Token- and Self-Policing-Based Scheduling for Multimedia Traffic Transmission Over WLANs

Polychronis Koutsakis, *Member, IEEE*

***Abstract*—The worldwide popularity of wireless local area networks (WLANs) calls for efficient solutions in scheduling multimedia (voice, data, and video) traffic transmissions. Enhanced distributed channel access (EDCA), which is the contention-based channel access function of IEEE 802.11e, is unable to guarantee priority access to higher priority traffic in the presence of significant traffic loads from low-priority users. In this paper, we propose the use of a token- and self-policing-based scheduling scheme, which not only addresses this problem but also prevents bursty video nodes from overusing the medium and tackles the problem of idle time due to large transmission opportunities (TXOPs).**

***Index Terms*—IEEE 802.11e, medium access control (MAC), MPEG-4 video, multimedia traffic, wireless local area networks (WLANs).**

I. Introduction

THE legacy IEEE 802.11 [1] standard uses two scheduling schemes. These are the distributed coordination function (DCF) and the point coordination function (PCF). Medium access in PCF is contention free. DCF is a random access scheme based on carrier-sense multiple access (CSMA) with collision avoidance. Hence, in DCF, a station transmits if it observes an idle medium. If a collision occurs, the station transmits again after choosing a backoff time between $[0, \mathrm{CW}-1]$, where CW is the contention window size.

An important disadvantage of DCF is its inability to provide quality of service (QoS) differentiation among different types of traffic. The number of multimedia applications is constantly increasing, and they have different and often contradictory QoS requirements. Hence, network service needs to be tailored to the characteristics and needs of each type of traffic. In this way, system throughput can increase, whereas QoS requirements are satisfied.

The IEEE 802.11e [2] enhances the legacy 802.11 medium access control (MAC) and improves the possibility of service differentiation among high- and low-priority traffic. This is achieved with the introduction of four access categories (ACs), i.e., background (BK), best effort (BE), video, and voice, in increasing priority order. The ACs are differentiated via the use of

1) Arbitration interframe space (AIFS), which is the minimum time interval that a station needs to sense that the medium is idle before transmitting. The difference between enhanced distributed channel access (EDCA, which is the contention-based channel access function of 802.11e) and DCF lies in the fact that AIFS differs in EDCA, depending on the AC. ACs with higher priority have a smaller AIFS than those with lower priority; therefore, they can contend earlier to gain access to the medium.
2) CW. The values of $\{\mathrm{CW}, \mathrm{CW}_{\min}$ and $\mathrm{CW}_{\max}\}$ again are used to favor higher priority over lower priority traffic. CW values are smaller for higher priority traffic; therefore, lower priority users need to wait longer to retransmit than high-priority users after a transmission failure.
3) TXOP, which is a bounded time interval indicating the maximum amount of time for which a terminal can initiate transmissions. TXOP is again different for each AC. A TXOP that is equal to zero means that the terminal can only transmit a single frame.

Although EDCA improves the performance of DCF in terms of being able to prioritize traffic, it still provides only *statistical* priority access; it cannot guarantee priority access to high priority traffic, particularly in the presence of significant traffic loads from low-priority users. Hybrid coordination function controlled channel access (HCCA, which is the polling-based channel access function of 802.11e) does provide guaranteed services and therefore outperforms EDCA when centralized access is possible. This is shown in numerous works in the literature in the recent past, e.g., [3], in spite of HCCA's overhead and complicated software architecture. However, recent work [4], [5] shows that the performance of EDCA can be clearly better than that of HCCA for variable-bit-rate video streams, particularly in multicollision domains where access points (APs) in neighboring basic service sets poll the stations in the overlapping area, resulting in collisions.

In this paper, we propose a new scheduling scheme for multimedia traffic over wireless local area networks (WLANs). Our scheme focuses on the case without a central controller and uses token passing and self-policing to provide guaranteed priority access to high-priority traffic. It is compared with EDCA and shown to significantly improve channel utilization, under both light and heavy videoconference traffic loads.

II. Related Work

The majority of existing work on MAC protocols for WLANs focuses on the transmission of integrated voice and data traffic. The problem of transmitting video traffic along with voice and data has received attention only in the past few years [6]–[11].

Manuscript received September 27, 2010; revised February 16, 2011, March 31, 2011 and June 2, 2011; accepted August 25, 2011. Date of publication September 19, 2011; date of current version December 9, 2011. The review of this paper was coordinated by Prof. P. Langendoerfer.

The author is with the Department of Electronic and Computer Engineering, Technical University of Crete, 73100 Chania, Greece (e-mail: polk@telecom.tuc.gr).

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.2011.2168572

Regardless of the types of traffic integrated, most of the work in the field [7]–[13] uses CWs to resolve the scheduling problem in WLANs, following the steps of the IEEE DCF and EDCA. An exception to this approach was proposed in [6], where the authors suggested that real-time traffic should be segregated from nonreal-time traffic and transmitted in a contention-free period. This approach, however, requires significant overhead for the transmission of polling frames from the AP to the stations transmitting real-time traffic. In addition, the authors in [6] and [11] proposed alternate approaches for defining the length of the TXOP. In [6], the TXOP is calculated based on the number of MAC service data units (MSDUs) in the current queue of each station. In [11], the authors use a window w of already known real queue length measurements (history information) to tune their estimation of the TXOP. However, both of these approaches are insufficient for bursty video traffic. The reasons are that the current queue length may be irrelevant to the size of the next video frame, leading to a quite false estimate, and history information does not provide an adequate estimation on the future behavior of the video source, particularly for short video sequences.

