Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2010, Article ID 489527, 15 pages
doi:10.1155/2010/489527

Research Article

Mathematical Analysis of EDCA's Performance on the Control Channel of an IEEE 802.11p WAVE Vehicular Network

Jose R. Gallardo, Dimitrios Makrakis, and Hussein T. Mouftah

Broadband Wireless and Internetworking Research Laboratory, School of Information Technology and Engineering, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5

Correspondence should be addressed to Jose R. Gallardo, gallardo@site.uottawa.ca

Received 3 July 2009; Accepted 5 February 2010

Academic Editor: Máirtín O'Droma

Wireless networks for vehicular environments are gaining increasing importance due to their ability to provide a means for stations on the roadside and radio units on board of vehicles to communicate and share safety-related information, thus reducing the probability of accidents and increasing the efficiency of the transportation system. With this goal in mind, the IEEE is currently developing the Wireless Access in Vehicular Environments (WAVE) IEEE 802.11p standard. WAVE devices use the IEEE 802.11's Enhanced Distributed Channel Access (EDCA) MAC protocol to compete for the transmission medium. This work proposes an analytical tool to evaluate the performance of EDCA under the specific conditions of the so-called control channel (CCH) of a WAVE environment, including the particular EDCA parameter values and the fact that all safety-critical data frames are broadcasted. The protocol is modeled using Markov chains and results related to throughput, frame-error rate, buffer occupancy and delay are obtained under different traffic-load conditions. The main analysis is performed assuming that the CCH works continuously, and then an explanation is given as to the considerations that are needed to account for the fact that activity on the CCH is intermittent.

1. Introduction

The transportation system has a significant impact on the well being of the society in general. Safe, timely and low-cost transportation of people and goods brings great benefits to commerce, industry, government, educational centers, families, and so forth. Moreover, because of its size, large number of components and complexity, the transportation system is extremely difficult to manage and supervise. This clearly indicates that it is necessary to develop intelligent systems able to enhance the transportation infrastructure in terms of safety, efficiency and cost. One way to do this is by creating a wireless communications network accessible by vehicles on the move and able to provide them with safety-critical information as well as with a gateway to the global Internet, hence to a greater amount of useful information.

With this goal in mind, the IEEE has taken on the responsibility of designing a new technology, known as WAVE (Wireless Access in Vehicular Environments), and is currently working towards defining a new standard referred to as IEEE 802.11p. The purpose of WAVE is to provide vehicle-to-roadside as well as vehicle-to-vehicle wireless communications, so that stations on the roadside (RSU) and mobile radio units located on board of vehicles (OBU) can share information related to road and traffic conditions and use it to improve the safety and efficiency of the transportation system. The overall architecture of WAVE is defined in IEEE standards 1609.1, 1609.2, 1609.3, 1609.4, and P802.11p. The latter is the current draft version of what will eventually become the IEEE 802.11p standard.

WAVE networks are supposed to use a dedicated frequency band that has been allocated by the US Federal Communications Commission (FCC), known as DSRC (Dedicated Short-Range Communications), and which is in the 5.9 GHz range. The WAVE concept includes seven channels within the DSRC band. One of them is known as the Control Channel (CCH) and the remaining six are known as Service Channels (SCH). Time is subdivided into

Table 1: EDCA parameters to be used in the CCH of WAVE networks.

Access category	CWmin	AIFSN
AC0	15	9
AC1	7	6
AC2	3	3
AC3	3	2

synchronization intervals of 100 ms each. During the first half of each interval (50 ms) it is mandatory for all stations (OBUs and RSUs) to listen to the CCH, unless they are transmitting on it. Therefore, really urgent messages have to be broadcasted over the CCH during the mandatory periods, so that they can be received by as many devices as possible.

There are different priorities for the messages sent over the CCH, depending on how critical they are for vehicle safety. For that reason, WAVE devices use the Enhanced Distributed Channel Access (EDCA) MAC protocol, defined in the 2007 version of the IEEE 802.11 standard [1], to orderly gain access to the transmission medium.

From what has been mentioned above we can conclude that, if we are interested in determining the degree of enhancement that WAVE networks can bring to the safety and efficiency of the transportation system, we have to analyze the performance of EDCA under the specific conditions of the CCH, including the particular EDCA parameter values proposed in the draft standard and the fact that all safety-critical data frames are broadcasted. Transmitting frames using the broadcast address means that neighboring nodes will not send back an acknowledgment even if frames are received correctly, which in turn indicates that the binary exponential backoff technique, commonly used to retransmit erroneous frames with a low collision probability, cannot be implemented.

Regarding the messages sent over the CCH, the highest priority is given to safety-related urgent information, which could be present at the RSU (e.g., information about accidents, obstacles, slippery or damaged roads, broken or missing traffic signs, etc.), or it could be information generated by cars (emergency vehicle approaching, vehicle with malfunctioning brakes, speeding over a certain limit, or posing any type of hazard to other vehicles). The second highest priority may be given for vehicles to advertise their presence to other vehicles, which is useful when visual detection is not possible (e.g., due to obstacles, uneven terrain, fog, heavy rain, sunshine, etc.). A lower priority may be given to non-urgent messages (e.g., informing that a vehicle needs help because it broke down, ran out of gas, or collided with an obstacle, but poses no risk to other vehicles) and, finally, the lowest priority may be given to messages aimed at establishing new non-safety-related connections over the service channels (e.g., to look for nearby hotels or to download a map).

Messages are then classified into different access categories (AC), where the lowest priority corresponds to AC0 and the highest to AC3. Table 1 shows the parameters to be used by each access category to compete for the CCH, as specified in the IEEE 802.11p draft standard. In that table, CWmin (minimum contention-window size) defines the interval from which the random backoff counter is selected when attempting the transmission of a frame following the collision-avoidance procedure; to be more specific, the counter will be selected from the set $\{0, 1, \ldots, \text{CWmin} - 1\}$ with a probability following a uniform distribution. Likewise, AIFSN (arbitration inter-frame space number) defines the length of the time interval that a given node has to sense the medium to decide if it is busy or idle; namely, the sensing time referred to as arbitration inter-frame space will be denoted by AIFS[AC i], $i \in \{0, 1, 2, 3\}$, and will be equal to the following expression:

$$\text{AIFS}[\text{AC}\, i] = \text{SIFS} + \text{AIFSN}[\text{AC}\, i] \cdot \sigma, \tag{1}$$

where SIFS is the duration of a short inter-frame space and we use σ to denote the duration of a slot. Since AIFSN is smaller for higher priorities, as shown in Table 1, so will the corresponding AIFS creating an advantage for higher-priority nodes to access the wireless medium.

The broadcasting of all safety-critical data frames and the fact that some nodes have advantages over others to access the medium are captured by our model in a very detailed way. The contribution of this work is then the proposal of an analytical tool to evaluate the performance of WAVE. This work improves and extends models that we have previously proposed cf. [2, 3] and other proposals described in the following sections.

The rest of this paper is organized as follows. Section 2 briefly describes similar proposals, including our own preliminary work, and explains how this new analytical tool is different from them. Section 3 presents our new model, including the details related to the behavior of the different access categories. Based on this model, section 4 includes equations that can be used to assess the performance of the protocol in terms of throughput, delay, losses and buffer occupancy, all of which are very important parameters in an environment, such as WAVE networks, in which the reliable and timely delivery of information is crucial. Section 5 gives specific numerical results under different scenarios. As mentioned before, nodes transmit information intermittently on the CCH. The main analysis is performed assuming that the CCH works continuously, and then an explanation is given in Section 6 as to the considerations that are needed to account for the fact that activity on the CCH is intermittent. Lastly, Section 7 presents the conclusions that can be drawn from this work.

2. Related Work

WAVE is a relatively new technology, still being defined by the IEEE as mentioned above, therefore there are not many publications analyzing its performance. The ground-breaking work published by Bianchi [4] is not directly applicable in this case since it does not analyze nodes with different priorities and assumes that nodes always have data frames to send (saturation condition). However, it is fair to say that our approach is nothing but an enhanced version of that methodology, designed to analyze the system under different traffic loads and to account for the fact that EDCA applies service differentiation among the various traffic categories.

The work published in [5] is a simplified version of [4] in the sense that the original model is modified to account for the fact that data frames in WAVE are transmitted in broadcast mode. In other words, the model only has one backoff stage to reflect the fact that there are no acknowledgements, hence frames are not retransmitted since there is no way to know if the initial transmission was successful or not. It therefore inherits the characteristic of being only applicable to the saturated case and to systems with only one access category.

Other authors have analyzed a WAVE environment using simulation, but their results are again only applicable to systems with a single access category [6–9].

The analysis presented in [10] is based on assumptions that are not properly justified. For instance, based on the initial assumption that the total number of stations is constant, the authors assume that the number of backlogged stations at any given time is binomially distributed. The authors further assume that the probability that a backlogged station transmits during any given slot is constant, and that the time needed to transmit a frame after it has arrived at the head of the queue is exponentially distributed. Additionally, [11] analyzes a situation similar to WAVE in that it deals with broadcast transmissions using CSMA/CA. It assumes slotted transmissions with fixed-size slots. However, it assumes that the probability β that a station starts transmitting during a given slot is a system parameter with a constant value, independent of the number of nodes (n), of the traffic generation rate of each node (λ), and of the packet length (L), which is not applicable to this type of systems. In addition to being based on unjustified assumptions that can produce inaccurate results, both [10, 11] only analyze the case of a system with a single access category.

