

Notes on the Inefficiency of 802.11e HCCA

C. Casetti, C.-F. Chiasserini, M. Fiore and M. Garetto

Dipartimento di Elettronica, Politecnico di Torino - Italy
E-mail: {casetti,chiasserini,fiore,garetto}@polito.it

Abstract— This paper addresses a simulation study of the QoS-compliant IEEE 802.11e WLAN protocol in a realistic wireless network scenario, featuring client/server file transfer applications as well Voice-over-IP (VoIP) multimedia traffic sources. We focus on the investigation of 802.11e EDCA and HCCA access procedures configuration issues. Simulation results are presented, showing the relevance of a well-tuned calibration of 802.11e parameters on system performances.

I. INTRODUCTION

During the last years, the development of network technologies has brought to an increase in the available connection bandwidth, with the direct consequence of the spreading of streaming audio and video applications. Such multimedia applications introduce new network requirements with respect to those previously asked for by file transfer applications: delays, jitters and packet losses are the parameters on which real-time applications impose quality thresholds in order to run smoothly. Therefore, Quality of Service support is claiming attention as one of the most looked-after features in both wired and wireless network environments.

IEEE 802.11 is the dominant standard for Wireless LANs, but its native QoS support has proved unable to satisfy multimedia requirements. Thus, the IEEE is currently developing a draft, 802.11e WLAN [1], that introduces QoS enhancements to the 802.11 WLAN standard through the so-called Hybrid Coordination Function (HCF). HCF includes two new medium access schemes, the contention-based Enhanced Distributed Channel Access (EDCA) and the contention-free HCF Controlled Channel Access (HCCA).

Analytical and simulation studies have been conducted in order to evaluate the performance of 802.11e access procedures [2], [3], [4], [5]. However, most works focus on a generic estimation of the improvements brought to multimedia traffic by 802.11e. Little attention has been paid instead to the configuration of EDCA and HCCA parameters to fully exploit IEEE 802.11e potential. The simulations we performed with *ns-2* focus on this aspect, also addressing the necessity to carefully consider whether to employ or not the HCCA mechanism, depending on the network scenario. The paper is structured as follows: in Section II an overview of the IEEE 802.11e draft is given and in Section III the network scenario is described. Section IV details the configuration of the 802.11e parameters, while in Section V simulation results are provided. Finally, in Section VI conclusions are drawn.

This work was supported through the PRIN TWELVE project and the FIRB VICOM project

II. IEEE 802.11E OVERVIEW

According to the 802.11e channel structure, the time is divided in Superframes, each made of two parts, the Contention Period (CP) during which EDCA is used, and the Contention Free Period (CFP) that is the phase during which HCCA is employed.

Like DCF, EDCA is based on the CSMA/CA mechanism and employs the concept of Inter Frame Space (IFS) as well as the backoff mechanism. Furthermore it introduces the concept of Transmission Opportunity (TXOP) and Access Category (AC). Whenever an 802.11e station (QSTA) seizes the channel, it is entitled to transmit one or more frames for a time interval named Transmission Opportunity; a TXOP is characterized by a maximum duration, called TXOP_Limit. Data traffic is classified into different priority level, named Access Categories: a QSTA operating under the EDCA function includes up to four MAC queues, each corresponding to an AC and acting as a separate instance of the CSMA/CA protocol (*virtual station*). ACs employ different sets of parameters to be used for contending the channel: (i) the Arbitration Inter Frame Spacing (AIFS[AC]), (ii) the Minimum and the Maximum Contention Window ($CW_{\min}[AC]$), $CW_{\max}[AC]$), (iii) and the TXOP_Limit[AC]¹. The higher the AC priority is, the smaller the AIFS[AC], $CW_{\min}[AC]$ and $CW_{\max}[AC]$ are. The larger the TXOP_Limit[AC], the greater the share of capacity of the AC.

In HCCA channel access handled by a central node, the so-called Hybrid Coordinator (HC). When the CFP must start, the HC waits for a channel idle period of PIFS duration, with PIFS shorter than any AIFS, and broadcasts a Beacon frame: QSTAs will then stay silent until the end of the CFP. The HC then starts to allocate TXOPs to the different stations, following a polling list that records the QSTAs that required to be polled for a particular AC during the CFP. Each TXOP is allocated by sending a QoS CF-Poll frame, specifying the destination AC and the TXOP duration. Once the polled station receives the QoS CF-Poll frame, the QSTA's stated AC can transmit until either it has data to send or the TXOP expires. When all stations in the list have been polled, or if there is not enough time to poll other QSTAs due to CFP duration limit, the HC ends the CFP broadcasting a CF-End frame, and the CP starts again.

