Available online at www.sciencedirect.com
ScienceDirect
Computer Communications 30 (2007) 2686–2695
computer communications
www.elsevier.com/locate/comcom

Adjustment mechanism for the IEEE 802.11 contention window: An efficient bandwidth sharing scheme

Yassine Chetoui [a,*], Nizar Bouabdallah [b]

[a] PRiSM Laboratory, UMR 8144, 45, Avenue des Etats-Unis, 78035 Versailles, France
[b] IRISA-INRIA, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France

Received 29 September 2006; received in revised form 7 June 2007; accepted 8 June 2007
Available online 14 June 2007

Abstract

In this paper, we propose a new solution to cope with the unfairness limitations of the Distributed Coordination Function (DCF) algorithm. Indeed, in current widely deployed IEEE 802.11b Wireless Local Area Networks (WLAN), the performance of all the competing access nodes are dramatically affected once the bit rate of one station degrades. This anomaly is due to the unfairness behavior of the DCF algorithm. To avoid this, our solution is based on multiple backoff windows principle. We demonstrate through both analytical models and simulations the efficiency of our proposal. Our results show that the proposed algorithm enables fair bandwidth sharing and increases significantly the total network throughput.

Keywords: 802.11 anomaly; Distributed Coordination Function (DCF) algorithm; Multiple Backoff Windows DCF-MB; Fair bandwidth sharing; Network throughput increase; Analytical models and simulations results

1. Introduction

IEEE 802.11 Wireless Local Area Network (WLAN) is now ubiquitous in access networks. The WLAN hotspots are widely deployed in residence, enterprise and public areas. In such networks the main concern of operators is ensuring fair and efficient sharing of the common bandwidth among competing access nodes while maximizing the network throughput.

Actually, access is arbitrated by the use of the Distributed Coordination Function (DCF) algorithm, which is based on Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) technique. This non centralized algorithm (i.e. DCF) strongly participates in the success of IEEE 802.11b thanks to its simplicity. Nevertheless, this basic concept presents two main drawbacks. First, DCF algorithm is unsuitable for Quality-of-Service (QoS) aware applications. In view of this, the IEEE 802.11 IETF Working Group is currently defining a new supplement to the existing legacy 802.11 MAC sub-layer in order to support QoS feature. The new 802.11e MAC [1] will therefore expand the 802.11 application domain by enabling new applications such as voice and video services.

The second drawback behind DCF algorithm is its throughput unfairness issues. This issue is known in literature as the 802.11b anomaly [2,3]. Accordingly, the throughput of each contending access node is drastically reduced once a station transmission bit rate decreases due to physical radio properties. Specifically, a node that is relatively far from the Access Point (AP) is subject to important signal fading and interference leading to repeated unsuccessful frame transmission. As a result, the current deployed IEEE 802.11b reacts by degrading the station bit rate from the nominal 11 Mbit/s rate to 5.5, 2 or 1 Mbit/s. Doing so, the station throughput as well as all the contending access nodes throughput are degraded due to the unfairness limitations of the CSMA/CA-based DCF algorithm. In other words, all the stations are

* Corresponding author. Tel.: +33619441804.
E-mail addresses: Yassine.chetoui@prism.uvsq.fr (Y. Chetoui), Nizar.bouabdallah@inria.fr (N. Bouabdallah).

0140-3664/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2007.06.006

penalized due to the position of one station. Indeed, the basic CSMA/CA scheme allows a fair access to the shared channel. In this regard, a station with a relatively low bit rate (e.g. 1 Mbit/s) captures the channel a longer period with respect to the remaining stations transmitting at 11 Mbit/s. This leads to a degradation of all the access nodes' throughput.

To alleviate this problem, we advise a new strategy based on multiple backoff windows concept. We refer to this technique as the DCF-MB (DCF Multiple Backoff). Considering our scheme, access nodes are classified into different sets according to their physical transmission bit rate (11, 5.5 or 1 Mbit/s). Moreover, each set will be characterized by each own backoff window.

The rest of this paper is organized as follows. In Section 2, we revise the related works presented in the literature, pointing out our position relative to these works. Section 3 analyses the DCF anomaly through simulation illustrations. In Section 4, we describe our proposed solution. Then, in Section 5, we introduce a mathematical model to assess the impact of the proposed DCF-MB solution on the network performances. Section 6, presents analytical and simulation results to evaluate the fairness introduced by our scheme as well as its impact on the total network throughput. Finally, conclusions are drawn in Section 7.

2. Related works

In the wireless literature, several studies dealt with the analysis of the unfairness behavior of the DCF protocol due to the basic CSMA/CA algorithm [2,3]. These works studied this concern without providing particular solutions. Specifically, the work in [2] analyzed theoretically the DCF anomaly by deriving simple expressions of the useful throughput. Furthermore, in [3], authors focus on the short-term unfairness of CSMA-based medium access protocol. They evaluated the short-term fairness degree through experimental and analytical methods.

Recently, there have been increasing research effort to overcome the 802.11 limitation [9,10]. Typically, these works propose to solve the 802.11 anomaly based on the SNR (signal-to-noise ratio) classes differentiation. Specifically, the authors in [9] provide each SNR node class with a distinct and fixed backoff value. This solution fixes the contention window interval size for each SNR class to 1, in the sense that $CW_{min} = CW_{max}$ for each SNR class (i.e. the nodes belonging to the same SNR class choose always the same backoff value equal to $CW_{min} = CW_{max}$). Particularly, the authors in [9] define two SNR classes: the good and low SNR classes, with, respectively, 31 and 1023 as backoff values.

In doing so, the authors in [9], allow the good SNR nodes to transmit more frequently than the low SNR nodes. As such, the throughout of good SNR nodes is no more deteriorated due to the low-speed transmitting nodes belonging the low SNR class. The basic idea behind this simple solution is interesting and can represent a good framework for further research. However, this solution is limited to the unrealistic case of two classes of SNR, whereas in reality much more SNR classes are defined. For instance, four SNR classes are defined in the 802.11b standard and eight SNR classes are defined in the 802.11g. Moreover, fixing the contention window interval size for each SNR class to 1, results in an increasing number of collisions in the network.