The idea of a token-based scheduling scheme, which practically eliminates collisions and hence increases channel utilization, was first proposed in [14], where a neat solution was presented for the transmission of voice and data traffic. The authors in [14] showed that the token-based scheduling approach achieved better results than DCF in terms of channel utilization, when integrating voice and data traffic over WLANs.

In this paper, we extend the work presented in [14] to study the more complex problem of integrating bursty video traffic with voice and data traffic over WLANs.

III. Proposed Scheduling Scheme

The work presented in [14] proposed an efficient token-based scheduling scheme for the transmission of voice and data traffic in a fully connected WLAN without a central controller, where all the nodes can hear each other. The addition of video traffic among the traffic types, which need to be integrated over the WLAN, makes the scheduling problem quite more complex. Our proposed scheme incorporates two of the three EDCA enhancements of DCF, i.e., AIFS and TXOP. We propose the following additions/changes to EDCA and [14], and we "translate" some of the ideas presented in [14] in the context of EDCA, as [14] was compared against DCF and did not include the EDCA enhancements.

A. Self-Policing for Video Users

There are three tokens in the system in our proposed scheme, as opposed to two in [14], since that work does not consider video traffic. These three tokens will be named "permission tokens" for the rest of this paper for reasons that will be explained here. The first permission token is circulated among voice nodes, the second among video nodes, and the third among data nodes. When a node holds the token, it will transmit its packet(s) when the channel is available. As in [14], a voice node transmits all its backlogged packets after obtaining the voice token. The portion of channel time unused by voice nodes is shared in our scheme by video nodes first and then by data nodes (BE and BK, respectively). A data node is assigned a maximum channel occupancy time, which is equal for all data nodes. During this time, the data node can transmit one or multiple packets, depending on its packet size and transmission rate. In our scheme, we assign a TXOP equal to zero for BK and BE traffic.

The proposed scheme works in a distributed manner; there is no central controller passing the tokens to others. The current token holder decides the next token holder. When a backlogged node holds the token, it piggybacks the token in its voice/data packet transmission and passes it to the next node. When a data token holder has no packet to transmit or a voice token holder changes from the ON-state to the OFF-state, the node directly passes the token to the next holder.

The aforementioned ideas, which were proposed in [14] for the transmission of voice and data packets, cannot be used for video nodes, however, as they do not take into account the burstiness of video traffic. If a video node was allowed to transmit all its backlogged packets, it could greedily occupy the channel for a significant amount of time, in case of a burst. On the other hand, if a strict TXOP value was defined for all video nodes (as for data nodes), this could lead to unfairness for a video node. For example, this could occur in the case of a node transmitting at a lower rate than its declared mean for a while and now needing to transmit a significantly larger video frame, e.g., a new I frame denoting a significant scene change.

To solve this problem, our scheme works as follows: When a video node obtains the video permission token, it does not transmit all its backlogged packets before sending the permission token to the next node. Instead, assuming nonselfish nodes, we propose the use of *self-policing* in each node, based on the accurate video traffic model presented in Section IV. Each node runs a jumping window (JW) policer [15], which is known in the literature and shown in [16] to be the most lenient traffic policing mechanism among other mechanisms studied. The JW mechanism uses windows of a fixed length T side by side through time. A new window immediately starts after the conclusion of the previous window. During a window, only K bytes (or packets) can be submitted by the source to the network. If a source attempts to transmit more than K bytes, the excessive traffic is dropped or marked as nonconforming, as in the case of the Token Bucket. The mechanism is implemented with the use of a transmission token counter. The number of transmission tokens is the equivalent of the number of packets that the user is allowed to transmit and should not be confused with the permission tokens, representing the turn of the user to transmit. In each new window, the associated packet counter is restarted with an initial value of zero.

In our study, we use a modification of the JW mechanism to implement a more dynamic mechanism: in the case that less than K bytes are transmitted by the source within one window, the token counter is not restarted but starts with an initial value equal to the remaining tokens. In addition, we use, from [16], our idea to generate tokens based on our video traffic model. Hence, our mechanism generates as many tokens as the model estimates that the user will need, instead of using a fixed (static) token generation rate based on the video user's declared mean rate. This dynamic approach was shown in [16] to outperform static traffic policing.

The JW mechanism is not used in our scheme to drop or mark excessive traffic but to control each user's TXOP. Therefore, each user's TXOP is equal to the time needed for the transmission of the number of packets that the user is "allowed" by its policer to send. This means that, in our proposed scheme, contrary to the approach of EDCA (for all types of traffic) and [14] (for voice/data traffic), *the TXOP is not the same for all video nodes*. To the best of our knowledge, the idea of using variable TXOPs per user, the value of which is controlled via a traffic-policing-like mechanism, is proposed for the first time in the relevant literature. This approach solves, as will be shown from our simulation results, the aforementioned problem of how to define TXOP for video nodes. It is combined in our work with the idea of token passing, which solves the well-known problem of EDCA, where a large number of stations from the same AC (and, hence, the same AIFS and TXOP) can lead to high collision probability and lower channel utilization.