¹The value of AIFS must be at least as long as the DIFS interval; the only exception is for the AP that can use an AIFS 30 μ s long.

TABLE I
DATA FLOWS LENGTH FOR TCP TRAFFIC SOURCES

Flow length (bytes)		
119	179	251
334	428	529
658	948	1650
2861	4706	8015
13681	26641	284454

III. SIMULATION SCENARIO

Simulations have been performed in a wired-cum-wireless scenario, where a fixed node is connected to an Access Point (AP), through an over-provisioned 55/Mbps wired link with a 2 ms propagation delay. The AP also acts as HC.

The IEEE 802.11e WLAN protocol is employed in the wireless portion of the network, composed of 20 wireless stations. Each station can exchange data only with the AP. A multi-rate environment is considered: wireless stations are either slow (2 Mbps) or fast (11 Mbps), with the AP switching its data transmission rate according to the destination station, sending data at 2 Mbps to slow stations and at 11 Mbps to fast stations. The selected scenario accounts for a situation where not all users enjoy the same channel conditions (e.g., because they are far from the AP) and therefore must use slower transmission rates. Those control frames that have to be received by all QSTAs are transmitted at the basic rate of 1 Mbps.

The wireless channel is modeled using an independent error model for each communicating pair of nodes, taking account of both short and long channel outages.

VoIP applications are used to load the network. The model chosen for VoIP calls simulates a two-speaker interaction: when one speaker talks, the other is silent. The duration of talk spurts is determined by a Pareto-distributed random variable with values ranging between 4 and 30 seconds. It follows that silences have the same distribution. During a talk spurt, the active source sends data following a real-life trace collected at the output of a voice codec. Two different voice codecs were used: a g.729A vocoder codec and a g.711 vocoder codec. In both cases the Voice Activity Detection (VAD) technique is applied, allowing for coding of the actual voice activity only and avoiding background noise. Resulting from the introduction of VAD, we observed an average encoding bit rate at the output of the g.729A of 2.8 Kbps, with frames of 20 bytes, generated every 20 ms, while at the output of the g.711 vocoder the observed rate was 28 Kbps, using frames of 240 bytes, generated every 30 ms.

VoIP traffic sources use the UDP protocol at transport layer, and UDP-carrying frames are timestamped when they enter the queue at the MAC layer. These frames may be preemptively discarded before transmission if they have exceeded a timeout (set to 400 ms, in our simulations). This avoids useless forwarding of time-critical data that are past their playout time.

Background FTP transfers using TCP are also simulated. To obtain more realistic results, the FTP over TCP sources have been modeled so as to mimic a client-server interaction, such

TABLE II
ACS CONFIGURATION

AC	AIFS	CW	Traffic
AC[0]	50 μ s	3:255	Downlink VoIP
AC[1]	50 μ s	7:511	Uplink VoIP
AC[2]	50 μ s	31:1023	Downlink TCP
AC[3]	50 μ s	63:4095	Uplink TCP

as Web browsing. Each wireless station behaves as a client, while the relative server is located at the fixed node.

The activities of client-server pairs are intertwined, meaning that when a client is active, the relative server is silent and vice versa. Active and silent periods span the transmission of data flows whose length in bytes is determined by a uniformly random selection of one of the values presented in Table I².

IV. CONFIGURATION OF 802.11E PARAMETERS

Traffic generated by VoIP applications are considered as multimedia time-bounded traffic, while TCP traffic is treated as best-effort traffic. as a consequence, higher priority ACs are associated to VoIP data flows, while the ACs with lower priority are assigned to TCP traffic, as in Table II. To provide fairness between uplink and downlink traffic, the latter has higher priority [6].

After extensive testing the AIFS parameter proved to introduce only marginal separation in terms of priority differentiation under the heavy-load traffic conditions we imposed to the network: the long backoff intervals caused by collisions made the contribution of the AIFS value almost negligible. As a consequence, we decided to keep the same AIFS parameter for all the ACs. ACs are then discriminated on a Contention Window basis. In particular, during our tests CW_{min} values proved to be much more effective than CW_{max} values in providing priority differentiation among traffic categories. This happens because CW_{min} is the basic value taken into account for the CW calculation, while CW_{max} is seldom used, since many retransmissions must occur before it is considered. The length of TXOPs allocated for multiple frames transmission during the CP has been set to a fixed value of 2.8 ms.

