EDCA and the HCCA. We have modified the wireless LAN MAC of OPNET to include the admission control algorithm and the signaling procedures, as explained above.

1) System Efficiency: We assume that each station carries a single traffic stream which requests a guaranteed rate of 5 Mb/s.[2] We also assume that all stations are required to transmit at 54 Mb/s for QoS guarantees and do not change their PHY rates. We increase the number of stations, starting from 1, until the wireless LAN cannot accommodate any more stations (or streams). For the EDCA case, we control the TXOP limit for airtime usage control. Since all streams have the same guaranteed rate ($g_i = 5$ Mb/s) and minimum PHY rate ($R_i = 54$ Mb/s), each station uses the same TXOP limit in this scenario. For the HCCA case, we follow the procedures in Section II.

[2]The average bit rate of a DVD-quality (MPEG-2) video is about 5 Mb/s.

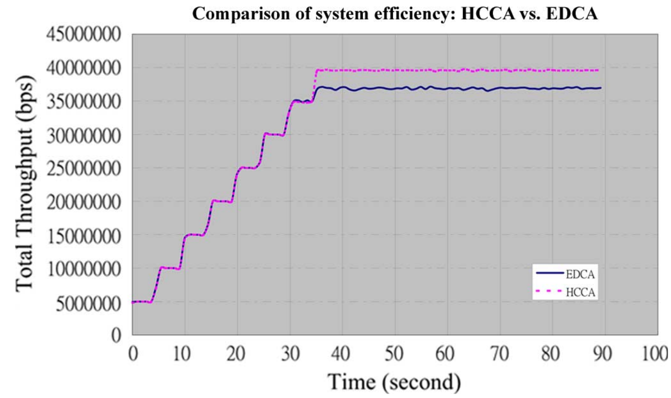

Fig. 6. Comparison of system efficiency, in terms of the total throughput, between the HCCA and the EDCA.[3]

Fig. 6 plots the total throughput under the HCCA and the EDCA. Since all stations request the same guaranteed rate, one can easily convert the total throughput to the total number of stations (i.e., streams) admitted into the wireless LAN. We increment the number of stations every 5 s in order to explicitly show the throughput received by individual streams. Prior to $t = 35$ s, every admitted stream gets exactly the 5-Mb/s guaranteed rate under both the HCCA and the EDCA. It shows that by using the enhanced EDCA, we can achieve the same QoS guarantees as using the polling-based HCCA.

After $t = 35$, the number of stations is increased to eight. The figure shows that using the EDCA can no longer guarantee the streams' QoS because it needs a total throughput of 40 Mb/s to support eight streams, but the wireless LAN can only provide about 37 Mb/s. In our OPNET simulation trace, we observe a lot of frame drops under the EDCA (not shown but can be inferred from the figure), starting at $t = 37$ s. However, in HCCA, all streams are still provided with the 5-Mb/s guaranteed rate. This result is expected because the HCCA uses the polling-based channel access (in contrast to the contention-based EDCA), thereby resulting in a higher efficiency. Based on the simulation results, one can also obtain the values of the effective airtime EA in (12). Because all streams are transmitted at the same PHY rate, the value of EA can be computed by

$$\mathrm{EA} = \frac{\text{system total throughput}}{\text{PHY rate}}. \tag{17}$$

Therefore, we have $\mathrm{EA} = 0.67$ under the EDCA and $\mathrm{EA} = 0.73$ under the HCCA. Although the value of EA varies under the EDCA (depending on the EDCA parameters used), it is always within the range between 0.65 and 0.68 in our simulation.

Although using the HCCA achieves a better efficiency, it only generates $0.06 = 0.73 - 0.67$ s more data-transmission time (within a one-second period) or about 3 Mb more data frames when all stations transmit at 54 Mb/s (the maximal PHY rate in the 802.11a PHY spec.). When stations use lower PHY rates, the small difference between the EA values of the EDCA

[3]A new station carrying a single stream is added to the wireless LAN about every 5 s and transmits at 54 Mb/s. The height of each "stair" in the figure is equal to a stream's guaranteed rate $= 5$ Mb/s.

and the HCCA results in an even smaller throughput difference. Therefore, one can expect that using the EDCA and the HCCA will generate a similar performance, particularly in terms of the total number of admissible streams.

The greatest advantage of using the HCCA for QoS guarantees is higher system efficiency (i.e., a higher EA value), which is due to the HCCA's contention-free nature. Due to this higher efficiency, the HCCA can provide more resource and may admit more traffic streams than the EDCA. However, there are several potential problems of using the HCCA, primarily due to its centralized access control.

1) As pointed out in the IEEE 802.11 standard, the operation of the polling-based channel access may require additional coordination to permit efficient operation in cases where multiple polling-based wireless LANs are operating on the same channel in an overlapping physical space. New standard supplements such as 802.11k are being developed to facilitate the required coordination, but additional operations such as monitoring the channel activity (via 802.11k) may incur control overhead, hence degrading the system efficiency. On the other hand, the EDCA does not need any coordination between wireless LANs using the same channel because the EDCA is designed to solve the channel sharing problem.
2) The HC in the HCCA needs to recompute the service schedule whenever a new traffic stream joins or an existing stream leaves the wireless LANs. However, the ACU in the EDCA assigns the appropriate EDCA parameters set to the new stream, and the existing streams may not need to make any adjustment,[4] as explained in the previous section.
3) The QoS of a traffic stream can only be guaranteed if the WSTA transmits at a (physical) rate higher than the negotiated minimum physical rate. If a station lowers its physical transmission rate (below the negotiated rate), the amount of airtime originally allocated to the stream (by the HC) may not suffice to support the required QoS even though the HC may still have enough unallocated resource to support that stream's QoS at this lower rate. Of course, the HC can temporarily allocate more airtime (by recomputing the service schedule) to support that stream's QoS at this lower rate. However, if more new streams request for QoS later, the HC needs to cut the stream's airtime allocation back to the originally negotiated amount since the HC needs airtime for new streams. However, using the EDCA will not require the AP or the ACU to reallocate airtime because WSTAs can automatically obtain the extra amount of airtime according to (12). Consider the previous example again. Stations 1 and 2 can actually halve their PHY rates and still meet the QoS requirements. In other words, the QoS can be automatically provided by the EDCA, regardless of the current physical transmission rate the station is operating, as long as the system airtime resource allows. The new streams will get the required amount of airtime as the airtime allocation is adjusted automatically according to (12).

[4]It depends on which airtime control methods of the EDCA are applied.

Having explained the advantages of using admission control and how to implement it in EDCA, we will now show how the quality of multimedia can be enhanced if the APP and PHY layers pass some information to the MAC layer. Although these techniques can be applied to both HCCA and EDCA of IEEE 802.11e, we concentrate the cross-layer implementation on IEEE 802.11e HCCA only because the HCCA is capable of providing tight QoS bounds and enhanced multimedia performance. The next section will highlight some useful knobs for multimedia that are available at the MAC layer and how these knobs can be tuned to obtain the optimal performance.

