

Analytical developments on QoS enhancements provided by IEEE 802.11 EDCA

Jeremy Gosteau, Mohamed Kamoun, Sebastien Simoens, Pietro Pellati
Motorola Labs, Paris
Email: jeremy.gosteau@motorola.com

Abstract—This paper quantifies the gain in throughput and Quality of Service (QoS) provided by the EDCA MAC adopted in the current IEEE 802.11e WLAN draft compared to the traditional DCF. In that purpose, an analytical model is developed to compute the throughput in the single and multi-user cases with a non ideal channel. This model is backed with system simulations in realistic deployment scenarios.¹ Results make a clear statement on the advantage of EDCA to guarantee some QoS. Final discussions provide inputs to further improve both the delivered QoS per user and the overall cell throughput for next generation of Wireless LANs.

I. INTRODUCTION

Most international bodies (IEEE802.11 in the US, ETSI BRAN in Europe and ARIB MMAC in Japan) put much effort to define a standard for Wireless Local Area Network (WLAN). To date, the available devices are all based on the Medium Access Control (MAC) specified by the IEEE802.11 working group. Since the first standard in 1997, most enhancements were made to increase the PHY data rate. And, indeed, the latest IEEE802.11a [1] and the ETSI BRAN HIPERLAN/2 [2] standards allow raw bit rates up to 54 Mbps (compared to the initial 2 Mbps). New modulation techniques and coding schemes allowed such progress. Besides, a clear concern arose to increase the actual throughput on top of the MAC layer and to provide end-to-end Quality of Service (QoS) and security. Therefore, modifications in layer 2 algorithms were necessary. HIPERLAN/2 proposes a centralized Time Division Duplex with Multiple Access (TDD/TDMA) scheme: this allows an optimization of the channel occupancy, which can increase the achievable throughput for applications [3], compared to the IEEE802.11 choice to use Carrier-Sense Multiple Access (CSMA) for its simplicity. Moreover, having a centralized process that manages smartly the allocation of the resources among the users in a TDD/TDMA scheme allows a better guarantee of QoS than a CSMA-based system. As a matter of fact, the latter introduces a new source of packet loss (besides noise and interference) which is collision.

In the IEEE802.11e standardization process (currently drafted [4]), the basic CSMA access scheme has been further improved to provide more QoS to the end users. To do so, two techniques manage the access to the channel which is partitioned in super-frames. A super-frame consists of a Contention Period (CP) and a Contention Free Period (CFP). During

the CP, the Enhanced Distributed Channel Access (EDCA) supports priority differentiation between stations and flows by using different back-off parameters for different traffic types. EDCA was previously known as Enhanced Distributed Coordination Function (EDCF) in earlier versions of the draft. An additional Hybrid Coordination Function (HCF), usually located in the Access Point (AP) on top of the EDCA, manages the access to the channel for all users during both the CP and the CFP.

This paper will not study the impact of HCF on the QoS as there are many ways of implementing it (depending mainly on how the scheduler manages the traffic specifications). The focus here is just to see how the EDCA performs compared to the DCF in terms of QoS. The technical interest behind this is the analytical model developed to obtain the throughput for such a system. This model is based on work performed by Bianchi [5] that uses a two dimensional Markov chain. But, if Bianchi focused his attention on DCF, this paper enlarges the model to the EDCA case and to an error-prone channel. This is key to get realistic results. Moreover, these results are compared to those obtained with a network simulator. The analytical developments performed here enable to distinguish this paper from previous works (such as [6]) on EDCA that were mainly based on simulation results. Other papers have also studied either the effect of channel errors on DCF ([7]) or the gain in QoS in ideal conditions ([8]).

The paper is organized as follows. After a brief description of the standard to understand how the DCF and the EDCA manage the channel, the single-user approach will explain the methodology to compute the throughput. This analysis is then expanded to the multi-user case. In both cases, illustrations are provided by running system simulations.

II. EDCA IEEE802.11E MAC PROTOCOL

The introduction mentioned that the EDCA basis is similar to DCF except that it supports priority differentiation between stations and flows by using different back-off parameters. This section describes the way EDCA manages the access to the channel. A preliminary knowledge of the CSMA technique managed by the DCF defined by the IEEE 802.11 [9] is nevertheless required. It must be mentioned at this point that in this paper, each station manages a single flow. This assumption avoids any confusion regarding the access to the channel (there is no contention nor virtual back-off among several flows within one station to gain the access to the channel).