As far as the HCCA procedure settings are concerned, static length of the TBTT, i.e., the time interval between two Beacon transmissions by the HC, of the CFP and of the TXOPs allocated by the HC resulted in poor performance as the network configuration changed. Thus, we implemented HCCA devising an original solution: a dynamical evaluation of the TBTT interval, of the CFP duration and of the TXOPs allocated during the CFP itself. According, to our solution, these parameters are computed depending on the voice codec employed (g.729A or g.711) and on the ratio of slow to fast stations in the wireless network. The duration values are obtained by computing the number of frames x that should

²These values were derived by measurements of real Internet traffic [8]. If the TCP source is a client, the selection of the file length is truncated to 2861 bytes to keep the client-server interaction more realistic since during a Web browsing session the client requests from the server more data than it transmits.

be transmitted by a wireless station during a polled TXOP, as follows:

$$BI = CFP + CP = x \cdot vocoder_framing \quad (1)$$

where BI is the TBTT interval or Beacon Interval, while CFP and CP stand for the durations of CFP and CP respectively. The $vocoder_framing$ parameter accounts for the duration of the framing time, that is to say the time needed to generate a voice frame, depending on the codec used. Imposing this condition, we state that the duration of BI must equal the time that the vocoder requires to generate x frames. Since we consider x as the number of frames transmitted during a polled TXOP, and since a wireless station is polled once every BI , (1) guarantees that during the polled TXOP the station will be able to transmit all the backlogged frames that the vocoder generated during the BI and that were not yet sent to the AP.

Moreover, we imposed that the CFP use one fifth of the overall superframe duration, so as to better balance the impact of the two access procedures. Therefore, we have

$$5 CFP = x vocoder_framing \quad (2)$$

The CFP parameter is then expanded to

$$\begin{aligned} CFP = & [N_f(Poll_f + SIFS) + N_s(Poll_s + SIFS)] + \\ & [3/2 (N_f/N_f + N_s) (TCP_f - Poll_f) + \\ & 3/2 (N_s/N_f + N_s) (TCP_s - Poll_s)] + \\ & 1/4 N_f [(x/2) (VoIP_f + 2SIFS + ACK_f) + \\ & Null_f + SIFS + Poll_f] + \\ & 3/4 N_f [Null_f + SIFS + Poll_f] + \\ & 1/4 N_s [(x/2) (VoIP_s + 2SIFS + ACK_s) + \\ & Null_s + SIFS + Poll_s] + \\ & 3/4 N_s [Null_s + SIFS + Poll_s] \end{aligned} \quad (3)$$

where N_f and N_s are the number of fast (data transmission rate of 11 Mbps) and slow (data transmission rate of 2 Mbps) wireless stations in the network, $Poll_f$, $Null_f$, ACK_f , TCP_f and $VoIP_f$ are the transmission times of QoS CF-Poll, QoS Null, QoS CF-Ack, TCP data and VoIP data frames for fast stations, while $Poll_s$, $Null_s$, ACK_s , TCP_s and $VoIP_s$ are the equivalent for slow stations.

Equation (3) models a CFP frame exchange that is very similar to the real behavior of the network we got from simulations. The first term within square brackets accounts for QoS CF-Polls that are sent during the polled CFP, N_f at fast transmission rate and N_s at slow transmission rate.

With the second term within square brackets we assume that on average three TCP packets are sent with QoS Data+CF-Poll or QoS Data+CF-Ack+CF-Poll frames every two CFPs and that such frames can be sent to either a fast or a slow QSTA according to a uniform probability distribution. Therefore higher the number of slow stations, higher the probability that the TCP packet will be sent to a slow station.

The third and fourth terms in square brackets model the actual VoIP frame exchange for fast stations: in the first case, occurring on average one fourth of times, the station has VoIP

data to send and $x/2$ frames are transmitted, while in the second case, occurring three times out of four, the station does not have data to send and, as a consequence, it just sends a QoS Null frame to the HC. The QoS CF-Poll that ends the polled TXOP is also considered. Finally, the fifth and sixth terms are the equivalent of the last two for slow stations.

