
A Scheduling Algorithm for Providing QoS Guarantees in 802.11e WLANs*

G. Cecchetti¹ and A. L. Ruscelli²

¹ CNIT - Scuola Superiore S. Anna g.cecchetti@sssup.it

² Scuola Superiore S. Anna a.ruscelli@sssup.it

Summary. In this paper we propose a scheduling algorithm for supporting Quality of Service (QoS) in an IEEE 802.11e network using the HCF Controlled Channel Access (HCCA) function. The is derived from Constant Bandwidth Server with Resource Constraints and adapted to wireless medium. It consists of a procedure to actually schedule transmission opportunities to HCCA flows with Quality of Service guarantees, in particular in the case of multimedia applications which present variable bit rate traffic.

Key words: Scheduling algorithm, Wireless LAN, Quality of Service.

1 Introduction

In recent years Wireless Local Area Networks (WLANs) are became very popular and the IEEE802.11 [1] has established as the world wide standard. At the same time, the continuous growth in the use of mobile devices that support multimedia applications and real-time services with strict latency/throughput requirements, such as multimedia video, VoIP (Voice Over IP), video-conference over a wireless channel, involves a great interest in the study of appropriate mechanisms to manage the wireless medium in order to achieve the expected *Quality of Service* (QoS). Recently approved IEEE 802.11e standard [2] for WLANs offers a complete set of primitives to provide delay guarantees while the previous IEEE 802.11b [3] was designated only for best effort services. However also IEEE 802.11e does not provide scheduling algorithms for packet transmission nor policies scheme for access control to the medium, leaving space to build blocks for a full Quality architecture.

Many research studies have evaluated the new standard employing analytical techniques [4] and simulations [5] [6], and they have demonstrated

* This work has been carried out in part within the framework of the VICOM project founded by the Italian Ministry of University and Research (MiUR) and in part within FRESOR EU project (Contract n. 034026)

the usefulness of the proposed mechanisms of 802.11e. Subsequent works have proposed several scheduling algorithms to improve the QoS provisioning [7] [8] [9] [10]. This improvement is necessary in particular in the case of VBR traffic for which the *reference* scheduler shows its limit. In fact it is particularly tailored for constant bit rate traffic.

In this paper we propose a new scheduling algorithm for the HCF Controlled Channel Access (HCCA) function of the IEEE 802.11e, namely the Wireless Constant Bandwidth Server with Resource constraints (WCBS-R). It does provide those flows that have been admitted to use the HCCA function with rate base guarantees. This algorithm is derived from real-time systems, and it is actually a modified version of Constant Bandwidth Server with Resource Constraints [11] [12]. The scheduling methodology adopted reserves a fraction of network bandwidth to each flow, assigning a suitable deadline to the server flow whenever the reserved time is consumed. Differently from the *reference* scheduler, WCBS-R is not based on periodic scheduling of fixed allocations but it manages dynamically the allocated capacity. Moreover the latter is made available for contention based access when it is not used by the HCCA flows. Through preliminary simulation results, we show that WCBS-R performs better than the reference scheduler, in terms of the capacity available for legacy DCF access, which is based on contention.

The rest of the paper is organized as follows. Section 2 introduces the IEEE 802.11e HCCA. In Section 4 we describe WCBS-R scheduling algorithm. Preliminary simulation results are discussed in Section 5 and conclusions are drawn in Section 6.

2 IEEE 802.11e protocol description

The new standard IEEE 802.11e introduces a new coordination function called the *Hybrid Coordination Function* (HCF) which multiplexes between two medium access modes: a distributed scheme called *Enhanced Distributed Channel Access* (EDCA) and a centralized scheme called *HCF Controlled Channel Access* (HCCA). To ensure compatibility with legacy devices, the standard allows the coexistence of DCF and PCF with EDCA and HCCA

2.1 Enhanced Distributed Channel Access

EDCA is a channel access mode which provides prioritized QoS and it enhances the original DCF by classifying traffic through the introduction of *Access Categories* (ACs), corresponding to different level of traffic priority. Each AC has its own transmission queue and its own set of channel access parameters. The most important ones are Contention Window (CW_{min} and CW_{max}), which sets backoff interval, and *Transmission Opportunity* (TXOP) limits which is the maximum duration for which a node can transmit after obtaining access to the channel. Using these parameters, when data arrives

from higher layers it is classified and placed in the appropriate AC queue. Then an internal contention algorithm is used to calculate the total backoff time for each AC. The AC with the smallest backoff time wins the internal contention and uses this backoff value to contend externally for the wireless medium. Nodes with higher priority can access the channel earlier than other nodes and prioritized flows have the advantage of longer channel access with their TXOP.