IV. Efficient Mapping of Prioritized (Scalable) Video Onto 802.11e TSPEC Parameters

The implementation of the simple scheduler explained in Section II is easy, but it can be quite inefficient for video streaming applications. This is because the video traffic varies over time and consists of frames/packets with considerably varying sizes and different delay constraints. Normally, one would consider the video as one stream and set the TSPEC parameters so that the MAC of IEEE 802.11e HCCA would do the admission control and scheduling as outlined previously. This often results in a lower number of video flows/applications admitted because not all layers are important for the proper reception of video. We termed this method in [15] as global TSPEC or global flow. To improve the overall system utilization, we introduce the subflow concept in which a video flow (bitstream) is divided into several subflows according to their relative priorities based on the overall distortion of the decoded video. The APP layer enables each subflow (priority layer) of the video to interface with the MAC as a separate flow. This ensures that more users can be admitted into the network while guaranteeing a minimum acceptable quality for the already admitted users. Also, by setting different delay bounds for different subflows, it allows the MAC to admit more users at the same video quality. The subflow concept can be implemented for each video coder for which the bitstream can be prioritized.

The proposed subflow concept can be implemented using any prioritized nonscalable (e.g., MPEG-2, H.264) or scalable video coding scheme. However, nonscalable video coding algorithms do not provide the same graceful degradation and adaptability to a large range of wireless channel conditions and power constraints as provided by state-of-the-art scalable coding schemes. Hence, although the concepts proposed in this paper can potentially be deployed with state-of-the-art nonscalable coding with or without bitstream switching or layered principles, this usually entails higher complexity and smaller granularity for real-time adaptation across the various layers of the OSI stack. Consequently, in this paper, we use a recently proposed scalable 3-D wavelet video coding based on motion compensated temporal filtering (MCTF) [22]. MCTF-based scalable video compression is attractive for wireless streaming applications since it provides on-the-fly adaptation to channel conditions, support for a variety of wireless receivers with different resource capabilities and power constraints, and easy prioritization of various coding layers and video packets. For this wavelet video coder, the subflow concept enables

TABLE III
TSPEC PARAMETERS FOR THE FLOWS/SUBFLOWS AND THEIR TXOP REQUIREMENTS

Flow	ρ (Mbps)	d (ms)	L (bytes)	σ (bytes)	P (Mbps)	g_i' (Mbps)	TXOP	User priority
Global Flow	2.048	200	2048	78858	2.915	2.295	7.5	5
Sub flow$_L$	0.501	200	2048	35586	0.534	0.522	1.9	5
Sub flow$_{t4}$	0.248	200	2048	18739	0.281	0.269	0.9	4
Sub flow$_{t3}$	0.402	200	2048	16941	0.508	0.439	1.4	3
Sub flow$_{t2}$	0.481	200	2048	9031	0.542	0.464	1.9	2
Sub flow$_{t1}$	0.416	200	2048	3581	0.430	0.391	1.4	1

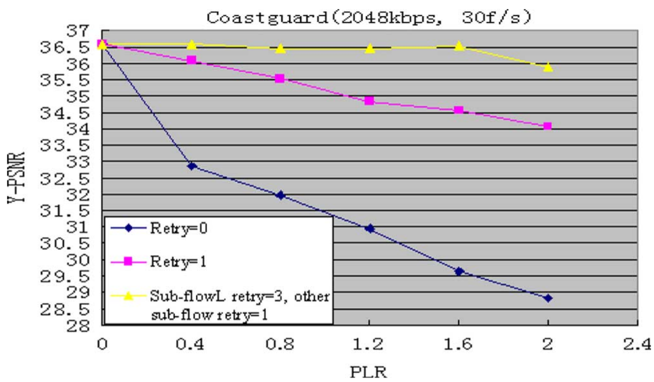

Fig. 7. PSNR performance comparison between GLOBAL TSPEC and subflow TSPEC for a given number of admitted stations (note that the PLRs of 0.4, 0.8, 1.2, 1.6, 2.0, and 2.4 correspond to 4%, 8%, 12%, 16%, 20%, and 24%, respectively).

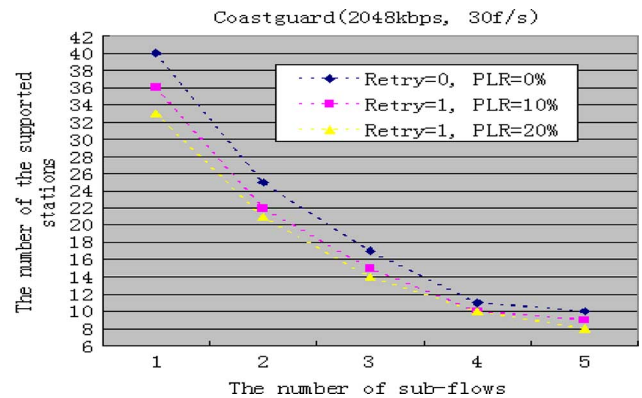

Fig. 8. Number of subflows versus the number of the stations supported by HCCA.

grouping together the various temporal subbands (frames at various temporal levels) independently. Each subflow has a different priority (determined by its distortion impact) and delay constraint (see [15] for more details). For instance, the low-pass temporal frame constitutes the highest priority subflow. A subflow has its own TSPEC parameters and is admitted independently by the ACU.

We will now show through simulations the performance of the video transmitted over IEEE 802.11a/e WLAN operating in the HCCA mode. All simulations use the Coastguard CIF sequence with a peak bit rate of 2048 kb/s and a maximum frame rate of 30 f/s, but similar results were obtained for alternative sequences. The sequences were encoded using the aforementioned MCTF-based video coder using a group of frames of 16 frames, a temporal level decomposition of four, and a spatial level decomposition of four. The low-pass temporal frames (LLLL) and other various high-pass frames (LLLH, LLH, LH, H) were grouped in individual subflows, resulting in a total of five subflows per video sequence. First, the global TSPEC was deployed for the video admission. Based on the TSPEC parameters for the video with global TSPEC, the TXOP requirement for each stream is found to be 7.5 ms. Since 80 ms of time in a beacon interval was allocated for streaming video using HCCA, a maximum of ten streams are admitted. In the second case, separate stream reservations are set up for different subflows with each flow having the same delay requirements. Simulations also assume that all stations see identical channel condition. All the TSPECs generated by the APP layer were converted into the corresponding time-based requirements. The simulation parameters are listed in the Table III.

Subsequently, a joint APP–MAC optimization needs to be performed that is aimed at maximizing the number of admitted WSTAs while providing a minimum acceptable video quality for each admitted station. This APP–MAC cross-layer optimization problem can be formulated as follows. Given the channel conditions PLR, the ACU (in a centralized scenario) or cooperating WSTAs (in a decentralized scenario) have to determine for each station the number of subflows the APP layer can transmit N_f, as well as their corresponding retransmission limits that guarantee the minimum multimedia quality Q required by the applications while maximizing the number of WSTAs N_s in the network. To answer this question, we use Figs. 7 and 8 to determine the number of subflows that have to be supported in order to maximize the number of connections while guaranteeing the minimum video quality. This can be computed offline, and then, based on the error rate performance, the admission controller at the MAC in conjunction with the APP layer can choose the optimal number of subflows that guarantees the minimum multimedia quality. To generate these results, we used the admission control strategy outlined in Section II.

Fig. 8 highlights the tradeoffs that can be performed between a higher quality for the admitted stations, i.e., a larger number of admitted subflows (along the X-axis) versus a higher number of admitted stations (the Y-axis). This example considered five subflows for each video stream emerging from a WSTA. When all the subflows are admitted, the number of stations admitted drops to nine and increases to 40 when only one subflow admission is made per WSTA. Note also that the number of admitted station is reduced when the admitted stations experience a worse channel condition (increased PLR), in response to which, they increase their retry limit per packet.