¹The authors wish to thank the other contributors who develop the network simulator notably the Motorola Labs team working in Sydney (MARC/WTL).

A. DCF basic procedure

The basic DCF is a two-way hand-shaking mechanism which consists in waiting for a random amount of time after Distributed IFS (DIFS) expiration: when the medium is sensed idle during a DIFS, every station willing to send a pending packet waits for a discrete random amount of time called back-off time.

The time after DIFS expiration is slotted in small intervals with a fixed duration σ . The back-off time is handled by the back-off counter which is initialized to a random integer chosen uniformly in the range $[0 : CW - 1]$, where CW is the Contention Window. If the medium is sensed idle for a slot time σ , the back-off counter is decremented. This decrementation stops whenever the medium becomes busy and can only resume if the medium stays idle during a DIFS. The station that generates the smallest back-off counter wins the access to the channel when this counter reaches 0.

To decrease further the collision probability, the value of CW is doubled after each transmission failure (ranging from CW_0 up to a pre-determined maximum CW_m).

Since wireless devices are not able to detect collision occurrence, all successfully received packets need to be acknowledged by an ACK packet. To avoid these packets colliding with other traffic, they are transmitted a Short IFS (SIFS) after successful decoding of its respective data packet. The SIFS is the smallest possible IFS. The mechanism is illustrated on figure 1.

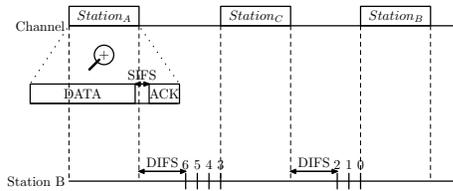


Fig. 1. DCF mechanism

B. EDCA procedure

The EDCA mechanism is able to resolve up to eight User Priorities (UP) by mapping them on four Access Categories (AC). Each AC can be assigned to multiple priorities. To do so, an AC is defined by three parameters: $CW_0[AC]$, $CW_m[AC]$ and $AIFS[AC]$. An AIFS (or Arbitrary IFS) behaves like a DIFS but its length is not fixed.

With these basic principles, EDCA provides some QoS that is analyzed later in this paper. But before going into these discussions, some primary analysis is performed in the single user case (one stream in the cell). It is mainly based on a work achieved by Bianchi [5] with adapted derivations and interpretations of the obtained results. As a matter of fact, in a single-user perspective, DCF and EDCA manage the channel in the same way. The difference with [5] is that, instead of having contention between stations (leading to collisions), the single user suffers from errors on the channel occurring with a constant Packet Error Rate of PER . This methodology will

be double-checked by using another method and by running system simulations.

III. THROUGHPUT CALCULATION FOR A UNIQUE CONNECTION

A. Definitions and assumptions

In what follows, the propagation delay is neglected compared to all other durations. As defined in [5], the time is discrete but the interval between two incrementations of the time axis (i.e. t and $t+1$) is not constant. It increases only when the channel is considered in idle state by the station. This defines a slot time as follows:

- σ when no packet is sent
- $T_s = T_{DATA}(P_{size}, \varphi) + T_{SIFS} + T_{ACK}(\varphi) + T_{AIFS}$ for a successful transmission
- $T_f = T_{DATA}(P_{size}, \varphi) + T_{AIFS}$ for a failed transmission due to an error on the channel with probability PER

Where P_{size} refers to the packet size and φ to the nominal bit rate.

It is also assumed that the station has always pending packets to transmit (saturation state) and that the ACKs suffer no error during their transmission.

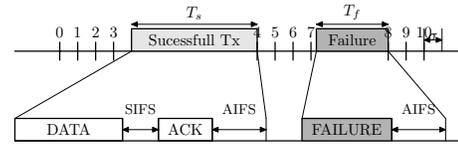


Fig. 2. Discrete time axis

In this paper, we assume that $AIFS[AC] = DIFS$ for every AC because we want to isolate the effect of the choice of CW_0 and CW_m in the prioritization. The collision window ranges from CW_0 to CW_m with:

$$CW_m = CW_0 * 2^m \quad (1)$$

B. Two dimensional channel modeling

With these assumptions, the channel seen by the station can be modeled with a two dimensional Markov chain [5]. The chain state is the couple $(s(t), b(t))$ with t , the discrete time defined above and where:

- $s(t)$ is the index of the contention window; if the station is making its first attempt to transmit a packet, $s(t) = 0$, otherwise, $s(t)$ is incremented; $s(t)$ ranges from 0 to m
- $b(t)$ is the value of the back-off counter; it is necessarily decremented from t to $t+1$ or recomputed when reaching 0; $b(t)$ ranges from 0 to $CW_i - 1$ (where i is the value taken by $s(t)$)

The stochastic process $(s(t), b(t))$ is indeed Markovian as $(s(t+1), b(t+1))$ depends only on $(s(t), b(t))$ and on the success or failure of a transmission attempt at time t .