In [12], the author proposes formulas to evaluate the throughput and the collision probability in a WAVE environment in which nodes always have frames to send (saturation condition). In the expression to calculate the throughput, the author assumes that backoff intervals are of constant length, equal to what would be the average value if the medium did not become busy during the countdown (i.e., if the backoff counter did not have to freeze during the transmission of a frame by another node). In addition to that, in the expression for the probability of collision, the author also assumes that nodes select the length of their backoff periods with the same probability, regardless of their access category. These unrealistic assumptions give results that, in general, will not be very accurate.

There are other contributions analyzing the performance of EDCA in general (e.g., [2, 13–15]). The results in [13–15] are only applicable to the saturated case (nodes always having frames to transmit), which is not a very relevant scenario in a WAVE environment, where fast delivery of messages is vital, hence working in saturation is not an option. Our previous work presented in [2] in turn assumes that the probability of collisions among nodes corresponding to different priorities is negligible. Even though in general this is a reasonable assumption for low-traffic conditions, in WAVE it is essential to obtain results as accurate as possible since, again, we need to make sure that the time taken to successfully deliver a safety-related critical message does not exceed the little time vehicle drivers have to react in the presence of danger. The present work does not make any of those assumptions; on the contrary, it models very explicitly the way EDCA differentiates among nodes corresponding to the different access categories.

Reference [3] is an earlier version of this work. In the present paper we improve the calculation of the service time, total delay and buffer occupancy by using renewal-process theory instead of the G/G/1 model that is not quite applicable in this case. We are also adding a section related to the considerations that are needed to account for the fact that activity on the CCH is intermittent.

3. Proposed Models

The models presented in this section are the same as those we originally proposed in [3]. For convenience and comprehensiveness we include them again here, augmenting their description and providing a more detailed analysis, especially in what relates to the derivation of the key equations included at the end of each subsection. We also improve in the present paper the calculation of the service time, total delay and buffer occupancy, as will be described in the next section devoted to performance analysis.

We propose a different model for each of the access categories AC1 through AC3. Each of the models shown in Figures 1–3 is a discrete-time Markov chain representing a single node competing for the medium through the EDCA MAC protocol. Table 2 displays the symbols used in all three models and explains their meanings. AC0 can also be modeled using the technique presented here, but its analysis is less tractable given the large number of states that the corresponding model has. In addition to that, traffic corresponding to AC0 is not vital for vehicle safety, which is the main focus of this paper, so we decided not to include it in this work.

In Figure 1, which represents a highest-priority node (AC3), the number $k \in \{0, 1, 2, 3\}$ in the lower states corresponds to the current value of the backoff counter. The node is only allowed to change state at very specific time instants. To be more precise, the node cannot change its state when the channel is detected busy. It will be allowed to change its state when a busy period ends and the channel remains idle for a period of length AIFS [AC3], as described in the IEEE 802.11 standard. If the channel continues being idle, the node can keep changing its state every slot, which is defined as a time period of length σ, as mentioned above. When the channel is detected busy again, the Markov chain will freeze once more until the next idle period. The time elapsed between instants at which the Markov chain is allowed to change state will be referred to as a *cycle*.

In Figure 2, which represents a second-highest-priority node (AC2), state (k, ℓ) is such that the number $k \in \{1, 2, 3\}$

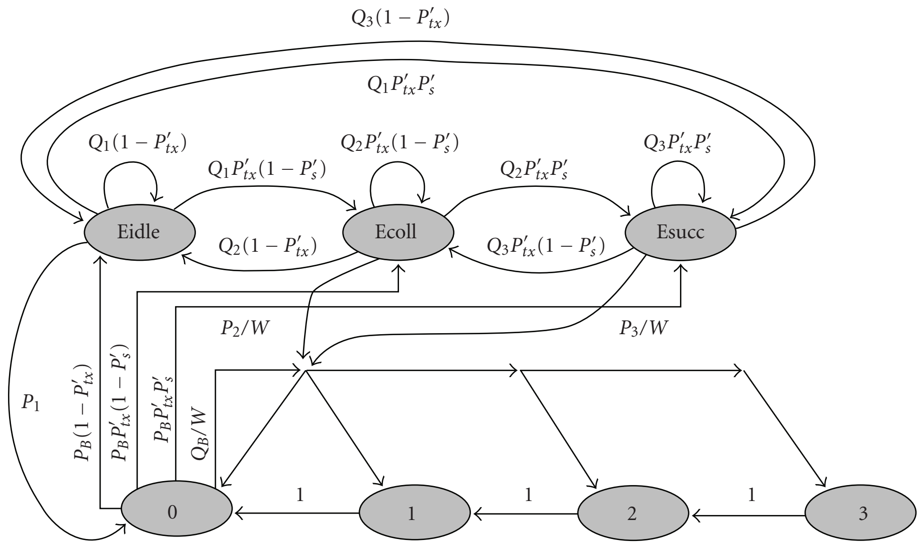


Figure 1: Markov model of a node corresponding to access category 3 (highest priority).

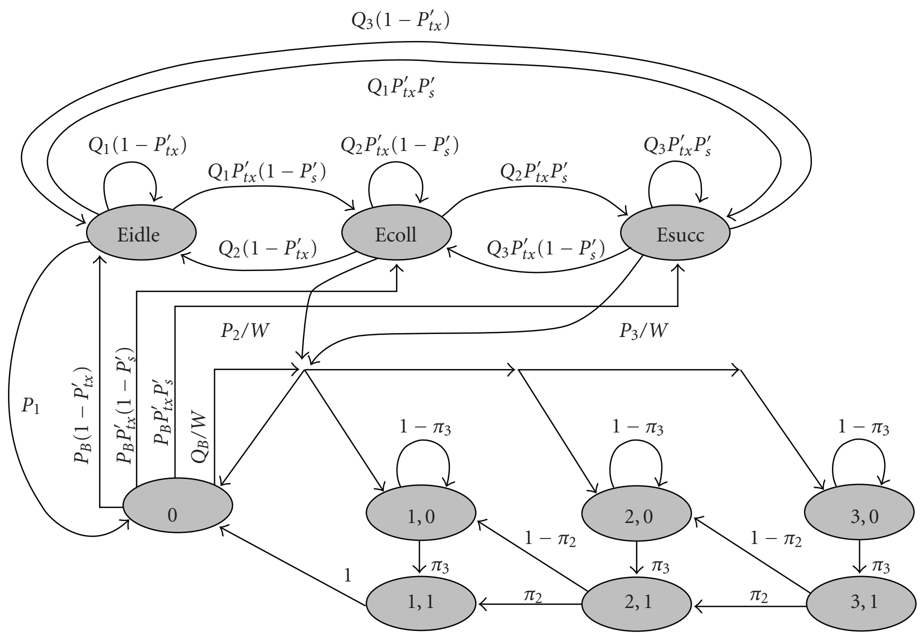


Figure 2: Markov model of a node corresponding to access category 2 (second-highest priority).

corresponds to the current value of the backoff counter and $\ell \in \{0, 1\}$ tells how many times a highest-priority node may have decreased its backoff counter since the most recent busy period. In other words, an AC2 node is always in state $(k, 0)$ when the medium is busy. Eventually the current transmission ends and the medium remains idle during AIFS [AC3], which is interpreted by AC3 nodes as the beginning of an idle period. If the medium stays inactive during an additional slot, then every backlogged AC3 node (if any) will decrease its backoff counter. If counting down once is enough for at least one of the AC3 nodes to reach $k = 0$ and decide to transmit, then the relevant AC2 node will go back to state $(k, 0)$ and remain there until the next idle period. If none of the AC3 nodes reaches $k = 0$, on the other hand, then the AC2 node will change its state to $(k, 1)$ indicating that AC3 nodes have counted down once during the current inactive period. Once in state $(k, 1)$, if the medium remains idle during an additional slot, then the AC2 node can reduce its backoff counter by moving into state $(k - 1, \ell)$. The value of ℓ will be zero or one, respectively, depending on whether the medium gets busy again or not. This time the transmitter can belong to AC2 or AC3. Strictly speaking, a value of ℓ equal to 1 indicates that the medium has been uninterruptedly idle for at least one slot.

Table 2: Symbols used in the models and their meanings.

Symbol	Meaning
β	Average number of frames in a burst, geometrically distributed
P_B	Probability that there are no more frames in current burst = $1/\beta$
Q_B	$1 - P_B$
$1/\lambda$	Average time between the end of a burst and the arrival of the next one, exponentially distributed
σ	Length of an idle cycle in which no node transmits
P_1	Probability of arrival of a traffic burst during an idle cycle = $1 - \exp(-\lambda\ \sigma)$
Q_1	$1 - P_1$
T_c	Length of a cycle in which there is a collision because two or more nodes attempt a transmission
P_2	Probability of arrival of a traffic burst during a busy cycle ending in a collision = $1 - \exp(-\lambda\ T_c)$
Q_2	$1 - P_2$
T_s	Length of a cycle in which a node successfully transmits a frame
P_3	Probability of arrival of a traffic burst during a busy cycle ending in the successful transmission of a frame = $1 - \exp(-\lambda T_s)$
Q_3	$1 - P_3$
P'_{tx}	Probability that at least one of the other competing terminals transmits during the current cycle
P'_s	Probability of a successful transmission given that at least one of the other competing terminals transmits
π_i	Probability that no node of priority i or higher will transmit during the current cycle
W	CWmin + 1

In Figure 3, which represents a node of the lowest priority we are analyzing (AC1), the meaning of state (k, ℓ) is similar to that in Figure 2. The value of ℓ goes now all the way up to 4 since, due to the difference in the respective AIFSN values (namely, AIFSN [AC1] = AIFSN [AC3] + 4), a highest-priority node may have counted down 4 times by the time a lowest-priority station decides that the medium is idle and is then free to start decreasing its own backoff counter. In this case, a value of ℓ equal to 1, 2 or 3 indicates that the medium has been uninterruptedly idle for *exactly* that many slots, while a value of ℓ equal to 4 indicates that the medium has been uninterruptedly idle for *at least* 4 slots.