When the PLR increases, the subflow concept exhibits another advantage: the ability to differentiate the subflow based on their importance and to correspondingly adjust the unequal

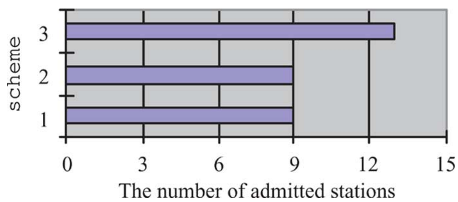

Fig. 9. Number of the admitted stations using different schemes.

error protection by adjusting the retry limit per subflow (see Fig. 7). In the global TSPEC case, the retry limit cannot be set to a value larger than 1 while admitting the entire traffic of the nine WSTAs. Hence, the resulting PLR for the admitted stations can be up to 20%, thereby resulting in a high PSNR impact. Alternatively, in the subflow case, the retry limit can be adaptively increased for individual subflows based on their distortion impact at the expense of transmitting less subflows per WSTA [23]. The granularity and prioritization provided by the subflow concept is exploited by the MAC layer to provide a more robust video transmission through conventional unequal protection techniques. For more details on these cross-layer strategies, see a variety of papers discussed in [23]. Fig. 7 illustrates that indeed, as expected, the PSNR performance is significantly better for subflow TSPEC, as compared to the global TSPEC solution.

Next, we show an additional advantage of using subflows besides providing prioritization. The various subflows also have different delays deadlines, and hence, this can be considered when admitting them individually (see [15] for more details). Fig. 9 shows the number of stations that can be supported by three different admission control schemes: two that do not consider the different delays for the various subflows and one that explicitly considers the delay deadline per subflow. In scheme 1, the global TSPEC is used, and a single delay deadline is used for all frames (i.e., 200 ms), in which the number of admitted streams/WSTAs is nine. A similar number was obtained in scheme 2 where all subflow TSPECs used the same delay constraint. In scheme 3, each subflow has a distinct delay constraint based on the dependencies among the various frames, thereby resulting in an increased number of stations.

V. Cross-Layer Protection Strategy

Having shown that the subflow concept can improve the overall system utilization, this section will present different cross-layer strategies that provide an improved multimedia quality for the admitted subflows. The cross-layer strategy proposed in this section is just illustrative to show how different cross-layer strategies can cooperate to improve the video quality [23]. The idea would be to map the APP layer subflows into MAC layer subflows (queues), thus providing the MAC the ability to adapt the unimportant subflows dynamically based on the channel condition. Our experiments consider a maximum of five subflows per multimedia connection (see [15] for more details). The cross-layer algorithm then chooses the optimal physical transmission rate based on the current channel condition using the so-called link adaptation algorithm. After choosing the optimal physical transmission rate for a given channel condition, the MAC layer computes the optimal frame size that would yield low frame error rates and pass that information to the APP layer, which would use that information to construct packets based on the recommended sizes. The MAC layer then computes the optimal retry limit that would be set on the fly on a per frame basis and recomputes the TXOP, if necessary, for the admitted TSPEC. The cross-layer algorithm is shown in Fig. 10. We will now outline the link adaptation algorithm in the next section.

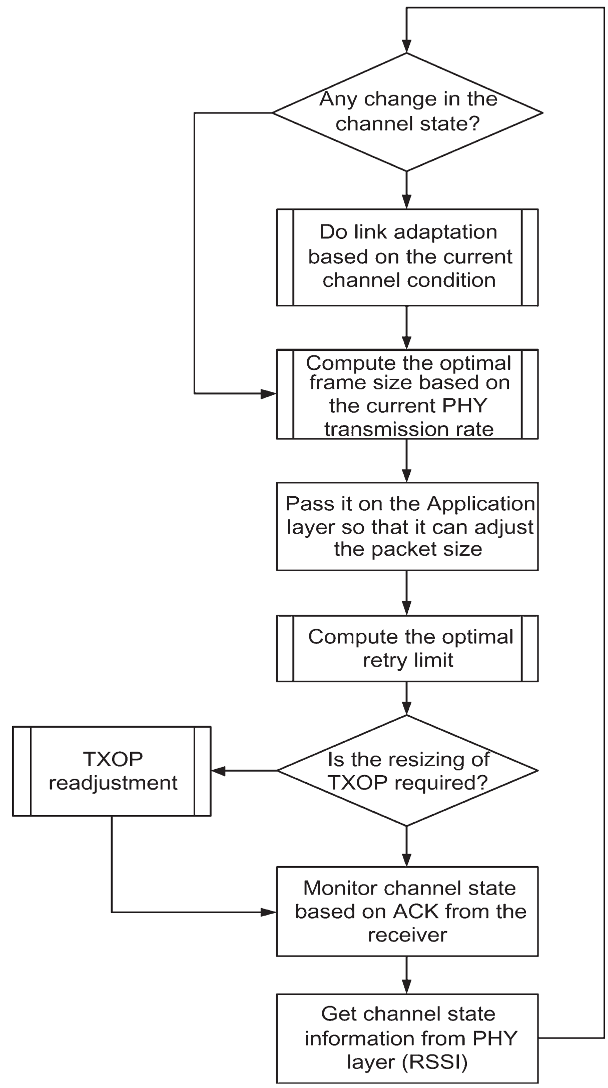

Fig. 10. Cross-layer algorithm that is performed at the MAC layer.

A. Link Adaptation Algorithm (MAC–PHY Collaboration)

As mentioned earlier, there are three different PHY layers for the IEEE 802.11 WLAN. To increase the throughput efficiency of the system under varying channel conditions, the WLAN devices need to adapt their transmission rate (or modulation rates) dynamically. In 802.11b, one can set rates 1, 2, 5.5, and 11 Mb/s, while it is possible to set eight different rates in IEEE 802.11a and IEEE 802.11g starting from 6 to 54 Mb/s. Each rate corresponds to a different modulation scheme with its own tradeoff between data throughput and distance between the stations. Fig. 11 shows the simulated performance in terms of

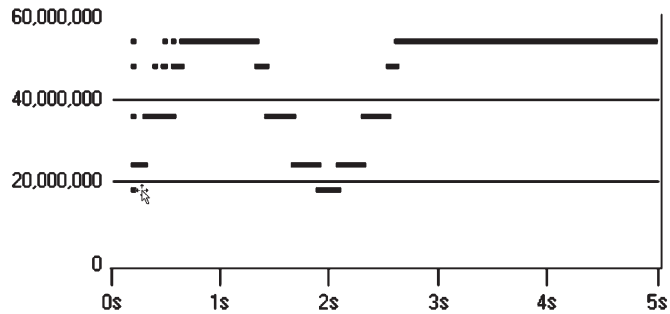

Fig. 11. Link adaptation algorithm for IEEE 802.11a.

the throughput for each modulation scheme available in IEEE 802.11b versus the SNR. Note that distance is related to SNR as SNR $\sim (1/\mathrm{dist}^{\alpha})$. Here, α is called the path attenuation factor. The higher transmission rate represents a more complex modulation scheme and, hence, offers a larger throughput, but it also has increased sensitivity to channel noise, thus providing a shorter operating range. Usually, one would want to extend the operating range as much as possible and, at the same time, to maximize the throughput. This can be done by matching the modulation scheme to the SNR or the received signal strength indicator (RSSI) available at the receiving WSTA MAC.