2.2 HCF Controlled Channel Access

HCCA provides a centralized polling scheme to allocate guaranteed channel access to traffic flows based on their QoS requirements. It uses a QoS-aware *Hybrid Coordinator* (HC) which is usually located at the *QoS Access Point* (QAP) in infrastructured WLANs and it provides polled access to the wireless medium. In order to be included in the polling list of the HC, a *QoS Station* (QSTA) must send a QoS reservation request to the QAP, using the special QoS management frame, *Add traffic Stream* (ADDTS), which contains the Traffic Specification (TSPEC), which includes the following parameters: as *mean data rate* (R_i), *nominal Service Data Unit (SDU) size* (N_i), *minimum PHY rate* (Γ_i), *delay bound* (D_i) and *maximum service interval* (MSI).

HC aggregates every TSPEC of QSTA TSs and determines the values of parameters needed by the transmission itself: *Service Interval* (SI) and TXOP. SI is the time duration between successive polls for the node and it is a submultiple of the 802.11e beacon interval duration. TXOP is the transmission duration of each node based on the mean application data rates of its requested flows. Before the calculation of the latter parameters, QAP has to verify if the admission of each TS does not compromise the service guarantees of the already admitted TSs and, if the specified TS is accepted, QAP sends a positive acknowledgement which contains also the service start time that indicates the time from when the QSTA is allowed to transmit frames relative to considered TS.

When there are admitted QSTAs which desire to access the medium, the QAP listen to the medium itself and, if it is idle for a PIFS, HC gains control of the channel and, within the *Controlled Access Phase* (CAP), it polls a single QSTA at turn, according to its polling list, generated by a scheduler.

3 Related work

Scheduling algorithms addressed to wireless networks have to take into account some limits due to wireless environment [13]. In particular, the wireless medium itself presents space and time varying characteristics, unlike what happens in wired networks, so wireless networks are subject to fast changes in *Signal to Interference plus Noise Ratio* (SINR) due to phenomena like *path loss*, *shadowing*, *multipath fading*, *signal attenuation* and *interference* [14].

This implies that the concept of fairness is difficult to apply. Furthermore, wireless resources (e.g. bandwidth and energy) are limited and that, jointly to need of lower computation complexity due to use of low-performance hardware, adds other constraints in the choice of algorithms.

Several studies has been done to verify performances of reference scheduler [5] [15] [6]. According to them, for every QSTA, fixed SI and TXOP based on mean values of the transmission parameters are useful for Constant Bit Rate TS, while they do not reflect the fluctuation of Variable Bit Rate TS. Particularly reference scheduler performances are evaluated using heterogeneous traffic stream like VoIP (G.711 codec), video stream (MPEG4 codec) and burst *best effort* data stream.

Some alternative algorithms introduce the following features: a) variable SI and/or TXOP, b) feedback based mechanism, c) queue length model.

3.1 Deadline-based algorithms

The results obtained in the field of scheduling real-time tasks in a multiprogrammed environment has been adapted to the context of HCCA scheduling.

In [16] the authors propose the SETT-EDD scheduling algorithm which limits the amount of time during which the stations control the wireless medium, it improves the performance of the scheduler and it enhances its flexibility. It uses the mean TXOP as a guideline for allocating time and uses a token bucket scheme of time units or TXOP timer to allow nodes to vary their TXOP over time according to their needs. The TXOP timer of station j increases at a constant rate equal to TD_j/mSI_j (where mSI_j is minimum SI of j^{th} QSTA), which corresponds to the total fraction of time the station can spend in polled TXOPs. The TXOP timer has a maximum value equal to MTD_j (where MTD_j is the Maximum Time Duration of j^{th} QSTA). The time spent by a station in a polled TXOP is deducted from the TXOP timer at the end of the TXOP. The station can be polled only when the value of the TXOP timer is greater than or equal to mTD_j , which ensures the transmission of at least one packet at the minimum PHY rate.

The authors also propose to change the service interval for each node based on the traffic profile and use Earliest Deadline First (EDF) to determine the polling order. If the due time to poll a station is t , the next poll shall be issued on a time t' that satisfies the relation: $t + mSI < t' < t + MSI$. Time instant $t + mSI$ is the instant after which the next poll can be done, equivalent to the release time in the real-time scheduling theory. Time instant $t + MSI$ is the maximum time by which the next poll has to be done, or deadline time.

