Exploiting Multimedia Frame Semantics and MAC-layer Enhancements for QoS Provisioning in IEEE 802.11e Congested Networks

Anastasios Politis
Dept. of Applied Informatics
University of Macedonia
Thessaloniki, Greece
anpol@uom.gr

Ioannis Mavridis
Dept. of Applied Informatics
University of Macedonia
Thessaloniki, Greece
mavridis@uom.gr

Athanasios Manitsaris
Dept. of Applied Informatics
University of Macedonia
Thessaloniki, Greece
manits@uom.gr

***Abstract*—Wireless Local Area Networks (WLANs) supporting modern streaming multimedia applications constitute a very challenging and rapidly changing field of research. Towards implementing effective multimedia wireless networks, the IEEE has published the "state of the art" IEEE 802.11e standard, which introduced a QoS-aware MAC-layer along with a series of efficiency enhancements. However, it has been proven inadequate in handling multimedia traffic optimally in periods of congestion. For the efficient support of multimedia applications in high load situations, numerous mechanisms have emerged, most of them focusing on altering the static nature of resource allocation specified in IEEE 802.11e. Nevertheless, traffic characteristics must be taken into consideration in order to achieve the highest gains. In this paper, an application-aware MAC-layer mechanism is developed that exploits multimedia frame semantics and existing MAC-layer enhancements to adequately cope with high congestion situations in IEEE 802.11e infrastructure networks. The proposed algorithm makes use of existing acknowledgment policies and adaptive resource allocation techniques depending on multimedia frame significance. The effectiveness of the algorithm is proven by means of simulations, where its functionality is evaluated and compared with other existing schemes.**

Keywords- WLANs, Multimedia, IEEE 802.11e, QoS, MAC-layer

I. INTRODUCTION

Wireless Local Area Networks (WLANs) have been established as one of the preferred network technologies by the majority of electronic equipment users. At the same time, networked multimedia applications have penetrated the market with a tremendous success. Hence, the combination of multimedia applications and WLANs has been an extremely interesting research topic for the networking scientific community. The ultimate goal is to design WLANs in a way to support efficiently the incorporated multimedia traffic.

In an attempt to address this challenge, the Institute of Electrical and Electronics Engineers (IEEE) has released a series of amendments, improving the functionality of the initial IEEE 802.11 WLAN standard [2]. The majority of these amendments focused on signal modulation techniques, in an attempt to provide high data rates at the physical (PHY) layer (IEEE 802.11a/b/g) [3], [4], [5]. However, it was soon discovered that MAC layer enhancements were also needed in order to efficiently utilize the available bandwidth. Furthermore, since multimedia applications demand certain and strict Quality of Service (QoS) levels, it is required that the IEEE 802.11 MAC layer is capable of traffic differentiation.

To this direction, several new mechanisms demonstrated an increased efficiency in multimedia applications support in WLANs, by providing prioritized access to different traffic flows and/or reducing MAC layer overhead [6], [7], [8]. Yet, the final act to these research efforts for providing multimedia support in WLANs, was the standardization of the IEEE 802.11e amendment by the IEEE Standards Committee [9]. Most of the enhancements provided by this standard are also included in the recently released IEEE 802.11n standard [10].

IEEE 802.11e specified a QoS-aware MAC layer protocol capable of service differentiation together with a series of MAC layer enhancements. According to the specification, a new coordinating function is introduced, namely the Hybrid Coordination Function (HCF). Two access methods are defined under HCF: the Enhanced Distributed Channel Access (EDCA) and the HCF Controlled Channel Access (HCCA). EDCA provides service differentiation and thus prioritized access to the wireless medium while HCCA is an enhanced version of legacy PCF (Point Coordination Function) with improved QoS features. Unfortunately, both HCCA and PCF mechanisms are rarely implemented in wireless networking products [11]. Therefore, our main concern focuses on the EDCA distributed channel access method.

Older and recent research studies revealed that EDCA functionality lacks adequate multimedia support in high load conditions in wireless infrastructure networks [1], [12]. This outcome is produced by a very common and critical issue present in these topologies, namely the downlink/uplink asymmetry problem. This phenomenon refers to the fact that, in general, downlink traffic (traffic destined to wireless stations) is, typically, considerably larger than the traffic destined to the wired network. In turn, the Access Point (called QAP in IEEE 802.11e terminology) becomes overcrowded and highly congested suffering from large queuing delays, buffer overflows and low throughput [1]. This has an immediate effect on QoS levels of the downlink multimedia flows. This phenomenon is mainly due to the static nature of resource allocation defined by the IEEE 802.11e standard.

In order to alleviate this problem, numerous solutions exist in the scientific bibliography, focusing on altering the static assignment of network resources to multimedia flows at the IEEE 802.11e MAC layer. However, as also noted in [13], such a layered approach to the QoS issue in multimedia networks leads to a simple and independent implementation, often achieving a suboptimal multimedia performance. The solution is the use of cross-layer techniques in order to achieve the highest possible efficiency. Roughly speaking, cross-layer design refers to protocol design by actively exploiting the dependence between protocol layers to obtain performance gains [14].

In this paper, following the work presented in [1], we confront the QoS degradation issue in congested IEEE 802.11e infrastructure networks by designing an application-aware MAC-layer mechanism, which exploits application level information in order to select the appropriate handling of the multimedia traffic at the MAC layer. The proposed mechanism is centralized and placed at the most congested node in the network (QAP) and its effectiveness is proven by means of simulation.

The rest of the paper is organized as follows: in Section II, a thorough overview of the EDCA and the acknowledgment policies defined in IEEE 802.11e is provided. In Section III, a performance comparison between the new acknowledgment schemes and the standard positive acknowledgment mechanism is given. This comparative study will aid the analysis and explanation of the proposed mechanism. Section IV identifies the primary reason for congestion in infrastructure WLANs. An overview of multimedia traffic characteristics and the quality metrics for voice and video applications is provided in Section V. Related work is outlined in Section VI, and our proposed semantic-aware MAC-layer mechanism is described in Section VII. Simulations results are discussed in Section VIII. The paper is concluded in Section IX with the final remarks.