B. Access Priority and Dynamic Token Passing for Multimedia Traffic

EDCA assigns the same AIFS, which is equal to 2, for the video and voice categories, whereas the values for BE and background traffic (BE) are 3 and 7, respectively. The fact that voice traffic is considered of higher priority than video traffic is expressed via the values of $\mathrm{CW}_{\min}$ and $\mathrm{CW}_{\max}$ for each type of traffic (smaller values for voice nodes). Since our proposed scheme does not use CWs, as it practically eliminates contention with the use of tokens, a different mechanism needs to be implemented to enforce voice priority. For this reason, we change the AIFS values of AC_VI and AC_BE, as shown in Table I.

Similar to the procedure followed for new voice nodes in [14], when a new video user enters the network, it waits for the channel to be idle for $T_{\mathrm{NEWVID}} < \mathrm{AIFS}(\mathrm{AC\_VI})$ and transmits. Video arrivals are Poisson distributed, and nodes broadcast a JOIN message and a LEAVE message when they arrive and depart from the network, respectively. In the pseudocode shown in Table II, we present the token-passing procedure in the case of a new video node, as well as in the case of token initialization.

If a video node leaves the WLAN or ends its transmission, the node sends a message to announce this and passes the token to the next video node. The previous video token holder makes note of this, so that it will not send the token to the departing video node again in the future. In addition, the same token initialization procedure as that previously described is followed if a node that has already transmitted "crashes" and does not send the LEAVE message.

In the case that a node (e.g., node D) is affected by localized noise and does not hear the token transmission, it needs to monitor channel activity upon returning to the "good state" because it has a different view of the current network status. If node D was not the one to which the token was sent, no further action is required. If, however, node D was the one that the token was sent to and did not correctly receive the token because of the localized noise, then the node (e.g., node C) that passed the token to node D monitors the activity of node D, and if no activity is observed, node C resends the token. A number of failed consecutive retransmissions due to a large duration of the localized noise leads node C to pass the token to the next node.

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Slot	20 µs
SIFS	10 µs
PHY preamble	192 µs
RTS frame size	20 bytes
CTS frame size	14 bytes
Token frame size	36 bytes
Channel Rate	11 Mbps
Basic Rate	2 Mbps
AIFS(AC_BK)	7
AIFS(AC_BE)	4
AIFS(AC_VI)	3
AIFS(AC_VO)	2
T_{NEWVID}	30 µs
TXOP(AC_BK)	0
TXOP(AC_BE)	0
TXOP(AC_VI)	Variable per node
TXOP(AC_VO)	1504 µs
Error Rates	10^{-1} and 10^{-5}

We use the idea proposed in [14] for proportional class differentiation, and we implement it for the two data ACs (BE and BK) with a differentiation ratio of 2:1. Other ratios were also used in our simulations, without affecting the nature of our results, which are presented in Section V. The data-token-passing process can be modeled by a stationary Markov chain, the steady-state probabilities of which ensure the existence of proportional class differentiation among data classes. With the Poisson arrival assumption, the packet arrival and departure at each data node can be modeled by an M/G/1 queue. A data token holder randomly chooses the next token holder based on the transition probabilities.

It is possible that a collision will happen when two or more video nodes enter the network and transmit simultaneously after the channel is idle for T_{NEWVID}. The p-persistent CSMA can be used, as in [14] for voice, to resolve rare video node collisions, which are shown via our simulations to occur with a very small (practically negligible) probability. Due to the fact that the number of video nodes in the network is much smaller than that of voice nodes, as video nodes demand significantly larger bandwidth, collisions between video nodes are much more rare than the rare collisions between voice nodes reported in [14]. Finally, we adopt from [14] the mechanism for recovery of lost tokens due to the unreliable wireless channel.

TABLE II
PSEUDOCODE FOR THE TOKEN-PASSING PROCEDURE

```
if (i=new video user)
        i waits for idle channel for T_NEWVID
        i transmits
        if (token exists) /* j=current video token holder, k=previous video token  holder, h=current  data token holder */
                j and h sense busy channel, defer transmissions
                k monitors channel, puts i as the next node to pass the token, in his list
                i finishes its self-policed transmission, monitors channel
                j waits for idle channel for AIFS(AC_VI)
                j transmits
                i puts j as the next node to which it will pass the token
        else            /* no token */
                i finishes its self-policed transmission
                i monitors channel, waits for AIFS(AC_VI) /* hears no transmission */
                i creates the first token
```

TABLE III
VIDEO TRACE STATISTICS AND HTTP TRAFFIC MODEL

Video Trace	**Mean**	**Standard Deviation**	**Peak**
Office Cam	400 Kbps	434 Kbps	2 Mbps
Lecture Room Cam	210 Kbps	182 Kbps	1.5 Mbps
N3 Talk	550 Kbps	329 Kbps	3.4 Mbps
ARD Talk	540 Kbps	346 Kbps	3.1 Mbps
Http Traffic Model			**Distribution**
Size of web request	360 bytes	106.5 bytes	Lognormal(5.84, 0.29)
Number of web requests per www session	25 pages	100 pages	Lognormal (1.8, 1.68)

IV. VIDEO TRAFFIC MODEL

In this section, we briefly present our model for single MPEG-4 videoconference traces, which is used in our scheme, as explained in Section III. In [17], we have investigated the possibility of modeling MPEG-4 videoconference traffic with quite a few well-known distributions. The results of our detailed statistical tests have shown that the use of the gamma distribution, which was the most commonly used in the literature for traces encoded with previous technology encoding schemes, is not a good choice. We found that the best fit among these distributions is achieved for all the studied traces with the use of the Pearson V distribution. However, the degree of goodness-of-fit for the Pearson V varied for all traces. The reason that the Pearson V distribution fit cannot be highly accurate is that the high autocorrelation between successive video frames in a videoconference trace can never be perfectly "captured" by a distribution generating frame sizes independently, according to a declared mean and standard deviation. Therefore, none of the fitting attempts, as good as they might be, can achieve perfect accuracy. Another important result from our work in [17] was that, to provide a good fit, I, P, and B frame sizes need to be separately modeled.