It has been shown that the proposed flexibility in the scheduler for voice and video traffic leads to significant reduction in average transmission delay (up to 33 percent) and packet loss ratio (up to 50 percent).

In [17] the authors combine the Earliest Deadline First (EDF) algorithm with the Stack Resource Policy (SRP), which are efficient policies for scheduling real-time periodic tasks in a multiprogrammed environment. When a TS

requests admission to the QAP, RTH computes a periodic timetable in which TSs are granted a fixed capacity. This allows TSs with different requirements to be scheduled efficiently, and therefore RTH admits more TSs than the sample scheduler. However, RTH cannot deal with VBR traffic.

3.2 Algorithms based on queue length estimation

Another approach to dealing with uplink VBR traffic is to estimate the length of uplink TS queues, and tune the length of TXOPs granted to them.

FHCF [18] tries to improve the fairness both of CBR and VBR flows by assigning variable TXOPs. These are computed using queue length. Actually FHCF is composed of two schedulers: the QAP scheduler estimates the varying queue length for each QSTA before the next SI and compares this value with the ideal queue length. The QAP scheduler uses a window of previous estimation errors for each TS in each QSTA to adapt the computation of the TXOP allocated to that QSTA. Because sending rate and packet size can change, this estimation can not be accurate. After this comparison QAP computes the additional requested time (positive or negative) for each TS of each QSTA and reallocates the corresponding TXOP duration. Then, the node scheduler located in each QSTA can redistribute the unused time among its different TSs since the TXOP is always allocated to a whole QSTA. It computes the number of packets to transmit in the TS and time required to transmit a packet according to its QoS requirements. Later, according to its allocated TXOP, it evaluates the remaining time that can be re-allocated. This is possible since each QSTA knows its TS queue size at the beginning of polling phase and it is able to estimate its queue length at the end of TXOP and the requested additional time for TS.

Performance study indicates that FHCF provides good fairness while supporting bandwidth and delay requirements for a large range of network loads and, because it uses to allocate TXOP the mean sending rate of VBR applications instead of the maximum sending rate usable for the standard HCF scheme, it may recover much time and more flows can accepted in HCCA. Furthermore, it is more efficient than the reference scheduler, admitting an higher number of traffic streams.

FBDS [19] assigns dynamically the TXOP according to queue length estimation while SI remains fixed. All the QSTAs which compose the communication system and its transmission queues are regarded as a system whose balance is perturbed by new incoming flows. The FBDS periodic scheduler, which uses HCF, behaves as a closed loop controller which restores this balance by bandwidth recovering. This is possible due queue length information sent by each QSTA through a 8-bit subfield of QoS Control Field. Moreover the closed loop system uses a discrete time model which permits to estimate queue length at beginning of new CAP phase and so it acts as compensation system against errors produced by channel perturbations not previewed by the scheduling algorithm.

This algorithm guarantees the delay bounds required by audio/video applications in presence of very broad set of traffic conditions and networks loads by using a control system action which ensures a maximum delay for queuing new frames.

4 WCBS-R scheduling protocol

The QAP schedules traffic streams using an algorithm derived from the soft real-time scheduling literature, with regard to the Constant Bandwidth Server with Resource Constraints (CBS-R) scheduling algorithm [12]. The CBS algorithm was modified to suit the needs of wireless traffic and named WCBS-R. It can be defined as follows:

1. A WCBS-R is characterized by a capacity c_i and by an ordered pair (Q_i, T_i) , where Q_i is the maximum capacity and T_i is the service interval of the TS_i . The ratio $U_i = Q_i/T_i$ is denoted as the TS_i bandwidth. At each instant, a fixed deadline d_i is associated with the TS_i . At the beginning $d_i = 0$.
2. Each served transmission chunk $H_{j,k}$ is assigned a dynamic deadline $d_{j,k}$ equal to the current TS_i deadline d_i .
3. Whenever a TS_i is served for transmission, the capacity c_i is decreased by the same amount.
4. When $c_i = 0$, the TS_i capacity is recharged at the maximum value Q_i and a new TS_i deadline is generated as $d_i = d_i + T_i$. Notice that there are no finite intervals of time in which the capacity is equal to zero.
5. A WCBS-R is said to be **active** at time t if there are pending transmissions (remember the budget c_i is always greater than 0); A WCBS-R is said to be **idle** at time t if it is not **active**.
6. When a transmission of TS_i arrives and the WCBS-R is **active** the request is enqueued in a queue of pending transmissions according to a given (arbitrary) discipline (e.g., FIFO).
7. When a flow TS_i arrives and the WCBS-R is **idle**, if $c_i \geq (d_i - r_i)U_i$ the scheduler generates a new deadline $d_i = r_i + T_i$ (where r_i is the arrival time of TS_i) and c_i is recharged at the maximum value Q_i , otherwise the scheduler generates a new deadline $d_i = \max(r_i + T_i, d_i)$ and the capacity becomes $c_i = c_i + (d_i^{new} - d_i^{old})U_i$.
8. When a TS finishes, the next pending transmission, if any, is served using the current capacity and deadline. If there are no pending transmission, the scheduler becomes **idle**.
9. At any instant, a TS is assigned the last deadline generated by the scheduler.
10. Whenever a served flow TS_i tries to access a critical section, if $c_i < \xi_i$ (where ξ_i is the duration of the *longest critical section* of TS_i such that $\xi_i < Q_i$), a capacity replenishment occurs, that is $c_i = c_i + Q_i$ and a new scheduler deadline is generated as $d_i = d_i + T_i$.

4.1 Admission Control

Let B_i denote the maximum duration of the transmission time of TS s with period longer than TS_i . A sufficient condition for a set of n TS each one characterized by a period T_i and capacity C_i , to be schedulable is as follows:

$$\frac{B_i}{T_i} + \sum_{j \leq i} \frac{Q_j}{T_j} \leq 1, \quad \forall i : 1 \leq i \leq n. \quad (1)$$

So, when admitting a stream i , the QAP has to calculate its Q_i and T_i , and to check if eqn. (1) holds. Given the TSPEC for i we define:

$$Q_{min} := \lceil \frac{R_i \cdot T_i}{N_i} \rceil, \quad Q_{max} := \lceil \frac{II_i \cdot T_i}{M_i} \rceil$$

where II_i is the peak data rate, M_i is the maximum SDU size for the i^{th} TSPEC. For WCBS-R, we use: $Q_i = Q_{min} + CWF(Q_{max} - Q_{min})$. For T_i we use the maximum service interval (MSI). The QAP keeps track of the allocated capacity, and when doing admission control it checks if a new stream would require more capacity than the system can provide.

If it can be admitted and it is a downlink stream no more actions than updating the currently used capacity have to be performed, otherwise, if it is an uplink one, the stream is added to the polling list, with a poll time p_i equal to the current time, so that it will be polled as soon as possible, on the next call to the scheduler.

4.2 Enqueueing a Packet

When a packet arrives, the QAP has to check if its associated stream i was already active. If it was not, it has to check if the remaining c_i can be given to the stream without exceeding the Q_i/P_i utilization of the medium, otherwise it has to postpone the deadline of the stream, replenishing its capacity.

4.3 Dequeueing a Packet

When in a CAP the QAP has to chose the next packet to send, it first updates the status of the stream being served, changing its capacity as needed, and updating its deadline if necessary. Then it checks if there are polling streams that can be added to the active list (i.e., their p_i is passed,) changing their state and requeueing them if necessary.

It then requeues the active stream if it has switched to a polling or idle state or if it is no more the one with the earliest deadline, selecting the next task in Earliest Deadline First (EDF) order.

If there are no active streams a CP is started. If the selected stream i is an uplink one the corresponding station is given a TXOP of c_i , otherwise the packet to be sent is extracted from the QAP queues.

5 Experimental results

In this section we analyze WCBS-R through simulation.

5.1 Simulation settings

The physical layer parameters are those specified by the High Rate Direct Sequence Spread Spectrum (HR-DSSS) [3], also known as 802.11b, and are reported in Table 1.

Table 1. MAC/PHY simulation parameters

Parameter	Value
SIFS (μs)	10
PIFS (μs)	30
DIFS (μs)	50
SlotTime (μs)	20
PHY header(μs)	192
Data rate (Mb/s)	11
Basic rate (Mb/s)	1
Bit error rate (b/s)	0

We focus on the system performance in ideal conditions so we assume that the channel is error-free, while MAC level fragmentation and multirate support are disabled. Furthermore we assume that all nodes can directly communicate with each other. Therefore, the hidden node problem and the packet capture are not taken into consideration and the RTS/CTS protection mechanism is disabled.