All three models implicitly assume that traffic is generated in bursts according to a Poisson process, and the number of frames generated in each burst has a geometric distribution. In other words, our model is equivalent to an ON-OFF process in which the ON and OFF periods are geometrically and exponentially distributed, respectively. It is worth pointing out that these distributions, one discrete and the other continuous, are the only ones that satisfy the memoryless condition, needed for Markov modeling. This is admittedly not the most accurate traffic model for all types of applications, but the advantage of using it is that it allows us to analyze the system under different traffic loads, as opposed to getting only results for the system in saturation, which is what all previous models do. Voice activity is an example of traffic that complies with this model, as explained in [16].

Each of the three models mentioned above includes three states that represent those periods of time during which the node has no traffic to transmit. There are three different *empty* states since, depending on the length of the period during which the Markov chain freezes (because the channel was idle during a slot, because an attempted transmission failed, or because a frame was successfully transmitted), the probability that the buffer has newly-generated frames at the end of the cycle is different. This is a direct consequence of our assumption that traffic is generated in bursts according to a Poisson process. To be more specific, E_{idle} represents a state in which the relevant node's transmission buffer is empty and none of the other nodes in the network attempts a transmission, E_{coll} represents a state in which the relevant node's transmission buffer is again empty and two or more of the other nodes attempt a transmission and collide, and finally E_{succ} represents a state in which the relevant node's transmission buffer is empty and exactly one of the other nodes successfully transmits a frame. The node remains in states E_{idle}, E_{coll}, and E_{succ} during periods of length σ, T_c, and T_s, respectively.

One important difference between our models and all models previously proposed in the literature is that the models proposed here only have one backoff stage, meaning that there are no retransmissions or binary-exponential increase of backoff periods. This is due to the fact that all transmissions over the CCH are broadcasted; hence there is no way to know if a frame transmission is successful or not since nodes will not send back an acknowledgment. Another major contribution of this work is the great level of detail in which priorities are explicitly modeled.

Each model includes the influence of other competing terminals through the following probabilities: π_j, $j \in \{1, 2, 3\}$, which is the probability that no node of priority j or higher will transmit during the current cycle; P'_{tx}, which is the probability that at least one of the *other* competing terminals transmits; and P'_s, which is the probability that there is a successful transmission given that at least one of the *other* competing terminals transmits. All of these probabilities depend on the access category $i \in \{1, 2, 3\}$ that the relevant node belongs to. This dependence is not explicitly shown in Figures 1–3 for the sake of simplicity, but it will be taken into account in the equations included below. If we denote by τ_j the probability that a node of priority j transmits during a cycle, and by N_j the number of priority-j nodes present in the system, then the following equations hold. As seen by nodes of AC2 and AC1:

$$\pi_3 = (1 - \tau_3)^{N_3}. \tag{2}$$

As seen by a node of AC2:

$$\pi_2 = (1 - \tau_3)^{N_3}(1 - \tau_2)^{N_2 - 1}. \tag{3}$$

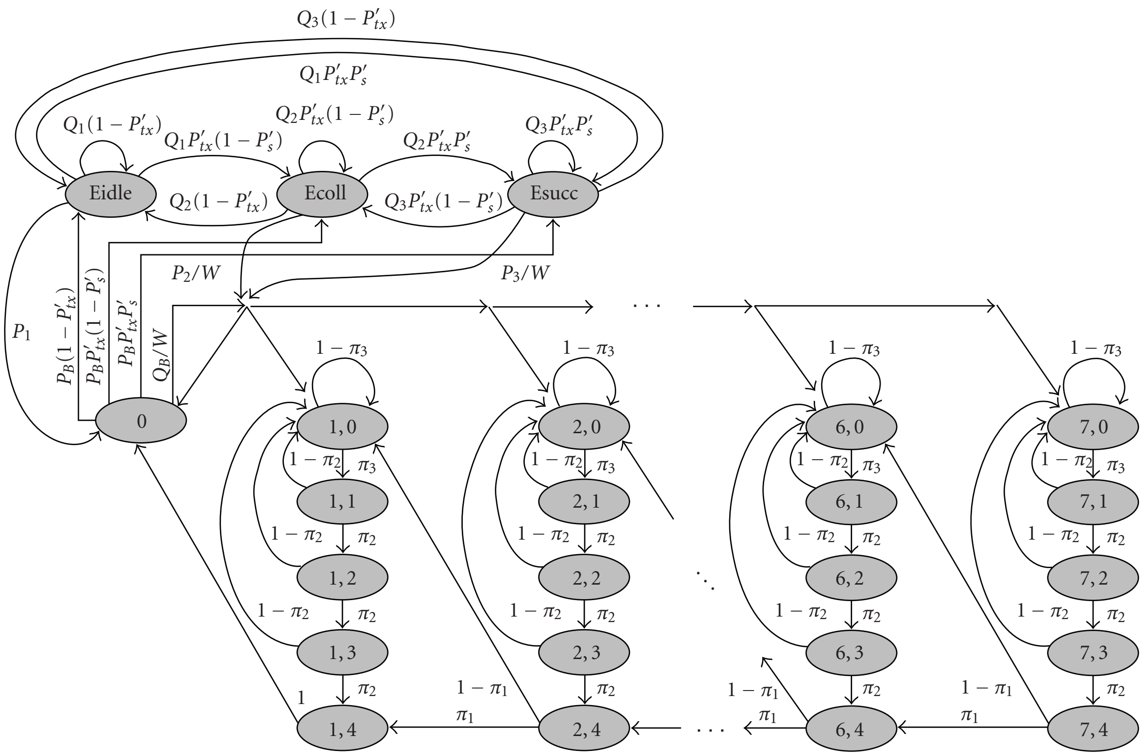


Figure 3: Markov model of a node corresponding to access category 1 (lowest priority).

As seen by a node of AC1:

$$\pi_2 = (1-\tau_3)^{N_3}(1-\tau_2)^{N_2}$$
$$\pi_1 = (1-\tau_3)^{N_3}(1-\tau_2)^{N_2}(1-\tau_1)^{N_1-1}. \tag{4}$$

Now, the value of P'_{tx} as seen by a node of access category i is

$$P'_{tx,i} = 1-(1-\tau_i)^{N_i-1}\prod_{\substack{j=1\\ j\neq i}}^{3}\left(1-\tau_j\right)^{N_j}. \tag{5}$$

Similarly, the value of P'_s as seen by a node of access category i is

$$p'_{s,i} = \frac{1}{p'_{tx,i}}\left[(N_i-1)\tau_i(1-\tau_i)^{N_i-2}\prod_{\substack{m=1\\ m\neq i}}^{3}(1-\tau_m)^{N_m} + \sum_{\substack{j=1\\ j\neq i}}^{3}N_j\tau_j\left(1-\tau_j\right)^{N_j-1} \times(1-\tau_i)^{N_i-1}\prod_{\substack{m=1\\ m\neq i,j}}^{3}(1-\tau_m)^{N_m}\right]. \tag{6}$$

In (6), the first term represents a successful transmission of a node of ACi and the second term represents a successful transmission of a node of an AC different from i. The presence of $P'_{tx,i}$ in the denominator comes from the fact that $P'_{s,i}$ is a conditional probability.

The symbols shown and described in Table 2 are used in the following subsections, which analyze individually each of the models presented in Figures 1 through 3.

3.1. Analysis of the Highest-Priority Model. Let b_k, $k \in \{0, 1, 2, 3\}$, denote the steady-state probability that the Markov chain in Figure 1 is in state k, and let b_{Eidle}, b_{Ecoll}, and b_{Esucc} each represent the steady-state probability that the Markov chain is in the corresponding empty state. To simplify notation, let us also define

$$W = \text{CWmin} + 1. \tag{7}$$

From Figure 1, if we apply the balance principle for the steady-state probabilities of the relevant Markov chain, we obtain the following equations:

$$b_k = \frac{b_0\cdot Q_B + b_{\text{Ecoll}}\cdot P_2 + b_{\text{Esucc}}\cdot P_3}{W} + b_{k+1}, \quad 1 \le k \le W-2 \tag{8}$$

$$b_{W-1} = \frac{b_0\cdot Q_B + b_{\text{Ecoll}}\cdot P_2 + b_{\text{Esucc}}\cdot P_3}{W}. \tag{9}$$

We can rewrite (8) as follows:

$$b_k = b_{W-1} + b_{k+1}, \quad 1 \le k \le W-2. \tag{10}$$

Using (10) recursively, we have:

$$\begin{aligned} b_k &= (W-k)\cdot b_{W-1} \\ &= \frac{(W-k)(b_0\cdot Q_B + b_{\mathrm{Ecoll}}\cdot P_2 + b_{\mathrm{Esucc}}\cdot P_3)}{W}, \\ &\qquad 1 \le k \le W-1. \end{aligned} \tag{11}$$

We can also see from Figure 1 that:

$$\begin{aligned} b_{\mathrm{Eidle}} &= \frac{b_0\cdot P_B\cdot \left(1-P'_{tx,3}\right)}{1-(A+B+C)} \\ b_{\mathrm{Ecoll}} &= \frac{b_0\cdot P_B\cdot P'_{tx,3}\cdot \left(1-P'_{s,3}\right)}{1-(A+B+C)} \\ b_{\mathrm{Esucc}} &= \frac{b_0\cdot P_B\cdot P'_{tx,3}\cdot P'_{s,3}}{1-(A+B+C)}, \end{aligned} \tag{12}$$

where $A = Q_1\cdot(1-P'_{tx,i})$, $B = Q_2\cdot P'_{tx,i}\cdot(1-P'_{s,i})$, and $C = Q_3\cdot P'_{tx,i}\cdot P'_{s,i}$. Additionally, the following equation has to be satisfied:

$$1 = \sum_{k=0}^{W-1} b_k + b_{\mathrm{Eidle}} + b_{\mathrm{Ecoll}} + b_{\mathrm{Esucc}}. \tag{13}$$