In [16], we studied the problem of traffic policing for H.263 [18] videoconference traffic transmission over wireless cellular networks. We have shown that the GBAR(1) model, with proper modifications to base it on the Pearson-V distribution, can be used to model single H.263 videoconference sources. For brevity reasons, we do not repeat the analysis of the model here; we need to emphasize, however, that contrary to the approach for H.263 traffic, three separate GBAR(1) models need to be constructed to accurately model single MPEG-4 [19] videoconference sources: one for each type of video frame of each video trace.

The simplicity of the model allows its practical applicability on a WLAN station without any significant computational complexity. The reason is that the model is an autoregressive process of order one; therefore, the data from the model can be easily derived with the use of just a few physically meaningful parameters: the mean, variance, and autocorrelation of the video traces.

V. RESULTS AND DISCUSSION

A. Simulation Setup

Voice traffic is represented by a two-state Markov model (on/off model). Active voice nodes transmit at a constant rate (ON-state), and inactive nodes (OFF-state) do not transmit. In line with the traffic characteristics digitized with the G.711 [20] coding standard, the voice packet interarrival period is 20 ms, and the packet size is 160 bytes. The interarrival time for BE data traffic is 7.5 ms, and the packet size is 1000 bytes. Hence, the voice rate is 64 kb/s, and the BE data rate is 1.07 Mb/s, respectively [6].

For BK, we adopt the HTTP traffic model from [21]. The packet size is 1000 bytes. The distributions of the random variables concerning the composition of web requests are presented in Table III.

We have used the high-quality coding version of four MPEG-4 video traces (from [22]). The trace statistics are also presented in Table III. The packet size is 1280 bytes [6]. A video node arriving in the network chooses one of the four traces with equal probability (25%). As in [23], we consider a two-state (Gilbert-Elliott) channel error model. We have studied our system for various values of packet error rates, ranging between 10^{-1} and 10^{-5}. (These values have been taken from [23] and [8], respectively.)

The simulation parameters are shown in Table I. The channel rate, for transmitting voice, video, and data packets, is 11 Mb/s, and the basic rate, for transmitting RTS, CTS frames, and token frames, is 2 Mb/s. Event-driven simulation is done in Matlab. Each simulation point is derived as the average of ten independent runs, each simulating 180 s of the channel time. For each new video node arrival, a 2-min sequence of the chosen video trace is used at random. The TXOP for AC_VI in EDCA, in our simulations, is 3008 μs.

We need to point out that our self-policing scheduling scheme is independent of the numerical values used in our simulation setup. These are simply used to validate our proposed solution.

B. Evaluation of Results

Fig. 1 shows the average video packet delay with six video nodes present in the network in the absence of voice traffic. It also presents the average voice packet delay for an increasing number of data nodes, with 30 voice nodes present in the network (case of a WLAN for an office environment), as well as a constant number of video nodes equal to 2. Our results show that, with the use of our scheme, which provides guaranteed priority to voice and video nodes over data nodes, the delays remain very low, i.e., about 1.5 ms for voice packets and between 12 and 12.5 ms for video packets. The minor fluctuation in video delay is due to the probabilistic choice among the different video traces. On the contrary, both delays significantly increase with the use of EDCA. Since our scheme shares with HCCA the concept of eliminating collisions (our scheme through token passing and HCCA through polling), we also provide in Fig. 1 a result comparison with HCCA. As shown in the figure, HCCA provides only marginally better results in terms of mean video packet delay, whereas the standard deviation of the video packet delay was shown from our results to be marginally lower in our scheme for high loads. The reason is that, when a poll is lost in HCCA, there is a certain delay until the same station is polled again. This result agrees with the respective results of [3] when comparing HCCA with EDCA. The standard deviation of the video packet delay for EDCA is again quite high, which indicates high jitter.

To show the efficiency of our proposed idea of variable TXOPs per video user, we present again in Fig. 2 the mean video packet delay results for our scheme and EDCA, which were shown in Fig. 1; this time, however, we compare them with the video packet delay results in the case where only token passing is used and TXOP is fixed. The values of TXOP used for our scheme are those corresponding to the following: 1) the time needed for a video user to transmit its mean video frame size; 2) the time needed for a video user to transmit a frame equal to the mean video frame size plus the standard deviation; and 3) the time needed for a video user to transmit its peak video frame size. The respective fixed values of TXOP for all video users in EDCA were those corresponding to the transmission of the following: 1) the mean of the average video frame sizes; 2) the mean of the (mean+standard deviation) video frame sizes; and 3) the mean of the peak video frame sizes. In all three cases, our scheme again outperforms EDCA, but the results are clearly worse than those achieved when we use variable TXOP in our scheme. The reason is that, with the use of our accurate video traffic model, the choice of TXOP is close to optimal; on the other hand, with a fixed TXOP value, there is always the problem of overallocating or underallocating time to a video user. The former results in unnecessary idle time and delays for the other video users, whereas, in the case of underallocation, the user loses the chance to transmit as many packets as possible when the channel is idle. This results in the "cutting" of the video frame into many separate transmissions, which may encounter significant delays due to the presence of other video users and of voice users. Our results conceptually agree with those in [24], where the authors reached the conclusion that allocating TXOP limit based on the burst size distribution can improve the network performance under bursty traffic, and that if the length of TXOP is adopted in a way to transmit the burst in a smaller number of TXOP service periods, the total network performance and the burst delay are noticeably improved. As expected, the relatively worse results among the three cases are produced for a TXOP equal to the time needed for the transmission of a peak video frame size, as the overallocation is constant. We need to point out that the results achieved by EDCA with the use of the default value for AC_VI (3008 μs) are close to those achieved by EDCA with a fixed TXOP corresponding to the mean of the (mean+standard deviation) frame sizes. The reason is that this fixed TXOP value is equal to 2780 μs, which is quite close to the default value.