II. OVERVIEW OF THE EDCA ACCESS METHOD

The EDCA mechanism defined by IEEE 802.11e is a modified DCF scheme designed to provide differentiated and distributed channel access. The service differentiation is distinguished between 8 different User Priorities (UPs), from 0 to 7, with 7 having the highest priority. Each frame from the higher layer arrives at MAC layer with a specific UP which is marked, afterwards, to its MAC header. An 802.11e STA (called QSTA), shall implement four Access Categories (ACs), from 0 to 3, with 3 having the highest priority. Hence a QSTA has four MAC queues, where each queue corresponds to an AC. Each AC is an enhanced variant of DCF and each frame is mapped to an AC according to its UP value as shown in Table I. The relative prioritization is described in the IEEE 802.1D specification [15].

The key feature of EDCA is that for each AC a different set of MAC parameters are assigned in order to achieve service differentiation. An AC uses *AIFS[AC]*, $CW_{min}[AC]$ and $CW_{max}[AC]$ instead of *DIFS*, CW_{min} and CW_{max} defined by legacy DCF. *AIFS* is the new Arbitration Inter-Frame Space introduced by IEEE 802.11e and is given by:

TABLE I. UP TO AC MAPPING

UP	802.1D Traffic Type (Acronym)	AC (AC Number)	IEEE 802.11e Designation
1	Background (BK)	AC_BK (0)	Background (BK)
2	Spare (-)	AC_BK (0)	Background (BK)
0	Best Effort (BE)	AC_BE (1)	Best Effort (BE)
3	Excellent Effort (EE)	AC_BE (1)	Best Effort (BE)
4	Controlled Load (CL)	AC_VI (2)	Video (VI)
5	Video (VI)	AC_VI (2)	Video (VI)
6	Voice (VO)	AC_VO (3)	Voice (VO)
7	Network Control (NC)	AC_VO (3)	Voice (VO)

$$AIFS[AC] = AIFSN[AC] \times T_{slot} + T_{SIFS} \quad (1)$$

AIFSN[AC] is a positive integer. The lower the value of *AIFS[AC]* the greater the priority for the contenting AC. Moreover, the value of the backoff timer for each AC is chosen randomly with a uniform distribution in the range [0,*CW[AC]*]. This gives the flexibility to ACs with higher priority to select a smaller contention window. In the case of simultaneous expiration of the backoff timers of two or more ACs belonging to the same QSTA, a virtual collision handler is responsible to grand access to the AC with the highest priority.

Another feature introduced by 802.11e is the concept of Transmission Opportunity (TXOP). This is defined as the interval of time in which a QSTA, after winning contention, has the right to initiate multiple frame transmissions, as long as the total transmission time does not exceed a limit called TXOP limit. This procedure is called Contention Free Burst (CFB) and is optional for a QSTA to utilize it. There is a set of default values specified by the IEEE 802.11e standard for the TXOP limit under the EDCA access mechanism. These values depend on the AC type and on the underlying physical layer and are depicted in Table II. The table reveals the prioritized access given on multimedia applications (video and voice) which use smaller values of *AIFSN*, *CWmin* and *CWmax*. Furthermore, the TXOP limit values provide multimedia applications with the CFB feature whereas background and best effort traffic are allowed to transmit single data frames before they re-enter the contention phase.

TABLE II. EDCA RELATED DEFAULT PARAMETER VALUES

AC	CW_{min}	CW_{max}	AIFSN	TXOP limit (μs)	
				802.11 802.11b	*802.11a 802.11g*
BK	CW_{min}	CW_{max}	7	0	0
BE	CW_{min}	CW_{max}	3	0	0
VI	$(CW_{min}+1)/2-1$	CW_{min}	2	6016	3008
VO	$(CW_{min}+1)/4-1$	$(CW_{min}+1)/2-1$	2	3264	1504

By allowing multiple frame transmissions after winning a contention, CFB reduces the number of backoff periods and the number of RTS/CTS frames exchanged, thus resulting in lower overhead. A CFB transmission chronicle is depicted in Fig. 1.

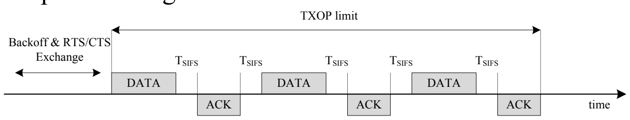


Figure 1. CFB timing structure.

It is a straightforward conclusion that assigning large values to TXOP limit will allow higher throughput and lower delays to a specific AC. As Table II depicts, the default TXOP limit values are statically assigned. Nonetheless, the IEEE 802.11e standard permits dynamic allocation of TXOP limit values.

Towards reducing the MAC layer overhead even further, IEEE 802.11e introduced two new acknowledgment policies besides the default DATA-ACK handshake. The standard acknowledgment policy has the disadvantage that a QSTA has to wait a significant amount of time before continuing with the transmission of the rest of its buffered frames. As depicted in Fig.1 each frame is required to be individually acknowledged before the QSTA may proceed with the next frame in its sequence. Hence, the actual subsequent data frame transmission commences not before the passing of $2 \cdot T_{SIFS}+T_{ACK}$ period of time. Accumulating all these waiting periods, the final amount of time dedicated to the exchange of control frames (ACK) may occupy a significant portion of the available TXOP limit assigned to the QSTA, depending on the ACK and data frame sizes as well as the SIFS period (which is PHY dependent). Suppose that a station has n data frames buffered and gained control of the channel. If their transmission times do not exceed the TXOP limit, then the total CFB transmission time can be expressed as follows:

$$T_{CFB}^{std} = \sum_{i=1}^{n} T_{DATA_i} + (2n-1)T_{SIFS} + nT_{ACK} \qquad (2)$$

In order for a QSTA to fully utilize the available TXOP limit, the Block Acknowledgment (BA) and the No Acknowledgment (NoACK) policies are defined under the IEEE 802.11e standard. These acknowledgment policies combined with the CFB feature may drastically improve channel utilization and MAC efficiency.

A. *Block Acknowledgment Policy*

The Block Acknowledgment scheme improves the MAC layer efficiency by aggregating multiple acknowledgments into a single frame. In this way, the control overhead imposed by the standard acknowledgment policy is reduced. The use of the BA mechanism is optional and is a subject of negotiation between the sender and receiver.

After gaining control of the channel, the sender may request the usage of the BA policy by transmitting an ADDBA (Add Block Acknowledgment) request frame to the receiver, who must acknowledge its reception. The receiver may accept or reject the proposal by issuing an ADDBA response frame. After acknowledging the response, the sender may proceed with a different acknowledgment policy if a rejection was indicated. In the case of a successful agreement between the sender and receiver regarding the usage of the BA policy, the sender will proceed with the transmission of its buffered frames in a CFB manner, without violating the assigned TXOP limit. Upon reception, the receiver shall not produce acknowledgment frames, until the reception of a Block ACK request (BAR) frame indicating the ending of the frame burst and the request of an aggregated acknowledgment frame by the sender. Afterwards, the receiver initiates the transmission of a Block ACK frame (BA) destined to the sender, indicating which frames were received correctly. At this moment, the receiver has two options: initiate an immediate Block ACK or a delayed Block ACK. The former option is suitable for low latency applications, while the latter is used by applications that tolerate moderate latency.