The cross-layer optimization problem can be formulated as follows. Given the channel conditions RSSI (e.g., in terms of SNR), determine the APP layer rate of the base layer R_{bl} and enhancement layer rate R_{el}, the MAC layer packet-size L and the PHY modulation strategy m that maximize the multimedia quality Q, i.e., find the optimal cross-layer strategy $S^{\mathrm{opt}}(\mathbf{x}) = \arg\max_S Q(S(\mathbf{x}))$ with $S(x) = \{R_{\mathrm{bl}}, R_{\mathrm{el}}, P, m\}$. Having explained the essential control knobs that can be tuned for optimal performance from the MAC perspective, we now turn our attention to the fairness in multirate wireless networks in general and show how this can influence the performance of WLAN system. This fairness is essential as it plays an important role in determining the quality of multimedia.

The link adaptation algorithm considered in this paper is based on the RSSI. The MAC uses the RSSI value of the received frames to select an appropriate link rate. The RSSI is proportional to the received power P_{recd} and is given by

$$\mathrm{RSSI} \propto P_{\mathrm{recd}}. \tag{18}$$

P_{recd} is related to the transmitted power by the following relation:

$$P_{\mathrm{recd}} \propto P_{\mathrm{trans}} \times \mathrm{dist}^{-\alpha}. \tag{19}$$

The attenuation factor assumes different values for different propagation environments. Essentially, the link adaptation algorithm at the MAC transmitter uses the RSSI value that it received at the $(k-1)$th instant to determine the rate R_i^k. Here, the superscript denotes the kth instant, and the subscript i denotes the rate $i (\in \{1, 2, \ldots, 8\})$. In the above example, we considered IEEE 802.11a PHY that has eight different modulation schemes. The link adaptation algorithm uses the latest signal strength to determine its future rate. Since there are eight modulation schemes, we will have eight different thresholds on which modulation scheme to choose for a given channel condition. Let the thresholds be denoted by Th_i. The transmission rate is R_i if the last received RSSI value is RSSI^{k-1} and is represented by the following relation: $\mathrm{Th}_i < \mathrm{RSSI}^{k-1} \leq \mathrm{Th}_{i+1}$. The boundary conditions are given by the highest and lowest transmission rates, which are chosen if the $\mathrm{RSSI}^{k-1} > \mathrm{Th}_7$ and $\mathrm{RSSI}^{k-1} \leq \mathrm{Th}_1$, respectively. Now, we can determine the probability that the transmission rate chosen is R_i at kth transmission as follows:

$$\begin{aligned} P_k(R_i) &= P\left(\mathrm{Th}_i < \mathrm{RSSI}^{k-1} \leq \mathrm{Th}_{i+1}\right) \\ &= 0.5 \left[\mathrm{erf}\left(\frac{\mathrm{Th}_{i+1} - \mathrm{RSSI}^{k-1}_{\mathrm{avg}}}{\sqrt{2}\sigma}\right) \right. \\ &\qquad \left. - \,\mathrm{erf}\left(\frac{\mathrm{Th}_{i} - \mathrm{RSSI}^{k-1}_{\mathrm{avg}}}{\sqrt{2}\sigma}\right)\right]. \end{aligned} \tag{20}$$

Now, we can easily write the transmission rate at any instance k as follows:

$$R = \sum_{i=1}^{8} P_k(R_i) R_i. \tag{21}$$

With the fixed or variable thresholds, one can easily adapt the above equation to maximize the throughput for a Gaussian channel. It is also easy to compute the outage probability that a rate chosen does not match with the channel condition resulting in frame error. The figure below shows the performance of the above mathematical model. Fig. 12 shows the maximum effective throughput obtained using different PHY mode selections for different SNR values. From the study in [24] and in Fig. 12(a) and (b), it is easy to see that the higher rate PHY modes result in better throughput performance in the high SNR range, while the lower rate PHY modes are better for the low SNR range. Another observation in Fig. 12 is that smaller packet sizes result in lower effective throughputs due to the fixed amount of MAC/PHY layer overheads for each transmission attempt. Consequently, the MAC can select the modulation strategy m at the PHY that maximizes the throughput. However, the modulation strategy m selected by the MAC–PHY is not always optimal for multimedia applications. The reason for this suboptimal performance is that the MAC-centric optimization focuses only on the throughput optimization and does not consider the resulting distortion impact. Hence, the impact on multimedia quality (distortion) needs to be explicitly considered for the cross-layer optimization.

B. Frame Length Adaptation as a Function of Channel Error Rate (MAC–PHY Collaboration)

Having selected the number of flows/subflows and optimal physical transmission rate, we will try to determine the optimal frame size that enhances the MAC layer throughput (or goodput). The MAC–PHY cross-layer optimization problem is formulated as follows. Given the channel condition, SNR, we have to determine the optimal frame size L^* that optimizes the MAC layer goodput $S^{\mathrm{opt}}(\mathbf{x}) = \arg\max_S \mathrm{Goodput}(S(\mathbf{x}))$ with $S(x) = \{L\}$.

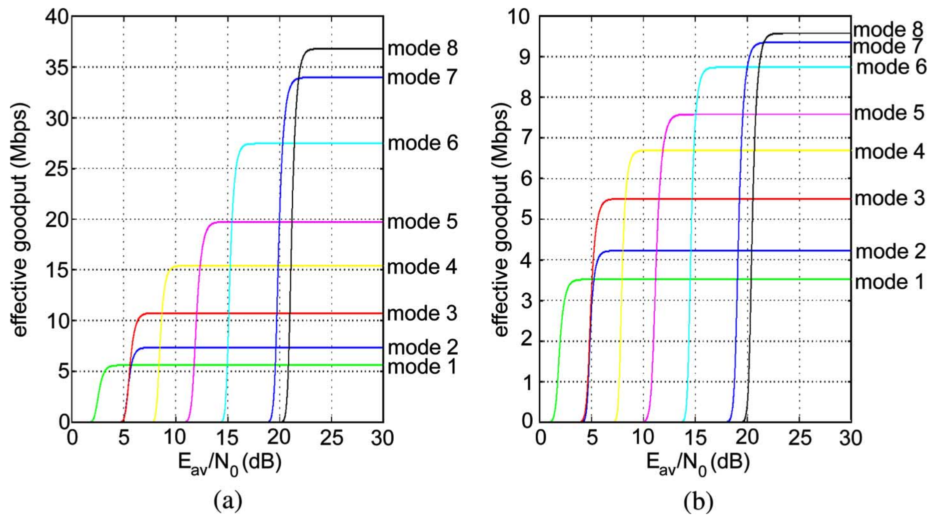

Fig. 12. Link adaptation algorithm for IEEE 802.11a that shows how frame length optimization is useful. (a) MSDU size: 2000 octets. (b) MSDU size: 200 octets.

The throughput (and hence the Goodput) of any wireless system is a function of channel error rate. For a given channel error rate, one can easily determine the optimal frame size that would keep the PLR below some thresholds. This optimization also depends on the modulation scheme that is chosen by the MAC layer.