From (11) through (13) we conclude:

$$\begin{aligned} \frac{1}{b_0} &= 1 + \frac{P_B}{1-(A+B+C)} + \frac{W-1}{2} \\ &\quad \cdot \left\{ Q_B + P_B\cdot P'_{tx,3}\cdot \left[\frac{P_2\cdot\left(1-P'_{s,3}\right) + P_3\cdot P'_{s,3}}{1-(A+B+C)}\right]\right\}. \end{aligned} \tag{14}$$

3.2. Analysis of the Second-Highest-Priority Model. Now, let $b_{k,\ell}$ represent the steady-state probability that the Markov chain in Figure 2 is in state (k,ℓ), for $k \in \{1, 2, 3\}$ and $\ell \in \{0,1\}$. The meanings of b_0, b_{Eidle}, b_{Ecoll}, and b_{Esucc} are the same as those defined in the previous section. Based again on the balance principle for the steady-state probabilities of the relevant Markov chain, we can see that the following equations are satisfied:

$$b_{k,1} = b_{k,0}\cdot \pi_3 + b_{k+1,1}\cdot \pi_2, \quad 1 \le k \le W-2. \tag{15}$$

Using (15) recursively, we have:

$$b_{k,1} = \pi_3 \sum_{i=0}^{W-k-2} b_{k+i,0}\cdot \pi_2^i + b_{W-1,1}\cdot \pi_2^{W-k-1}. \tag{16}$$

But we can also obtain from Figure 2 that:

$$b_{W-1,1} = b_{W-1,0}\cdot \pi_3 \tag{17}$$

which turns (16) into the following expression:

$$b_{k,1} = \pi_3 \sum_{i=0}^{W-k-1} b_{k+i,0}\cdot \pi_2^i, \quad 1 \le k \le W-1. \tag{18}$$

Likewise, we can also obtain the following equations:

$$\begin{aligned} b_{k,0} &= \frac{b_0\cdot Q_B + b_{\mathrm{Ecoll}}\cdot P_2 + b_{\mathrm{Esucc}}\cdot P_3}{W} + b_{k,0}\cdot(1-\pi_3) \\ &\quad + b_{k+1,1}\cdot(1-\pi_2), \quad 1 \le k \le W-2, \end{aligned} \tag{19}$$

$$\begin{aligned} b_{W-1,0} &= \frac{b_0\cdot Q_B + b_{\mathrm{Ecoll}}\cdot P_2 + b_{\mathrm{Esucc}}\cdot P_3}{W} + b_{W-1,0}\cdot(1-\pi_3) \\ &= \frac{b_0\cdot Q_B + b_{\mathrm{Ecoll}}\cdot P_2 + b_{\mathrm{Esucc}}\cdot P_3}{W\cdot \pi_3}. \end{aligned} \tag{20}$$

From (18) and (19), we obtain:

$$\begin{aligned} b_{k,0} &= \frac{b_0\cdot Q_B + b_{\mathrm{Ecoll}}\cdot P_2 + b_{\mathrm{Esucc}}\cdot P_3}{W\cdot \pi_3} \\ &\quad + (1-\pi_2)\sum_{i=0}^{W-k-2} b_{k+i+1,0}\cdot \pi_2^i, \quad 1 \le k \le W-2. \end{aligned} \tag{21}$$

Putting together (20) and (21) we find that, for $1 \le k \le W-1$:

$$\begin{aligned} b_{k,0} &= [(W-k)-(W-k-1)\cdot \pi_2]\cdot b_{W-1,0} \\ &= [(W-k)-(W-k-1)\cdot \pi_2] \\ &\quad \cdot \left(\frac{b_0\cdot Q_B + b_{\mathrm{Ecoll}}\cdot P_2 + b_{\mathrm{Esucc}}\cdot P_3}{W\cdot \pi_3}\right). \end{aligned} \tag{22}$$

Substituting these values into (18) gives, for $1 \le k \le W-1$:

$$\begin{aligned} b_{k,1} &= (W-k)\cdot \pi_3\cdot b_{W-1,0} \\ &= (W-k)\cdot\left(\frac{b_0\cdot Q_B + b_{\mathrm{Ecoll}}\cdot P_2 + b_{\mathrm{Esucc}}\cdot P_3}{W}\right). \end{aligned} \tag{23}$$

Again, the sum of all probabilities has to equal one, which in this case can be expressed as the following condition:

$$1 = b_0 + \sum_{k=1}^{W-1}\left(b_{k,0}+b_{k,1}\right) + b_{\mathrm{Eidle}} + b_{\mathrm{Ecoll}} + b_{\mathrm{Esucc}}. \tag{24}$$

Therefore, from (22), (23) and (24), we find that, for a second-highest-priority station, we have:

$$\begin{aligned} \frac{1}{b_0} &= 1 + \frac{P_B}{1-(A+B+C)} \\ &\quad + \frac{W-1}{2}\left[\left(\frac{1+\pi_3}{\pi_3}\right) - \left(\frac{W-2}{W}\right)\cdot\left(\frac{\pi_2}{\pi_3}\right)\right] \\ &\quad \cdot \left[\frac{P_B\cdot P'_{tx,2}\cdot\left[P_2\left(1-P'_{s,2}\right) + P_3\cdot P'_{s,2}\right]}{1-(A+B+C)} + Q_B\right]. \end{aligned} \tag{25}$$

3.3. Analysis of the Lowest-Priority Model. In this case, the expressions relating the different probabilities are the following:

$$b_{k,1} = b_{k,0} \cdot \pi_3, \quad 1 \leq k \leq W-1 \tag{26}$$

$$b_{k,2} = b_{k,1} \cdot \pi_2 = b_{k,0} \cdot \pi_2 \cdot \pi_3, \quad 1 \leq k \leq W-1 \tag{27}$$

$$b_{k,3} = b_{k,2} \cdot \pi_2 = b_{k,0} \cdot \pi_2^2 \cdot \pi_3, \quad 1 \leq k \leq W-1 \tag{28}$$

$$b_{k,4} = b_{k,3} \cdot \pi_2 + b_{k+1,4} \cdot \pi_1, \quad 1 \leq k \leq W-2. \tag{29}$$

Using (29) recursively, we have:

$$b_{k,4} = \pi_2 \sum_{i=0}^{W-k-2} b_{k+i,3} \cdot \pi_1^i + b_{W-1,4} \cdot \pi_1^{W-k-1}. \tag{30}$$

But we can also obtain from Figure 3 that:

$$b_{W-1,4} = b_{W-1,3} \cdot \pi_2. \tag{31}$$

Substituting (31) and (28) into (30), results in the following expression:

$$b_{k,4} = \pi_2^3 \cdot \pi_3 \sum_{i=0}^{W-k-1} b_{k+i,0} \cdot \pi_1^i, \quad 1 \leq k \leq W-1. \tag{32}$$

The following equations are also obtained from Figure 3 for $1 \leq k \leq W-2$:

$$\begin{aligned} b_{k,0} = {} & \frac{b_0 \cdot Q_B + b_{\text{Ecoll}} \cdot P_2 + b_{\text{Esucc}} \cdot P_3}{W} + b_{k,0} \cdot (1-\pi_3) \\ & + \left(b_{k,1} + b_{k,2} + b_{k,3}\right) \cdot (1-\pi_2) + b_{k+1,4} \cdot (1-\pi_1), \end{aligned} \tag{33}$$

$$\begin{aligned} b_{W-1,0} = {} & \frac{b_0 \cdot Q_B + b_{\text{Ecoll}} \cdot P_2 + b_{\text{Esucc}} \cdot P_3}{W} + b_{W-1,0} \cdot (1-\pi_3) \\ & + \left(b_{W-1,1} + b_{W-1,2} + b_{W-1,3}\right) \cdot (1-\pi_2) \\ = {} & \frac{b_0 \cdot Q_B + b_{\text{Ecoll}} \cdot P_2 + b_{\text{Esucc}} \cdot P_3}{W \cdot \pi_2^3 \cdot \pi_3}. \end{aligned} \tag{34}$$

Following a procedure similar to the one used in the previous section, in which equations are recursively simplified, we obtain from (32), (33), and (34) that, for $1 \leq k \leq W-1$:

$$\begin{aligned} b_{k,0} & = \left[(W-k) - (W-k-1) \cdot \pi_1\right] \cdot b_{W-1,0} \\ & = \left[(W-k) - (W-k-1) \cdot \pi_1\right] \\ & \quad \cdot \left(\frac{b_0 \cdot Q_B + b_{\text{Ecoll}} \cdot P_2 + b_{\text{Esucc}} \cdot P_3}{W \cdot \pi_2^3 \cdot \pi_3}\right). \end{aligned} \tag{35}$$

And that:

$$\begin{aligned} b_{k,4} & = (W-k) \cdot \pi_2^3 \cdot \pi_3 \cdot b_{W-1,0} \\ & = (W-k) \cdot \left(\frac{b_0 \cdot Q_B + b_{\text{Ecoll}} \cdot P_2 + b_{\text{Esucc}} \cdot P_3}{W}\right). \end{aligned} \tag{36}$$