If the Block ACK frame indicates unacknowledged data frames, the sender shall retransmit the lost frames in this or a later TXOP. Otherwise, the BA mechanism is terminated from the sender with a DELBA (Delete Block Acknowledgment) frame which must be acknowledged by the receiver. Moreover, after a timeout of inactivity, the BA agreement may be torn down automatically.

According to IEEE 802.11e, the use of the BA policy is permitted if the following conditions are satisfied:

- A protective mechanism is used (such as HCCA or RTS/CTS) in order to reduce the possibility of other stations transmitting during the TXOP. If no protective mechanism is used, then the first frame of the burst should be acknowledged individually to help the other stations to update their Network Allocation Vectors accordingly.
- The sender may not transmit more frames than the receiver has indicated to be able to buffer.
- All frame transmissions are limited by the TXOP. However, the sender may split frames using this mechanism across several TXOPs.

Assuming that, the receiver has successfully accepted the BA scheme, the Block ACK is immediate, the RTS/CTS protection mechanism is used and all frame exchanges are within the TXOP limit, then the timing of the BA procedure can be modeled as in Fig. 2.

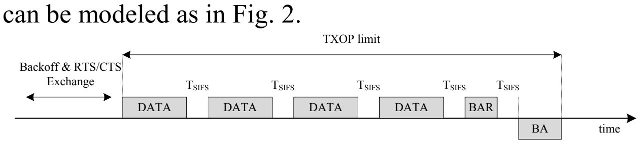


Figure 2. CFB timing structure with BA policy.

If a station transmits n data frames using the BA mechanism, the total CFB time is given by:

$$T_{CFB}^{BA} = T_{BAR} + T_{BA} + \sum_{i=1}^{n} T_{DATA_i} + (n+1)T_{SIFS} \qquad (3)$$

The overhead of individually acknowledging each frame in the sequence is replaced by the BAR and BA frames exchange after concluding the data frames transmission. Furthermore, for a large number of data frames the SIFS periods are almost halved compared to the standard acknowledgment policy.

B. No Acknowledgment Policy

The concept of NoACK policy is fairly simple: for every data frame received, the receiver does not produce an acknowledgment packet, thus the overhead imposed by the acknowledgment frames is completely eliminated. The benefit of exploiting the NoACK policy on packet delay is straightforward. However, in this way the MAC-level recovery mechanism is suppressed and the reliability of the traffic is reduced due to the probability of lost frames from interference, collisions or time-varying channel conditions. To cope with that, the IEEE 802.11e standard proposes that a protective mechanism is used (such as HCCA or RTS/CTS) to reduce the probability of another QSTA transmitting during the TXOP. Similarly to the paradigm given in Fig. 2, the total CFB time consumed to transmit n data frames with the NoACK policy may be written as:

$$T_{CFB}^{NoACK} = \sum_{i=1}^{n} T_{DATA_i} + (n-1)T_{SIFS} \tag{4}$$

Eq. 4 shows the great overhead cost reduction but at the expense of reduced reliability. Hence, the NoACK policy resembles a UDP-like behavior at the MAC layer.

III. Comparative Study of IEEE 802.11e Acknowledgment Policies

In order to facilitate the description and analysis of the proposed mechanism in this paper, this section provides a simple comparative study of the Standard (StdACK), BA and NoACK acknowledgment policies. The efficiency improvement of the new acknowledgment schemes defined under IEEE 802.11e is calculated in terms of total CFB transmission time and compared to the CFB transmission time obtained when using the StdACK policy.

More specifically, a comparison of total CFB times is provided containing different number of equally-sized data frames for various payloads, using Equations 2, 3 and 4. In order to accomplish that, we need to specify the values of T_{DATA}, T_{ACK}, T_{SIFS}, T_{BAR} and T_{BA}. T_{SIFS} is a constant time period and its value depends on the underlying physical technology. Every data and control (ACK, BAR and BA) frame is charged with a physical and MAC overhead. The physical overhead, T_{PHY}, is constant and comprised by a PLCP preamble and a PLCP header. The MAC overhead is frame type dependant and consists of the MAC header and the FCS field. Depending on the physical layer used, the physical overhead has different sizes. For example, in 802.11b the overhead is 192 µs (when using the long preamble) while in 802.11g the overhead is reduced to 20 µs. The MAC overhead for frames carrying data depends on whether the transmission is directed in the same Basic Service Set or in the Extended Service Set, to or from Access Points etc.

The physical layer divides data from the MAC layer into a series of symbols for transmission. Each symbol encodes a certain number of bits, L_{SYM}, depending on the transmission rate selected and then it is transmitted at a prescribed symbol rate, $1/T_{SYM}$. Hence generalizing, T_{DATA}, T_{ACK}, T_{BAR} and T_{BA} may be derived from the following equation:

$$T = T_{PHY} + \left\lceil \frac{L_{MPDU} \cdot 8}{L_{SYM}} \right\rceil \cdot T_{SYM} + T_{SIGNAL} \tag{5}$$

L_{MPDU} is the size of the MAC Protocol Data Unit (MPDU) measured in bytes and is composed by the payload, p, (or MAC Service Data Unit – MSDU) and the MAC overhead (L_{MAC}) and can be expressed as $L_{MPDU}=L_{MAC}+p$. T_{SIGNAL} is an additional time extension for encoding purposes, applicable only to IEEE 802.11g PHYs.

A. Assumptions

In this comparative study the IEEE 802.11g specification was assumed as the underlying physical layer. IEEE 802.11g specifies actually four physical layers defined as Extended Rate Physicals (ERP's) [5]. These layers make use of the DSSS, OFDM or both modulation methods in order to provide IEEE 802.11a data rates in the 2.4 GHz band and backward compatibility with legacy IEEE 802.11b systems. For this study, the so called ERP-OFDM physical layer was assumed which is used when all stations in a BSS are IEEE 802.11g compliant. Table III summarizes the physical characteristics for the ERP-OFDM PHY, operating at 54Mbps.