By substituting (3) in (2), we can calculate the wanted value of x . Once x has been evaluated, the durations of Beacon Interval, of the CFP and of the polled TXOPs for fast and slow stations are again obtained in the following way:

$$BI = ceil(x) vocoder_framing \quad (4)$$

$$CFP = BI/5 \quad (5)$$

$$CFP_TXOP_f = ceil(x/2) (VoIP_f + 2SIFS + ACK_f) \quad (6)$$

$$CFP_TXOP_s = ceil(x/2) (VoIP_s + 2SIFS + ACK_s) \quad (7)$$

Equations (6) and (7) guarantee the same transmission possibilities to fast and slow stations. As a matter of fact, in accordance with the third and fifth terms of (3) both kinds of station are allowed to send a number $x/2$ of frames during the polled TXOPs, that are shorter for fast stations and longer for fast ones. This is because we assume the QSTAs to be able to transmit during the CP at least the other half of the x frames generated by the vocoder during the Beacon Interval, since the CP is 4 times longer than the CFP. Notice that increasing the number of slow stations increments the value of x , with longer TBTT intervals and polled TXOPs. As a final remark, notice that the TXOP durations calculated this way are only associated to the polled TXOPs, that is to say TXOPs allocated during the CFP: the $TXOPLimit$ value used for EDCA is set to 2.8 ms and does not vary.

V. SIMULATION RESULTS

We implemented the 802.11e access procedures on the *ns-2* [7] simulator and configured them according to what stated in the previous Section. Simulations were run for a total of 20 stations and a number of slow stations ranging from 0 to 20. Vcoders employed at VoIP traffic sources can be g.711 or g.729A, depending on the simulation set. All simulations were run for 1000 seconds, so to allow the system to reach a steady state.

A first set of simulations features the g.711 vocoder as the multimedia traffic generator at the VoIP source nodes. We first performed a benchmark test, using the IEEE 802.11b protocol at the MAC layer. The resulting average delays plot is shown in Figure 1, where the good and medium quality thresholds of 150 ms and 400 ms, respectively, are reported. Four curves can be distinguished, representing delays for downlink VoIP, uplink VoIP, downlink TCP and uplink TCP. It appears clear that 802.11b cannot guarantee the fulfillment of multimedia traffic requirements, due to high downlink delays.

Figure 2 presents the same average delays obtained employing the IEEE 802.11e EDCA access procedure. The improvement brought by the service differentiation is noticeable, with VoIP delays that drop below the medium quality threshold with