In this case, the requirement that probabilities must add up to 1 leads to the following expression:

$$1 = b_0 + \sum_{k=1}^{W-1} \sum_{i=1}^{4} b_{k,i} + b_{\text{Eidle}} + b_{\text{Ecoll}} + b_{\text{Esucc}}. \tag{37}$$

Finally, substituting (26) through (28) and (35) through (36) into (37), we find that for a lowest-priority station:

$$\begin{aligned} \frac{1}{b_0} = {} & 1 + \frac{P_B}{1-(A+B+C)} + \frac{W-1}{2} \\ & \times \left\{\left[1 - \frac{(W-2) \cdot \pi_1}{W}\right]\right. \\ & \quad \left. \cdot \left(\frac{1 + \pi_3 + \pi_2 \cdot \pi_3 + \pi_2^2 \cdot \pi_3}{\pi_2^3 \cdot \pi_3}\right) + 1\right\} \\ & \cdot \left[\frac{P_B \cdot P'_{tx,1} \cdot \left[P_2\left(1 - P'_{s,1}\right) + P_3 \cdot P'_{s,1}\right]}{1-(A+B+C)} + Q_B\right]. \end{aligned} \tag{38}$$

4. Performance Analysis

In all models analyzed in the previous sections, note that a node will transmit a frame when it reaches state 0. Therefore, the probability τ_i that a priority-i node transmits during a given cycle is equal to the corresponding value of b_0. This fact plus (14), (25), and (38) allow us to numerically calculate the values of τ_i, $i \in \{1, 2, 3\}$. However, in order to be able to use (5) and (6), it is necessary to know the values of τ_j for all priorities j different from i. When we start performing our calculations, we clearly do not know any of those values. To solve this problem, we propose to use initially the following approximation to estimate those τ_j that have not yet been calculated:

$$\tau_j \approx \min\left\{\frac{\rho_j}{\rho_i} \cdot \tau_i, 0.5\right\} = \min\left\{\frac{\lambda_j \cdot \beta_j}{\lambda_i \cdot \beta_i} \cdot \tau_i, 0.5\right\}. \tag{39}$$

In (39), ρ_j is the traffic load generated by a priority-j terminal, for which λ_j is the burst generation rate and β_j is the average number of frames per burst. The rationale behind this approximation is that, if the system is not extremely congested, the percentage of cycles that each node will be transmitting is nearly proportional to the number of frames that it has to transmit per unit time. The deviation comes from the unpredictability associated to the selection of the random number of backoff periods before transmitting each frame and from the advantages that some nodes have over other nodes due to the different priorities. The expression on the left side of the brackets may be greater than or equal to one, which are values that would give anomalous results. To avoid this, the approximation is forced to have values smaller than one; specifically, in our case we use an upper bound equal to 0.5. This bound is not restricted to have any particular value, but we found that the selected value works well in terms of the algorithm's rate of convergence.

When all values of τ_j have been calculated, the process is restarted, but this time the most recently calculated values are used instead of (39). This process is recursively repeated until the new values and the previous ones are sufficiently close. In our case, the relative difference has to be smaller than 0.1%.

4.1. Throughput Calculation. When the values of τ_i, $i \in \{1, 2, 3\}$, have been calculated, we can compute $P_{tx,i}$ and $P_{s,i}$, which are respectively the probability that at least one priority-i terminal transmits, and the probability that there is a successful transmission given that at least one priority-i terminal transmits. Namely:

$$P_{tx,i} = 1 - (1 - \tau_i)^{N_i}$$
$$P_{s,i} = \frac{1}{P_{tx,i}} N_i \cdot \tau_i \cdot (1 - \tau_i)^{N_i - 1} \prod_{\substack{j=1 \\ j \neq i}}^{3} \left(1 - \tau_j\right)^{N_j}. \tag{40}$$

This will in turn allow us to compute the throughput achievable by priority-i terminals as follows:

$$S_i = \frac{P_{tx,i} \cdot P_{s,i} \cdot E[P_i]}{E[\text{cycle}]}. \tag{41}$$

In the previous equation, $E[P_i]$ is the average frame payload size and $E[\text{cycle}]$ is the average duration of each cycle, which is given by

$$E[\text{cycle}] = \sum_{j=1}^{3} P_{tx,j} \cdot P_{s,j} \cdot T_s + P_{\text{notx}} \cdot \sigma + \left(1 - P_{\text{notx}} - \sum_{j=1}^{3} P_{tx,j} \cdot P_{s,j}\right) \cdot T_c. \tag{42}$$

In turn, σ represents the duration of an empty cycle in which no node transmits, T_c is the average duration of a cycle in which there is a collision because two or more nodes attempt a transmission, and T_s is the average duration of a cycle in which a node successfully transmits a frame. P_{notx} denotes the probability that no station transmits, regardless of priority, and is given by

$$P_{\text{notx}} = \prod_{j=1}^{3} \left(1 - \tau_j\right)^{N_j}. \tag{43}$$

The parameters σ, T_c and T_s depend on the physical layer. We are assuming a scenario compatible with the IEEE 802.11a standard, with 10 MHz channel spacing, and running at 6 Mbps. We also assume that the propagation delay δ is equal to 1 μs. Table 3 summarizes the parameter values assumed. With them we can calculate:

$$T_s = \text{PHY\_header} + \text{MAC\_header} + E[P] + \text{AIFS}[\text{AC}\,3] + \delta$$
$$T_c = T_s + (\text{EIFS} - \text{AIFS}[\text{AC}\,3]) = T_s + (\text{SIFS} + \text{ACKtime}). \tag{44}$$

Table 3: Physical layer parameters.

Parameter	Value
PHY header	
PLCP Preamble	32 μs
PLCP Signal	8 μs
PLCP Service	16 bits
PLCP Tail	6 bits
MAC header	288 bits
ACK size	400 bits
Frame payload E[P]	4096 bits
Channel bit rate	6 Mbps
Propagation delay (δ)	1 μs
Slot (σ)	13 μs
SIFS	32 μs

4.2. Frame-Error-Rate Calculation. The frame error rate (FER) is defined as the probability that a frame transmission attempt fails due to collisions. The expression to calculate FER for a priority-i node is as follows:

$$\text{FER}_i = 1 - (1 - \tau_i)^{N_i - 1} \prod_{\substack{j=1 \\ j \neq i}}^{3} \left(1 - \tau_j\right)^{N_j}. \tag{45}$$

4.3. Service Time Calculation. The average service time for a priority-i node, defined as the time elapsed from the moment a frame arrives at the head of its queue until it is finally transmitted, is given by

$$E[X_i] = E[n_{x,i}] \cdot E[\text{cycle}] + T_s, \quad i \in \{1, 2, 3\}, \tag{46}$$

where $n_{x,i}$ is the number of cycles the frame has to back off, waiting in the buffer for its turn to be transmitted.

The following equation applies to AC3:

$$E[n_{x,3}] = \sum_{k=0}^{\text{CWmin}} k \cdot \Pr[\text{backoff} = k] = \frac{\text{CWmin}}{2} = 1.5. \tag{47}$$

In the previous equation we are using the fact that $\Pr[\text{backoff} = k]$ is the same for all values of k, and equal to $1/(\text{CWmin} + 1)$.

For AC2, the calculation is slightly more complicated. If we denote by m_k, $k \in \{1, 2, 3\}$, the geometrically-distributed random variable that represents the number of cycles that an AC2 node will remain in state $(k, 0)$, unable to decrease its backoff counter because higher-priority stations do not allow it, then we have that:

$$
\begin{aligned}
E[n_{x,2}] = {} & \Pr[\text{backoff} = 1] \cdot (E[m_1] + 1) \\
& + \Pr[backof = 2] \\
& \cdot [(E[m_2] + 2) \cdot \pi_2 + (E[m_2] + E[m_1] + 2) \\
& \quad \cdot (1 - \pi_2)] \\
& + \Pr[\text{backoff} = 3] \\
& \cdot \big[(E[m_3] + 3) \cdot \pi_2^2 + (E[m_3] + E[m_2] + 3) \\
& \quad \cdot \pi_2(1 - \pi_2) + (E[m_3] + E[m_1] + 3) \\
& \quad \cdot (1 - \pi_2) \cdot \pi_2 + (E[m_3] + E[m_2] + E[m_1] + 3) \\
& \quad \cdot (1 - \pi_2)^2\big]
\end{aligned}
\tag{48}
$$

Taking into account that:

$$
E[m_k] = \sum_{n=1}^{\infty} n \cdot (1 - \pi_3)^{n-1} \cdot \pi_3 = 1/\pi_3, \quad k \in \{1, 2, 3\}
\tag{49}
$$

then we obtain:

$$
E[n_{x,2}] = \frac{3}{4}\left(2 + \frac{2}{\pi_3} - \frac{\pi_2}{\pi_3}\right).
\tag{50}
$$

Similarly, the calculation of $n_{x,1}$ for AC1 can be done as follows. If we now denote by $\widehat{m}_k$, $k \in \{1, 2, \ldots, 7\}$, the random variable that represents the number of cycles that an AC1 node will remain trapped in states of the form (k, ℓ), $\ell \in \{0, 1, 2, 3\}$, unable to reach state $(k, 4)$ and therefore unable to decrease its backoff counter, then we have that:

$$
\begin{aligned}
E[n_{x,1}] = {} & \sum_{k=1}^{7} \Pr[\text{backoff} = k] \cdot \sum_{j=0}^{k-1} \binom{k-1}{j} \\
& \cdot \left[(j+1)E[\widehat{m}_1] + k\right] \cdot (1 - \pi_1)^j \cdot \pi_1^{k-j-1}.
\end{aligned}
\tag{51}
$$

In the previous equation the expression $E[\widehat{m}_1]$ is used to represent any value $E[\widehat{m}_k]$, $k \in \{1, 2, \ldots, 7\}$, since all of them are equal. It can be proved that:

$$
E[\widehat{m}_k] = \frac{1 + \pi_3 + \pi_2 \cdot \pi_3 + \pi_2^2 \cdot \pi_3}{\pi_2^3 \cdot \pi_3}, \quad k \in \{1, 2, \ldots, 7\}.
\tag{52}
$$

This result can be obtained by analyzing each column of states in Figure 3 in isolation, as if it were an absorbing Markov chain having $(k, 4)$, $k \in \{1, 2, \ldots, 7\}$, as the absorbing state [17]. This partial Markov chain is shown in Figure 4. $E[\widehat{m}_k]$ will then be the average number of steps needed to reach state $(k, 4)$ for the first time assuming that the initial state is $(k, 0)$.