TABLE III. IEEE 802.11G PHYSICAL CHARACTERISTICS

T_{PHY}	20 µs
T_{SYM}	4 µs
L_{SYM}	216 bits
T_{SIFS}	10 µs
T_{SIGNAL}	6 µs

TABLE IV. MPDU SIZES FOR DATA AND CONTROL FRAMES

Type of Frame	$L_{MPDU}=L_{MAC}+p$	No. of OFDM Symbols
Data	28+128 Bytes	6
	28+512 Bytes	20
	28+1024 Bytes	40
	28+1500 Bytes	57
ACK	14 Bytes	1
BAR	24 Bytes	1
BA	152 Bytes	6

Regarding the size of the L_{MPDU}, Table IV summarizes the different sizes for data and control frames. Furthermore, the table reveals the number of OFDM symbols for each frame to be transmitted at the rate of 54 Mbps. L_{MAC} for the data frame is always 28 Bytes, as long as the frame is directed to a station belonging at the same BSS [2]. Four values are assumed as payload size in data frames, namely 128, 512, 1024 and 1500 Bytes. All L_{MPDU} sizes for control frames are taken from [2] and [9].

Furthermore, the following series of assumptions are made in order to simplify the comparative study:

- The channel is error-free, meaning that all frames are received correctly.
- There are no collisions present.
- Regarding the BA scheme, the immediate BA mechanism is used.
- No protective mechanism such as RTS/CTS, CTS-to-self or HCCA is used.
- Packets are not fragmented.
- A CFB is comprised of equal-sized data frames.

B. Results and Analysis

The relative improvement of the total CFB transmission time was used as a measure for comparing the different acknowledgment policies and is defined as the improvement of CFB transmission time achieved by the BA and NoACK mechanisms relative to the CFB transmission time experienced by the usage of the StdACK policy. This metric was obtained from Equations 2, 3 and 4 which were verified via simulations with the OPNET simulation tool [16]. Fig. 3 displays the relative improvement obtained by increasing the number of equally-sized frames, *n*, in CFB for different payload sizes, *p*. The solid lines are the values obtained by simulations. It must be noted that the default TXOP limit values during simulations were adjusted accordingly, in order to include the different number of frames in the CFB.

A first comment on the displayed outcome may be the observed improvement reduction of both BA and NoACK schemes as payload size increases. This is an expectable finding since larger data frames exhibit large transmission delays and thus the control frames exchange in the StdACK policy occupies a smaller percentage of the total CFB transmission time, thus reducing its margin between the CFB times of BA and NoACK policies.

Another observation from the depicted graphs is the negative improvement achieved by the BA policy on all four cases for $n \leq 2$. This is also an expected result since the BA mechanism uses a large-sized control frame (Block ACK frame) for data acknowledgment. For a small number of data frames this BA frame increases the total CFB transmission time of the BA policy.

As an overall results conclusion, it can be stated that using the BA and NoACK policies for large frames does not provide significant performance improvement, while exploiting these mechanisms for applications with small and constant-sized data frames (such as VoIP applications) leads to a significant lessening of CFB transmission times. This reduction of CFB transmission times of an AC enables the queue to transmit more frames in the remaining portion of the TXOP limit or release the channel sooner for another competing AC to capture it, thus achieving greater intra and inter-AC efficiency.

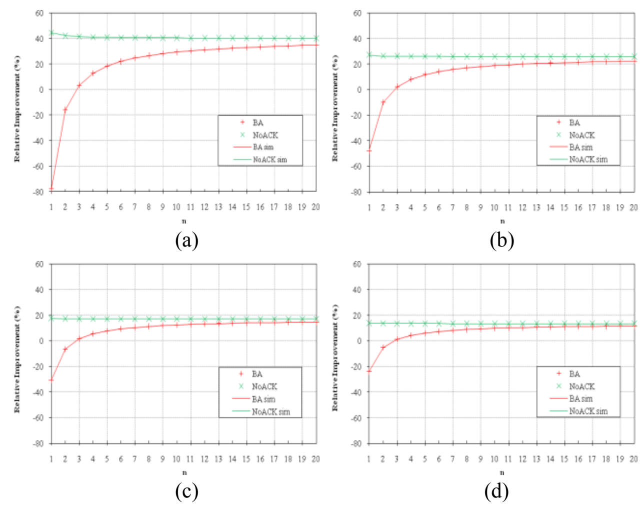


Figure 3. Relative Improvement of BA and NoACK policies for payload sizes (a) 128 Bytes, (b) 512 Bytes, (c) 1024 Bytes and (d) 1500 Bytes.

IV. DOWNLINK/UPLINK ASYMMETRY

One very common and critical issue in infrastructure WLANs is the downlink/uplink asymmetry. This refers to the fact that the downlink traffic (traffic transmitted from the Access Point to the QSTAs) is in most cases considerably larger than the uplink traffic (traffic transmitted from QSTAs to the Access Point). An AC in an Access Point (QAP) which serves all downlink traffic receives the same access priority with the AC in a QSTA which serves the uplink traffic. This leads to unfairness problem in which the QAP ACs suffer from large queuing delays, buffer overflows and low throughput [1], [12].

Since QAP accumulates all downlink traffic, there must be a centralized mechanism to allocate dynamically the needed channel resources to ACs in the QAP. A way to allocate these resources is to adapt the TXOP limit of the ACs in the QAP depending on their queue size. This method has been proven to be extremely beneficial in terms of channel efficiency and application performance. However, also noted by [17], little work has been done in the literature regarding VBR video traffic (such as streaming video) which exhibits time varying characteristics.

V. MULTIMEDIA TRAFFIC CHARACTERISTICS

This section summarizes the multimedia traffic characteristics. We focus on VoIP and MPEG streaming video, since they exhibit an increased popularity on both real applications and network related studies. The most frequently used quality metrics of these applications are also described.

A. VoIP

The traditional voice encoder is the G.711, which uses Pulse Code Modulation (PCM) to generate 8 bits samples per 125 μs, and leads to a minimum bandwidth requirement of 64 Kbps for each traffic flow. New voice encoding schemes have been implemented in order to drastically reduce bandwidth reduction, but at the cost of additional coding delay. Popular techniques include the G.729A and G.723.1 codecs. G.729A [18] is one of the most commonly used codecs in VoIP applications, due to its lower bandwidth requirements (8 Kbps) and acceptable complexity. Unless silence compression techniques are used, VoIP codecs typically produce constant bit rate streams with low frame sizes (≤160 Bytes).

