Saturated Throughput Analysis of IEEE 802.11e Using Two-Dimensional Markov Chain Model

Lixiang Xiong and Guoqiang Mao
School of Electrical and Information Engineering
The University of Sydney
NSW 2006 Australia
Email: {xlx, guoqiang}@ee.usyd.edu.au

***Abstract*— IEEE 802.11e standard has been recently published to introduce quality of service (QoS) support to the conventional IEEE 802.11 wireless local area network (WLAN). Enhanced Distribution Channel Access (EDCA) is used as the fundamental access mechanism for the medium access control (MAC) layer in IEEE 802.11e. In this paper, a novel Markov chain model with a simple architecture for EDCA performance analysis under the saturated traffic load is proposed. Compared with existing analytical models, the proposed model considers more features of EDCA. Firstly, the effect of using different arbitration interframe spaces (AIFSs) is analyzed. Secondly, the possibility that a station's backoff procedure may be suspended due to transmission from other stations is considered. We consider that the contention zone specific transmission probability caused by using different AIFSs can affect the occurrence of the backoff suspension state. Based on the proposed model, saturated throughput of EDCA is obtained. Simulation study is performed, which demonstrates that the proposed model has better accuracy than other models.**

I. INTRODUCTION

IEEE 802.11 wireless local area network (WLAN) [1] has been widely used for high speed wireless Internet access. Unfortunately the IEEE 802.11 WLAN is based on the best-effort service model, and its access mechanism for the medium access control (MAC) layer, Distributed Coordination Function (DCF), can not satisfy the demand for better quality of service (QoS) support from multimedia applications. Thus IEEE 802.11e Enhanced Distribution Channel Access (EDCA) [2] is developed.

IEEE802.11e standard classifies traffic into four Access Categories (ACs), i.e., voice, video, best effort and background. AC based traffic prioritization is implemented by using a combination of AC specific parameters, which include arbitration interframe space (AIFS), the minimum contention window (CW_{min}), the maximum contention window (CW_{max}) and transmission opportunity (TXOP) limit. The working procedure of EDCA can be found in [2, clause 9], and it has been discussed in many related publications [3]–[15]. Thus, we will not describe the complete EDCA procedure in this paper. Some details of EDCA should be noted because they are important for our later analysis:

- Every time the channel returns idle from a transmission, the station must sense the channel idle for a complete AIFS or EIFS (extended interframe space) from the end of the last busy channel to start or resume its routine backoff procedure. The selection of AIFS or EIFS depends on the type of the last channel busy event. If the last channel busy event is an unsuccessful transmission (e.g., a collision), the station will wait an EIFS, otherwise it must wait an AIFS. If a transmission from other stations occurs before the completion of AIFS/EIFS, the station must wait another complete AIFS/EIFS after the channel returns idle. So long as the channel is not idle for a compete AIFS/EIFS, the station keeps suspending its backoff procedure. To avoid any confusion, in this paper, we use the term "backoff suspension" to represent the process that a station must sense the channel idle for a complete AIFS/EIFS before it can start a new backoff procedure or resume the suspended backoff procedure. Moreover, we use the term "backoff slot" to represent a time slot in which at least one station can possibly complete its backoff procedure and start a transmission.
- In the case that a collision occurs, colliding stations (i.e., stations involved in the collision) need to wait a further ACK timeout duration to detect the collision, and then they will wait an AIFS before starting another backoff procedure. The sum of the ACK timeout duration and an AIFS is equal to an EIFS [16]. Non-colliding stations (i.e., stations not involved in the collision) will wait an EIFS after a collision [17, clause 9.2.5.2, pp.77-79].
- A station decrements its backoff counter by one at the beginning of each backoff slot during the backoff procedure. This means that the backoff counter decrement decision is made at the end of the previous idle backoff slot, independent of whether the channel is busy or not in the current backoff slot. Furthermore, every time the station leaves the backoff suspension state after completing an AIFS/EIFS, its non-zero backoff counter will be decreased by one at the beginning of the immediately following backoff slot, and this decrement is independent of the channel status in the backoff slot [2, clause 9.9.1.3, pp.81-83], [18].
- When the backoff counter is decreased to zero at the beginning of a backoff slot, the station will start its transmission at the beginning of the next backoff slot provided that there is no transmission from other stations in the current backoff slot. Otherwise the station will enter into a backoff suspension state to wait a complete idle

AIFS/EIFS and start its transmission at the beginning of the immediately following backoff slot [2, clause 9.9.1.3, pp.81-83], [18].

To investigate the performance of EDCA, an accurate analytical model is necessary. In addition to the effect of different CW ranges that has been well investigated, we consider some important factors that should be carefully handled for accurately analyzing EDCA.

Firstly, the effect of using different AIFSs should be considered. In this paper, we assume that each station carries traffic from an AC only for the sake of simplify, thus stations can be classified into different sets based on their AIFS values. The stations in the same set have the same AIFS value. Different sets of stations will wait different AIFSs (or related EIFSs) in the backoff suspension state before they may access the channel. Thus stations with smaller AIFS may access the channel while other stations with larger AIFS are still in the backoff suspension state. We consider the time period from the end of the last channel busy event can be classified into different intervals, referred to as contention zones. Stations will have different transmission probability probability in each zone.

Secondly, the possibility of backoff suspension should be analyzed. As mentioned earlier, before the start of a new backoff procedure and every time the channel turns busy during the backoff procedure, the station may enter into a backoff suspension state. We consider that the occurrence of backoff suspension is uncertain since it depends on the channel status affected by the activities of other stations. The occurrence of backoff suspension can affect the performance of EDCA.