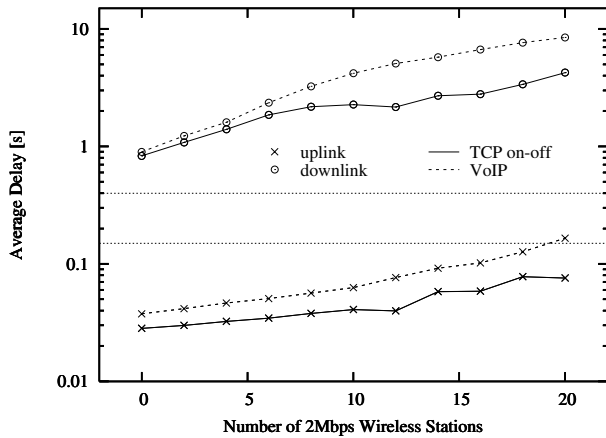


Fig. 1. 802.11b DCF - Delay versus the number of 2 Mbps QSTAs with g.711

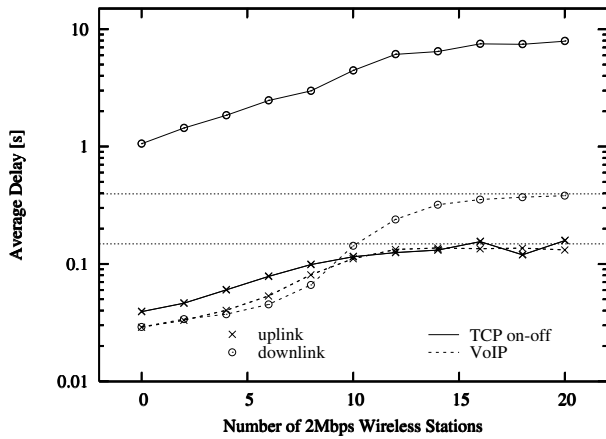


Fig. 2. 802.11e EDCA - Delay versus the number of 2 Mbps QSTA with g.711

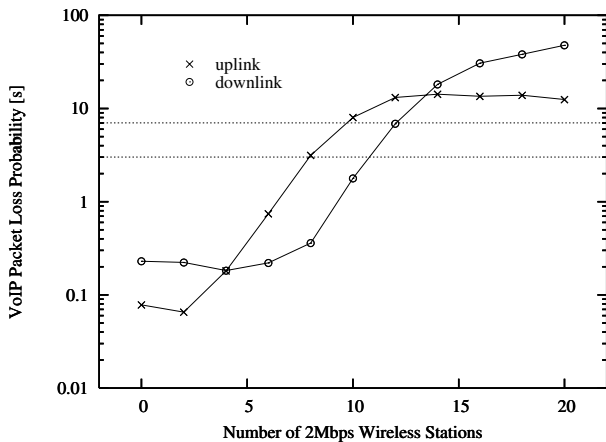


Fig. 3. 802.11e EDCA - VoIP drop probability at MAC layer versus the number of 2 Mbps QSTAs with g.711

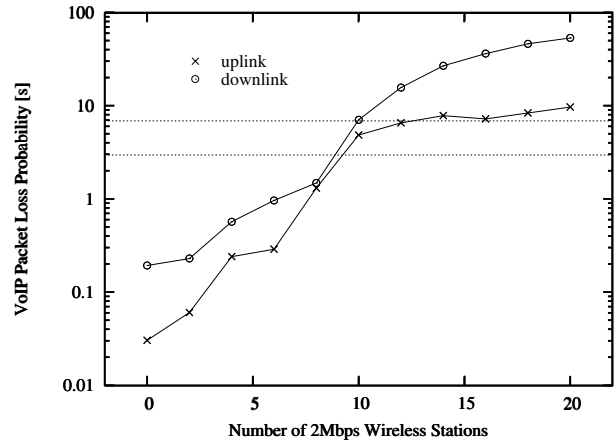


Fig. 4. 802.11e EDCA & dynamic HCCA - VoIP drop probability at MAC layer versus the number of 2 Mbps QSTAs with g.711

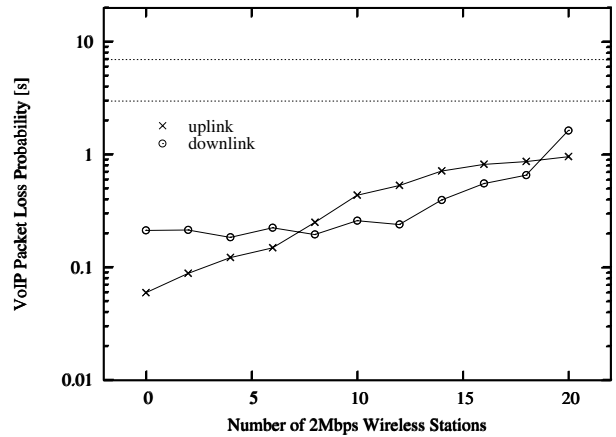


Fig. 5. 802.11e EDCA - VoIP drop probability at MAC layer versus the number of 2 Mbps QSTAs with g.729A

any number of slow QSTAs. With 10 or less slow QSTAs, the quality of VoIP delays becomes good. In Figure 3 the percentage of packets dropped at MAC layer is reported: good performance is reached for losses lower than 3%, while and medium quality is obtained with losses lower than 7%. Drop percentage, taking account for frames dropped because of full buffers as well as for drops due to exceeded timeout, prove to be a most restrictive parameter with respect to the delay, with good results with no more than 8 slow QSTAs and poor quality otherwise.

The introduction of the HCCA mechanism with static parameters does not bring any noticeable improvement: packet loss, is just slightly reduced, still providing good QoS with 8 or less slow QSTAs and poor QoS otherwise. However, employing the dynamic HCCA configuration scheme, the percentage of packets dropped at the MAC layer, shown in Figure 4, is reduced, and now medium quality is obtained with 10 slow QSTAs. Thus, good quality is still achieved with 8 or less slow QSTAs, but now we can guarantee a medium QoS also with 10 slow QSTAs in the network.