The voice applications requirements are stringent. Their demand for assured quality real-time communication restricts the maximum tolerable one-way delay to 100 – 150 ms. Furthermore, the jitter imposed by the network must remain at low values (maximum 50 ms). These strict delay requirements lead to the need of QoS provision by the underlying network.

The most popular performance metric of VoIP applications used in multimedia networking studies is the Mean Opinion Score (MOS) [19]. According to this method, the perception quality of a VoIP call is determined by a single numerical value from 1 to 5, with 1 representing the lowest and 5 the highest quality. Table V presents typical MOS values for implementations of G.711, G.729A, and G.723.1 codecs.

TABLE V. MOS VALUES OF G.711, G.729A AND G.723.1 CODECS

Codec	Data Rate	Frame Size	MOS
G.711	64 Kbps	160 Bytes	4.3
G.729A	8 Kbps	20 Bytes	3.7
G.723.1	5.3 Kbps	20 Bytes	3.62

From Table V, the existence of a trade-off between lowering the required data rate and the perceived quality becomes obvious. Data rate reduction requires higher complexity algorithms which, in turn, produce a lower quality outcome.

B. *Streaming Video*

The main principle of MPEG encoding is inter- and intraframe coding. It distinguishes between three frame types, namely I, P and B-frames. I-frames are completely intra-coded, P-frames are predicted from previous I or P-frames, and B-frames depend on both previous I or P-frames and forward I or P-frames. Frames are arranged in so-called Group of Pictures (GoP). The sequence of frames from a given I-frame up to and including the frame preceding the next I-frame forms one GoP. A GoP pattern is determined by the total number of frames, N, comprising it and the number of B-frames, M, enclosed by successive P-frames. Thus the notation $G_N B_M$ is used to symbolize the GoP pattern of a video sequence. Typical GoP patterns include: G_6B_2, G_9B_2, $G_{12}B_2$ and $G_{15}B_2$, depending on the required video quality [20]. Fig. 4 depicts a G_9B_2 GoP pattern and the forward and backward references that exist between the frame types.

I-frames contain by far the most information, thus they exhibit the lowest compression ratios. By exploiting temporal redundancies, P-frames achieve higher compression rates than I-frames. Since B-frames are predicted from both previous and following frames they are appointed as the frames with the highest compression ratios.

In contrast with VoIP applications, MPEG video frames are distinguished by their semantics. I-frames are identified as the most significant frame type, since their absence will render a GoP completely undecodable. On the other hand, B-frames are not needed for the decoding of any other frame, thus they are appointed with the lowest significance. P-frames have a variable significance. There is a distinction between the semantics of P-frames, rooting from their relative position in a GoP sequence. Considering the GoP presented in Fig. 4, a possible loss on the first P-frame in the sequence will have a negative chain effect on 89% of the GoP. Similarly, losing the second P-frame will influence the decoding process of 55% of the GoP. The P-frames significance becomes even higher when scalable video is considered, where both I and P-frames are needed to provide a basic video quality [21].

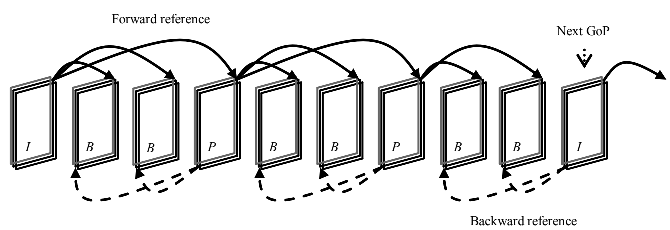


Figure 4. MPEG GoP coding structure.

The standard method for assessing the perceived video quality is to calculate the Peak Signal to Noise Ratio (PSNR) between the original (transmitted) and the received (possibly distorted) image. It is a differential metric which is determined image-wise and yields a quality indicator for each received image of the video sequence. Symbolizing with f the original image, with f' the distorted image and assuming an $m \times n$ image size, the PSNR is determined as:

$$PSNR = 20\, log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right) \tag{6}$$

$$MSE = \frac{1}{m \times n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[f(i,j) - f'(i,j)\right]^2 \tag{7}$$

MAX is the maximum possible value of a pixel (255 for an 8 bit pixel). *MSE* is the Mean Square Error and calculates the difference between each pixel of the original and distorted picture. Typical values for video compression lies between 30 to 50 dB with higher values preferred over lower ones.

VI. RELATED WORK

There is a growing research literature on semantic-aware QoS provisioning in WLANs supporting multimedia traffic. The most related to our work are briefly reviewed in this section.

In [22], Ksentini et al. propose a QoS cross-layer architecture based on both application and MAC layer features for improving H.264 video transmission over IEEE 802.11e networks. The mechanism relies on a data partitioning technique at the application layer and an appropriate QoS mapping at the IEEE 802.11e MAC layer. More specifically, the authors map the application layer video generated slices to appropriate ACs at the MAC layer according to their significance. AC_VO, AC_VI and AC_BE are used for this purpose while AC_BK is left for serving all other traffic. Furthermore, the retry count parameter at the MAC layer is exploited to unequally protect the high priority information against lower significance frames.

The semantic-aware mechanism presented in [23], follows a similar approach for MPEG-4 video transmission in IEEE 802.11e networks. This scheme introduced a single-video multilevel queue by assigning I-frames to AC_VO, P-frames to AC_VI, B-frames to AC_BE and non-video frames to AC_BK.

In [24], the authors follow a different approach on cross-layer design for H.264 video traffic transmission. At first, they determine the significance of a video frame by using a method called first order estimation. According to this method, the PSNR of all packets in the video sequence is determined by intentionally dropping selected frames. Thus, a frame is more important when its PSNR value is lower. Afterwards, the packets are placed to ACs in the MAC layer according to the access waiting time of an AC and its priority. Hence, the AC with the lowest waiting time is selected for serving a particular video packet.

In [25], Goel and Sarkar propose a mechanism that resides in the interface between LLC and MAC layers to provide QoS for streaming video traffic. The essence of this scheme is to mark I-frames of a video sequence as the Most Valuable Video Packet (MVVP) and en-queue these frames to a higher priority queue called Video Friendly Queue (VFQ) in the interface between LLC and MAC layers. Other frames are en-queued in the so-called Interface Queue (IFQ) and receive FIFO treatment. Whenever frames need to be send to the MAC layer the VFQ receives priority against IFQ. In this way, preferential treatment is provided to MVVP frames ensuring that they get highest priority which minimizes delay.