In this paper, we propose a novel Markov chain model for EDCA performance analysis. The proposed model has a simple architecture and considers more features of EDCA. Both the effects of using different AIFSs and backoff suspension are considered, which gives a more accurate analysis.

The rest of this paper is organized as follows: section II gives a brief introduction to some related publications; section III introduces the proposed analytical model; saturated throughput of EDCA is analyzed in section IV; simulation study is performed in section V; finally section VI concludes the paper.

II. Related Work

Some analytical models for EDCA have been proposed in the literature [3]–[15]. Most of them use Markov chain approach [5]–[15]. However, there are some researchers trying to obtain a closed-form expression for collision probability and saturation throughput using elementary probability theory directly [3], [4]. We refer to the approach in [3], [4] as non-Markov approach. The major problem with the non-Markov approach is in order to obtain a closed-form solution using elementary probability theory, significant simplification and approximation have to be made, thus they can not fully capture the complexity of EDCA, including the effects of using different AIFSs and backoff suspension. For example, the possibility that the backoff procedure of lower priority stations may be consecutively interrupted by transmission from higher priority stations is not considered in [4], and the effect of using different AIFSs is ignored in [3].

Compared with the non-Markov approach, the Markov-chain approach has a disadvantage that a closed-form solution is difficult to obtain. However, a well designed Markov chain model can capture the complexity of EDCA more easily. Using Markov chain to analyze EDCA performance was originally started by a Markov chain model developed by Bianchi for analyzing legacy DCF [19]. In [19], two stochastic processes are used to construct a two-dimensional Markov chain model for modeling DCF. One stochastic process is used to represent the backoff counter, the other is used to represent the number of consecutive retransmissions. A similar approach is used in many papers for EDCA modeling [5]–[15]. However we consider that some limitations exist among them.

In [5]–[10], some Markov chain models are developed based on that in [19], such as the post-collision analysis presented in [5], which considers the effect of using different AIFSs, the delay analysis in [7], and the Z-transform approach in [8]. But a common problem exists among them: the possibility of backoff suspension is ignored.

Compared with those in [5]–[10], models presented in [11]–[13] consider backoff suspension. In [11], [12], the backoff suspension is modeled by adding a transition for each state, which starts and ends at the same state. It represents that the backoff procedure is suspended in the corresponding backoff slot. In [13], the backoff suspension is modeled by using some extra states to represent the possible backoff suspension that may occur in the corresponding backoff state. However, some flaws exist among them. Firstly, the difference between backoff suspension state and backoff state is not considered in [11], [12]. In the backoff suspension state, a station must wait a complete idle AIFS/EIFS before it can decrease its backoff counter and move to the next state; while in the backoff state, a station only needs to wait an idle backoff slot in order to decrement its backoff counter and move to next state. Secondly, the mandatory backoff suspension state before the start of a new backoff procedure is not considered in [11], [12]. Finally, the contention zone specific transmission probability caused by using different AIFSs is not included in [13].

Some Markov chain models consider both the effect of different AIFS and the effect of backoff suspension [14], [15]. In [14], a three-dimensional Markov chain model is used for the lower priority traffic with a larger AIFS, where the third dimension is a stochastic process representing the possible backoff suspension. In [15], an extra stochastic process is used in its 3-dimensional Markov chain model to represent the elapsed backoff slots since the end of a transmission. The third dimension used in [14], [15] leads to some extra states, which are used to represent the possible backoff suspension, and the effect of using different AIFSs is included when analyzing the transition probabilities between those states. However, some limitations exist in [14], [15] in addition to that a complex Markov chain architecture is used. Firstly, it is assumed in

[15] that a station will keep retransmitting until the frame has been successfully transmitted. The possibility that it may be dropped after reaching the maximum retransmission limit is not included. Secondly, the Markov model for high priority traffic in [14] does not consider the possibility that the backoff procedure of a station with high priority traffic may also be suspended by transmission from other stations.

Furthermore, some details of EDCA, such as the backoff counter is decremented in advance at the beginning of a backoff slot and the exact time point when a transmission is started after the backoff counter reaches zero, are simply ignored or not correctly analyzed in the Markov chain models in [5]–[9], [11], [12], [14], [15]. This may be caused by the fact that IEEE 802.11e standard was not finished yet when those papers were published. A Markov chain model that carefully considers these effects will improve the accuracy of EDCA performance analysis.

III. A MARKOV CHAIN MODEL FOR EDCA PERFORMANCE ANALYSIS

In this section, we shall present the proposed analytical model of EDCA. Firstly, the basic Markov chain model is proposed. Secondly, the transition probabilities for the proposed Markov chain model are analyzed, where the effect of the contention zone specific transmission probability differentiation caused by using different AIFSs is considered. Finally, a solution for the proposed Markov chain model is obtained. In our analysis, the following assumptions are made.