From (51) and (52) we obtain the following equation:

$$
E[n_{x,1}] = \frac{7}{8}\left[4 + \frac{(4 - 3 \cdot \pi_1)(1 + \pi_3 + \pi_2 \cdot \pi_3 + \pi_2^2 \cdot \pi_3)}{\pi_2^3 \cdot \pi_3}\right].
\tag{53}
$$

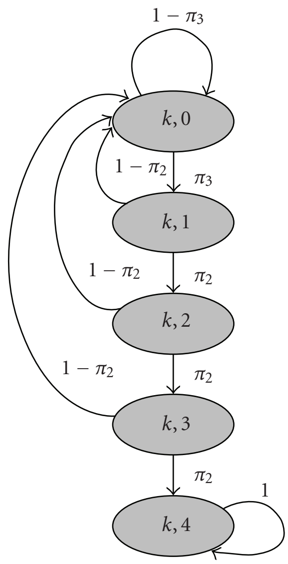
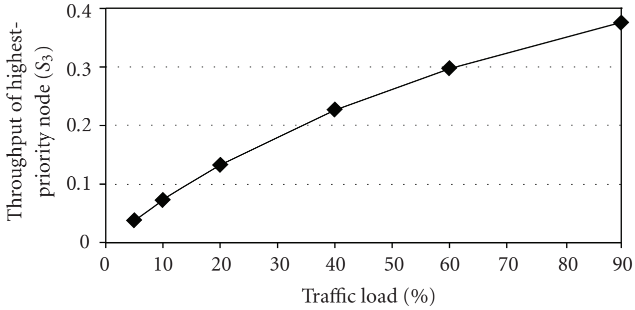


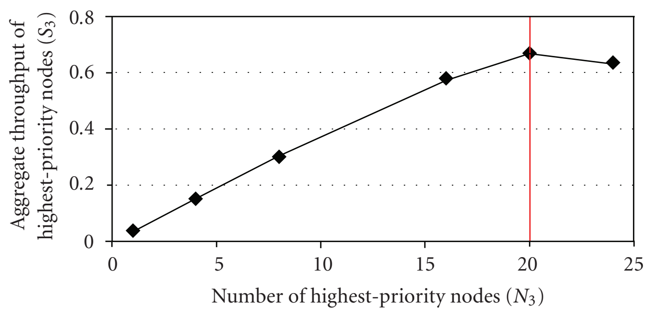
Figure 4: Absorbing Markov chain used to calculate service time for access category 1 (lowest priority).

4.4. Total Delay Calculations. The total delay is defined as the time elapsed from the moment a frame is generated as a result of a packet being received from the application layer to the time when it is finally transmitted. In order to calculate the average total delay that will be experienced by frames, it may be observed in Figures 1, 2 and 3 that the transmission of each burst, from its generation to the time when the buffer is again empty, is independent from all other bursts, which constitutes a renewal process. Taking this into account, we can calculate the average total frame delay by analyzing a single burst. Let us assume that ϕ is the number of frames in the burst being analyzed, that $X_i(n)$ is the service time of the nth frame generated by a node of access class i, and that D_{burst} is the cumulative delay that results from adding the total delay experienced by every frame in the burst. In what follows we will use the fact that, for access class i, $E[\phi] = \beta_i$. Then, the average total delay that will be experienced by frames of access class i is given by

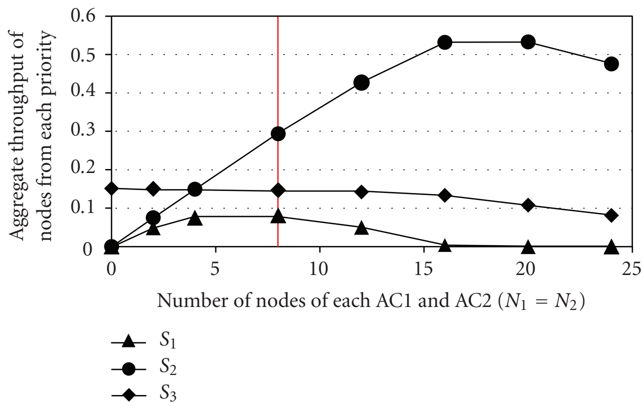
$$
\begin{aligned}
E[D_i] &= \frac{E[D_{\text{burst}}]}{E[\phi]} \\
&= \frac{1}{\beta_i} \sum_{k=1}^{\infty} E[D_{\text{burst}} \mid \phi = k] \cdot \Pr[\phi = k] \\
&= \frac{1}{\beta_i} \sum_{k=1}^{\infty} E\left[\sum_{j=1}^{k} \sum_{n=1}^{j} X_i(n)\right] \cdot \Pr[\phi = k] \\
&= \frac{1}{\beta_i} \sum_{k=1}^{\infty} \left[\sum_{j=1}^{k} \sum_{n=1}^{j} E[X_i]\right] \cdot \Pr[\phi = k] \\
&= \frac{E[X_i]}{\beta_i} \cdot \sum_{k=1}^{\infty} \left[\frac{k(k+1)}{2}\right] \cdot \Pr[\phi = k] \\
&= \frac{E[X_i]}{2 \cdot \beta_i} \cdot \{E[\phi^2] + E[\phi]\} = \beta_i \cdot E[X_i].
\end{aligned}
\tag{54}
$$



(a) Highest-priority throughput with one node generating different traffic loads



(b) Highest-priority throughput with a varying number of highest-priority nodes



(c) Throughput with a varying number of lower-priority nodes

Figure 5: Throughput results.

The last equality comes from the fact that ϕ is assumed to be a geometrically-distributed random variable, as mentioned earlier.

4.5. Buffer Occupancy Calculation. Once we have calculated the average total delay, we can also determine the average buffer occupancy for access category i ($\mathcal{O}_i$) by using Little's theorem, as follows:

$$\begin{aligned} E[\mathcal{O}_i] &= \lambda_{\text{eff}} \cdot E[D_i] \\ &= \lambda_{\text{eff}} \cdot \beta_i \cdot E[X_i], \quad i \in \{1,2,3\}, \end{aligned} \tag{55}$$

where λ_{eff} is the effective frame arrival rate, given by the following expression:

$$\begin{aligned} \lambda_{\text{eff}} &= \lambda_i \cdot \beta_i \cdot \Pr[\text{Buffer is empty}] \\ &= \lambda_i \cdot \beta_i \cdot \frac{b_0 \cdot P_{B,i}}{1-(A+B+C)} = \frac{\lambda_i \cdot b_0}{1-(A+B+C)}. \end{aligned} \tag{56}$$

Putting together (55) and (56), we obtain:

$$E[\mathcal{O}_i] = \frac{\lambda_i \cdot \beta_i \cdot E[X_i] \cdot b_0}{1-(A+B+C)}, \quad i \in \{1,2,3\}. \tag{57}$$

In the previous equation, b_0, A, B, and C have to be values calculated for the relevant access category i.

5. Numerical Analysis

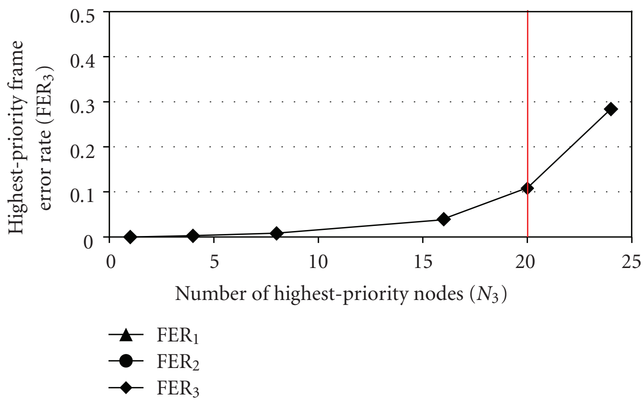
Different scenarios were analyzed.

(i) A single highest-priority station generating different traffic loads, from 5% to 90% of the channel actual capacity, as explained below.

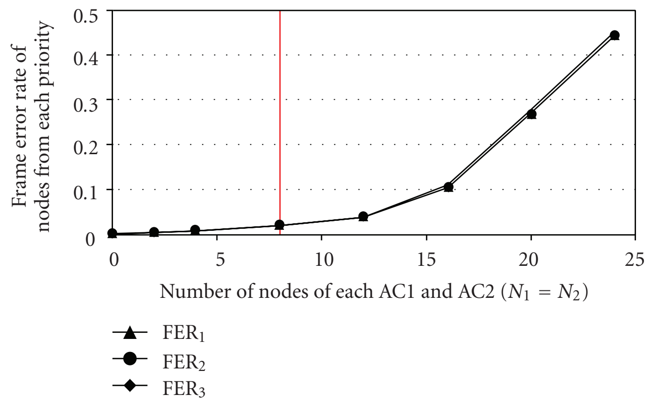
(ii) A varying number of highest-priority stations, each generating a constant amount of traffic equal to 5%. In this case, the number N_3 of stations of AC3 varies from 1 to 24, hence the offered traffic varies from 5% to 120%.