To overcome the 802.11 anomaly, the authors in [10] provide each SNR class with its specified maximum packet size. In other words, the authors specify the maximum packet size that can be transmitted by each access node according to its associated SNR class. Typically, the low SNR classes are provided with low values of maximum packet sizes. As such, the amount of traffic transmitted by the low SNR classes decreases as opposed to that transmitted by the high SNR classes, which increases. This solution is shown to be effective, but it entails two main limitations. First, applying this approach involves several modifications to the existing 802.11 protocol and induces thus some protocol complexities. Specifically, a new layer is required at each access station to fragment the packets according to the maximum packet size associated to each SNR class. Furthermore, such fragmentation operation may be not desired in shared wireless access networks since it impacts the access delay and deteriorates the resource utilization efficiency. Typically, low SNR classes are attributed very small maximum packet size threshold. Such a threshold on packet sizes may be too small compared to the header size, which results in unsatisfactory goodput for this class of nodes. Moreover, the packet fragmentation increases the number of packets contending for the channel access, increasing thus the collision probability in the network.

On the other hand, unfairness engendered by the TCP utilization in IEEE 802.11 WLAN was extensively addressed in [4–6]. Nonetheless, the proposed solutions are TCP-specific and are not adapted to our case of study.

Recently, some service differentiation schemes have been proposed for the IEEE 802.11 DCF to support QoS feature [1,7,8]. The basic idea consists in providing a priority scheme for the DCF. The differentiation is simply achieved through varying the amount of time a station would sense the channel to be idle and the length of the contention window for a backoff. Such methods give an access priority for the shared medium to hosts with stringent QoS requirements but without resolving the above-mentioned unfairness issue.

In this study, we adapt these priority mechanisms to achieve fairness. As a key distinguishing feature from existing literature, we provide an effective solution to the unfairness concern with minor changes in the DCF algorithm.

3. IEEE 802.11 DCF anomaly

The IEEE 802.11b standard defines two access methods: the DCF technique, which is based on the CSMA/CA protocol, and the centralized Point Coordination Function (PCF). Unlike DCF, the PCF method provides free collision access via a central arbitration by a point coordinator, which resides in the AP. Even though, the PCF method is barely implemented in today's products due to its complexity. In contrast, DCF thanks to its simplicity is the main reason of the tremendous growth in IEEE 802.11 installation.

As stated before, the DCF access method is based on the CSMA/CA principle. Accordingly, a host wishing to transmit a frame senses the channel activity until an idle period equal to Distributed Inter Frame Space (DIFS) is detected. Then, the station waits for a random backoff interval before transmitting. The backoff time counter is decremented in term of time slots as long as the channel is sensed free. The counter is suspended once a transmission is detected on the channel. It resumes with the old remaining backoff interval when the channel is sensed idle again for a DIFS period. The station transmits its frame when the backoff time becomes zero.

If the frame is correctly received, the receiving host sends an Acknowledgement (ACK) frame after a Short Inter Frame Space (SIFS). If the sending host does not receive this ACK frame, a collision is assumed to have occurred. In this case, the sending host attempts to send the frame again when the channel is free for a DIFS period augmented by the new backoff calculated as follows.

For each new transmission attempt, the backoff interval is uniformly chosen from the range $[0, CW]$ in term of slot of times. At the first transmission attempt of a frame, CW equals the initial backoff window size, CW_{min}. Following to each unsuccessful transmission, CW is doubled until a maximum backoff windows size value CW_{max} is reached. Once the frame is successfully transmitted, the CW value is reset to CW_{min}. Fig. 1 illustrates the DCF mechanism.

In essence, the DCF algorithm ensures equal access to the shared medium among all the contending stations. However, equal access probability does not guarantee a fair medium occupancy among all the hosts. Specifically, a station moving away from the AP may result in the degradation of its nominal bit rate (i.e. 11 Mbit/s) to 1 Mbit/s. In this case, it captures the channel for a period 11 times longer than the period required by a station close to the AP to transmit the same frame. In this regard, this kind of access policy may not be desired since it is extremely penalizing for all the stations. In addition, this issue affects the total network throughput.

To illustrate this anomaly, we consider the simple example of three station-access network. The three contending

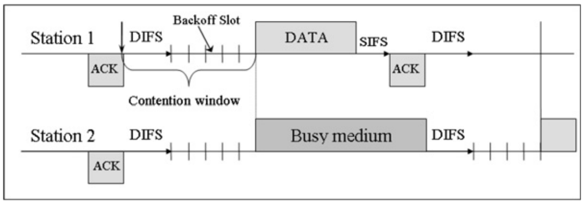


Fig. 1. The basic access mechanism of IEEE 802.11 CSMA/CA-based DCF.

Table 1
Parameter of IEEE 802.11b

PLCP preamble	18 bytes
PLCP header	6 bytes
ACK	14 bytes
Class bit rate	1, 5.5, 11 Mbit/s
DIFS	50 μs
SIFS	10 μs
Backoff slot of time	20 μs
CW_{min}	31
CW_{max}	1023

access stations are situated at different distances from the AP. Accordingly, the first station, which is the closest node to the AP, transmits at a bit rate equal to 11 Mbit/s. The second station transmits at 5.5 Mbit/s and the third station at 1 Mbit/s. We assume that packets arrive with the same rate at each station buffer level according to a Poisson process. In this example, the arrival rate is set high enough, so that, there is always at least one frame in each host buffer.

Moreover, in our simulation, station 2 is activated at $t_1 = 10$ s and station 3 is activated at $t_2 = 40$ s. This scenario enables us to check the network throughput evolution. We note that the DCF parameter settings used in our simulations are depicted in Table 1.

As depicted in Fig. 2, during the first 10 s, only station 1 is activated. Its useful throughput is maximal and attains 6.41 Mbit/s, which represents nearly 0.6 of its transmission bit rate (11 Mbit/s). This difference is mainly due the backoff delay, SIFS and DIFS free periods left on the medium for each frame transmission.

Once the station 2 is activated, the first station throughput logically reduces. But, this reduction is dramatic since the new useful throughput 2.42 Mbit/s is less than the half of the old throughput (6.41 Mbit/s). Moreover, we point out that both stations present the same throughput although their different bit rates. Indeed, the throughput of the first station is decelerated due the relatively low bit rate of station 2. This is typically due to the CSMA/CA policy, which allows fair access probability between both

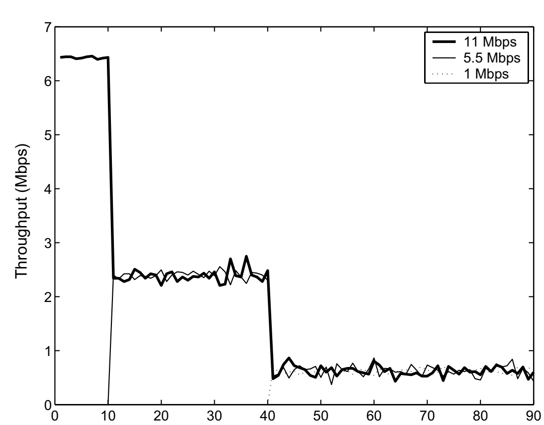


Fig. 2. Throughput of each node according to its class bit rate: DCF case.

stations but does not ensure a fair medium occupancy in term of time. In this case, station 2 occupies the channel twice more time than the first station. As a result, station 1 is unfairly penalized as well as the total network throughput, which significantly decreases as it passes from 6.41 to 4.84 Mbit/s.