According to [5], the transition probabilities are given by:

$$\begin{cases} P\{(i, k)|(i, k+1)\} = 1 & (k, i) \in ([0..CW_i - 2], [0..m]) \\ P\{(0, k)|(i, 0)\} = \frac{(1-PER)}{CW_0} & (k, i) \in ([0..CW_0 - 1], [0..m]) \\ P\{(i, k)|(i-1, 0)\} = \frac{PER}{CW_i} & (k, i) \in ([0..CW_i - 1], [1..m]) \\ P\{(m, k)|(m, 0)\} = \frac{PER}{CW_m} & k \in [0..CW_m - 1] \end{cases} \quad (2)$$

From (1) and (2), it is possible to derive the probability that a transmission occurs (whether successful or not):

$$P_{tr} = \frac{2}{(CW_0 + 1) + CW_0 \frac{1-(2PER)^m}{1-2PER} PER} \quad (3)$$

This result is similar to the one obtained in [5]. In the present case, instead of having collisions due to other stations' transmissions (translated into a collision probability p), the channel "delivers" incorrectly the packet with a probability PER . From the station point of view, the situation is similar.

C. Normalized throughput computation

The normalized throughput is defined by:

$$\rho = \frac{E[\text{duration to transmit payload in a slot time}]}{E[\text{duration of a slot time}]} = \frac{E_1}{E_2} \quad (4)$$

where the "slot time" is defined in III-A and

$$E_1 = T_{DATA}(P_{size}, \varphi) \cdot (1 - PER) \cdot P_{tr} \quad (5)$$

$$E_2 = (1 - P_{tr})\sigma + P_{tr}((1 - PER) \cdot T_s + PER \cdot T_f) \quad (6)$$

This gives after simplifications:

$$\rho = \frac{T_{DATA}(P_{size}, \varphi)}{\frac{\sigma}{2} \left(\frac{CW_0 - 1}{1 - PER} + CW_0 \frac{PER}{1 - PER} \frac{1 - (2PER)^m}{1 - 2PER} \right) + T_s + \frac{PER}{1 - PER} T_f} \quad (7)$$

If the denominator is expanded in the form of a Taylor series as a function of PER , (7) becomes:

$$\rho = \frac{T_{DATA}(P_{size}, \varphi)}{\sum_{k=0}^{\infty} \alpha_k PER^k + T_s} \quad (8)$$

where the first terms are logically explained (the generic expression for α_k is not derived because it is not relevant in this context):

- $\alpha_0 = (CW_0 - 1) \frac{\sigma}{2}$ the back-off time before any transmission
- $\alpha_1 = T_f + (2CW_0 - 1) \frac{\sigma}{2}$ the time to transmit the packet erroneously a first time followed by a new back-off, occurring with probability PER
- $\alpha_2 = T_f + (4CW_0 - 1) \frac{\sigma}{2}$ the time to retransmit the packet erroneously a second time followed by another back-off, occurring with probability PER^2
- T_s the time to retransmit it successfully

In other words, equation (7) or (8) could have been obtained with a simpler approach, but it was necessary and interesting to adopt the same approach as Bianchi for two reasons. First, we assess the methodology with an error-prone channel. Second, this method can be re-used for a larger number of stations with EDCA and an error-prone channel.

D. Illustration by system simulation

System simulations have been run in the 802.11a context [1] to corroborate the model presented in the previous subsection. In that purpose, the following parameters and assumptions are considered:

- PER varying from 0 to 50 % - in reality, to avoid the PER being higher than a few percents, some Link Adaptation algorithms are used to enable the system to switch to a more robust physical mode when necessary;
- the ACKs are considered error free;
- packet size of 1024 bytes;
- physical layer bit rates ranging from 6 to 54 Mbps as shown on Figure 3;
- $CW_0 = 16$ and $m = 6$;
- $\sigma = 9 \mu s$, $T_{SIFS} = 16 \mu s$, $T_{DIFS} = T_{SIFS} + 2\sigma$;
- the other numerical values (for the computation of T_{ACK} and T_{DATA}) are available in the standard [1].