- Traffic load is saturated. That is, traffic is always backlogged at each station.
- Only two ACs are considered: AC A and AC B. AC A has higher priority than AC B and $AIFS[A] < AIFS[B]$. However, our analysis can be readily extended to include more than two ACs.
- Each station carries traffic from one AC only. Thus a station may be referred to as a priority A station or a priority B station depending on the AC of the traffic it carries.
- Only one frame is transmitted in each TXOP.
- A WLAN with a fixed number of stations is considered. The number of stations for AC A and AC B is denoted by n_A and n_B respectively. n_A and n_B are known numbers.
- The transmission probability of a station in a generic backoff slot is a constant, which is determined by its AC only. This is an assumption widely used in the area [5], [8]–[14]. The transmission probabilities of a priority A station and a priority B station in a generic backoff slot are represented by τ_A and τ_B respectively. The values of τ_A and τ_B are unknown, which need to be determined.
- The probability that a transmission from a station experiences a collision within a contention zone is a constant, which is determined by its AC only.
- The radio channel is ideal. That is, there are no noise, no external interference and hidden station problems.

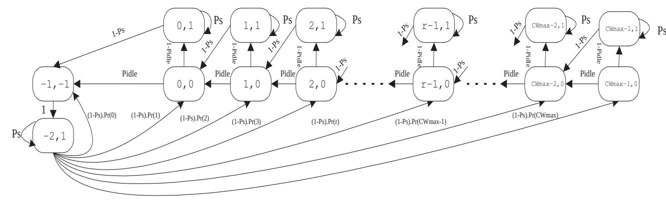


Fig. 1. The Markov chain model for modeling the backoff procedure for a station of a specific AC

A. A Discrete Time Two-Dimensional Markov Chain Model

1) The Basic Markov Chain Model: Fig. 1 illustrates the proposed discrete time two-dimensional Markov chain model. It models the channel contention procedure for a station of a specific AC. Two stochastic processes exist within the Markov chain model. The first process, denoted by $w(t)$, represents the value of the backoff counter during the backoff procedure of the station at time t. The second process, denoted by $v(t)$, indicates whether the station is in a backoff suspension state or not at time t. Unlike the Markov chain models in most related publications, the stochastic process representing the number of consecutive retransmissions is not used in our model and it is considered in the transition probabilities of the Markov chain model by weighting the probabilities of multiple consecutive retransmissions, as discussed later.

A logical time scale is used in the proposed Markov chain model, similar as that in [15]. In this logical time scale, between two consecutive logical time points, "t" and "t+1", either of two events may occur: (i) an idle backoff slot elapses; (ii) a transmission starts or ends. Here the term "logical" is used to denote that we ignore the amount of time used in a transmission and have the time period after a transmission slotted. A similar method has been widely used to construct the Markov chain model for EDCA performance analysis [5]–[15].

The two-dimensional Markov process $\{w(t),\ v(t)\}$ determines each state (r, z) in the Markov chain model, where r represents the value of the backoff counter with a range [0, $CW_{max} - 1$][1], and z denotes whether the station is located in either the backoff suspension state ($z = 1$) or its routine backoff procedure ($z = 0$). Here state (r, 1) represents backoff suspension state caused by transmission from other stations during the backoff procedure. Two special states are created: the first one, $(-2, 1)$, represents the backoff suspension state before the start of a new backoff procedure; the other one, $(-1, -1)$, represents the transmission procedure of the station. CW_{max} is the maximum contention window size of the AC under consideration, which is a known constant.

2) Transitions: The one-step transition probabilities for the Markov chain model are described as follows.

[1] $(CW_{max} - 1)$ instead of CW_{max} is used by considering the backoff counter decrement rule, as discussed earlier.

- If the channel is idle in a backoff slot, the system may move from state (r, 0) to state (r-1, 0) and the backoff counter is decreased by one:

$$P\{(r-1,0)|(r,0)\} = P_{idle}, \text{for } 1 \le r \le CW_{max}-1,$$

where P_{idle} is the probability that the channel is idle in a backoff slot. For the special case that r equals to zero, the station will start a transmission and enter into the state (-1,-1):

$$P\{(-1,-1)|(0,0)\} = P_{idle}.$$

- If the channel turns busy in a backoff slot due to transmission from other stations, the system will move to the backoff suspension state (r, 1) and wait a complete idle AIFS/EIFS interval. Meanwhile, the backoff counter is unchanged:

$$P\{(r,1)|(r,0)\} = 1-P_{idle}, \text{for } 0 \le r \le CW_{max}-1,$$

where $1-P_{idle}$ is the probability of channel being busy.
- If the channel becomes idle after the completion of transmission from other stations and remains idle for an AIFS/EIFS interval in the corresponding backoff suspension state, the backoff counter is decreased by one at the end of the backoff suspension as described in section I, and the system may move from the suspension state (r, 1) to state (r-1, 0):

$$P\{(r-1,0)|(r,1)\} = 1-P_s, \text{for } 1 \le r \le CW_{max}-1,$$

where P_s is the probability that there is at least one transmission during the AIFS/EIFS interval in the corresponding backoff suspension state, so that the station can not leave the backoff suspension state. For the special case that r equals to zero, the station will leave the corresponding backoff suspension state and enter into the state (-1, -1) to start a transmission:

$$P\{(-1,-1)|(0,1)\} = 1-P_s.$$

- If at least one transmission from other stations occurs before the completion of an AIFS/EIFS interval, the system will remain in the backoff suspension state (r, 1):

$$P\{(r,1)|(r,1)\} = P_s, \text{for } 0 \le r \le CW_{max}-1.$$

- When the system finally reaches the state (-1, -1), the station will start a transmission. After the transmission, it will start another backoff procedure for the next transmission which can be a retransmission in the case that the previous transmission encounters a collision, or a new transmission. The new backoff procedure starts with the backoff suspension state (-2,1):