This anomaly is more pertinent when station 3 is activated. In this case, the useful throughput of each station is limited to only 0.57 Mbit/s and the total throughput becomes 1.71 Mbit/s.

4. The proposed solution: DCF-MB

To relieve this issue, we advise a method that ensures a fair channel sharing in term of time occupancy among the contending nodes instead of ensuring fair access probability. To achieve this, we give different access priority to different hosts according to their transmission bit rate classes (11, 5.5 or 1 Mbit/s). Let us revisit the example of Section 3. As stated before, thanks to its relatively high transmission bit rate (i.e. 11 Mbit/s), station 1 sends the same frame two times faster than station 2 and 11 times faster than station 3. In view of this, station 1 has to access the channel two more times than station 2 and 11 more times than station 3 in order to obtain a fair occupancy of the medium.

This aim can be simply accomplished with centralized systems such as the PCF technique by allocating more time to the high-priority classes. Nonetheless, such centralized methods are not deployed due to their complexity. On the other side, one possible solution to achieve this, while keeping the DCF algorithm, is to use a priority scheme. Such a scheme can be easily designated with minor changes in DCF.

The key idea behind our proposal is to provide each class bit rate C_i with its associated initial contention window size $\mathrm{CW}_{\min}(i)$ for backoff procedures. Specifically, $\mathrm{CW}_{\min}$ (1) associated to class C_1 (i.e. 11 Mbit/s) is set equal to 31 as specified in the standard. Moreover $\mathrm{CW}_{\min}(i)$ of class C_i is derived as follows:

$$\mathrm{CW}_{\min}(i) = \mathrm{CW}_{\min}(1)\frac{r_1}{r_i} \tag{1}$$

where, r_i denotes the bit rate of class C_i.

Specifically, in our study, we assume three classes of stations. According to (1), we get $\mathrm{CW}_{\min}(1) = 31$, $\mathrm{CW}_{\min}(2) = 60$ and $\mathrm{CW}_{\min}(3) = 330$, which are the window sizes of classes C_1 (11 Mbit/s), C_2 (5.5 Mbit/s) and C_3 (1 Mbit/s), respectively.

Doing so, we guarantee, for instance, that the average backoff counter timer of class C_1 is the half of that of class C_2 (5.5 Mbit/s). Hence, we ensure that class C_1 stations access the medium two more times than C_2 stations.

Finally, we underline that the main advantage of this method is its simplicity. It requires minor modifications in the existing DCF. Indeed, each station modulates its contention window size according to its current physical bit rate. This decision is taken locally, at the station level, without requiring any extra communications with the AP, keeping thus the simplicity and the distributed feature of DCF.

5. Performance analysis

In this section, we present mathematical models for both DCF and DCF-MB schemes. Solving these models, we derive the total network throughput as well as the throughput per station. To achieve this, we first calculate the collision probability in such networks caused by the multiple access nodes. Then, we derive the expression of the average time required by the network to send a frame. Based on these results, we simply get the network throughput.

In this study, we assume N stations contending to access the common data channel. As before, the N stations are divided into M sets of classes $C_1, \ldots, C_M$ according to their physical bit rates $r_i (i = 1, \ldots, M)$. Moreover, we denote by n_i the number of stations belonging to class C_i. Note that nodes belonging to class C_1 have the maximal bit rate, while those belonging to C_M have the lowest bit rate.

5.1. Basic assumptions

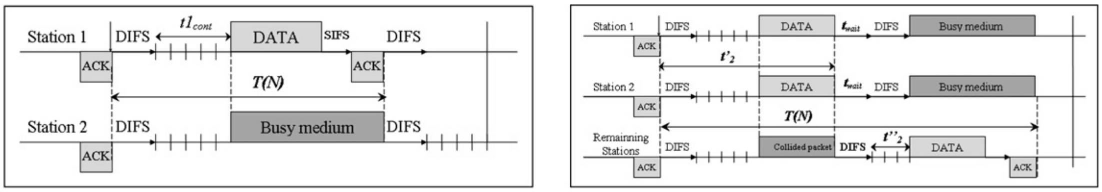
We use the following classical assumptions in our study:

- The number of transmissions that are subject to multiple successive collisions is negligible. This assumption, denoted henceforth by assumption 1, is widely used in literature to simplify the analytical models.
- Accordingly, following to a successful transmission, we can also assume that the backoff $B_i (i = 1, \ldots, N)$ of each access station takes a value in $[0, \mathrm{CW}_{\min}]$. This second assumption holds since we omit successive collisions occurrence as explained in [2]. The accuracy of these approximations is justified, as it will be demonstrated in the next section, through the perfect match between the analytical and simulation results.
- There is always at least one frame in each host buffer (saturation condition).
- The parameter settings in our study are listed in Table 1.

5.2. Probability of collision

Let us now calculate the probability of collision occurrence $P_c(N)$, among the N competing access nodes, during the next transmission cycle (TC). The TC is defined as the average time required by the network to send successfully a frame. Before we delve in calculations, as a first main contribution, we outline that our method gives simple expression and more accurate results of the collision probability than [2].