Figure 3 plots both the curves obtained from equation (8) and the output of the system simulator. This validates the analytical model presented in III-C.

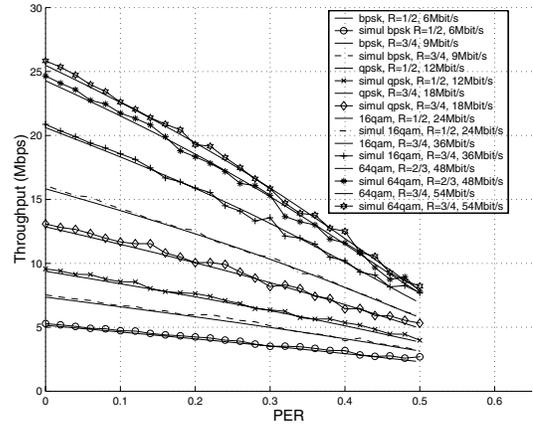


Fig. 3. Throughput for a single connection on an error-prone channel

IV. THROUGHPUT CALCULATION FOR MULTIPLE CONNECTIONS

A. Definitions and assumptions

In this section n users are considered in the cell. They all access an error-prone channel with EDCA. Each user manages a single connection. Up to N_p priorities (denoted j) can be used and therefore, a user is defined by three parameters: $CW_{0,j}$, $CW_{m,j}$ and $AIFS_j$. For the sake of simplicity, the AIFS selected for any priority is equal to a DIFS and m , that defines the maximum contention window size, is identical for all users. The only remaining parameter that differentiates the users in terms of priority is $CW_{0,j}$. And clearly, the smaller the $CW_{0,j}$, the higher the priority j .

Let n_i be the number of users of priority j ($j \in [1..N_p]$) and $\sum_{j=1}^{N_p} n_j = n$. It is assumed that all the n_j stations sending

data at priority j see the same collision probability P_{c_j} over the channel.

Although it would be possible to consider variable parameters for each station, we assume that all stations transmit packets of the same size (P_{size}) at the same raw bit rate (characterized by φ) on the same error-prone channel (all non colliding packets are received erroneously with probability PER). This is incorrect in real deployments (PER typically ranging from 1 to 20 % for all stations) but this allows to control the respective weights of failures due to noise and to collisions in the overall transmission failure probability.

If the time axis is defined as in III-A (figure 2), T_f now corresponds to a failed transmission due to noise with probability PER or to a collision occurrence on the channel with probability P_{c_j} . Let p_j be the failure probability viewed by the station when a packet is transmitted: either a collision is sensed on the channel (with probability P_{c_j}) or there is no collision but the packet is erroneous (with probability PER). Hence:

$$p_j = P_{c_j} + (1 - P_{c_j})PER \quad (9)$$

We indeed assume that a collision systematically results in a failure for the colliding packets. In the following, errors due to thermal noise will simply be called errors as opposed to collisions.

It is still assumed that all users are in saturation state and that the ACKs suffer no error.

B. Two dimensional channel modeling

The same Markov chain as in III-B can be considered with our assumptions. The transitions probabilities proposed in the set of equations (2) are re-used but PER is replaced by p_j . From these equations, the transmission probability of a station of priority j is given by:

$$\tau_j = \frac{2}{(CW_{0,j} + 1) + CW_{0,j} \frac{1 - (2p_j)^m}{1 - 2p_j} p_j} \quad (10)$$

In addition to that, the collision probability P_{c_j} depends also on the other stations' transmission probabilities:

$$P_{c_j} = 1 - \frac{\prod_{l=1}^{N_p} (1 - \tau_l)^{n_l}}{1 - \tau_j} \quad (11)$$

Combining equations (9) and (11) provides the failure probability p_j as a function of the transmission probabilities τ_j :

$$p_j = 1 - \frac{\prod_{l=1}^{N_p} (1 - \tau_l)^{n_l}}{1 - \tau_j} + \frac{\prod_{l=1}^{N_p} (1 - \tau_l)^{n_l}}{1 - \tau_j} PER \quad (12)$$