In [26], a dynamic mapping algorithm of MPEG-4 video frames is proposed. According to this algorithm, the video frames are allocated to ACs according to their significance and network load. When the size of AC_VI reaches a certain threshold, the newly arrived frame is mapped to a lower priority AC (AC_BK or AC_BE). The choice of the AC is determined by the frame significance.

It is clear that the entire semantic-aware mechanisms presented, exploit a mapping technique to allocate video frames to ACs either statically [22], [23], [25] or dynamically [24], [26]. However, with the exception of [26], they disregard the QoS issues of VoIP traffic by allocating voice frames to ACs with lower priority or mixing them with video traffic.

Furthermore, MAC-layer mechanisms, such as dynamic TXOP limit tuning and acknowledgment policies, are left completely unexploited. The usage of MAC-layer strategies may improve the system efficiency, and thus multimedia application performance, dramatically.

VII. THE PROPOSED SCHEME

Extending the work presented in [1], we propose a semantic-aware MAC-layer mechanism that falls into the cross-layer mechanisms category. The proposed scheme exploits multimedia frame semantics to decide an appropriate MAC-layer strategy for handling these frames. The mechanism is centralized and intends to improve EDCA performance at the QAP in times of congestion. Furthermore, only the functionality of multimedia ACs is affected, leaving the rest of the ACs (AC_BK and AC_BE) uninfluenced.

The essence of the proposed algorithm is to map multimedia frames into AC_VI or AC_VO according to their significance. To this direction, two categories of multimedia frames are introduced: High Priority Multimedia Frames (HPMF) and Low Priority Multimedia Frames (LPMF). I and P video frames are indicated as high significance frames and tagged as HPMFs, while B-frames and voice packets as LPMFs. Every category is linked to a specific AC: HPMF to AC_VO and LPMF to AC_VI. Such a distinction among the multimedia frames can easily be accomplished by manipulating the UP of the multimedia frame:

```
If UP of packet i ∈ (4, 5, 6)
    {
        If UP of packet i = 4||5 && packet_type ∈ (I, P)
                UP of packet i = 6
        ElseIf UP of packet i = 6
                UP of packet i = 4
    }
```

At this point, the appropriate MAC-layer strategy must be selected for both AC_VI and AC_VO. HPMFs belonging to AC_VO, are treated with the maximum protection by using the standard acknowledgment policy. However, regarding the high sizes of these frames, we apply a TXOP limit adaptation algorithm, ensuring the periodical relaxation of this queue.

The TXOP limit adaptation algorithm for the AC_VO traffic class is calculated every Service Interval (SI) which is defined as the time between the start of two subsequent TXOPs. At the beginning of the SI the actual queue length is calculated and the average frame size of all packets contained is determined. Then the TXOP limit is computed as the time needed to successfully transmit the en-queued frames:

$$TXOP\_limit = N \times T_{frame} \tag{8}$$

where N is the number of frames contained in the AC_VO queue at the start of the SI and T_{frame} the successful frame transmission time with the average payload size L. The T_{frame} is computed as:

$$T_{frame} = T_L(r) + T_{ACK}(r) + T_{SIFS} \tag{9}$$

where $T_L(r)$ and $T_{ACK}(r)$ are the transmission times of the data and acknowledgment frames respectively for a specific

PHY data rate r and accounting PHY and MAC overhead. T_{SIFS} is the Short Inter-frame Space. Eq. 9 does not contain any contention waiting periods (*AIFS* and backoff time) since at the start of the SI the contention is already won by the AC_VO.

Regarding the AC_VI queue, which holds all the LPMF frames, we propose that no TXOP limit adaptation takes place, in order to keep complexity as low as possible. However, taking into consideration the low significance of B-frames, the loss tolerance of voice frames and the low sizes of the LPMFs (compared to HPMFs), we propose the usage of the NoACK scheme as the acknowledgment policy of this queue. By doing so, we aim at the reduction of the transmission times of LPMFs (as described in Section III) at the cost of an increased loss probability. Nevertheless, in congested networks the usage of this policy is beneficial in retaining medium quality voice calls, as noted in [27]. Furthermore, the loss of B-frames is acceptable, to a certain degree, since their absence will not significantly reduce the video quality.

VIII. Simulations and Results

In order to evaluate the proposed algorithm the OPNET network simulator was used [16]. In this section we provide a description of the simulation scenarios after which the results that were obtained are analyzed and explained.

A. Setup

We considered an infrastructure IEEE 802.11e network with a QAP and four QSTAs in the QBSS. Eight G.729A (20-Bytes frames transmitted every 20ms) encoded VoIP streams (UP=6) were traversing the network: four in the downlink direction and four in the uplink direction. Four MPEG-4 streaming video flows (UP=4) were destined to the QSTAs from the wireline network. Finally, two HTTP connections (UP=0) were representing best effort traffic. The real video trace “Highway”, available from [28], was used as the transmitted video sequence. The trace was encoded in MPEG-4 CIF (Common Intermediate Format) with a GoP pattern G_9B_2 (IBBPBBPBB), 2000 frames, frame rate of 30 frames per second and 67 seconds duration. The video sequence exhibits a mean bit rate of 0.41 Mbps and a peak bit rate of 1.89 Mbps.

The wireless channel was assumed to be error-free, hence no packets were lost due to fading effects. The PHY data rate was set to 11 Mbps. There were three simulation scenarios: the first scenario applies standard EDCA default values as depicted in Table II, the second applies an implementation of a cross-layer mechanism similar to [25] (named Cross), and the third scenario implements the modified version of EDCA according to the proposed algorithm. The Cross implementation allocates I-frames to AC_VO, while all other multimedia frames (P, B video frames and voice frames) are served by the lower priority AC_VI.

All scenarios had a total simulation time of 120 sec. The starting times of each application with respect to the starting time of simulation run are depicted in Table VI.

TABLE VI. Application Timing Characteristics

Application	Start Time (sec)	Stop Time (sec)	Duration (sec)
VoIP	5	105	100
Video Streaming	15	80	65
HTTP	7	End of Simulation	113

As depicted in Table VI, for 65 seconds all applications coexist in the network, creating a highly congested period at the QAP.

To compare the EDCA performance obtained by all three scenarios, four performance metrics were considered: overall network application-level throughput (goodput), overall network application-level end-to-end delay, average PSNR for video streaming and MOS for VoIP applications.