A collision occurs when two or more backoff counters B_i $(i = 1, \ldots, N)$ of different stations expire at the same time. As we neglect multiple successive collisions occurrence, during a TC, a frame can be either successfully transmitted



(a) Case of a successful transmission from the first attempt (b) Case of a successful transmission after a first failed attempt

Fig. 3. The transmission cycle (TC) of a packet: $T(N)$.

from the first attempt [Fig. 3(a)] or following to a first collision [Fig. 3(b)]. Hence, a collision can only occur at the beginning of the TC with a probability $P_c(N)$, while all the nodes' backoffs B_i $(i = 1,\ldots,N)$ vary between $[0, \mathrm{CW}_{\min}]$ (refer to assumption 2). Hence, the probability of collision $P_c(N)$ can be written as follows:

$$P_c(N) = \Pr\{\overline{U}\} = \sum_{k=0}^{\mathrm{CW}_{\min}} \Pr\{X = k, \overline{U}\} \tag{2}$$

where, the random variable X denotes $(\min_{i \in \langle 1,N \rangle} B_i)$ and the event $\overline{U}$ is defined as follows:

$$\begin{aligned}\overline{U} &= \left\{\exists i, j \in \langle 1, N\rangle, i \neq j, B_i = B_j = X\right\} \\ &= \{\text{Collided transmission}\}.\end{aligned} \tag{3}$$

The event $\{X = k, \overline{U}\}$ simply implies that the backoff counter becomes zero for the first time in k slots for at least two stations, which leads to a collision occurrence. Thus, $\Pr\{X = k, \overline{U}\}$ can be derived as follows:

$$\Pr\{X = k, \overline{U}\} = \sum_{i=2}^{N} \binom{N}{i} \frac{(\mathrm{CW}_{\min} - k)^{N-i}}{(\mathrm{CW}_{\min} + 1)^{N}} \tag{4}$$

5.3. Model definition and resolution for the classic DCF scheme

In this section we evaluate the network throughput considering the classical DCF protocol, where all the access nodes have the same initial contention window size (i.e. $\mathrm{CW}_{\min} = 31$). To achieve this, we first derive the average time $T(N)$ of an TC. It is the mean time required to successfully transmit a packet. $T(N)$ can be written as follows:

$$T(N) = t_{\mathrm{tr}} + t_{\mathrm{ov}} + t_{\mathrm{cont}}(N) \tag{5}$$

where t_{tr} is the transmission time of the successfully transmitted data packet. It can be calculated as follows:

$$t_{\mathrm{tr}} = \frac{\mathrm{PDU}}{N} \sum_{i=1}^{M} \frac{n_i}{r_i}. \tag{6}$$

Note that, for the sake of simplicity, in this study, we consider fixed-size PDU (packet data unit) traffic. On the other side, t_{ov} is a constant overhead, which can be simply deduced from Fig. 3, and thus given by:

$$t_{\mathrm{ov}} = \mathrm{DIFS} + 2 \times t_{\mathrm{PLCP}} + \mathrm{SIFS} + t_{\mathrm{ACK}}. \tag{7}$$

Moreover, $t_{\mathrm{cont}}(N)$ represents the average time spent in contention procedure. In other words, it is the extra time lost due to the collision occurrence. Hereafter, we focus on $t_{\mathrm{cont}}(N)$ calculations. As stated before, we neglect in our study the successive collisions occurrence. Doing so, we distinguish between two cases:

- *Case 1*: The data packet is transmitted successfully by one of the N access nodes from the first attempt (i.e. following a successfully transmitted packet) [see Fig. 3(a)].
- *Case 2*: The data packet is transmitted successfully by one of the N access nodes following to a first collision occurrence on the medium [see Fig. 3(b)].

5.3.1. Case 1

This case happens with a probability $(1 - P_c(N))$. In this case, $t_{\mathrm{cont}}(N) = t1_{\mathrm{cont}}(N)$ is simply the average backoff time spent by the transmitting node, denoted by node j, before accessing to the data channel [see Fig. 3(a)]. According to assumption 2, all the access nodes' backoff counters take values in $[0, \mathrm{CW}_{\min}]$ at the beginning of an TC. Moreover, as the data packet is successfully transmitted, the transmitting node j has certainly the minimum backoff value among the N competing access nodes (i.e. $X = B_j$). In addition $\forall i \neq j$, we have $B_i > B_j$. Let U denote that event:

$$\begin{aligned}U &= \left\{\exists! j \in \langle 1, N\rangle, B_j = X\right\} \\ &= \{\text{Successfully transmission}\}.\end{aligned} \tag{8}$$

Note that

$$\Pr\{U\} = 1 - \Pr\{\overline{U}\}. \tag{9}$$

Doing so, $t1_{\mathrm{cont}}(N)$ can be expressed as follows:

$$t1_{\mathrm{cont}}(N) = E[X|U] \text{ Slots} \tag{10}$$

where,

$$E[X|U] = E[X, U]/\Pr\{U\}. \tag{11}$$

Moreover, $E[X, U]$ can be written as follows:

$$E[X, U] = \sum_{k=0}^{\mathrm{CW}_{\min}} k \Pr\{X = k, U\} \tag{12}$$

where $\Pr\{X = k, U\}$ can be simply derived based on (4):

$$\Pr\{X=k,U\}=\Pr\{X=k\}-\Pr\{X=k,\overline{U}\}=\binom{N}{1}\frac{(\mathrm{CW}_{\min}-k)^{N-1}}{(\mathrm{CW}_{\min}+1)^{N}} \tag{13}$$

5.3.2. Case 2

In this case, the data packet is successfully transmitted by one of the access nodes after a first failed attempt. Such case happens with a probability $P_c(N)$. $t_{cont}(N)=t2_{cont}(N)$ is therefore the sum of the time spent from the beginning of the TC until the transmission of the collided data frame (t'_2) and the average backoff time required by the new transmitting node j to access to the channel in order to transmit correctly a new packet (t''_2) [see Fig. 3(b)]. Hence, we get:

$$t2_{cont}(N)=t'_2+t''_2, \tag{14}$$

and we have:

$$t'_2=\mathrm{DIFS}+E[X|\overline{U}]\ \mathrm{Slots}+E[t'_{tr}|\overline{U}] \tag{15}$$

where $E[X|\overline{U}]$ is the average backoff time of the collided stations. It can be simply derived using the fact that $E[X|\overline{U}]=E[X,\overline{U}]/\Pr\{\overline{U}\}$. Doing so, we have:

$$E[X,\overline{U}]=\sum_{k=0}^{\mathrm{CW}_{\min}}k\Pr\{X=k,\overline{U}\} \tag{16}$$

where, $\Pr\{X=k,\overline{U}\}$ is given by (4).