$$P\{(-2,1)|(-1,-1)\} = 1.$$

- If transmission from other stations occurs before the completion of an AIFS/EIFS interval in the backoff suspension state (-2,1), the system must remain in this state:

$$P\{(-2,1)|(-2,1)\} = P_s.$$

- If the channel remains idle for a complete AIFS/EIFS interval in the backoff suspension state (-2,1), the station may start a new backoff procedure with an initial backoff counter r. As described in section I, the backoff counter will be decremented by one at the end of the backoff suspension state, and the system will move directly to state (r-1,0):

$$P\{(r-1,0)|(-2,1)\} = (1-P_s)Pr(r), \quad \text{for } 1 \le r \le CW_{max},$$

where $Pr(r)$ is the probability that the station starts a new backoff procedure with a random initial backoff counter r. For the special case that the initial backoff counter is zero, the station may start a transmission after the backoff suspension state is completed:

$$P\{(-1,-1)|(-2,1)\} = (1-P_s)Pr(0).$$

All the aforementioned transition equations related parameters are AC specific and they will be analyzed later.

3) System Equations: Let $b_{(r,z)}$ be the steady probability of state (r, z) in the Markov chain model. The following relations can be obtained due to the regularity of the Markov chain:

$$\begin{cases} b_{(-2,1)} = b_{(-1,-1)}/(1-P_S), \\ b_{(CW_{max}-1,0)} = b_{(-2,1)}(1-P_s)Pr(CW_{max}), \\ b_{(r,1)} = b_{(r,0)}/(1-P_s), \text{for } 0 \le r \le CW_{max}-1, \\ \\ b_{(r,0)} = b_{(-2,1)}Pr(r+1)/(1-P_s) \\ \quad\quad + b_{(r+1,0)}P_{idle} + b_{(r+1,1)}(1-P_s), \\ \text{for } 1 \le r \le CW_{max}-2, \end{cases} \tag{1}$$

and

$$\sum_{r=0}^{CW_{max}-1} b_{(r,0)} + \sum_{r=0}^{CW_{max}-1} b_{(r,1)} + b_{(-1,-1)} + b_{(-2,1)} = 1. \tag{2}$$

Since the state (-1, -1) represents the transmission procedure of the station, the corresponding steady state probability $b_{(-1,-1)}$ should be equal to the transmission probability τ:

$$b_{(-1,-1)} = \tau, \tag{3}$$

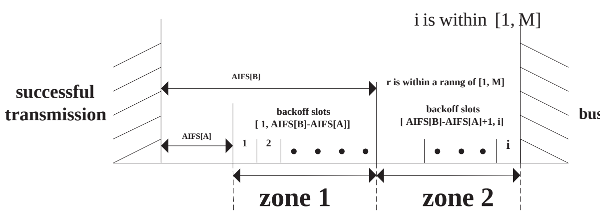
where τ is an unknown AC specific constant to be solved.

B. Unknown Parameters in Transition Equations

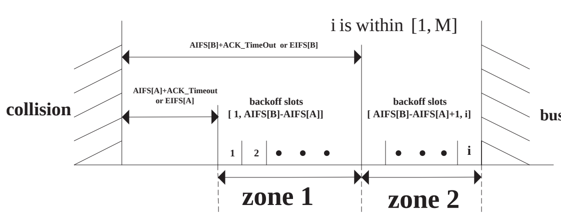
In this section, the unknown parameters in the transition equations shown in the last section are analyzed. It is organized as follows. Firstly, a new Markov chain model is used for analyzing the effect of the contention zone specific transmission probability caused by using different AIFSs, which is also used in [5]. Secondly, using the new Markov chain model, the average collision probability p and the transition probability P_{idle} that the channel remains idle are obtained. Thirdly, the transition probability P_s that the station remains in the backoff suspension state is obtained. Finally, the transition probability $Pr(r)$ is analyzed by creating a new Markov chain model.

1) A Markov Chain Model for Analyzing the Effect of the Contention Zone Specific Transmission Probability: Fig. 2 depicts the number of consecutive backoff slots between two successive transmissions in the WLAN. As shown in Fig. 2, no station can transmit during the first AIFS[A]/EIFS[A] time interval from the end of the busy channel. During the backoff slots in the range of [1, AIFS[B]-AIFS[A]] after AIFS[A]/EIFS[A], referred to as zone 1, priority A stations which have completed their AIFS[A]/EIFS[A] may begin their backoff procedure and transmit, while priority B stations are still waiting for the completion of their AIFS[B]/EIFS[B] and can not transmit. During the backoff slots in the range of [AIFS[B]-AIFS[A]+1, i], referred to as zone 2, priority B stations also begin their backoff procedure and may transmit by contending with priority A stations. Here i is bounded by M, which is the maximum number of possible consecutive back slots between two successive transmissions in the WLAN:

$$M = min(CW_{maxA}, AIFS[B] - AIFS[A] + CW_{maxB}).$$



(a) after a successful transmission



(b) after a collision

Fig. 2. Backoff slot distribution between two successive transmissions in the system

From Fig. 2, a new discrete time one-dimensional Markov chain model can be created, which is shown in Fig. 3. The stochastic process in this Markov chain model represents the number of consecutive backoff slots between two successive transmissions in the WLAN. The state (i) in the Markov chain model represents i consecutive backoff slots from the end of the previous transmission in the WLAN.