B. Simulation Results and Analysis

Fig. 5 shows the overall network application-level throughput (goodput) in packets/sec for both VoIP and video streaming applications. The video streaming application enjoys a large improvement in throughput performance as show casted in Fig. 5(a). The overall goodput is almost leveled at 120 packets/sec (for all four video streams) while both the standard EDCA and the implemented Cross-layer scheme exhibit a significant performance degradation due to congestion at the QAP. As far as VoIP goodput is concerned (Fig. 5(b)), looking carefully at the graph one can observe that during the presence of the four video flows, VoIP performance for standard EDCA and Cross exhibits large oscillations while the proposed algorithm produces a smoother graph.

Analyzing the overall application level end-to-end delay, Fig. 6(a) and (b) reveal a significant improvement on the packet delay that both multimedia applications receive. Specifically, the VoIP application, that has strict delay requirements, receives better QoS by applying the proposed algorithm. Video packet delay oscillates well below the delay produced by the other schemes as depicted in Fig. 6(a).

Finally, all the benefits of applying the proposed algorithm are revealed from Fig. 7(a) and (b), where the average PSNR of streaming video and the MOS values for VoIP applications are plotted. The majority of PSNR values are well above 35 dB, indicating an excellent quality of the received video streams. Regarding the VoIP applications, both the EDCA and the Cross schemes are outperformed by the proposed algorithm with acceptable MOS values for downstream calls.

As an overall simulation results conclusion, it can be stated that the proposed algorithm clearly produces a significant EDCA performance improvement. The exploitation of the multimedia packet semantics combined with the appropriate MAC-layer enhancements is capable of dealing with high congestion periods in infrastructure IEEE 802.11e WLANs.

IX. CONCLUSIONS

This paper addresses the challenge of transporting multimedia traffic over IEEE 802.11e congested WLANs. The multimedia frame semantics are exploited to select an appropriate MAC-layer strategy. A TXOP limit adaptation scheme is used together with acknowledgment policies in order to relax the ACs in the congested QAP, and at the same time protect high significance multimedia frames. The proposed semantic-aware algorithm is proven extremely beneficial in terms of application level throughput, end-to-end delay and QoS metrics for video and voice applications.

REFERENCES

[1] A. Politis, I. Mavridis and A. Manitsaris, “Enhancing Multimedia Traffic Performance in IEEE 802.11e Networks”, *in Proc. of the Sixth International Conference in Wireless and Mobile Communications (ICWMC 2010)*, IEEE Press, Valencia, Spain, pp. 125-130, September 2010.

[2] ISO/IEC and IEEE Std., “Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications”, 1999.

[3] IEEE Std. 802.11a, “High-Speed Physical Layer in the 5 GHz Band”, 1999.

[4] IEEE Std. 802.11b, “Higher-Speed Physical Layer (PHY) Extension in the 2.4 GHz Band”, 1999.

[5] IEEE Std. 802.11g, “Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band”, 2003.

[6] F. Cali, M. Conti, and E. Gregori, “IEEE 802.11 Protocol: Design and Performance Evaluation of an Adaptive Backoff Mechanism”, *IEEE Journal on Selected Areas in Communications,* Vol. 18, No. 3, pp. 1776–1784, 2000.

[7] N. Vaidya, P. Bahl, and S. Gupta, “Distributed Fair Scheduling in a Wireless LAN”, in *Proc. Sixth Annual International Conference on Mobile Computing and Networking (MOBICOM)*, Boston, MA, pp. 167-178, 2000.

[8] J. Sobrinho and A. Krishnakumar, “Quality-of-Service in Ad hoc Carrier Sense Multiple Access Networks”, *IEEE Journal on Selected Areas in Communications*, Vol. 17, No. 8, pp. 1353–1368, 1999.

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

[10] IEEE Std. 802.11n, “Enhancements for Higher Throughput”, 2009.

[11] M. M. Rashid, E. Hossain, V. K. Bhargava, “Controlled Channel Access Scheduling for Guaranteed QoS in 802.11e-Based WLANs”, *IEEE Transactions on Wireless Communications*, Vol. 7, No.4, pp. 1287-1297, 2008.

[12] A. Andreadis and R. Zambon, “QoS Enhancement for Multimedia Traffics with Dynamic TXOPlimit in IEEE 802.11e”, *in Proc. of the 18th IEEE International Symposium in Personal, Indoor and Mobile Radio Communications (PIMRC 07)*, Athens, Greece, pp. 16-22, 2007.

[13] M. van der Schaar and D. Sai Shankar, “Cross-layer wireless multimedia transmission: Challenges, principles and new paradigms”, *IEEE wireless Communications*, Vol. 12, No. 4, pp. 50- 58, 2005.

[14] C. Bouras, A. Gkamas and G. Kioumourtzis, “Challenges in Cross Layer Adaptation for Multimedia Transmission”, *in Proc. of the IADIS International Conference Wireless Applications and Computing 2007*, Lisbon, Portugal, pp. 129 – 133, 2007.

[15] IEEE Std 802.1D “IEEE Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges”, Revised version, 2004.

[16] OPNET Technologies Inc., http://www.opnet.com, accessed 13/1/2011.

[17] H. Liu and Y. Zhao, “Adaptive EDCA Algorithm Using Video Prediction for Multimedia IEEE 802.11e WLAN”, *in Proc. of the 2nd International Conference on Wireless and Mobile Communications* (ICWMC), Bucharest, Romania, 2006.

[18] ITU-T Annex A to Recommendation G.729, “Coding of Speech at 8 Kbit/s using Conjugate Structure Algebraic-Code-Excited Linear Prediction (CS-CELP)”, November 1996.

[19] ITU-T Recommendation P.800, “Methods for subjective determination of transmission quality”, 1996.

[20] A. Lazaris, P. Koutsakis and M. Paterakis, “A new model for video traffic originating from multiplexed MPEG-4 videoconference streams”, *Performance Evaluation*, Vol. 65, No. 1, pp. 51-70, 2008.

[21] P. Seeling, M. Reisslein and B. Kulapala, “Network Performance Evaluation Using Frame Size and Quality Traces of Single-Layer and Two-Layer Video: A Tutorial”, *IEEE Communications Surveys and Tutorials*, Vol.6, No.2, pp. 58-78, Third Quarter 2004.

[22] A. Ksentini, M. Naimi and A. Gueroui, “Toward an improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture”, *IEEE Commun. Mag.*, Vol. 44, No. 1, pp.107-114, January 2006.