We run a second set of simulations using g.729A voice

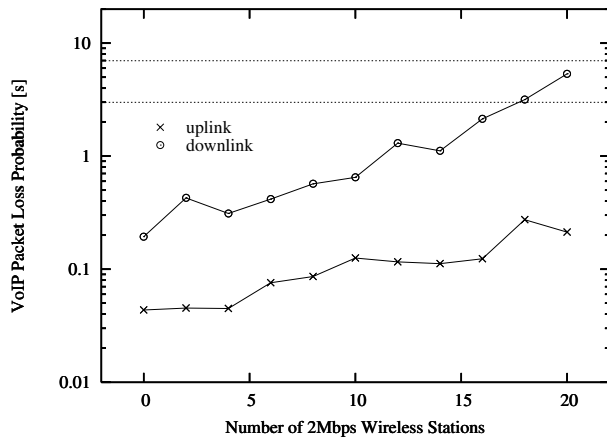


Fig. 6. 802.11e EDCA & HCCA - VoIP drop probability at MAC layer versus the number of 2 Mbps QSTAs with g.729A

codecs as VoIP sources. As already said, the g.729A vocoder has less strict requirements with respect to g.711, resulting in lower quality of voice coding. The percentage of frames discarded at the MAC layer appears to be again the most hard to handle, especially because of the 400 ms timeout we imposed for multimedia data flows, which causes VoIP frames to be automatically discarded at the MAC layer. Figure 5 presents the results obtained with EDCA only. System performances are very good: losses are lower than the 3% threshold independently from the number of slow QSTAs.

However, tests on the same network, conducted employing the complete IEEE 802.11e scheme, show that adding the HCCA access procedure to EDCA does not bring any advantage. In fact, performance are worsened by the introduction of HCCA, even if the dynamic configuration of the parameters, which instead proved to be effective with the g.711 vocoder, is used. This can be noticed by comparing Figure 6, which depicts the percentage of packets lost when also HCCA is used, against Figure 5, where only EDCA is employed. We can conclude that the multimedia traffic generated by the g.729A sources can be efficiently handled by the network when it implements EDCA alone. The reason for the worsened performance in presence of HCCA is that some of the QSTAs transmit all of their VoIP traffic during EDCA, making unuseful the HC's poll during the CFP: in these cases, the HCCA procedure just brings unuseful overhead.

VI. CONCLUSIONS

We run simulation tests of the IEEE 802.11e medium access procedures on a realistic scenario involving VoIP and background best-effort traffic, focusing on EDCA and HCCA configurable parameters tuning.

The presented results led us to conclude that i) the voice codec employed highly affects the maximum number of sustainable wireless stations, ii) the HCCA mechanism should be used only if EDCA is not sufficient, as in the case of high quality voice coding, and iii) a proper configuration of HCCA is needed to obtain noticeable improvements in system performance.

REFERENCES

- [1] IEEE 802.11 WG, Draft Supplement to Standard Part II: Wireless MAC and PHY Layer Specifications: MAC Enhancements for Quality of Service, IEEE 802.11e Draft 5.0, July 2003.
- [2] G. Bianchi and I. Tinnirello, "Analysis of Priority Mechanisms Based on Differentiated Inter Frame Spacing in CSMA-CA," *IEEE VTC-Fall*, Orlando, FL, Oct. 2003.
- [3] A. Grilo and M. Nunes, "Performance Evaluation of IEEE 802.11e," *IEEE PIMRC'02*, Lisboa, Portugal, Sep. 2002.
- [4] S. Choi, J. del Prado, S. Shankar and S. Mangold, "IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation," *IEEE ICC 2003*, Anchorage, Alaska, May 2003.
- [5] A. Lindgren, A. Almquist and O. Schelén, "Quality of Service Schemes for IEEE 802.11 Wireless LANs - An Evaluation," *MONET, Special Issue on Performance Evaluation of QoS Architectures in Mobile Networks*, Vol. 8, No. 3, June 2003, pp. 223-35.
- [6] C. Casetti and C.-F. Chiasserini, "Improving Fairness and Throughput for Voice Traffic in 802.11e EDCA," *IEEE PIMRC 2004*, Barcelona, Spain, Sep. 2004.
- [7] ns-2 - <http://www.isi.edu/nsnam/ns/>
- [8] tstat - <http://tstat.tlc.polito.it/>