To derive an analytical expression, we assume that the noise over the wireless channel is white Gaussian with spectral density $N_0/2$. Although this model is not a realistic assumption, we believe that the error performance analysis based on the Gaussian channel model will also hold for more realistic and complicated channel models [24]. The probability of error in a frame of size L using the modulation scheme m is a function of bit error probability p_b^m. Assuming an independent bit error rate (BER), the expression for frame error probability is given by

$$P_e^m(L) = 1 - (1 - p_b^m)^L. \tag{22}$$

Let the number of overheads that are added to the frame be represented by O. These overheads are fixed for every frame, and hence, we can get a simple expression for throughput [25] as

$$S = \frac{L}{L + O_m} * R * P_s^m(L). \tag{23}$$

Differentiating the above expression with respect to L and equating it to zero, we have

$$L^* = \frac{-O_m + \sqrt{O_m^2 - \frac{O_m}{2\log\left(1-p_b^m\right)}}}{2}. \tag{24}$$

The second differential of the (24) is negative, suggesting that the above expression for the optimal frame length maximizes the throughput of the IEEE 802.11e system for a fixed physical modulation scheme as well as bit rate. As indicated above, the MAC overhead O is fixed at 432 B. This number 432 arises by counting the time duration taken in transmitting the PLCPPreamble + PLCPHeader($= 20\ \mu$s), MACHeader($=$ 36 B), ACKduration (ACK frames are always transmitted using the lowest physical transmission rate, and their transmission duration is equal to 28 μs), PIFSTime($= 25\ \mu$s) and SIFSTime($=$ 16 μs), and translating the bits/second that would have been lost if all those times were utilized in transmitting the data at a physical transmission rate. Table IV illustrates the optimal frame sizes for different BER using 54-Mb/s PHY rate.

TABLE IV
OPTIMAL FRAME SIZES FOR DIFFERENT CHANNEL BERs

p_b	L^*
0.000006	2248.5
0.000008	1945.2
0.00001	1738.3
0.00003	997.15
0.00005	768.98
0.0001	539.35

Note that the values mentioned in Table IV are obtained with the sole purpose of maximizing the MAC throughput. Now, one needs to check if this optimal frame length determined by the MAC is really beneficial in improving the multimedia quality at the APP layer. In order to evaluate the quality of the multimedia, we pass these optimal frame sizes to the APP layer so that it can tune its frame sizes accordingly. The resultant frame size exactly meets the expectation on the frame sizes that optimize the throughput at the MAC layer. This scheme is compared with a simple scheme that keeps the frame size constant, and the multimedia PSNR is evaluated. This is summarized in Table V. The video sequence considered for illustration here is Coastguard. The sequence was coded using a Group of Pictures of 16 frames, with four spatial and four temporal levels. The fixed frame size experiment consisted of a frame of size equal to 1000 B. The experiment is conducted 12 times, and then, the PSNR mean and variations are calculated over these 12 runs. The D in this experiment is the difference between the maximum and minimum PSNR values that were obtained from the above 12 experiments.

It is clear in Table V that the optimal frame size chosen by the MAC enhances the multimedia quality. We verified the above

TABLE V
PSNR AND FRAME LOSS COMPARISONS FOR FIXED AND OPTIMAL FRAME LENGTHS WITHOUT RETRIES (COASTGUARD)

	$p_b = 0.00001$		$p_b = 0.00003$		$p_b = 0.00005$		$p_b = 0.0001$	
	L	L^*	L	L^*	L	L^*	L	L^*
Avg. PSNR	28.103	31.201	25.434	25.434	23.093	25.116	19.121	22.714
D	3.772	10.826	11.421	11.421	10.98	9.861	9.267	9.563

TABLE VI
PSNR AND FRAME LOSS COMPARISONS FOR FIXED AND OPTIMAL FRAME LENGTHS WITHOUT RETRIES (STEFAN)

	$p_b = 0.00001$		$p_b = 0.00003$		$p_b = 0.00005$		$p_b = 0.0001$	
	L	L^*	L	L^*	L	L^*	L	L^*
Avg. PSNR	30.619	30.369	24.466	24.466	21.782	24.795	20.156	22.281
D	2.617	2.2011	5.717	5.717	5.801	3.487	6.593	1.821

result by considering a variety of different video sequences, with different texture and motion characteristics. In Table VI, we present results for the Stefan sequence and can conclude that they are indeed similar to those reported in Table V.

C. Retry Limit Adaptation (MAC–PHY Collaboration)

Having optimized the number of subflows, physical transmission rates, and frame sizes, we now try to adapt the retry limit at the MAC that enhances the quality of multimedia. Setting a retry limit for an application is very important in IEEE 802.11e WLAN. The previous study [26] on setting the retry limit does not consider the delay bound parameter as set by the application. The application-like wavelet video passes the probability of packet drop, delay bound, and TSPEC element to the MAC, which, in turn, computes the retry limit. Also, one can set the retry limit separately for subflows in the SVC as each subflow negotiates its own TSPEC with the MAC. This provides unequal error protection for the different flows if the retry limits are different.

Based on the TSPEC, we calculate the effective bandwidth g. We first relate the packet drop probability p_d to the channel frame error probability p_e. This is given by

$$p_d = p_e^{N_{\text{retry}}+1}. \tag{25}$$

Here, N_{retry} is the retry limit. The above equation is valid if and only if there is no delay bound. Using N_{retry} from the above equation, we can find the average number of transmissions for a single frame until it is successfully transmitted or it is dropped after reaching the retry limit. This is given by

$$\begin{aligned} N_{\text{avg}} &= 1(1-p_e) + 2p_e(1-p_e) + 3p_e^2(1-p_e) + \cdots \\ &\quad + N_{\text{retry}} p_e^{N_{\text{retry}}-1}(1-p_e) + (N_{\text{retry}}+1)p_e^{N_{\text{retry}}} \\ &= \frac{1-p_e^{N_{\text{retry}}+1}}{1-p_e}. \end{aligned} \tag{26}$$

Then, the average time to transmit this frame using the guaranteed rate is given by

$$T_{\text{avg}} = \frac{1-p_e^{N_{\text{retry}}+1}}{1-p_e}\frac{L}{g}. \tag{27}$$

TABLE VII
OPTIMAL AND NONOPTIMAL RETRY LIMITS FOR DIFFERENT CHANNEL BERs

p_b	N^*_{retry}	N_{retry}
0.000006	10	3
0.00001	13	4
0.00003	20	5
0.00005	24	6
0.0001	30	7

Now, since we know that the delay bound as requested by this application is d, we need to satisfy the following equation:

$$N_{\text{retry}} = \frac{d}{T_{\text{avg}}}. \tag{28}$$

By using (25)–(28), we can get the value of retry limit as a function of delay bound, frame error probability, and guaranteed rate as

$$N^*_{\text{retry}} = \frac{d}{\frac{1-p_d}{1-p_e}\frac{L}{g}}. \tag{29}$$

Equation (29) provides a direct relation of the retry limit at the MAC layer as a function of application drop rate, channel error, and the delay bound. To obtain the optimal retry limits for the video sequences like Coastguard and Stefan sequences discussed in the previous section, we consider that this sequence can have a maximum tolerable delay of 200 ms. We evaluate the optimal retry limit given in (29) and (25). These retries are compared with a fixed retry for evaluating the multimedia quality.