Fig. 3 shows that the increase in the number of video nodes does not affect the voice packet delay with the use of our proposed scheme, as the delay remains about 1.7 ms. However, it has a significant impact when EDCA is used. The standard deviation of the voice packet delay is also presented in the figure for both schemes. A constant number of ten data nodes was used in the simulations from which these results were derived. In addition, the results presented in Figs. 1–3 have been derived for a channel with a low packet error rate of 10^{-5}.

Fig. 4 shows our results on the channel utilization achieved by our proposed scheme and EDCA, respectively. Channel utilization is the ratio of the system throughput versus the channel rate. With the use of self-policing, our scheme tackles the problem of idle time due to large TXOPs; this, combined with the use of token passing to avoid collisions, results in a very significant increase in the achieved channel utilization in comparison with EDCA. As shown in the figure, even for a high packet error rate of 10^{-1}, our scheme provides higher channel utilization than EDCA does for a channel with a low packet error rate of 10^{-5}. All the results shown in the figure have been derived as averages over extensive simulations, which covered ten scenarios and ten independent runs for each scenario. In each scenario, the system traffic load shown in the x-axis was generated with a specific mixture of voice, video, BE, and BK traffic (e.g., in one scenario, the mixture was 25% voice, 15%

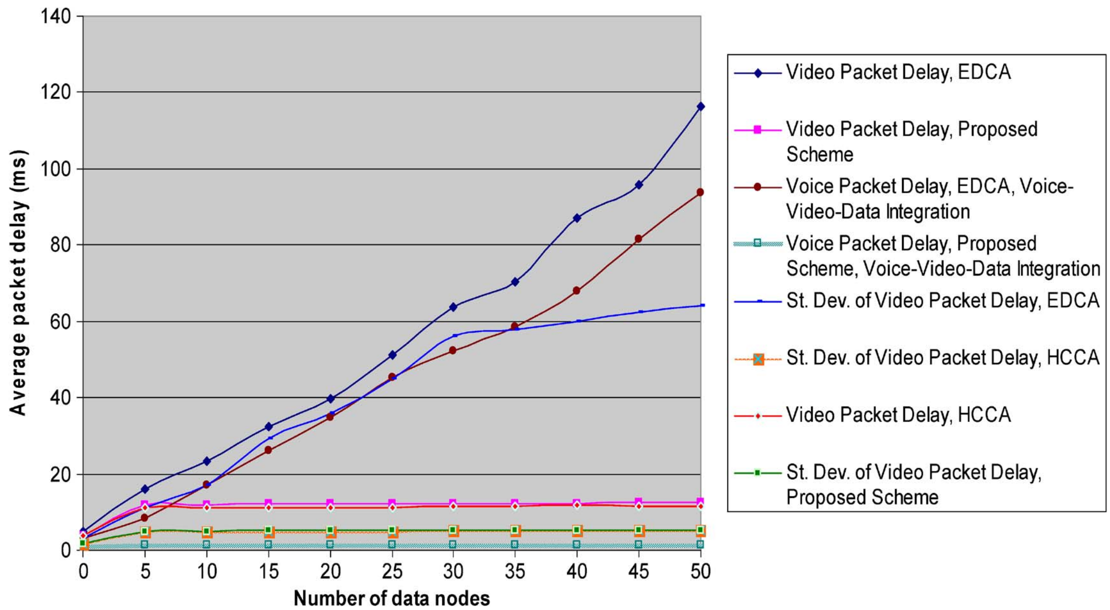

Fig. 1. Average voice and video packet delays versus the number of data nodes.