(iii) A fixed number N_3 of highest-priority stations, set to 4, and a varying number of lower-priority stations. In this case, N_1 and N_2, which represent the number of stations of AC1 and AC2, respectively, are assumed equal and vary from 1 to 24. Each station generates a constant amount of traffic equal to 5%, hence the offered traffic varies from 30% to 260%.

We estimate the load that each node puts into the system as $\lambda_i \cdot \beta_i \cdot T_S$, $i \in \{1,2,3\}$, which is the average number of frames generated per time interval of length T_S. When this value is 1 we consider that the load is 100% since the system is being required to service one frame every T_S, which is the average duration of a cycle in which a node successfully transmits a frame, as explained earlier. In all cases β_i is assumed to be equal to 5 and λ_i is varied to achieve the desired load.

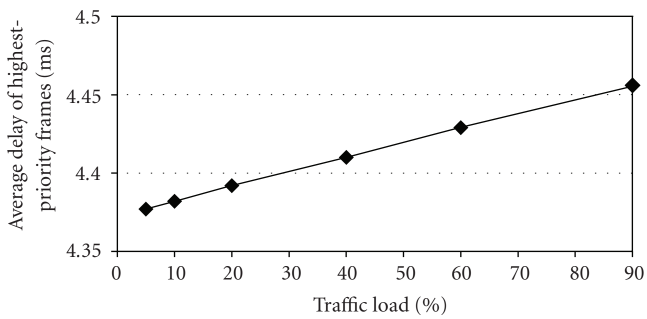


(a) Highest-priority FER with a varying number of highest-priority nodes

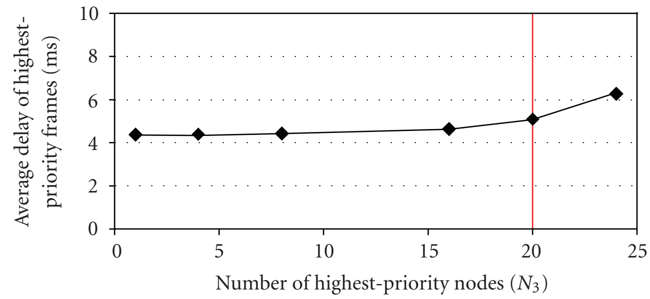


(b) FER with a varying number of lower-priority nodes

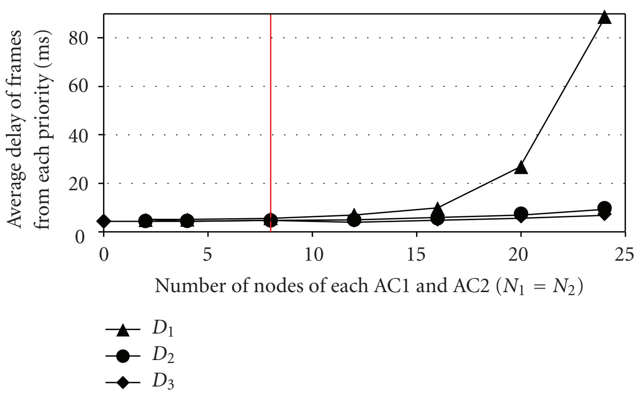
FIGURE 6: Frame-error rate results.



(a) Total highest-priority delay with one node generating different traffic loads



(b) Total highest-priority delay with a varying number of highest-priority nodes



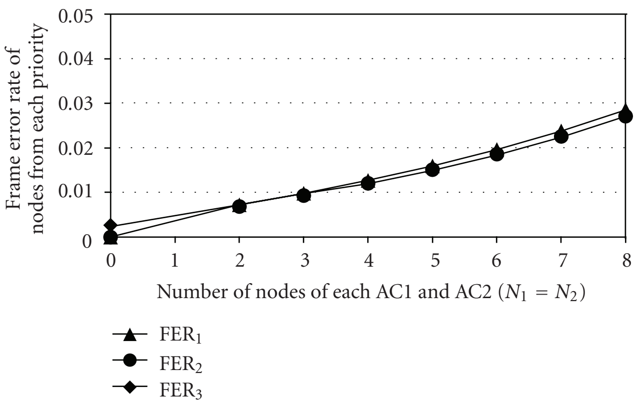
(c) Total delay with a varying number of lower-priority nodes

FIGURE 7: Total delay results.

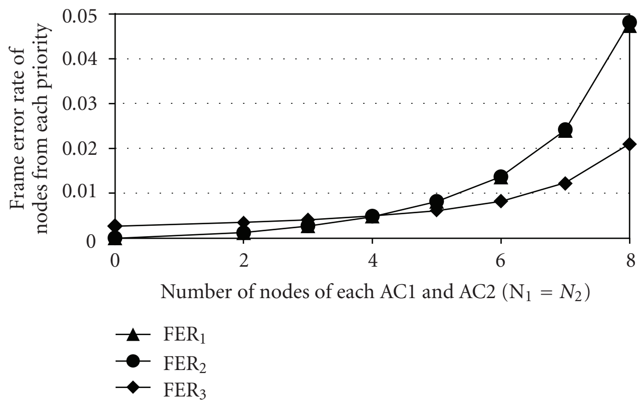
We present results related to throughput, losses, and total delay for frames corresponding to all three access categories.

We can see from Figure 5(a) that, in the case of a single highest-priority station using the network, the throughput keeps increasing even for offered loads of 90%. The price to pay is an increase in frame delays, as shown in Figure 7(a), but not in a significant way. In this case there are no frame losses since there are no other stations to collide with.

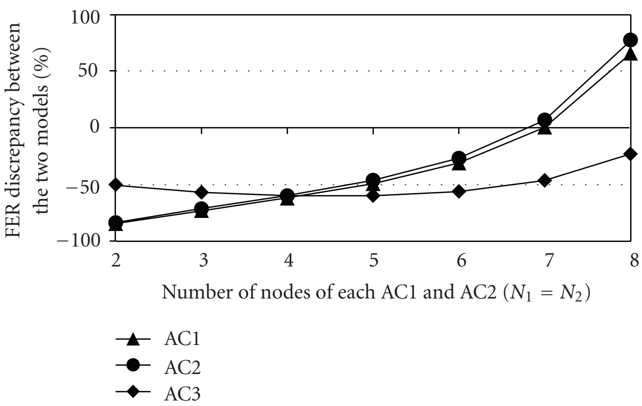
When there are only highest-priority nodes, the system can work relatively well for loads of up to 40% (8 competing nodes), in the sense that the throughput increases linearly and the delay and losses do not increase noticeably, as displayed in Figures 5(b), 6(a) and 7(b). After that point, the throughput keeps increasing linearly for load values of up to 80%, but the FER goes above 4% and increases very sharply for loads above the saturation point (marked by a vertical line in the figures).



(a) FER results with model proposed in this paper



(b) FER results with model proposed in previously-published paper



(c) Comparison of FER results

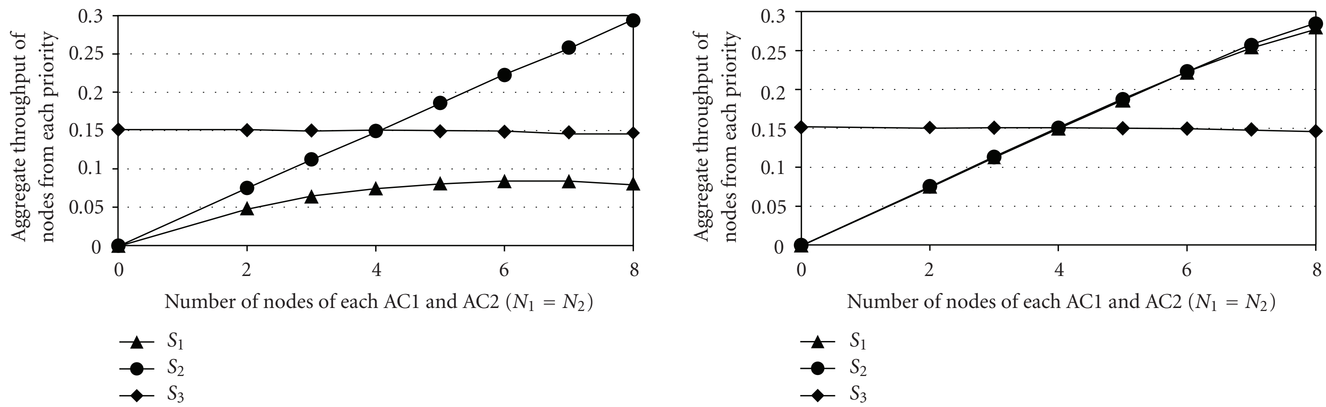
Figure 8: Comparison of FER results.

When we have all types of nodes, on the other hand, the highest-priority throughput is not affected noticeably even when the system reaches saturation, as shown in Figure 5(c). The throughput of AC2 nodes keeps increasing until the load offered by AC3 and AC2 nodes reaches 100%, meaning that the load offered by AC1 nodes does not have a perceptible impact on it. When we compare the FER corresponding to AC3 nodes in Figures 6(a) and 6(b), we can see that the load introduced by lower-priority nodes does not have as bad an impact as compared to the case in which competition is among AC3 nodes only (e.g., compare saturation points), but still performance is degraded beyond acceptable levels for loads of 60% and above.