Let us now consider the effect of retry limit on the stations multimedia quality as well as the overall system utilization. We first calculate the optimal frame size for a given channel condition. We then apply the fixed retry limit and compare it with the optimal retry limit that takes both delay and loss demanded by the application into consideration and the nonoptimal retry limit that does not depend on the delay bound demanded by the application. This is shown in Table VII to evaluate the quality of multimedia. Table VIII summarizes the quality of multimedia for a particular BER and optimal frame size evaluated from the previous section. In Table VIII, it is clear that the optimal retry limit achieves zero frame loss but

TABLE VIII
COMPARISON OF PSNR QUALITY OF MULTIMEDIA USING OPTIMAL FRAME LENGTH WITH DIFFERENT RETRY LIMITS

	$p_b = 0.000006$			$p_b = 0.00001$			$p_b = 0.00003$		
	N_{retry}^{fixed}	N_{retry}	N_{retry}^{*}	N_{retry}^{fixed}	N_{retry}	N_{retry}^{*}	N_{retry}^{fixed}	N_{retry}	N_{retry}^{*}
PSNR	35.531	35.531	36.485	35.527	35.924	36.485	35.036	35.994	36.485
Frame loss	0.9	0.9	0	1.2	0.5	0	1.8	0.3	0

	$p_b = 0.00005$			$p_b = 0.0001$		
	N_{retry}^{fixed}	N_{retry}	N_{retry}^{*}	N_{retry}^{fixed}	N_{retry}	N_{retry}^{*}
PSNR	33.783	36.191	36.485	33.781	36.308	36.485
Frame loss	1.9	0.4	0	4.1	0.8	0

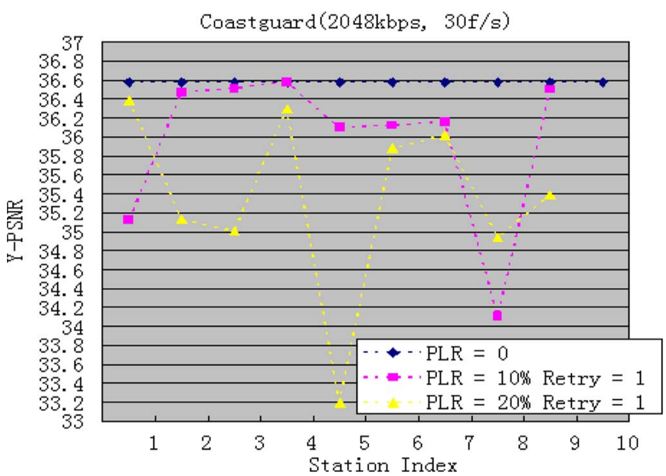

Fig. 13. PSNR performance of different video streams in the case of global TSPEC reservation. When PLR $= 10\%$ and retry $= 1$, ten stations are admitted; when PLR $= 20\%$ and retry $= 1$, nine stations are admitted.

is considered an overengineered solution. On the contrary, the nonoptimal retry limit achieves good performance in terms of PSNR and loss, and hence, it would be better to consider the nonoptimal retry limit at the MAC layer. This is a result of MAC layer trying to optimize the retry limit without considering the APP layer characteristics such as priorities, compression scheme, and encoding rate.

Fig. 13 also shows the system performance using global TSPEC but with the retry limit set to one to overcome 10% error rate. Now, the bandwidth required by each stream increases to $1.11(= 1/(1-0.9))$ times the original bandwidth. This results in increased TXOP allocation per station, and hence, the number of streams/WSTAs admitted is reduced by one. A similar reduction happens when the frame error rate is 20%, which is illustrated in Fig. 13.

VI. AIR TIME FAIRNESS

Having explained the different cross-layer protection strategies, we will look into the important concept of fairness and how it plays an important role in determining the video quality in a multirate environment. IEEE 802.11 standard supports more than one physical transmission rate. IEEE 802.11a can support eight different rates, namely, 54, 48, 36, 24, 18, 12, 9, and 6 Mb/s while IEEE 802.11b can support four different rates, namely, 11, 5.5, 2, and 1 Mb/s. As explained in Section V-A, depending on the channel conditions, particularly their distance from the QAP, WSTA may choose different transmission rates (i.e., "link adaptation") in order to increase the probability of successful transmission. For example, station i may choose 2 Mb/s to transmit/receive data frames to/from the QAP while station j chooses 11 Mb/s. Defining a "fair share of resource" among the stations in such a network is a very challenging problem, because serving an equal amount of traffic from individual stations with different transmission rates requires allocation of different amounts of airtime. In other words, a fair share of system throughput is no longer synonymous with a fair share of airtime in a system that supports multitransmission rates. Since no such location-dependent transmission rates exist in wired networks, applying the existing scheduling algorithms without considering this property may result in a station's abuse of radio resource. Multimedia applications suffer most from airtime unfairness due to their delay sensitivity, which rules out rate and congestion control. For instance, assume that both WSTAs are running broadcast-quality standard-definition video streaming sessions. WSTA2 has dropped its PHY rate to 1 Mb/s due to moving away from the QAP. Thus, WSTA2 can no longer sustain its video streaming application. However, the bandwidth fairness causes the throughput of WSTA1 to drop to the same rate as WSTA2 even though nothing has changed for WSTA1, leading to unfair resource usage. This observation has led us to the use of air or time fairness [12], in which each flow may use a different transmission rate. Let us now categorize the effects of time unfairness analytically. Let WSTAs 1 and 2 have continuously backlogged traffic. Assume the frame sizes of individual streams be fixed to size L.

The objective of fair scheduling [13] is to provide multimedia applications with different amounts of "work" (resources) proportional to their requirements in terms of bandwidth, delay, and packet-loss rates. Usually, "work" is measured by the amount of data transmitted (either in number of bytes or packets/frames) during a certain period of time. Let $W_i(t_1, t_2)$ be the amount of video flow i's traffic served in a time interval (t_1, t_2), and let ϕ_i be its corresponding weight based on its requirements. Then, an ideal fair scheduler (i.e., the Generalized Processor Scheduler [13]) for N WSTAs (and their flows) can be defined as follows:

$$\frac{W_i(t_1,t_2)}{W_j(t_1,t_2)} \geq \frac{\phi_i}{\phi_j}, \qquad j = 1, 2, \ldots, N \tag{30}$$

for any multimedia flow i that is continuously backlogged during (t_1, t_2) [backlogged means that flow i has frames in its buffer during the specified time interval (t_1, t_2)]. If all multimedia flows are transmitted at a fixed rate, we can obtain, from (30)

$$\frac{W_i(t_1,t_2)}{t_2 - t_1} \geq \frac{\phi_i}{\sum_j \phi_j} r \tag{31}$$

where r is the physical transmission rate or the total channel capacity. Thus, each multimedia flow i is guaranteed to have the throughput given by (31), regardless of the states of the queues and frame arrivals of the other flows.