Afterwards, $E[t'_{tr}|\overline{U}]$ is the transmitting time of the collided PDUs. As we assume that all the nodes send identical packets' size, the collision duration is simply the transmission time required by the collided node with the lowest bit rate. Let the random variable Y denote that bit rate (i.e. the lowest bit rate among the $N_c(N_c\geqslant 2)$ collided nodes). $E[t'_{tr}|\overline{U}]$ can be therefore written as follows:

$$E[t'_{tr}|\overline{U}]=\sum_{i=1}^{M}\frac{\mathrm{PDU}}{r_i}\Pr\{Y=r_i|\overline{U}\}. \tag{17}$$

By conditioning on the number of collided nodes $N_c=n$, we get:

$$E[Y=r_i|\overline{U}]=\sum_{n=2}^{N}\Pr\{Y=r_i|N_c=n,\overline{U}\}\Pr\{N_c=n|\overline{U}\} \tag{18}$$

where,

$$\Pr\{Y=r_i|N_c=n,\overline{U}\}=\frac{\binom{\sum_{j=1}^{i}n_j}{n}-\binom{\sum_{j=1}^{i-1}n_j}{n}}{\binom{N}{n}} \tag{19}$$

Furthermore, we have:

$$\Pr\{N_c=n|\overline{U}\}=\frac{\Pr\{N_c=n,\overline{U}\}}{\Pr\{\overline{U}\}}=\frac{\Pr\{N_c=n\}}{\Pr\{\overline{U}\}} \tag{20}$$

where,

$$\Pr\{N_c=n\}=\sum_{k=0}^{\mathrm{CW}_{\min}}\binom{N}{n}\frac{(\mathrm{CW}_{\min}-k)^{N-n}}{(\mathrm{CW}_{\min}+1)^{N}}. \tag{21}$$

Doing so, we get the expression of $E[t'_{tr}|\overline{U}]$. Moreover, substituting (16) and (17) in (15), we obtain the expression of t'_2.

Let us now focus on the calculation of t''_2. As we mentioned before, a collision can occur only when $N_c(N_c\geqslant 2)$ stations send data packets at the same time. The N_c collided stations perceive the collision as they do not receive the ACK frame from the receiving nodes after $T_{\mathrm{ACK}}+\mathrm{SIFS}=t_{\mathrm{wait}}$ units of time. On the other side, the remaining $N-N_c$ nodes, which did not participate in the collision, detect immediately the collision occurrence as they receive a collided data frame and they instantaneously attempt again to access the channel. In this case, the residual backoff counters of these $N-N_c$ nodes take values in $[0,\mathrm{CW}_{\min}]$.

On the other hand, the backoff windows of the N_c collided stations double. Accordingly, the backoff counters of the collided stations take values in $[0,(2\times\mathrm{CW}_{\min})]$. However, these stations have to wait for a period of time approximately equal to 11 slots corresponding to $t_{\mathrm{wait}}=t_{\mathrm{ACK}}+\mathrm{SIFS}$ before they try again to access to the data channel. Hence, starting from the collision occurrence, the backoff counters of the N_c collided stations vary between $[t_{\mathrm{wait}},t_{\mathrm{wait}}+(2\times\mathrm{CW}_{\min})]$, whereas the remaining nodes' backoff counters vary between $[0,\mathrm{CW}_{\min}]$.

Let the random variable X' denote $(\min_{i\in\langle 1,N\rangle}B_i)$ and U' be the following event:

$$U'=\{\exists!j\in\langle 1,N\rangle,B_j=X'\}=\{\text{Successfully transmission}\}. \tag{22}$$

We recall that we aim at calculating t''_2, which is the average backoff time required by the network to successfully transmit a new packet after a first failed attempt. t''_2 can be therefore written as:

$$t''_2=E[X',U'|\overline{U}] \tag{23}$$

which leads to:

$$t''_2=\sum_{k=0}^{t_{\mathrm{wait}}+(2\times\mathrm{CW}_{\min})-1}k\Pr\{X'=k,U'|\overline{U}\}. \tag{24}$$

In order to calculate t''_2, we have first to derive the expression of $\Pr\{X'=k,U'|\overline{U}\}$. To achieve this, three cases are to be distinguished according to the value of X' (in term of time slots).

(a) $X'=k<t_{\mathrm{wait}}$: In this case, the host j that accesses the medium is one the $N-N_c$ noncollided stations. Using the theorem of total probability, we get:

$$\Pr\{X'=k,U'|\overline{U}\}=\sum_{n=2}^{N-1}\Pr\{X'=k,U',N_c=n|\overline{U}\}. \tag{25}$$

This yield to

$$\Pr\{X' = k, U'|\overline{U}\} = \sum_{n=2}^{N-1} \Pr\{X' = k, U'|N_c = n, \overline{U}\} \times \Pr\{N_c = n|\overline{U}\} \tag{26}$$

where, $\Pr\{N_c = n|\overline{U}\}$ is already given by (20). Moreover, since the transmitting node j did not participate in the previous collision, we have:

$$\Pr\{X'=k,U'|N_c=n,\overline{U}\} = \binom{N-n}{1} \frac{(\mathrm{CW_{min}}-k)^{N-n-1}}{(\mathrm{CW_{min}}+1)^{N-n}}. \tag{27}$$

(b) $t_{\mathrm{wait}} \leqslant X' = k \leqslant \mathrm{CW_{min}}$: In this case the host j that accesses the channel may be either one of the N_c stations, which already participated in the first collision or belongs to the $N - N_c$ remaining ones. Accordingly, we distinguish between two sub-cases:

Sub-case (b.1). The transmitting host j already participated in the first collision. Such event is denoted by C. In this case, we have:

$$\Pr\{X' = k, U', C|\overline{U}\} = \sum_{n=2}^{N} \Pr\{X' = k, U', C|N_c = n, \overline{U}\} \times \Pr\{N_c = n|\overline{U}\} \tag{28}$$

where,

$$\Pr\{X' = k, U', C|N_c = n, \overline{U}\} = \binom{n}{1} \frac{(\mathrm{CW_{min}} - k)^{N-n}(2 \times \mathrm{CW_{min}} + t_{\mathrm{wait}} - k)^{n-1}}{(\mathrm{CW_{min}} + 1)^{N-n}(2 \times \mathrm{CW_{min}} + 1)^{n}} \tag{29}$$

and $\Pr\{N_c = n|\overline{U}\}$ is already given by (20).