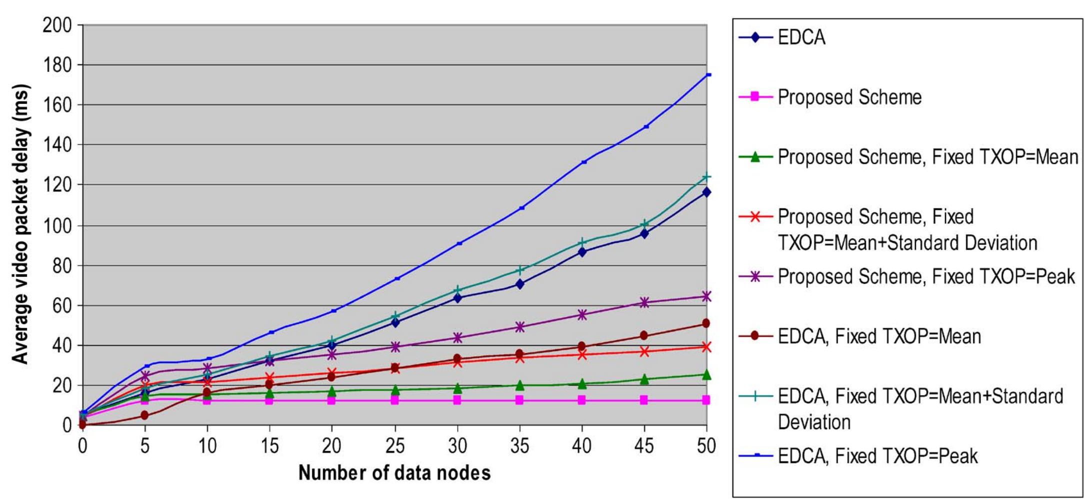

Fig. 2. Average video packet delays versus the number of data nodes for fixed TXOPs.

video, 30% BE, and 30% BK). The voice traffic load ranged, in all the scenarios, between 10% and 60%. The video traffic and BE and BK loads ranged between 10% and 70%. In addition, HCCA is again shown to provide only marginally better results in the case of high traffic loads. The preceding results, in Figs. 1 and 4, for HCCA do not take into consideration the problems of HCCA's complicated software architecture and performance in multicollision domains, which were discussed in Section I.

The results shown in Fig. 5 focus on the video packet delay encountered by individual video streams. Given that, in this work, we do not consider the problem of providing guaranteed QoS to all types of multimedia traffic, a measure of fairness needs to be used to show the merits of our proposed scheme. We use Jain's fairness index [25]. Once again, as in Fig. 1, the idea of using variable TXOP for video users, based on the use of an accurate video traffic model, is shown to clearly outperform EDCA, as well as the implementations of the proposed scheme with fixed TXOP. Again, the worst results for fixed TXOP are produced for a TXOP equal to the time needed for the transmission of a peak video frame size, for the reasons explained in the discussion on Fig. 2. For high traffic loads, the choice of a fixed TXOP equal to the time needed for the transmission of a peak video frame size leads to a very skewed delay distribution: Our simulations have shown that almost two thirds of the video users experience high video packet delay, whereas one third (those who get the token earlier) experience less than half the average video packet delay. This result is also confirmed by Jain's index in Fig. 5.

VI. Conclusion

We have proposed, for the first time in the relevant literature (to the best of our knowledge), a scheduling scheme using token-passing and self-policing for the integration of voice, videoconference, and data traffic over WLANs. Our scheme introduces significant-in-essence but easy-to-implement

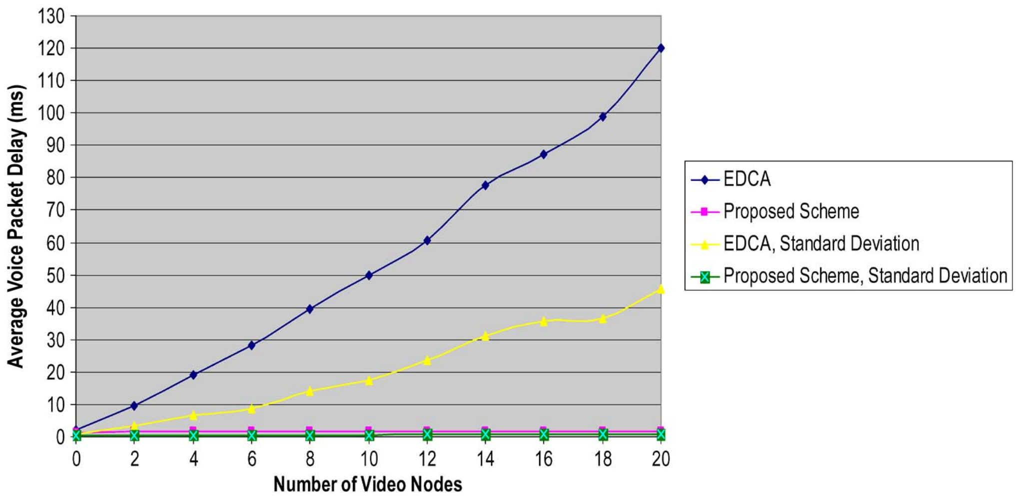

Fig. 3. Average voice packet delay versus the number of video nodes.

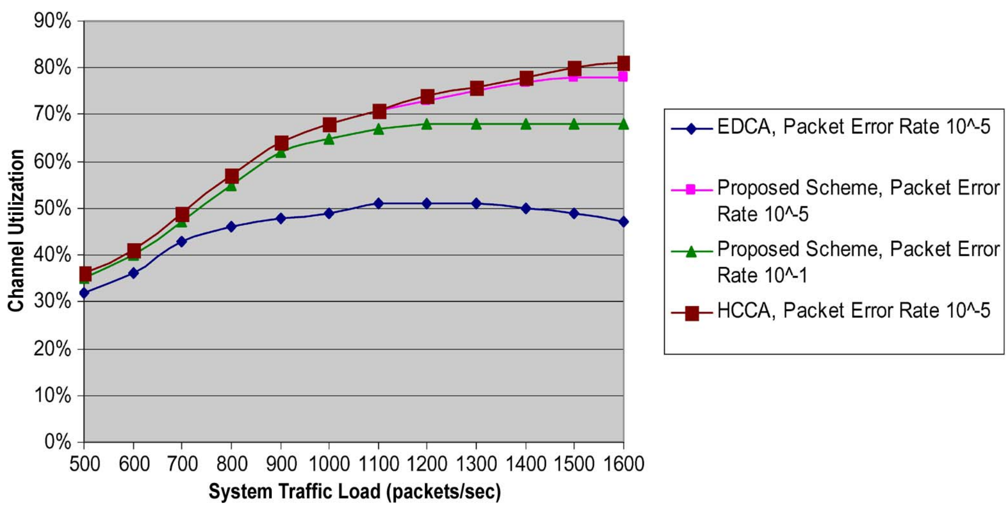

Fig. 4. Channel utilization versus the system traffic load.