This Markov chain is described by its one-step transition probabilities as follows:

- In zone 1, if any transmission from priority A stations occurs while the system is in state (i), the system will

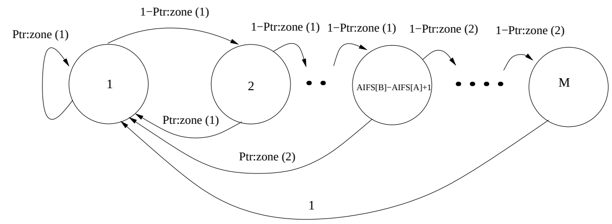


Fig. 3. The Markov chain model for modeling the number of consecutive backoff slots between two successive transmissions in the WLAN

move from state (i) to state (1):

$$P\{(1)|(i)\} = P_{tr:zone(1)}, \quad \text{for } 1 \le i \le (AIFS[B] - AIFS[A]),$$

where $P_{tr:zone(1)}$ is the probability that at least one transmission from priority A stations occurs in a backoff slot in zone 1, and it equals to $(1 - (1 - \tau_A)^{n_A})$.

- If no transmission occurs the system will move from state (i) to state (i+1) with a probability $1 - P_{tr:zone(1)}$:

$$P\{(i+1)|(i)\} = 1 - P_{tr:zone(1)}, \quad \text{for } 1 \le i \le AIFS[B] - AIFS[A].$$

- In zone 2, both priority A stations and priority B stations begin their backoff procedure and may transmit. A transmission from either priority A or priority B stations can cause the system to return to state (1):

$$P\{(1)|(i)\} = P_{tr:zone(2)}, \quad \text{for } AIFS[B] - AIFS[A] + 1 \le i \le M,$$

where $P_{tr:zone(2)}$ is the probability that there is at least one transmission in a backoff slot in zone 2, and it equals to $(1 - (1 - \tau_A)^{n_A}(1 - \tau_B)^{n_B})$.

- If no transmission occurs the system will move from state (r) to state (r+1) with a probability $1 - P_{tr:zone(2)}$:

$$P\{(i+1)|(i)\} = 1 - P_{tr:zone(2)}, \quad \text{for } AIFS[B] - AIFS[A] + 1 \le i \le \ M - 1.$$

- When the system reaches the last state (M), a transmission will definitely occur. Thus the system will return to state (1) with a probability 1:

$$P\{(1)|(M)\} = 1.$$

Using above transition equations and considering the fact that the sum of the steady state probabilities of the Markov chain equals to 1, the steady state probability $s_{(i)}$ can be solved, which is given in (4) and (5).

2) Average Collision Probability p and Transition Probability P_{idle} that the Channel Remains Idle: The zone specific transmission probability caused by using different AIFSs is considered by using an average transmission probability in our analysis related to the proposed Markov chain model shown in Fig. 1. The average transmission probability is obtained by

$$s_{(1)} = \{\frac{1-(1-P_{tr:zone(1)})^{AIFS[B]-AIFS[A]}}{P_{tr:zone(1)}} + (1-P_{tr:zone(1)})^{AIFS[B]-AIFS[A]}\frac{1-(1-P_{tr:zone(2)})^{M-AIFS[B]+AIFS[A]}}{P_{tr:zone(2)}}\}^{-1}. \quad (4)$$

$$\begin{cases} s_{(i)} = (1-P_{tr:zone(1)})^{i-1}s_{(1)}, \text{for } 2 \leq i \leq AIFS[B]-AIFS[A]+1, \\ s_{(i)} = (1-P_{tr:zone(2)})^{r-AIFS[B]+AIFS[A]-1}s_{(1)}(1-P_{tr:zone(1)})^{AIFS[B]-AIFS[A]}, \\ \text{for } AIFS[B]-AIFS[A]+2 \leq i \leq M). \end{cases} \quad (5)$$

weighting the transmission probabilities in different contention zones.

For a transmission started by a station in a backoff slot, collision may occur if one or more other stations start a transmission in the same backoff slot. The corresponding collision probability is determined by the composition of contending stations. In zone 1, only priority A stations can transmit and cause collisions. In zone 2, both priority A stations and priority B stations can transmit and collide with each other. Thus the collision probability for priority A stations should be contention zone specific, which can be obtained by

$$\begin{cases} p_{A:zone(1)} = 1-(1-\tau_A)^{n_A-1}, \\ p_{A:zone(2)} = 1-(1-\tau_A)^{n_A-1}(1-\tau_B)^{n_B}, \end{cases}$$

For a priority A station in the backoff counter count-down procedure, it sees an "idle" backoff slot when no other stations start a transmission in the same backoff slot. Considering the contention zone specific transmission probability, the contention zone specific probability that a priority A station sees an idle backoff slot can be obtained by

$$\begin{cases} P_{idleA:zone(1)} = (1-\tau_A)^{n_A-1}, \\ P_{idleA:zone(2)} = (1-\tau_A)^{n_A-1}(1-\tau_B)^{n_B}. \end{cases}$$

Thus, the average collision probability for a priority A station can be obtained as the sum of the weighted contention zone specific collision probability:

$$p_A = \sum_{i=1}^{M} s_{(i)}p_{A:zone_i}, \quad (6)$$

where $p_{A:zone_i}$ is the contention zone specific collision probability in the i^{th} backoff slot. Depending on whether the i^{th} slot belongs to zone 1 or zone 2, $p_{A:zone(1)}$ or $p_{A:zone(2)}$ should be used. $s_{(i)}$ is the steady state probability, which is obtained from (4) and (5).