Sub-case (b.2). The transmitting host j did not participate in the first collision. Such event is denoted by $\overline{C}$. In this case, we have:

$$\Pr\{X' = k, U', \overline{C}|\overline{U}\} = \sum_{n=2}^{N} \Pr\{X' = k, U', \overline{C}|N_c = n, \overline{U}\} \times \Pr\{N_c = n|\overline{U}\} \tag{30}$$

where,

$$\Pr\{X' = k, U', \overline{C}|N_c = n, \overline{U}\} = \binom{N-n}{1} \frac{(\mathrm{CW_{min}} - k)^{N-n-1}(2 \times \mathrm{CW_{min}} + t_{\mathrm{wait}} - k)^{n}}{(\mathrm{CW_{min}} + 1)^{N-n}(2 \times \mathrm{CW_{min}} + 1)^{n}}. \tag{31}$$

Putting both sub-cases together, we get the expression of $\Pr\{X' = k, U'|\overline{U}\}$ when ($t_{\mathrm{wait}} \leqslant X' = k \leqslant \mathrm{CW_{min}}$) as follows:

$$\Pr\{X' = k, U'|\overline{U}\} = \Pr\{X' = k, U', C|\overline{U}\} + \Pr\{X' = k, U', \overline{C}|\overline{U}\}. \tag{32}$$

(c) $\mathrm{CW_{min}} < X' = k < t_{\mathrm{wait}} + 2 \times \mathrm{CW_{min}}$: This case happens only when all the N access nodes participated in the first collision (i.e. $N_c = N$). Thus, we have:

$$\Pr\{X' = k, U'|\overline{U}\} = \Pr\{X' = k, U', N_c = N|\overline{U}\}. \tag{33}$$

This leads to:

$$\Pr\{X' = k, U'|\overline{U}\} = \Pr\{X' = k, U'|N_c = N, \overline{U}\} \times \Pr\{N_c = N|\overline{U}\} \tag{34}$$

where,

$$\Pr\{X' = k, U'|N_c = N, \overline{U}\} = \binom{N}{1} \frac{(2 \times \mathrm{CW_{min}} + t_{\mathrm{wait}} - k)^{N-1}}{(\mathrm{CW_{min}} + t_{\mathrm{wait}})^{N}} \tag{35}$$

and $\Pr\{N_c = n|\overline{U}\}$ is already given by (20).

Finally, using (24), (25), (32) and (34), we simply derive t_2'' and thus we get the expression of $t2_{\mathrm{cont}}$ by means of (14). Doing so, we derive the expression of $t_{\mathrm{cont}}(N)$, which is given by:

$$t_{\mathrm{cont}}(N) = (1 - P_c(N))t1_{\mathrm{cont}}(N) + P_c(N)t2_{\mathrm{cont}}(N). \tag{36}$$

By summing (6), (7) and (36), as depicted in (5), we get easily the average time of a TC (i.e. $T(N)$). Based on this result, the total useful network throughput can be derived as follows:

$$\mathrm{Th_{total}} = \frac{\mathrm{PDU}}{T(N)}. \tag{37}$$

Moreover, according to the classical DCF mechanism, all the access nodes exhibit the same useful throughput although their different physical bit rates. Thus, the useful throughput per node is:

$$\mathrm{Th_{node}} = \frac{\mathrm{Th_{total}}}{N}. \tag{38}$$

5.4. Model definition and resolution for the proposed DCF-MB scheme

In this section, we consider our proposed DCF-MB scheme, which is introduced to alleviate the unfairness issue of the classical DCF mechanism. To achieve this, we provide each class $C_i (i = 1, \ldots, M)$ with its associated initial contention window size $\mathrm{CW_{min}}(i)$ as described in (1).

In the following we derive the analytic expressions of the total useful network throughput as well as the useful throughput for each station according to its class bit rate C_i. It is worthwhile to note that the analysis of the proposed scheme, with regard to the total useful network throughput, is equivalent to the study of the classical DCF with N_{virtual} access nodes instead of N, and with an initial contention window size $\mathrm{CW_{min}} = \mathrm{CW_{min}}(M)$. Moreover, N_{virtual} is defined as follows:

$$N_{\mathrm{virtual}} = \sum_{i=1}^{M} n_{\mathrm{virtual}}(i) \tag{39}$$

where,

$$n_{\mathrm{virtual}}(i) = n_i \frac{\mathrm{CW_{min}}(M)}{\mathrm{CW_{min}}(i)}. \tag{40}$$

To illustrate this, let revisit the scenario of Section 3. According to our DCF-MB scheme, a class C_1-station accesses the medium 11 times more than a C_3-station, since $CW_{min}(1) = CW_{min}(3)/11$. Hence, from a C_3-station perspective of view, the system is equivalent to a network with $n_{virtual}(1) = 11 \times n_1$ stations of class C_1 with the same initial contention window size [i.e.$CW_{min}(1) = CW_{min}(3)$].

Therefore, we can derive straightforwardly the total useful network throughput based on (5) and (37) by substituting N with $N_{virtual}$. Then, the useful throughput per node according to its class $C_i (i = 1, \ldots, M)$ is given by:

$$\mathrm{Th}_{node}(i) = \frac{\mathrm{Th}_{total}}{N_{virtual}} \times \frac{n_{virtual}(i)}{n_i}. \quad (41)$$

6. Performance evaluation

In this section, we analyze the impact of our proposed DCF-MB scheme on the network performances using both analytical and simulation approaches. To achieve this, we develop our own event-driven simulator. Note that simulations are used to assess the accuracy of the proposed analytical models.

For the sake of simplicity, we consider along this section, $M = 3$ classes of bit rate C_1, C_2 and C_3 with $r_1 = 11$ Mbit/s, $r_2 = 5.5$ Mbit/s and $r_3 = 1$ Mbit/s, respectively.

In order to gauge the efficiency of our proposal, we first apply our DCF-MB scheme using the same scenario of Section 3 and the results are reported in Fig. 4. Recall that according to this scenario, we have three stations belonging to different class bit rates (C_1, C_2 and C_3), where station 2 is activated at $t_1 = 10$ s and station 3 is activated at $t_2 = 40$ s.

Fig. 4 shows that the initial useful throughput of station 1 (i.e. 6.41 Mbit/s) is divided by two when station 2 is activated and is divided by three when station 3 joins the network. Unlike the classical DCF (see Fig. 2), thanks to our method, the performance of station 1 only depends on the number of sharing access stations and no more on their relative positions with respect to the AP. In other word, the fact that station 2 transmits at 5.5 Mbit/s or more or less does not really affect the station 1 throughput. According to our scheme, the utilization time of the medium is equally shared among the different stations. Moreover, each station uses its proportion of time according to its transmission bit rate. In this regard, using a low bit rate, the station will transmit less without penalizing the remaining contending stations. Based on Fig. 4, we can observe that the useful throughput of station 1 is the double of the station 2 throughput and 11 times the station 3 throughput.