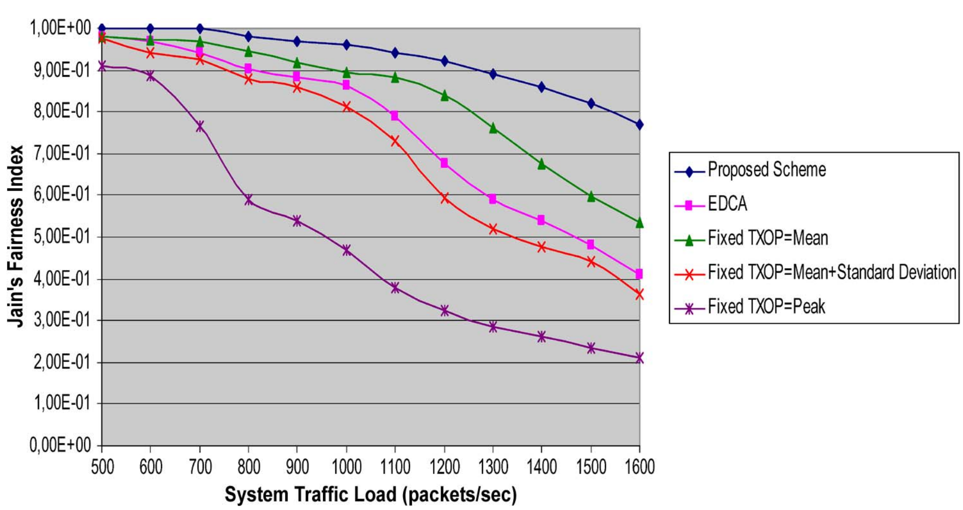

Fig. 5. Fairness index versus traffic load.

modifications of the EDCA to ensure proper prioritization among different ACs. Most importantly, our scheme practically eliminates contention and TXOP idle time and hence leads to a significant increase in channel utilization when compared to EDCA. It also provides guaranteed priority to voice traffic over all traffic types and to video traffic over data traffic.

We believe that the proposed scheme is suitable for multimedia applications in general, possibly with a few modifications to fine-tune it to the parameters of each type of multimedia traffic. The scheme has been shown in this work to perform well for three major types of multimedia traffic: voice, data, and MPEG-4 videoconference video. It can easily be extended to any type of video traffic as long as an accurate video traffic model exists. The only significant change that may have to be incorporated into the scheme would be associated with the case when data from an urgent data application would need to be transmitted (e.g., telemedicine data). In that case, our scheme's approach for data traffic could lead to unsatisfactory results, as this type of traffic cannot be treated as BE or BK traffic, i.e., of lesser priority compared with voice and video. Therefore, in that case, data users with urgent traffic would certainly have to acquire a smaller AIFS and possibly use variable TXOPs, similarly to the approach proposed here for video traffic.

One limitation of this work is that, as in [14], the work of which it enhances, we consider a fully connected WLAN without a central controller, where all the nodes can hear each other. Therefore, it does not address the well-known hidden-node problem, which would affect our scheme, as it uses token passing. Quite a few approaches, such as the increase in transmission power from the nodes and the use of omni-directional antennas, have been proposed in the literature to solve the hidden-node problem, but this issue is out of the scope of this work. In addition, similarly to [14], we provided a "closed solution," in the sense that nodes are preconfigured with specific network parameters. In the case that another network shares the same medium and stations from that network do not comply with our proposed network parameters, the problem can be resolved with wireless bridging.

In future work, we will focus on the problem of providing guaranteed QoS to all types of multimedia traffic. We need to point out that the IEEE 802.11n amendment [26] offers a significant increase in WLAN throughput, through the use of a multiple-input–multiple-output scheme. However, the change in throughput does not affect the reasons for which our scheme excels in comparison with EDCA under the same physical-layer (PHY) conditions; hence, our scheme will continue to provide better results. Still, in the case of a PHY allowing for much higher throughput, it will be easier for our scheme and for EDCA to provide guaranteed QoS to all types of multimedia traffic.

REFERENCES

[1] IEEE 802.11 Working Group (1999-07-15), Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Std. 1999-07-15, 1999.

[2] Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 8: Medium Access Control Quality of Service Enhancements, IEEE Std. 802.11e-2005, 2005.

[3] L. W. Lim, R. Malik, P. Y. Tan, C. Apichaichalermwongse, K. In addition,o, and Y. Harada, "A QoS Scheduler for IEEE 802.11e WLANs," in *Proc. IEEE CCNC*, 2004, pp. 199–204.