The set of equations provided in (10) and (12) can be

arranged in a vectorial form:

$$\begin{aligned} \mathbf{T} &= \begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_{N_p} \end{bmatrix} = \mathcal{F} \left(\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{N_p} \end{bmatrix} \right) \\ \mathbf{P} &= \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{N_p} \end{bmatrix} = \mathcal{G} \left(\begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_{N_p} \end{bmatrix} \right) = \mathcal{G}(\mathcal{F}(\mathbf{P})) \end{aligned} \quad (13)$$

Solving this set of equations can be done by noticing that \mathbf{P} is a fixed point of $\mathcal{G} \circ \mathcal{F}$. For example, \mathbf{P} can be obtained with a gradient algorithm by ensuring adequate initial conditions and by noticing that $\|\mathcal{G} \circ \mathcal{F}(\mathbf{P}) - \mathbf{P}\|^2$ is locally convex around its null point.

C. Normalized throughput computation

The same definition of the throughput as in (4) is used, where:

$$E_1 = T_{packet}(P_{size}, \varphi) \cdot P_s \cdot P_{tr} \quad (14)$$

$$E_2 = (1 - P_{tr})\sigma + P_{tr}(P_s \cdot T_s + (1 - P_s) \cdot T_f) \quad (15)$$

P_s is the probability of success providing a packet is transmitted and P_{tr} is the probability of transmission. They are computed as follows:

$$P_s = \sum_{j=1}^{N_p} P_{s_j} = \frac{(1 - PER) \sum_{j=1}^{N_p} n_j \tau_j \frac{\prod_{l=1}^{N_p} (1 - \tau_l)^{n_l}}{1 - \tau_j}}{P_{tr}} \quad (16)$$

$$P_{tr} = 1 - \prod_{l=1}^{N_p} (1 - \tau_l)^{n_l} \quad (17)$$

Where P_{s_j} is the probability that one among the n_j stations of priority j has a successful transmission.

All these equations enable the computation of the normalized cell throughput ρ . The normalized throughput for user j is given by

$$\rho_j = \frac{\rho}{n_j} \frac{P_{s_j}}{P_s} \quad (18)$$

By setting $N_p = 1$ and $n = 1$, equation (7) is obtained.

D. Illustration by system simulation

In the example, 20 users are divided into two categories of priority. The higher priority users have a $CW_{0,2}$ of 32, whereas the others have a $CW_{0,1}$ of 64. In figure 4, the number of stations in each category ranges from 0 to 20. The PER is supposed null. The overall cell throughput remains nearly constant no matter what the number of stations in each category is. If the number of lower priority stations is equal to the number of higher priority stations, the throughput for the higher priority stations is almost twice as much as the throughput of the others.

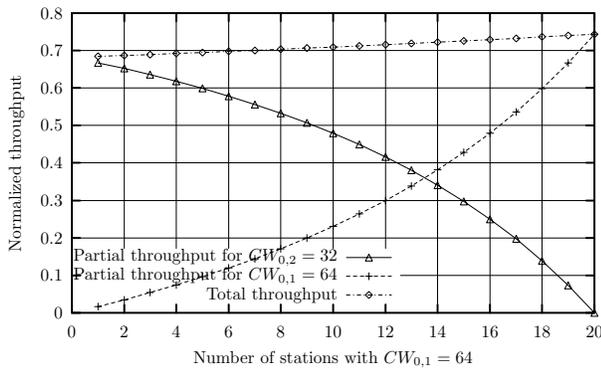


Fig. 4. EDCA throughput with 20 stations

Figure 5, where the throughput is displayed per user, illustrates the same phenomenon: a high priority station wins the channel about twice as much as the lower priority does. And this is an illustration of the actual QoS provided by EDCA.