We will now evaluate the total throughput degradation due to WSTAs employing link adaptation (e.g., different PHY rates) in the WLAN network. Consider n WSTAs (with all stations having the same frame size), with $n_i(\sum_{i=1}^{8} n_i = n)$ WSTAs operating at, e.g., PHY mode $i(= 1, \ldots, 8)$, the throughput degradation can be determined as [14]

$$\text{Throughput} = \frac{1}{\frac{1}{n}\left(\sum_{j=1}^{8} \frac{n_j}{R_j}\right)} \tag{32}$$

where WSTAs have different transmission rates R_j due to the different PHY modes or other deployed cross-layer optimization strategies that cause this unfairness. Equation (32) can be modified to represent the influence of frame size adopted by different stations having different transmission rates. Let us assume that a station chooses a frame of size L_i corresponding to the physical transmission rate R_i. The above equation that represents the unfairness caused by different transmission rates but having the same frame size is modified to represent the unfairness caused because of different frame size as well as different transmission rates as

$$\text{Throughput} = \frac{\sum_{j=1}^{8} n_j L_j}{\frac{1}{n}\left(\sum_{j=1}^{8} \frac{n_j L_j}{R_j}\right)}. \tag{33}$$

The above equation can be further generalized when each WSTA has k_i different frame sizes as

$$\text{Throughput} = \frac{\sum_{i=1}^{8} \sum_{j=1}^{k_i} n_i L_{ij}}{\frac{1}{n}\left(\sum_{i=1}^{8} \sum_{j=1}^{k_i} \frac{n_{ij} L_{ij}}{R_j}\right)}. \tag{34}$$

To solve this problem, we propose the concept of time fairness [14], [15]. In this concept, each WSTA is given a fair share of time proportional to the requirements mentioned in, e.g., their TSPEC, rather than guaranteeing the bandwidth. This time allocation (e.g., allocated to a stream at admission time) removes the unfairness due to deploying different cross-layer strategies. Equation (30) can be thus rewritten to impose time fairness as

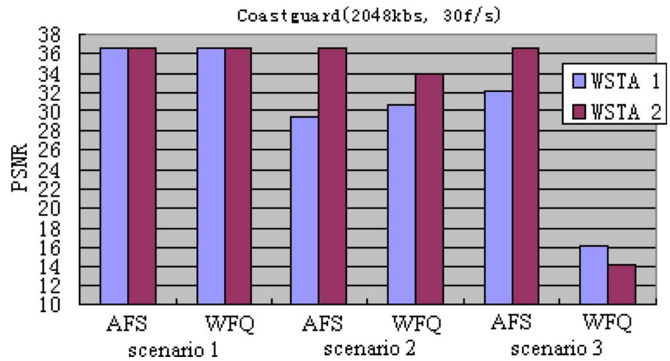

Fig. 14. PSNR performance of AFS and WFQ in different scenarios.

$$\frac{T_i(t_1,t_2)}{T_j(t_1,t_2)} \geq \frac{\phi_i}{\phi_j}, \qquad j = 1, 2, \ldots, N \tag{35}$$

where T_i and T_j represent the time allocated to the streams i and j, respectively. We consider the transmission time allocated to each flow, not the amount of traffic served during that interval, as we want to eliminate the effect of different transmission rates (i.e., the potential unfairness it may cause). From (35), we can obtain

$$T_i(t_1,t_2) = \frac{\phi_i}{\sum_j \phi_j}(t_2 - t_1). \tag{36}$$

Equation (36) states that a flow, which is continuously backlogged in the arbitrary time period (t_1, t_2), is guaranteed to receive a certain portion of transmission time given by the above equation rather than the bandwidth guaranteed by (31).

The bandwidth fairness concept is translated into a practical algorithm called weighted fair queuing (WFQ), and the time fairness is translated to an implementable algorithm called air fair scheduling (AFS). The advantages of the proposed AFS [15], as opposed to the conventional WFQ, are highlighted in Fig. 14. The air or time fairness concept can be employed for every video coder. In this paper, we show the benefits of air/time fairness using a wavelet video coder. In the first simulation scenario, the same cross-layer strategies, resulting in the same transmission rate and no losses, are deployed for both WSTAs, and the performance is compared using WFQ and AFS. WFQ and AFS result in the same PSNR values as shown by the first four bars in the left-hand side of Fig. 14. Consider the second simulation scenario: WSTA1 experiences more frame errors because of interference and fading. The PLR has increased, and it takes on average 50% more time to transmit a frame from WSTA1 than from WSTA2. In the conventional WFQ, this would mean that the "start-of-service time" of frames in WSTA2 is deferred resulting in QoS violation and dropping of packets at the MAC layer. This directly affects the PSNR performance as most of the higher priority packets are dropped for both the WSTAs. Using AFS, the stream between the WSTA1 and AP alone is affected, yielding low PSNR, whereas the other WSTA2 is not affected because of WSTA1's channel error condition. In the third simulation scenario, the WSTA1 moved far away from the AP, and the cross-layer strategy switched the PHY mode to a more robust modulation scheme. Since the physical transmission rate of WSTA1 has dropped, it would take more time to transmit the frame, and the same problem of deferred start-of-service happens for both WSTAs in case of WFQ. However, the AFS isolates the channel and differential transmission rates to WSTA1, thus guaranteeing the multimedia performance.

The throughput degradation can be clearly measured when different WSTAs in WLAN deploy different PHY modes by simply subtracting (32) from (31). This unfairness is caused by stations having lower data rates compared to stations having higher data rates. However, if we measure the throughput as the time goes to infinity, the throughput carried by multirate WLAN and unirate WLAN is the same. Thus, the performance of the WLAN network is impacted if there are WSTAs that operate in multirate. In order to solve this problem, the concept of time fairness is introduced. In this concept, each WSTA is given a fair share of time proportional to the requirements mentioned in TSPEC rather than guaranteeing the bandwidth as the means proportional to its weight share. This time allocation removes the unfairness caused to in the WLAN system because of differential PHY rates. This is the allocated TXOP to a stream at admission time.

VII. Conclusion

In this paper, we addressed the important problem of delay-sensitive transmission of video over IEEE 802.11a/e WLANs. We outline the admission control scheme that uses a new parameter called "channel burstiness" to model the first-order characteristics of the time-varying wireless channel and then use effective bandwidth calculations to admit the optimal number of video streams. We showed how different access mechanisms in IEEE 802.11e can provide the necessary QoS for the admitted video applications. Additionally, we highlight the differences between the two modes, EDCA and HCCA, in the IEEE 802.11e WLAN system and show the relative merits and demerits. We also introduced the concept of subflow concept that enhances the quality of video at the receiver in the presence of errors. Then, we illustrate a simple cross-layer strategy using subflows at the APP, frame length, retry limit at the MAC, and modulation schemes at the PHY to improve the quality of multimedia. We showed through analysis and simulation that controlling the contention-based access in IEEE 802.11e is simpler to realize in real products and how different cross-layer strategies used in polled access lead to a larger number of stations being simultaneously admitted and/or a higher video quality for the admitted stations. Finally, we introduce the concept of air/time fairness that enables a fair division of resources among competing WSTAs when different cross-layer strategies are deployed.

References

[1] *IEEE 802.11b: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard, 1999.

[2] *Draft Supplement to Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS),* IEEE 802.11e/D10.0, Nov. 2004.

[3] D. Majumdar, G. Sachs, I. V. Kozintsev, and K. Ramchandran, "Multicast and unicast real-time video streaming over wireless LANs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 524–534, Jun. 2002.

[4] Y. Pei and J. Modestino, "Multi-layered video transmission over wireless channels using an adaptive modulation and coding scheme," in *Proc. IEEE ICIP*, Thessaloniki, Greece, Oct. 2001, pp. 1009–1012.

[5] Y. Shan and A. Zakhor, "Cross layer techniques for adaptive video streaming over wireless networks," in *Proc. IEEE ICME*, 2002, vol. 1, pp. 277–280.