[4] Y. He, J. Sun, R. Yuan, and W. Gong, "A reservation based backoff method for video streaming in 802.11 home networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 332–343, Apr. 2010.

[5] Y. He, X. Ma, J. Sun, R. Yuan, and W. Gong, "Deterministic backoff: Towards efficient polling for 802.11e HCCA in wireless home networks," in *Proc. IEEE ICC*, 2010, pp. 1–6.

[6] H. Zen, D. Habibi, A. Rassau, and J. Wyatt, "Adaptive segregation-based MAC protocol for real-time multimedia traffic in WLANs," in *Proc. 15th IEEE ICON*, 2007, pp. 461–466.

[7] R. Haywood, S. Mukherjee, and X.-H. Peng, "Investigation of H.264 video streaming over an IEEE 802.11e EDCA wireless testbed," in *Proc. IEEE ICC*, 2009, pp. 1–5.

[8] S. G. Sitharaman and K. M. Anantharaman, "Impact of retransmission delays on multilayer video streaming over IEEE 802.11e wireless networks," in *Proc. 2nd Int. Conf. Commun. Syst. Softw. Middleware*, 2007, pp. 1–12.

[9] Y. P. Fallah, P. Nasiopoulos, and H. Alnuweiri, "Scheduled and contention access transmission of partitioned H.264 video over WLANs," in *Proc. 50th IEEE GLOBECOM*, 2007, pp. 2134–2139.

[10] J. Lee and M. Kang, "Design of a dynamic bandwidth reallocation scheme for hot-spot video stream transmission over the IEEE 802.11 WLAN," in *Proc. IEEE TENCON*, 2006, pp. 1–4.

[11] P. Ansel, Q. Ni, and T. Turletti, "FHCF: A simple and efficient scheduling scheme for IEEE 802.11e wireless LAN," *Mobile Netw. Appl. J.*, vol. 11, no. 3, pp. 391–403, Jun. 2006.

[12] I. Aad and C. Castelluccia, "Differentiation mechanism for IEEE 802.11," in *Proc. IEEE INFOCOM*, 2001, pp. 209–218.

[13] M. Barry, A. T. Campbell, and A. Veres, "Distributed control algorithms for service differentiation in wireless packet networks," in *Proc. IEEE INFOCOM*, 2001, pp. 582–590.

[14] P. Wang and W. Zhuang, "A token-based scheduling scheme for WLANs supporting voice/data traffic and its performance analysis," *IEEE Trans. Wireless Commun.*, vol. 7, pt. 1, no. 5, pp. 1708–1718, May 2008.

[15] E. P. Rathgeb, "Modeling and performance comparison of policing mechanisms for ATM networks," *IEEE J. Sel. Areas Commun.*, vol. 9, no. 3, pp. 325–334, Apr. 1991.

[16] P. Koutsakis, "Dynamic versus static traffic policing: A new approach for videoconference traffic over wireless cellular networks," *IEEE Trans. Mobile Comput.*, vol. 8, no. 9, pp. 1153–1166, Sep. 2009.

[17] A. Lazaris, P. Koutsakis, and M. Paterakis, "A new model for video traffic originating from multiplexed MPEG-4 videoconference streams," *Perform. Eval. J.*, vol. 65, no. 1, pp. 51–70, Jan. 2008.

[18] ITU-T Recommendation, H.263, ver. 3, 1996.

[19] T. Sikora, "MPEG digital video-coding standards," *IEEE Signal Process. Mag.*, vol. 14, no. 5, pp. 82–100, Sep. 1997.

[20] ITU-T Recommendation G.711 (1988) Amendment 12009, Aug. 2009.

[21] P. Tran-Gia, D. Staehle, and K. Leibnitz, "Source traffic modeling of wireless applications," *Int. J. Electron. Commun.*, vol. 55, no. 1, pp. 27–36, 2001.

[22] F. H. P. Fitzek and M. Reisslein, "MPEG-4 and H.263 video traces for network performance evaluation," *IEEE Netw.*, vol. 15, no. 6, pp. 40–54, Nov./Dec. 2001.

[23] G. Min, J. Hu, W. Jia, and M. E. Woodward, "Performance analysis of the TXOP scheme in IEEE 802.11e WLANs with bursty error channels," in *Proc. IEEE WCNC*, 2009, pp. 1–6.

[24] S. Rashwand and J. Misic, IEEE 802.11e EDCA under bursty traffic—How much TXOP can improve performance, *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1099–1115, Mar. 2011. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5699947

[25] R. Jain, *The Art of Computer Systems Performance Analysis*. Hoboken, NJ: Wiley, 1991.

[26] IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput.

Polychronis Koutsakis (M'09) received the Diploma in electrical engineering in 1997 from the University of Patras, Greece, and the Ph.D. degree in electronic and computer engineering from the Technical University of Crete, Greece, in 2002.

He was a Visiting Lecturer with the Department of Electronic and Computer Engineering, Technical University of Crete, from 2003 to 2006. From July 2006 to December 2008, he was an Assistant Professor with the Electrical and Computer Engineering Department, McMaster University, Hamilton, ON, Canada. In January 2009, he joined the Department of Electronic and Computer Engineering, Technical University of Crete, as an Assistant Professor

Dr. Koutsakis has authored more than 90 peer-reviewed papers on the design, modeling, and performance evaluation of computer communication networks. He serves as Editor for the *IEEE Communication Surveys and Tutorials Journal* and the *Elsevier Computer Communications Journal*.