Similarly, the average probability P_{idleA} that a priority A station in the backoff procedure sees an idle backoff slot can be obtained by

$$P_{idleA} = \sum_{i=1}^{M} s_{(i)}P_{idleA:zone_i}, \quad (7)$$

where $P_{idleA:zone_i}$ is the contention zone specific probability for a priority A station in the i^{th} backoff slot. Depending on whether the i^{th} slot belongs to zone 1 or zone 2, $P_{idleA:zone(1)}$ or $P_{idleA:zone(2)}$ should be used.

For a priority B station, all of its backoff slots are located in zone 2, where all stations may transmit. Thus its average collision probability can be simply obtained by

$$p_B = 1-(1-\tau_A)^{n_A}(1-\tau_B)^{n_B-1}, \quad (8)$$

and so is the average probability that a priority B station has an idle backoff slot:

$$P_{idleB} = (1-\tau_A)^{n_A}(1-\tau_B)^{n_B-1}. \quad (9)$$

3) The Transition Probability P_s of Remaining in the Backoff Suspension State : A station suspending its backoff procedure may leave the backoff suspension state if the channel remains idle for a complete AIFS/EIFS from the end of the last channel busy event. Any transmission from other stations during this time interval can stop the station from leaving the backoff suspension state.

For a priority A station, the required time interval for leaving the backoff suspension state is a complete idle AIFS[A]/EIFS[A] interval. As illustrated in Fig. 2, no transmission is possible in this time period. Thus, the probability P_s for a priority A station remaining in the backoff suspension state is zero:

$$P_{sA} = 0. \quad (10)$$

For a priority B station, the required time interval for leaving the suspension state is a complete idle AIFS[B]/EIFS[B] interval. According to Fig. 2, the backoff slots in zone 1 are part of AIFS[B]/EIFS[B], where transmission from priority A stations is possible. Thus, the probability Ps for a priority B station remaining in the suspension state can be obtained as

$$P_{sB} = 1-((1-\tau_A)^{n_A})^{AIFS[B]-AIFS[A]}. \quad (11)$$

4) Transition Probability $Pr(r)$: The backoff counter is drawn randomly from the range [0, CW] and the CW value is determined by the AC specific CW_{min} and CW_{max} values as well as the number of previous consecutive retransmissions. Therefore the probability of obtaining a specific backoff counter value r is related to the number of previous consecutive retransmissions. The Markov chain model shown in Fig. 1 does not explicitly consider the effect consecutive retransmissions.

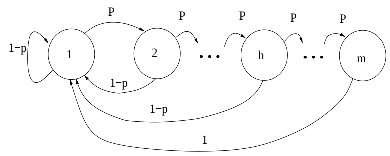


Fig. 4. The Markov chain model for modeling the number of the consecutive retransmissions of a station

Instead, its effect is considered in the probability $Pr(r)$ of obtaining a specific backoff counter r by weighting the probability of the number of consecutive retransmissions.

For obtaining $Pr(r)$, a discrete time one-dimensional Markov chain model is created, as shown in Fig. 4. The stochastic process in this Markov chain model represents the number of consecutive retransmissions (including the first transmission of the frame) for a station at time t. Thus state (k) represents that the station is performing the k^{th} consecutive retransmission. In this Markov chain, state (h) represents the h^{th} consecutive retransmission in which the CW value reaches CW_{max} for the first time, and state (m) represents the m^{th} consecutive retransmission, which is the maximum retransmission limit. Both h and m are constants determined by the WLAN standard.

The activity of the Markov chain shown in Fig. 4 is governed by its one-step transition probabilities as follows:

- If the k^{th} retransmission is unsuccessful, the system will move from state (k) to state (k+1) with a probability p:
$$P\{(k+1)|(k)\} = p, \text{for } 1 \leq k \leq m-1,$$
where p is the AC specific average collision probability, which can be obtained from (6) or (8).
- If the k^{th} retransmission is successful, the system will move from state (k) to state (1) with a probability $1-p$ and the station will start transmitting a new frame:
$$P\{(1)|(k)\} = 1-p, \text{for } 1 \leq k \leq m.$$
- when the maximum retransmission limit m is reached, the station will begin the first transmission of a new frame no matter whether the m^{th} consecutive retransmission is successful or not. Thus the system will return to state (1) with a probability 1:
$$P\{(1)|(m)\} = 1.$$

Using above transition equations and considering the fact that the sum of the steady state probabilities of the Markov chain equals to 1, the steady state probability $d_{(k)}$ can be obtained:

$$d_{(k)} = p^{k-1}(1-p)/(1-p^m), \text{ for } 1 \leq k \leq m.$$

Since the backoff counter is a random integer uniformly distributed in the range $[0, CW]$, the probability of obtaining a specific backoff counter value from this range should be $\frac{1}{1+CW}$. Thus, the AC specific probability $Pr(r)$ of obtaining a specific backoff counter r can be obtained as the sum of the probability of obtaining a specific initial backoff counter r in the t^{th} consecutive retransmission weighted with the probability of the occurrence of the k^{th} consecutive retransmission:

$$Pr(r) = \sum_{k=1}^{m} \frac{d_{(k)}c_{(r)}}{CW(k)+1},$$

where CW(k) is the corresponding CW value in the k^{th} consecutive retransmission; and $c_{(r)}$ indicates whether the specific value r is included in the corresponding range [0, CW] or not (if yes, $c_{(r)}$ is 1, otherwise it is zero).