Furthermore, it is interesting to note that Fig. 4 shows a perfect match between the analytical and simulation results. Indeed, analytical curves practically coincide with simulation ones, which exhibits the accuracy of our models. This remark holds for all the remaining results of this section.

As explained before, one of the major concerns with DCF is the drastic degradation of the total network throughput due to the relatively far away stations with respect to the AP. Fig. 5 confirms this issue, where the total throughput significantly degrades once station 2 and 3 join the network. Fig. 5 also shows that our DCF-MB scheme alleviates this issue. Indeed, the increase of sharing nodes degrades less significantly the total network throughput when using DCF-MB. Specifically, when the number of sharing nodes is 3, the throughput obtained thanks to our DCF-MB is 4.21 Mbit/s whereas it is limited to 1.71 Mbit/s with the classical DCF.

Note that the slight decrease of the total throughput with DCF-MB when station 2 and 3 are activated is due two main reasons. First, station 2 and 3 transmit at relatively low bit rates with respect to station 1 during their utilization of the medium, which reduces the total network throughput. Moreover, increasing the number of access stations increases the collision probability among different nodes' frames, leading thus to increasing bandwidth waste.

In this context, Fig. 6 shows the total network throughput evolution with the network density. We refer by the network density as the same number of access nodes n_i

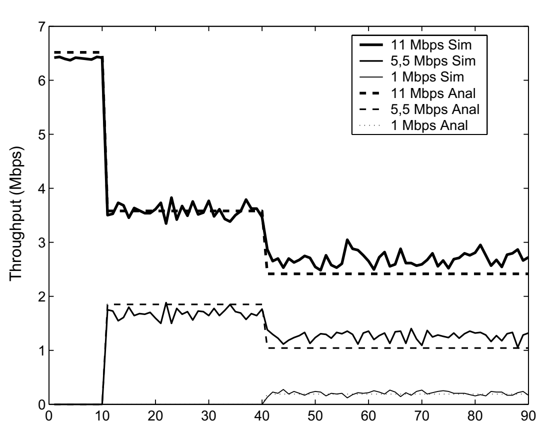


Fig. 4. Throughput of each node according to its class bit rate: DCF-MB case.

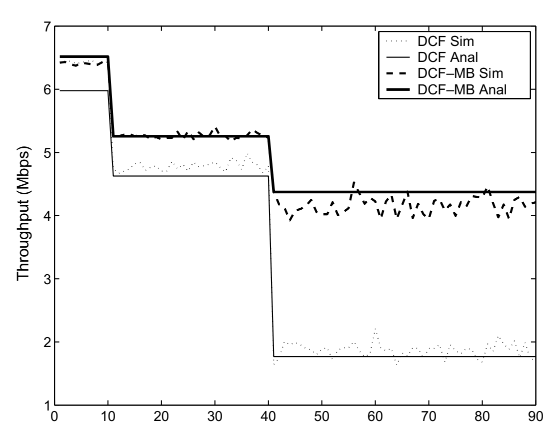


Fig. 5. Evolution of the total useful throughput of the network.

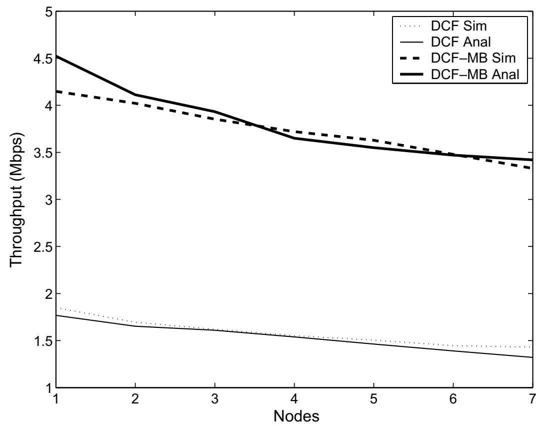


Fig. 6. Impact of the network density on the total network throughput.

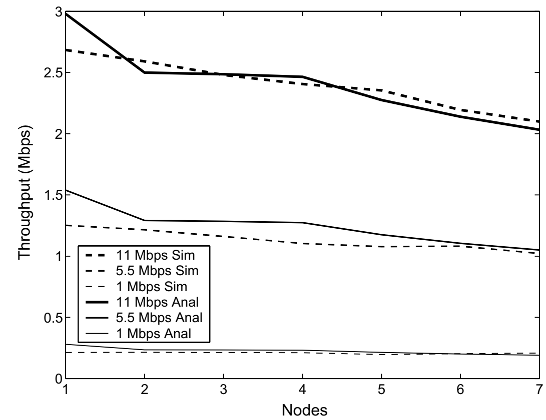


Fig. 8. Impact of the network density on the useful throughput of each class bit rate: DCF-MB case.

composing each bit rate class $C_i (i = 1,\ldots,3)$. In this case, the network density varies from 1 to 7, that is, the total number of access nodes varies from 3 to 21. Again, Fig. 6 shows that the total network throughput decreases with the increase of access nodes for both cases (DCF and DCF-MB). Moreover, this figure exhibits once more the significant gain introduced by our method.

Figs. 7 and 8 show the average useful throughput of each class bit rate for the DCF and DCF-MB cases, respectively. Fig. 7 shows that, using the classical DCF, all the access stations have the same useful throughput although their different transmitting bit rates. Moreover, the useful throughput of each class is very low (less than 0.6 Mbit/s). This is typically due to the limitations of the classical DCF.

On the other hand, enabling our DCF-MB scheme, this issue is alleviated. Fig. 8 shows that the throughput is fairly distributed among different classes. In addition, the throughput per class significantly increases. Specifically, when density is equal to 1, a station belonging to class C_1 benefits from a throughput around 2.5 Mbit/s, whereas the same station has a throughput less than 0.6 Mbit/s when DCF-MB is disabled.

In Fig. 9 the network density is set equal to 1. In other words, the network is composed of three stations belonging to classes C_1, C_2 and C_3. Fig. 9 depicts the evolution of the total network throughput with the arrival rate of frames. Recall that the frames arrive to each node level according to a Poisson process. Fig. 9 shows that both DCF and DCF-MB behaves similarly when the network load is low. On the other side, increasing the arrival rate, a network using the classical DCF is rapidly saturated with a maximal network throughput of 1.71 Mbit/s. In contrast, enabling DCF-MB, the network throughput attains 4.21 Mbit/s.