[23] A. M. Jama, S. Issa and O. O. Khalifa, “Performance evaluation of MPEG-4 video transmission over IEEE 802.11e”, *IJCNS*, Vol. 2, No. 5, pp. 11-15, May 2010.

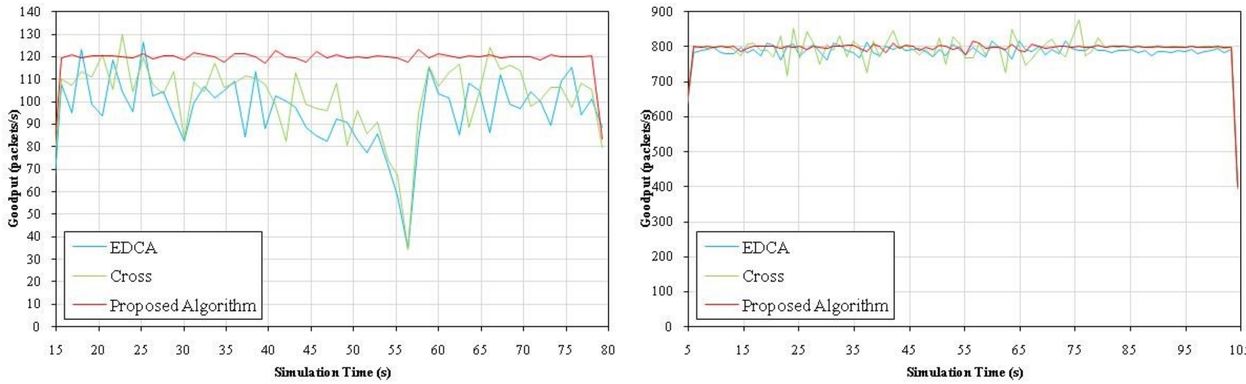
[24] C.-H. Mai, Y.-C. Huang and H.-Y. Wei, “Cross-Layer Adaptive H.264/AVC Streaming over IEEE 802.11e Experimental Testbed”, *in Proc. of the 71st IEEE Conference on Vehicular Technology (VTC 2010)*, Taipei, Taiwan, May 2010.

[25] R. Goel and M. Sarkar, “Enhancing QoS of streaming videos over WLANs”, *in Proc. of IAENG WCECS ’08*, Berkeley, California, 2008.

[26] C.-H. Ke, C.-H. Lin, C.-K. Shieh, N. Chilamkurti and S. Zeadally, "A Novel Cross-Layer Architecture for MPEG-4 Video Stream over IEEE 802.11e Wireless Network", *in Special Issue of International Journal of Telecommunications Systems (SPRINGER)*, Vol. 42, No. 3-4, December 2009.

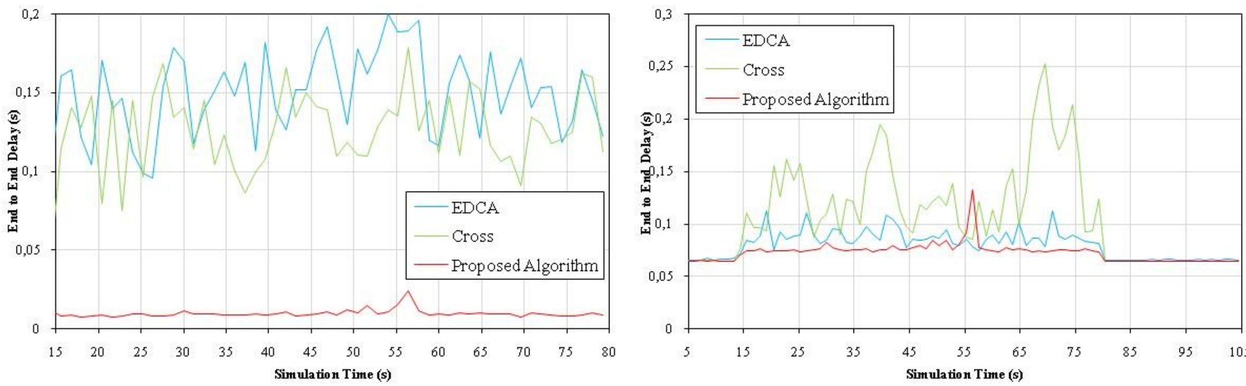
[27] J. Barcelo, B. Bellalta, A. Sfairopoulou, C. Cano, M. Oliver, “No Ack in IEEE 802.11e single-hop ad-hoc VoIP networks”, *in Proc. of MED-HOC-NET ‘08*, Mallorca, Spain, 2008.

[28] Video Traces Research Group, http://trace.eas.asu.edu/, accessed 13/1/2011.



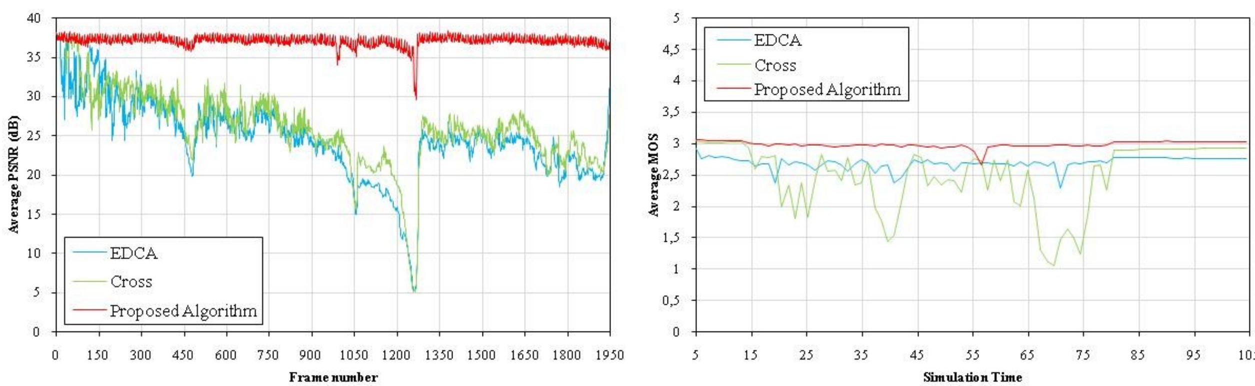
(a) (b)

Figure 5. Overal network application-level throughput (goodput): (a) Video, (b) VoIP.



(a) (b)

Figure 6. Overall network application-level end-to-end delay: (a) Video, (b) VoIP .



(a) (b)

Figure 7. QoS metrics: (a) Average PSNR, (b) Average MOS.