Based on the earlier analysis, an expression for the AC specific probability $Pr(r)$ can be obtained:

$$\Pr(r) = \begin{cases} \sum_{k=1}^{h-1} \frac{d_{(k)}}{2^{k-1}CW_{min}+1} + \sum_{k=h}^{m} \frac{d_{(k)}}{CW_{max}+1}, \\ \text{for } 0 \leq r \leq CW_{min}, \\ \\ \sum_{k=j}^{h-1} \frac{d_{(k)}}{2^{k}CW_{min}+1} + \sum_{k=h}^{m} \frac{d_{(k)}}{CW_{max}+1}, \\ \text{for } 2^{j-1}CW_{min}+1 \leq r \leq 2^{j}CW_{min}, \\ \quad \text{and } 1 \leq j \leq h-1, \\ \\ \sum_{k=h}^{m} \frac{d_{(k)}}{CW_{max}+1}, \\ \text{for } 2^{h-1}CW_{min}+1 \leq r \leq CW_{max}, \end{cases} \tag{12}$$

where CW_{min} and CW_{max} are AC specific and known.

C. Summary

Finally, this section summarizes the relationship of earlier analysis.

In section III-A, a novel Markov chain model is created for each AC in the WLAN, which is shown in Fig. 1. The system equations (1) and (2), and (3) for the Markov chain model are obtained. The unknown AC specific transition probabilities for the Markov chain model are analyzed in section III-B, including (6)-(12). By using these equations, two non-linear equations about τ_A and τ_B can be constructed, and the values of τ_A and τ_B can be obtained from the equations.

IV. SATURATED THROUGHPUT ANALYSIS FOR EDCA

In this section, we shall use the earlier model to analyze the saturated throughput of EDCA. We consider that the throughput is equal to the ratio of the effective payload to the time required for transferring the effective payload. The Markov chain model shown in Fig. 3 is used to obtain the throughput. This Markov chain model represents the number of consecutive backoff slots between two successive transmissions in the WLAN. Two possible events may occur in a backoff slot: (i) at least one transmission occurs in the backoff slot, with a probability of $P_{tr:zone(1)}$ or $P_{tr:zone(2)}$ respectively; (ii) no transmission occurs in the backoff slot with a probability of $(1-P_{tr:zone(1)})$ or $(1-P_{tr:zone(2)})$ respectively. For the first possibility that at least one transmission occurs, it can be furthermore classified into two possibilities.

At first, it may be a successful transmission from either a priority A station or a priority B station. The corresponding contention zone probability for a successful transmission can be obtained by

$$\begin{cases} P_{sucA:zone(1)} = n_A\tau_A(1-\tau_A)^{n_A-1}, \\ P_{sucA:zone(2)} = \\ \qquad n_A\tau_A(1-\tau_A)^{n_A-1}(1-\tau_B)^{n_B}, \\ P_{sucB:zone(1)} = 0, \\ P_{sucB:zone(2)} = \\ \qquad n_B\tau_B(1-\tau_B)^{n_B-1}(1-\tau_A)^{n_A}. \end{cases}$$

Secondly, it may be a collision. That is, two or more stations start transmitting in the same backoff slot. The corresponding contention zone specific collision probability can be obtained by

$$\begin{cases} P_{col:zone(1)} = P_{tr:zone(1)} \\ \qquad - P_{sucA:zone(1)} - P_{sucB:zone(1)}, \\ P_{col:zone(2)} = P_{tr:zone(2)} \\ \qquad - P_{sucA:zone(2)} - P_{sucB:zone(2)}. \end{cases}$$

Therefore, the average effective payload for priority A stations can be obtained as:

$$E[A] = \sum_{i=1}^{M} P_{sucA:zone(i)} s_{(i)} E[P],$$

where $E[P]$ is the payload size of a frame, and $s_{(i)}$ can be obtained from (4) and (5). $E[P]$ is considered as a known constant. The effective payload for priority A station measures the effective amount of priority A traffic that is transmitted between two successive transmissions.

Similarly, the average effective payload for priority B stations can be obtained by

$$E[B] = \sum_{i=1}^{M} P_{sucB:zone(r)} s_{(i)} E[P].$$

The average time duration between two successive transmission can be obtained by

$$EL = \sum_{i=1}^{M} s_{(i)} \{ (P_{sucB:zone(i)}$$
$$+ P_{sucA:zone(i)})Ts + P_{col:zone(i)}Tc + P_{idle:zone(i)}aTimeSlot\},$$

where Ts and Tc are time required for a successful transmission and a collision respectively. They can be obtained by

$$Ts = H + P + SIFS + ACK + AIFS_{min},$$

and

$$Tc = H + P + EIFS_{min},$$

where H is the time required for transmitting the physical layer header and the MAC layer header of a frame, P is the time required for transmitting the data payload of a frame, ACK is the duration for transmitting an ACK frame, $AIFS_{min}$ is the minimum AIFS used in the WLAN, and $EIFS_{min}$ equals to $(SIFS + ACK + AIFS_{min})$.

Finally, the throughput for each station of each AC can be obtained by

$$\begin{cases} Throughput_A = E[A]/EL/n_A, \\ Throughput_B = E[B]/EL/n_B. \end{cases}$$

V. Simulation Study

In this section, theoretical analysis presented in the earlier sections is validated using simulation. Simulation is conducted using OPNET [20].