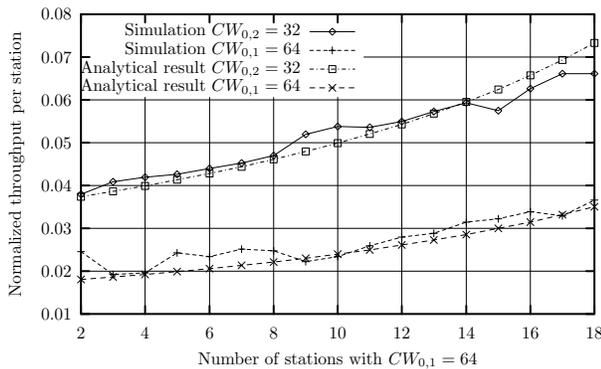


Fig. 5. EDCA normalized throughput per station

This can be checked by using the equations developed in the previous subsections (IV-B and IV-C), since the throughput ratio between two stations of different priority is expressed as follows (from (10), (16) and (18)) in the two priority case:

$$\frac{\rho_2}{\rho_1} = \frac{CW_{0,1} \frac{1-p_1(1+(2p_1)^m)}{1-2p_1} - 1}{CW_{0,2} \frac{1-p_2(1+(2p_2)^m)}{1-2p_2} - 1} \quad (19)$$

If $CW_{0,i}(1-p_i) \gg 1$ and if $p_1 \approx p_2 \ll 1$, equation (19) can be approximated by:

$$\frac{\rho_2}{\rho_1} \approx \frac{CW_{0,1}}{CW_{0,2}} \quad (20)$$

Besides, equation (20) shows that the influence of the PER in the QoS provided by a smaller contention window is limited. Figure 6 illustrates this remark: the throughput ratio between two stations of different priorities is plotted against the PER. And indeed, the PER hardly affects the gain in throughput. It can be noticed though, that an increasing PER defavors the stations of higher priority.

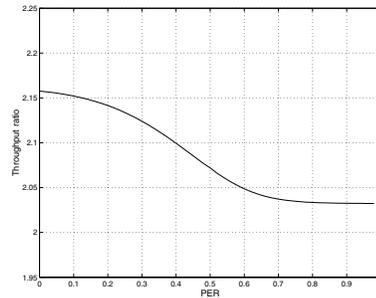


Fig. 6. Ratio between high and low priority throughputs

V. CONCLUSION

This paper has presented an analytical approach to compute the cell and user throughput in an 802.11 EDCA network under realistic channel conditions. This approach has been validated by simulation results. They both show the interest in introducing EDCA in terms of QoS. Indeed, by adding some differentiation amongst the traffic, EDCA statistically provides relative QoS reliability. This aspect can be used with non real-time applications classified in a priority scale.

However, EDCA cannot guarantee a service for real-time traffic. To perform that, a centralized access is preferable. HIPERLAN/2 propose a solution with a TDD/TDMA access technique. The contention free access of IEEE 802.11e HCF is able to control the access over the whole super-frame duration by taking into account the user traffic specification.

REFERENCES

- [1] IEEE 802.11a, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer specifications (PHY) - High Speed Physical Layer in the 5GHz band," IEEE Standards Department, New York, IEEE Std 802.11a-1999, Jan. 1999.
- [2] ETSI Normalization Committee, "Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) Layer," European Telecommunications Standards Institute, Sophia-Antipolis, Valbonne, France, Norme ETSI, document RTS0023003-R2, Feb. 2001.
- [3] A. Doufexi, S. Armour, M. Butler, A. Nix, D. Bull, J. McGeehan, and P. Karlsson, "A comparison of the HIPERLAN/2 and IEEE 802.11a wireless LAN standards," *IEEE Communications Magazine*, vol. 40, pp. 172-180, May 2002.
- [4] IEEE 802.11 Task Group e, "Draft Supplement to IEEE Std 802.11 - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer specifications (PHY) - Medium Access Control (MAC) Enhancements for Quality of Service (QoS)," IEEE Standards Department, New York, IEEE Std 802.11e/D6.0, Nov. 2003.
- [5] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distribution Coordination Function," *IEEE Journal on Selected Areas in Communications*, 2000.
- [6] S. Choi, J. del Prado, S. Mangold, and S. Shankar, "IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation," in *Proceedings of the Int. Conf. on Communications*, Anchorage, Alaska, USA, May 2003.
- [7] J. Yeo and A. Agrawala, "Packet Error Model for the IEEE 802.11 MAC Protocol," *Proceedings of IEEE Personal Indoor Mobile Radio Communications Conf.*, Sept. 2003.
- [8] J. Zhao, Q. Zhang, Z. Guo, W. Zhu, and Y.-Q. Zhang, "Throughput and QoS Optimization in IEEE 802.11 WLAN," *3G Wireless 2002 and WAS 2002*, 2002.
- [9] IEEE 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," IEEE Standards Department, New York, Standard, Nov. 1997.