Finally, we conclude this section by studying the impact of our scheme on the collision in the network. In such network, a collision between two stations occurs when their associated backoff counters expire at the same time. Fig. 10 depicts the collision probability according both DCF and DCF-MB schemes. To achieve this, we use the same scenario used in Section 3. According to Fig. 10, we can observe that the collision probability reduces when DCF-MB is enabled. This is a direct result of the utilization of different contention windows for the different classes of stations.

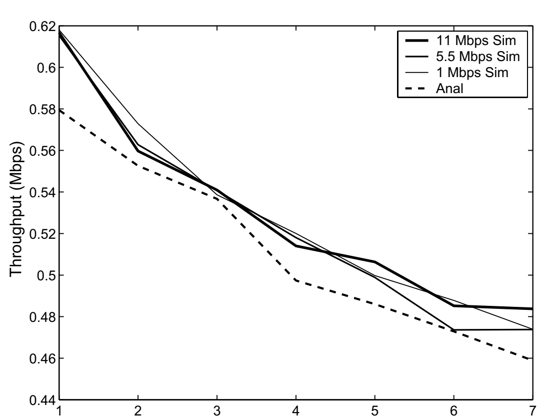


Fig. 7. Impact of the network density on the useful throughput of each class bit rate: DCF case.

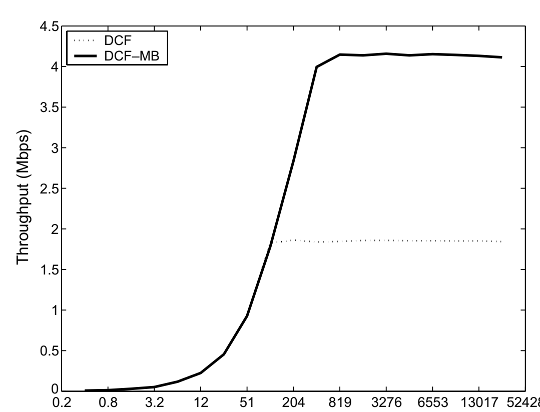


Fig. 9. Evolution of the total network throughput as a function of the arrival rate $\times 10^{-6}$.

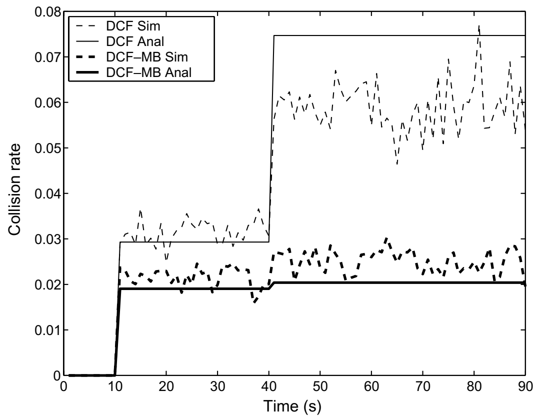


Fig. 10. Comparison between DCF and DCF-MB collision rate.

7. Conclusions

In this paper, we proposed an improvement of the existing DCF scheme in order to cope with its unfairness limitations. We advised the introduction of relative priorities among different access stations according to their physical transmission bit rate. To achieve this, we used different contention window sizes for each class of bit rate. Finally, we motivated the use of the proposed scheme since it allows achieving fairness among contending access nodes while improving the total network throughput.

References

[1] IEEE Std 802.11e/D13.0, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Amendment: Medium Access Control (MAC) Quality of Service (QoS) Enhancements, January 2005.
[2] M. Heusse, F. Rousseau, G. Berger-Sabbatel, A. Duda, Performance anomaly of 802.11b, in: Proceedings of IEEE Infocom 03, March 2003, San Francisco, USA, 2003.
[3] C. Koksal, An Analysis of Short-term Fairness in Wireless Media Access Protocols, in: Proceedings of ACM SIGMETRICS, 2000.
[4] G. Holland, N.H. Vaidya, Analysis of TCP Performance Over Mobile Ad Hoc Networks, in: Proceedings of IEEE/ACM MobiCom'99, August 1999, pp. 219–230.
[5] K. Tang, M. Correa, M. Gerla, Effects of ad hoc MAC layer medium access mechanisms under TCP, CM/Bal MONET 6 (4) (2001) 317–329.
[6] C.L. Barrett, M.V. Marathe, D.C. Engelhart, A. Sivasubramaniam, Analyzing the Short-term Fairness of IEEE 802.11 in Wireless Multihop Radio Networks, MASCOTS'02, October 2002, Fort Worth, TX, USA, 2002.
[7] S. Choi, J. del Prado, S. Shankar, and S. Mangold, IEEE 802.11e Contention-Based Channel Access (EDCF) Performance Evaluation, in: Proceedings of IEEE ICC 03, May 2003, Anchorage, Alaska, USA, 2003.
[8] I. Aad, C. Castelluccia, Differentiation Mechanisms for IEEE 802.11, in: Proceedings of IEEE Infocom 01, April 2001, Anchorage, AK, USA, 2001.
[9] O.C. Branquinho, N. Reggiani, D.M. Ferreira, Mitigating 802.11 Mac Anomaly Using SNR to Control Backoff Contention Window, in: Proceedings of IEEE ICWMC, 2006.
[10] Yoo et al., Eliminating the Performance Anomaly of 802.11b, in: Proceedings of the Fourth International Conference on Networking (ICN 2005), April 17–21, 2005, Reunion Island, France, 2005.

Yassine Chetoui, received a B.S. degree in network computer systems in 2004, and an M.S. degree in network and computer science from the University of Paris VI, France, in 2005. Currently, he is working towards the Ph.D. degree in network and computer science at the University of Versailles, France. His research interests include wireless networking, network security and performance evaluation.

Nizar Bouabdallah, received the B.S. degree in telecommunications engineering from Ecole Superieur des Communications (Sup'Com), Tunis, Tunisia, in 2001, and the M.S. and Ph.D. degrees in network and computer science from the University of Paris VI, Paris, France, in 2002 and 2004, respectively. He joined Alcatel Research Laboratories, Marcoussis, France, in 2002, while working on his Ph.D. degree. In 2005, he was with the North Carolina State University, Raleigh, NC, USA, as a Postdoctoral Fellow. He is currently a researcher at INRIA (Institut National de Recherche en Informatique et en Automatique). Since February 2007, he has been a Visitor Researcher at the School of Computer Science, University of Waterloo, Waterloo, ON, Canada. His research interests include optical networking, wireless and sensor networks, performance evaluation, network planning and modeling, as well as control and management architectures.