We have implemented the proposed W-CBS in the ns-2 network simulator [20], using the HCCA implementation framework described in [21]. Then we compared the results with respect of reference IEEE 802.11e standard scheduler. The analysis has been carried out using the method of independent replications. Specifically we ran independent replications of 600 seconds each with 100 seconds warm-up period until the 95% confidence interval is reached for each performance measure. Confidence intervals are not drawn whenever negligible.

Then we compared the results with respect of reference IEEE 802.11e standard scheduler.

5.2 Admission control analysis

We first evaluated the performance of W-CBS in terms of the admission control limit.

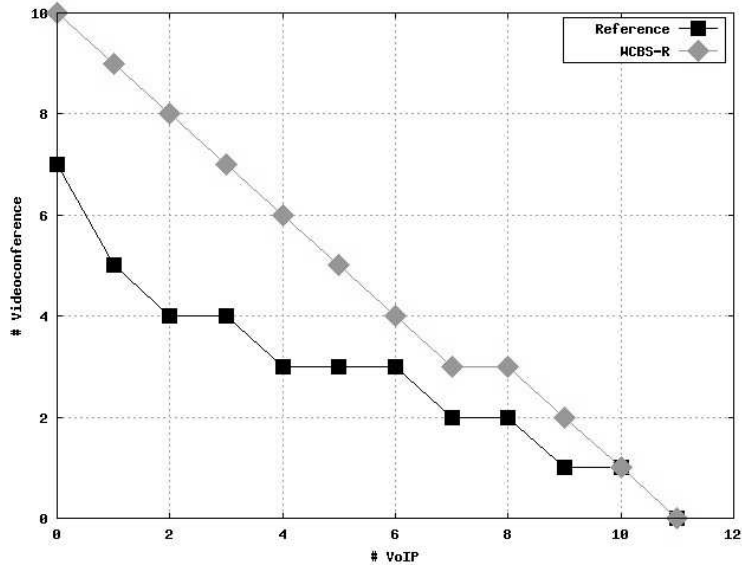


Fig. 1. Admission control. Number of admitted videoconference TSs against the number of admitted VoIP G.711 TSs.

Fig. 1 shows the number of admitted videoconference TSs, as a function of the number of admitted VoIP G.711 TSs. In both cases, the sample scheduler curve lies significantly below the WCBS-R curve. This behavior confirms that the sample scheduler cannot efficiently accommodate TSs with different TSPECs. In fact, firstly, it polls TSs with $\Delta i \neq SI$ more often than needed, by setting the scheduling duration to the smallest TS period. Secondly, it overestimates the capacity needed by TSs.

5.3 Data Throughput

We evaluate a scenario with four stations with mixed CBR and VBR traffic. To do so, we set up an increasing number of QSTAs, from 0 to 4, each having a bi-directional VoIP TS and bi-directional Video Conference TS. The delay bound of VoIP is set to 20 ms and that of VC TSs to 33 ms.

The Fig. 2 shows the throughput achieved by stations with data traffic against the number of stations with bi-directional VoIP and VC sessions. Stations with data traffic operate in asymptotic conditions, i.e. they always have a frame to transmit. The packet length of data traffic is constant and equal to 1500 bytes. If there are not any stations with CBR and VBR TSs, the data throughput is maximum and WCBS-R behaves in a very similar way to the reference scheduler. Otherwise, if there are TSs with significantly different delay bound requirements, such as the VoIP and VC TSs, the MAC overhead of the reference scheduler is higher than that with framework scheduler and, therefore, the throughput achievable by data traffic is much lower.

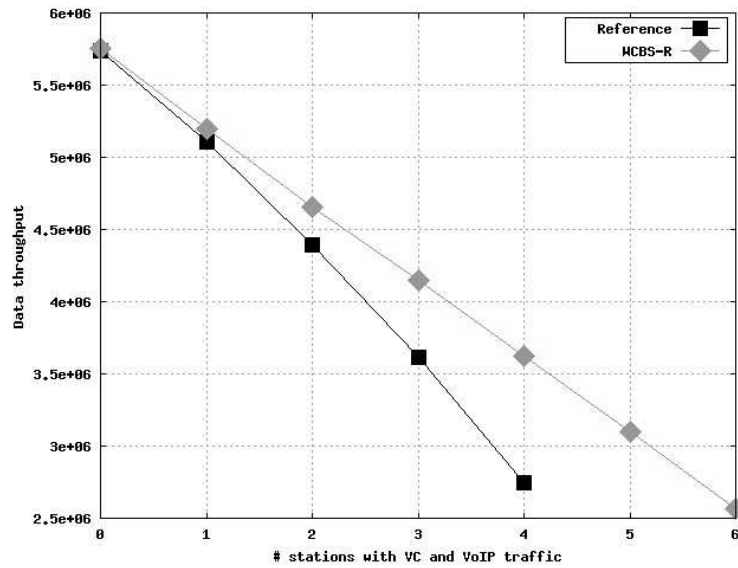


Fig. 2. Throughput of stations.