[6] W. Kumwilaisak, T. Hou, Q. Zhang, W. Zhu, C.-C. Jay Kuo, and Y.-Q. Zhang, "A cross-layer quality of service mapping architecture for video delivery in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1685–1698, Dec. 2003.

[7] J. Shin, J. G. Kim, J. Kim, and C.-C. J. Kuo, "Dynamic QoS mapping control for streaming video in relative service differentiation networks," *Eur. Trans. Telecommun.*, vol. 12, no. 3, pp. 217–230, May/Jun. 2001.

[8] B. Girod, M. Kalman, Y. Liang, and R. Zhang, "Advances in channel-adaptive video streaming," *Wirel. Commun. Mob. Comput.*, vol. 2, no. 6, pp. 549–552, Sep. 2002 (Invited Paper).

[9] A. Reibman and M.-T. Sun, Eds., *Compressed Video Over Networks*. New York: Marcel Dekker, 2000.

[10] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Commun. Mag.*, vol. 41, no. 10, pp. 74–80, Oct. 2003.

[11] P. Ansel, Q. Ni, and T. Turletti, "An efficient scheduling scheme for IEEE 802.11e," in *Proc. IEEE Model. Optimization Mobile, Ad Hoc Wireless Netw. Workshop (WiOpt)*. Cambridge, U.K.: Univ. Cambridge, Mar. 24–26, 2004.

[12] S. Shankar N, "Method, access point and program product for providing bandwidth and airtime fairness in wireless networks," World Patent No. WO2005011199, May 3, 2006.

[13] A. K. Parekh and R. G. Gallager, "A generalized process sharing approach to flow control in integrated services networks—The single node case," in *Proc. IEEE INFOCOM*, 1992, pp. 915–924.

[14] S. Shankar N, R. Chen, R. Schmitt, C. T. Chou, and K. G.Shin, "Air fair scheduling for multimedia transmission over multi rate wireless LANs," in *Proc. SOIA*, Eindhoven, The Netherlands, 2004.

[15] S. Shankar N, Z. Hu, and M. van der Schaar, "Cross-layer optimized transmission of wavelet video over IEEE 802.11a/e WLANs," in *Proc. IEEE Packet Video*, Irvine, CA, Dec. 2004.

[16] M. van der Schaar and S. Shankar N, *Cross-layer Design for Wireless Multimedia Transmission*, Jun. 2005. accepted book chapter.

[17] M. Zorzi, R. R. Rao, and L. B. Milstein, "On the accuracy of a first-order Markov model for data transmission on fading channels," in *Proc. IEEE ICUPC*, Nov. 1995, pp. 211–215.

[18] D. Wu and R. Negi, "Effective capacity-based quality of service measures for wireless networks," *ACM Mob. Netw. Appl.*, vol. 11, no. 1, pp. 91–99, Feb. 2006.

[19] C. Liu and J. Leyland, "Scheduling algorithms for multiprogramming in hard real-time environment," *J. ACM*, vol. 20, no. 1, pp. 46–61, Jan. 1973.

[20] C. T. Chou, S. Shankar N, and K. G. Shin, "Per-stream QoS in the IEEE 802.11e Wireless LAN: An integrated airtime-based admission control and distributed airtime allocation," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005, pp. 1584–1595.

[21] C. T. Chou, S. Shankar N, and K. G. Shin, "Contention-based airtime usage control in multirate IEEE 802.11 wireless LANs," in *IEEE/ACM Trans. Netw.*, vol. 14, Dec. 2006, pp. 1179–1192.

[22] J.-R. Ohm, M. van der Schaar, and J. W. Woods, "Interframe wavelet coding—Motion picture representation for universal scalability," *Signal Process. Image Commun.*, vol. 19, no. 9, pp. 877–908, Oct. 2004.

[23] M. van der Schaar and S. Shankar N, "Cross-layer wireless multimedia transmission: Challenges, principles and new paradigms," *IEEE Wireless Commun. Mag.*, vol. 12, no. 4, pp. 50–58, Aug. 2005.

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

[25] M. Schwartz, *Telecommunication Networks: Protocols, Modeling, and Analysis*. Reading, MA: Addison-Wesley, 2005.

[26] Q. Li and M. van der Schaar, "Providing adaptive QoS to layered video over wireless local area networks through real-time retry limit adaptation," *IEEE Trans. Multimedia*, vol. 6, no. 2, pp. 278–290, Mar. 2004.

[27] A. Grilo, M. Macedo, and M. Nunes, "Scheduling algorithm for QoS support in IEEE 802.11e networks," *IEEE Wireless Commun. Mag.*, vol. 10, no. 3, pp. 36–43, Jun. 2003.

Sai Shankar N (M'99–SM'04) received the Ph.D. degree from the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, in the area of ATM networks.

In 1999, he joined Philips Research, Eindhoven, The Netherlands, where he served as a Research Scientist, working on various problems involving hybrids, fiber, coaxial cable (IEEE 802.14) networks, and IP protocols. In 2001, he joined Philips Research, Briarcliff Manor, NY, where he worked in the area of wireless LANs/UWB, cognitive radios, and cooperative communications. He was an active contributor of the wireless LAN standard and has submitted more than 15 proposals that have shaped QoS related issues in the IEEE 802.11e Standard. He is also an active participant in the Ultra-Wideband (UWB) Working Group of the WiMEdia Alliance and has contributed to shaping the new MAC at the Multiband OFDM Alliance (MBOA) Forum. He has been the reviewer for almost all important journals in the area of networking and has chaired many conferences. Currently, he is with Qualcomm, Inc., San Diego, CA, researching UWB, cognitive radios, and cooperative communications. He has authored more than 45 conference and journal papers and is the holder of more than 40 patents.

Dr. Shankar N was awarded the German Fellowship, DAAD, from the Department of Mathematics, University of Kaiserslautern, Kaiserslautern, Germany, to work on queueing approaches in manufacturing in 1998. He was nominated as one of the five finalists for the Innovator of the Year category by EETimes.

Mihaela van der Schaar (SM'04) received the Ph.D. degree from Eindhoven University of Technology, Eindhoven, The Netherlands, in 2001.

Prior to joining the faculty of the Electrical Engineering Department, University of California at Los Angeles, on July 1, 2005, she was a Senior Researcher with Philips Research in the Netherlands and in the U.S. between 1996 and June 2003, where she led a team of researchers working on multimedia coding, processing, networking, and streaming algorithms and architectures. From July 1, 2003 to July 1, 2005, she was an Assistant Professor with the Electrical and Computer Engineering Department, University of California, Davis. She has published extensively on multimedia compression, processing, communications, networking, and architectures. She is the holder of 27 granted U.S. patents and has several more pending.

Dr. van der Schaar was elected as a member of the Technical Committee on Multimedia Signal Processing of the IEEE Signal Processing Society. Since 1999, she has been an active participant in the ISO Motion Picture Expert Group (MPEG) standard to which she made more than 50 contributions and received three ISO recognition awards. She chaired for three years, the *ad hoc* group on MPEG-21 Scalable Video Coding, and also cochaired the MPEG *ad hoc* group on Multimedia Test-bed. She was an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA and the *SPIE Electronic Imaging Journal* from 2002 to 2005. Currently, she is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS. She received the NSF CAREER Award in 2004, the IBM Faculty Award in 2005, and the Best Paper Award for her paper published in 2005 in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.