As we can see from Figure 7(c), the delay is not increased dramatically for AC2 and AC3 even when the offered load goes far beyond saturation; it increases from approximately 4 ms to 7.35 and 9.78 ms, respectively. The reason for this is the fact that, even though there is a great level of competition among stations, the feedback usually provided by acknowledgments (or the lack thereof) to detect congestion is non-existent in this case since all frames are broadcasted. The delay experienced by frames corresponding to AC1 does increase when the load is high, but in this case it is due to the fact that there are many frames of a higher priority being transmitted, which leaves few opportunities for the channel to be idle for long-enough periods of time in order for AC1 stations to decrease their backoff counters; as an extension of the previous result we argue that, if the traffic load offered by AC3 nodes were higher, AC2 frames would also experience noticeable delay increases.

In what follows, we include additional results to clarify the advantages of the method presented in this paper over the one we proposed in [2], which is the only one comparable to the present work in terms of comprehensiveness of QoS-related results and consideration of traffic loads below and above saturation, as mentioned in Section 2.

Figures 8(a) and 8(b) show the FER results obtained with our new proposed method and with equations from [2], respectively, for traffic loads that range from low to high values (saturation is reached when $N_1 = N_2 = 8$). Results in Figure 8(a) are similar to those in Figure 6(b), but we are including them again to offer a closer look and to allow a clearer comparison. We can see in Figure 8(c), which plots the discrepancy between the two methods, that the work presented in [2] underestimates the value of FER by at least 50% in most cases for the highest-priority traffic and also by large percentages for the other two access categories. This is a direct consequence of the fact that [2] assumes that



(a) Throughput results with model proposed in this paper

(b) Throughput results with model proposed in previously-published paper

Figure 9: Comparison of throughput results.

the probability of collision between frames corresponding to different access categories is negligible, when it clearly is not. Figure 9 compares throughput results. We can observe in this figure that the method from [2] overestimates the throughput that AC1 nodes can achieve. We do not include a comparison of results related to delay since they are not very different. The reason for that is the fact, already mentioned, that the delay varies very little for broadcast frames when the traffic load is below saturation.

6. Considerations to Account for Intermittent Activity on the CCH

As mentioned before, activity on the CCH is intermittent, meaning that time is subdivided into synchronization intervals of 100 ms each and we are analyzing what happens during the first half of each of these intervals, which are used by stations to transmit safety-critical messages. The results presented so far are obtained assuming a system that works continuously. To account for intermittent activity, we have to make some adjustments, as explained below.

Throughput represents the fraction of time that the channel can be used to successfully transmit higher-layer information. Intermittence causes for the channel capacity to be reduced in half, which implies that the actual throughput is also reduced in half. For instance, if our equations give a value of 10% for the throughput, we can interpret that as if 5% of the 6 Mbps channel rate is being used for successful transmissions; in other words, higher-layer information will be flowing at an average rate of 0.3 Mbps.

Frame error rate calculations are not altered by intermittence, which means that our results are valid to evaluate the performance of an actual network working as specified by the IEEE 802.11 draft standard.

The actual service time, generically denoted by X_{act}, can be expressed as follows:

$$X_{\text{act}} = X + \left\lfloor \frac{Y + X}{T} \right\rfloor \cdot T. \qquad (58)$$

In the previous equation, X represents the service time that would be observed in a system without pauses, T denotes an interval of 50 ms, and $Y \in [0, T]$ is a uniformly-distributed random variable that represents the time instant, within an active period, at which the relevant frame arrives at the head of the queue to start being serviced. The second term in (58) corresponds to the idle periods that have to be added to the service time to account for intermittence.

When a typical service time X is comparatively smaller than T, which is what our results show in all unsaturated cases, then most frames will not include pauses in their service times, except for those few ones that happen to start service at a time close to the end of an active period. From this we can conclude that the expected value of the second term in (58) will represent a small fraction of $E[X]$. In the worst case, when the system is working in saturation and a typical X is large compared to T, the expected value of X_{act} will approach a value equal to $2 \cdot E[X]$. This tells us that even in a saturated environment, our equations give values that are not too far from what will be observed in a real system, and a factor between 1 and 2, depending on the level of congestion, can be used to account for the deviation.

An adjustment similar to (58) is also valid for the total delay experienced by frames, given by (54).

Our buffer occupancy calculations are not altered by intermittence, meaning that our results can be safely used in a real environment.

7. Conclusion

This paper describes and analyzes a model for the EDCA MAC protocol, taking into consideration the specific conditions of the control channel of an IEEE 802.11p WAVE vehicular network. To be more specific, the model captures the details as to how EDCA establishes different priorities among stations and that safety-related information is transmitted using the broadcast address.

The proposed model is based on discrete-time Markov chains and our analysis allows a comprehensive performance evaluation of this type of networks. Namely, we present results related to throughput, losses, buffer occupancy and delays, all of which are very important quality-of-service metrics, especially in a vehicular environment in which a short time can make a great difference when there are dangerous situations.

The main differences between this work and what is already available in the literature are:

(a) in our case there are no unreasonable simplifying assumptions,

(b) the protocol priorities are very explicitly modeled,

(c) the traffic load can be varied,

(d) the total frame delay is calculated, and

(e) an explanation is given as to how to take into account the fact that the activity in the CCH is intermittent.

Paragraphs (a) and (b) tell us about the accuracy of our proposed model. Paragraph (c) is paramount in a WAVE environment since, in terms of practical value, it only makes sense to evaluate the performance of the system with low traffic loads, and not under congestion. Regarding paragraph (d), to the best of our knowledge results here are the first ones published that relate to total frame delay. Other publications obtain at best results related to service time, defined as the time that a frame has to wait from its arrival at the head of the queue until it finishes being transmitted; however, that parameter only gives partial information about the system performance since queueing delays are overlooked. Finally, our work would have been incomplete without the feature described in (e), since the CCH works intermittently.

References

[1] IEEE Std 802.11-2007, "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," June 2007.

[2] J. R. Gallardo, S. C. Cruz, D. Makrakis, and A. Shami, "Analysis of the EDCA access mechanism for an IEEE 802.11e-compatible wireless LAN," in *Proceedings of the 13th IEEE Symposium on Computers and Communications (ISCC '08)*, pp. 891–898, Marrakech, Morocco, July 2008.

[3] J. R. Gallardo, D. Makrakis, and H. T. Mouftah, "Performance analysis of the EDCA medium access mechanism over the control channel of an IEEE 802.11p WAVE vehicular network," in *Proceedings of the IEEE International Conference on Communications (ICC '09)*, Dresden, Germany, June 2009.

[4] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[5] X. Chen, H. H. Refai, and X. Ma, "Saturation performance of IEEE 802.11 broadcast scheme in Ad hoc wireless LANs," in *Proceedings of the 66th Vehicular Technology Conference (VTC '07)*, pp. 1897–1901, Baltimore, Md, USA, September 2007.

[6] K. Bilstrup, E. Uhlemann, E. G. Ström, and U. Bilstrup, "Evaluation of the IEEE 802.11p MAC method for vehicle-to-vehicle communication," in *Proceedings of the 68th Semi-Annual IEEE Vehicular Technology (VTC '08)*, Calgary, Canada, September 2008.

[7] B. S. Gukhool and S. Cherkaoui, "IEEE 802.11p modeling in NS-2," in *Proceedings of the 33rd IEEE Conference on Local Computer Networks (LCN '08)*, pp. 622–626, Montreal, Canada, October 2008.

[8] Y. Wang, A. Ahmed, B. Krishnamachari, and K. Psounis, "IEEE 802.11p performance evaluation and protocol enhancement," in *Proceedings of the IEEE International Conference on Vehicular Electronics and Safety (ICVES '08)*, pp. 317–322, Columbus, Ohio, USA, September 2008.

[9] T. Murray, M. Cojocari, and F. Huirong, "Measuring the performance of IEEE 802.11p using ns-2 simulator for vehicular networks," in *Proceedings of IEEE International Conference on Electro/Information Technology (EIT '08)*, pp. 498–503, Ames, Iowa, USA, May 2008.

[10] A. Vinel, V. Vishnevsky, and Y. Koucheryavy, "A simple analytical model for the periodic broadcasting in vehicular ad-hoc networks," in *Proceedings of IEEE Globecom Workshops (GLOBECOM '08)*, New Orleans, La, USA, November-December 2008.

[11] A. Rao, A. Kherani, and A. Mahanti, "Performance evaluation of 802.11 broadcasts for a single cell network with unsaturated nodes," in *NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*, vol. 4982 of *Lecture Notes in Computer Science*, pp. 836–847, Springer, Berlin, Germany, 2009.

[12] S. Eichler, "Performance evaluation of the IEEE 802.11p WAVE communication standard," in *Proceedings of the 1st IEEE International Symposium on Wireless Vehicular Communications (WiVEC '07)*, Baltimore, Md, USA, September-October 2007.

[13] G. Bianchi and I. Tinnirello, "Analysis of priority mechanisms based on differentiated inter frame spacing in CSMA-CA," in *Proceedings of the 58th Vehicular Technology Conference (VTC '03)*, pp. 1401–1405, Orlando, Fla, USA, October 2003.

[14] J. Hui and M. Devetsikiotis, "A unified model for the performance analysis of IEEE 802.11e EDCA," *IEEE Transactions on Communications*, vol. 53, no. 9, pp. 1498–1510, 2005.

[15] X. Ling, K.-H. Liu, Y. Cheng, X. Shen, and J. W. Mark, "A novel performance model for distributed prioritized MAC protocols," in *Proceedings of the 50th Annual IEEE Global Telecommunications Conference (GLOBECOM '07)*, pp. 4692–4696, Washington, DC, USA, November 2007.

[16] D. J. Goodman and S. X. Wei, "Efficiency of packet reservation multiple access," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 1, part 2, pp. 170–176, 1991.

[17] C. M. Grinstead and J. L. Snell, *Introduction to Probability*, American Mathematical Society, Providence, RI, USA, 2nd edition, 1997.