6 Conclusions

In this paper we have defined a new scheduling algorithm alternative to the reference scheduler to integrate a QoS support in IEEE 802.11e wireless networks with soft real-time guarantees. The scheduling algorithm, named WCBS-R, is based on CBS-R Real-Time algorithm which permits to dynamically manages the medium resources and it supports variable packet size and variable bit rate traffic streams. The simulation analysis shows that WCBS-R efficiently services QoS traffic with different delay bounds, thus saving capacity which is made available to contention-based data traffic and outperforming the reference scheduler.

References

1. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.
2. *802.11E-2005 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 8: Medium Access Control (MAC) Quality of Service Enhancements*, IEEE Std. 802.11e, 2005.
3. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification: Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, IEEE Std. 802.11b, 1999.

4. H. Zhu and I. Chlamtac, "An analytical model for IEEE 802.11e EDCF differential services," in *Proc. 12th International Conference Computer Communication and Networks*, Dallas, TX, Oct 2003.
5. A. Grilo and M. Nunes, "Performance evaluation of IEEE 802.11e," in *Proc. PIMRC 2002*, vol. 1, Lisboa, Portugal, Sep 2002, p. 511–517.
6. J. Cowling and S. Selvakennedy, "A detailed investigation of the IEEE 802.11e HCF reference scheduler for VBR traffic," in *13th International Conference On Computer Communications And Networks (ICCCN 2004)*, Chicago, US, Oct 2004.
7. H. Fattah and C. Leung, "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wireless*, vol. 9, no. 5, pp. 76–83, Oct 2002.
8. S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," *IEEE/ACM Trans. Net.*, vol. 7, no. 4, pp. 473–489, Aug 1999.
9. S. Tsao, "Extending earliest due-date scheduling algorithms for wireless networks with location dependent errors," in *Proc. IEEE VTC. 2000*, Boston, MA, Sep 2000.
10. A. Grilo, M. Macedo, and M. Nunes, "A service discipline for support of IP QoS in IEEE 802.11 networks," in *Proc. PWC 2001*, Laapenranta, Finland, Aug 2001.
11. L. Abeni and G. Buttazzo, "Integrating multimedia applications in hard real-time systems," in *IEEE Real Time-Systems Symposium*, Dec. 1998, pp. 4–13.
12. M. Caccamo and L. Sha, "Aperiodic servers with resource constraints," in *IEEE RealTime Systems Symposium*, Dec. 2001.
13. Y. Cao and O. L. Li, "Scheduling algorithms on broad-band wireless networks," *Proceeding of the IEEE*, vol. 89, pp. 76–87, Jan. 2001.
14. Y. L. Boudec and T. Thiran, "A short tutorial on network calculus i: Fundamental bounds in communication networks," in *IEEE ISCAS*, Geneva, Switzerland, May 2000, pp. IV–93–IV–96.
15. S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, and L. Stibor, "IEEE 802.11e wireless LAN for quality of service," in *Proc. European Wireless (EW2002)*, 2002.
16. A. Grilo, M. Macedo, and M. Nunes, "A scheduling algorithm for QoS support in IEEE 802.11e networks," *IEEE Wireless Communications*, pp. 36–43, jun 2003.
17. C. Cicconetti, L. Lenzini, E. Mingozzi, and G. Stea, "Efficient provisioning of real-time qos guarantees in iee 802.11e wlans," in *European Wireless 2006*, Athens, April 2006.
18. P. Ansel, Q. Ni, and T. Turletti, "An efficient scheduling scheme for IEEE 802.11e," in *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2004.
19. A. Annese, G. Boggia, P. Camarda, L. A. Grieco, and S. Mascolo, "Providing delay guarantees in IEEE 802.11e networks," in *59th IEEE Semiannual Vehicular Technology Conference, VTC Spring*, 2004.
20. *Network Simulator 2*, <http://www.isi.edu/nsnam/ns/>.
21. C. Cicconetti, L. Lenzini, E. Mingozzi, and G. Stea, "A software architecture for simulating iee 802.11e hcca," in *3rd IPS MoMe*, Warsaw (Poland), Mar 2005.