The parameters used in the simulation are shown in Table I. Four ACs are used in the simulation and their parameters are consistent with those defined in [2, Table 20df, p.49]. Two scenarios are simulated. In the first scenario, two ACs, i.e., voice and video, are used. This scenario is designed to investigate the effect of using different CW sizes since a common AIFS but different CW sizes are used by AC[voice] and AC[video] respectively. In the second scenario, two ACs, i.e., best effort and background, are used. The purpose of this scenario is to investigate the effect of using different AIFSs, since a common CW size but different AIFSs are used by AC[best effort] and AC[background] respectively. In both scenarios, there are equal number of stations in each AC.

TABLE I

WLAN parameter setting

Frame payload size	8000 bits
data rate	1Mbp/s
Payload data rate	1Mbp/s
Time slot	20 μs
SIFS	10 μs
Maximum retransmission limit	7
AIFSN	voice and video:2, best effort:3, background:7
CW_{min}	voice:3, video:15 best effort and background:15
CW_{max}	voice:7, video:31 best effort and background:1023

Fig. 5 shows the simulation results and theoretical results obtained from the proposed model for the first scenario. The throughput of a station in a specific AC under different number of stations is shown. It is shown in the figure that theoretical results obtained from the proposed model generally agree very well with simulation results, especially when the number of stations is large. However, when the number of stations is small, there is obvious discrepancy between theoretical results and simulations results. This discrepancy is attributable to the assumption made in the analysis that the transmission probability of a station in a generic backoff slot is a constant. As pointed out in [19], this assumption is more accurate when the number of stations is larger. As shown in the figure, by using different CW_{min} and CW_{max}, traffic is successfully classified into two different classes. Traffic with a smaller CW_{min} and CW_{max} can have better quality of service. When the number of stations in each AC is small, the difference in throughput for each AC is significant. When the number of stations in each AC increases, the difference in throughput decreases and throughput of both ACs decreases significantly because of more stations contending for bandwidth.

Fig. 6 shows the simulation results and theoretical results obtained from the proposed model for the second scenario. Again theoretical results obtained from the proposed model accurately matches the simulation results, especially when the

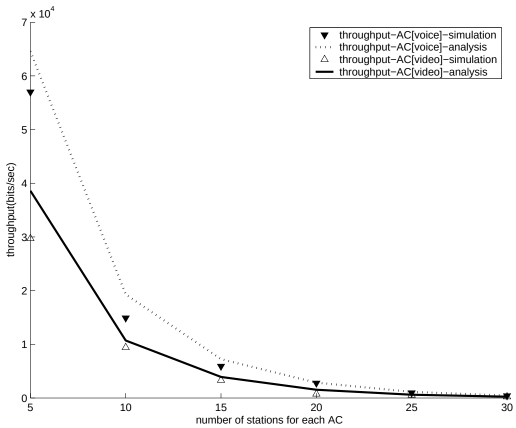


Fig. 5. Simulation and analysis results for AC[voice] and AC[video]

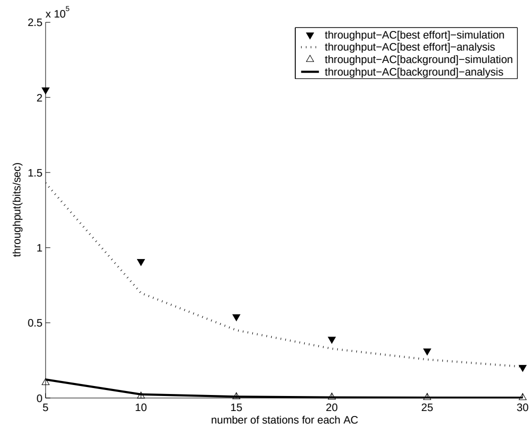


Fig. 6. Simulation and analysis results for AC[best effort] and AC[background]

number of stations is large. As shown in the figure, by using different $AIFS$s, traffic is successfully classified into two different classes, and this difference appears more obvious than that in the first scenario. Traffic with a smaller $AIFS$ can have better quality of service. It should be noticed that when the number of stations in each AC increases, the lower priority traffic belonging to the background AC may be starved.

The effect of using different AIFSs and CW sizes on traffic prioritization observed in the simulation results as well as theoretical results can be readily explained. Use of different AIFSs introduces the contention zone specific transmission probability. Lower priority station may be excluded for transmission in some contention zone, which causes some higher priority stations monopolize transmission opportunities and bandwidth. However, use of different CW sizes will only result in longer delay for lower priority stations and lower priority stations can still get the opportunity to transmit. Moreover, as shown in Fig. 5, when the number of voice and video stations increases, the throughput of both ACs drops severely. The reason is both AC[voice] and AC[video] have small AIFS and CW values. This enables stations to have a high transmission probability at a backoff slot, and accordingly their transmission will suffer a high collision probability when the number of stations is large. Therefore the majority of the available bandwidth is wasted on collision instead of successful transmission.

Finally, the results obtained in this paper is compared with those in [13]. Considering that multiple ACs in one station are used in [13], we slightly revise its analytical model (more specific, we revise equations (8) and (9) in [13]) so that it is consistent with the single AC in one station in our proposed model. The comparison is shown in Fig.7 and Fig.8. As shown in Fig.7 and Fig.8, the proposed model can generally achieve more accuracy than that in [13]. This result is expected as the proposed model in this paper incorporates more features of EDCA into analysis that that in [13]. The zone specific transmission probability caused by using different AIFSs is not considered in [13], where Kong $et\ al.$ consider that time slots within each AIFS/EIFS interval suffer the interruption caused by transmission from other stations at a same probability.

VI. CONCLUSION

In this paper, a novel Markov chain model for EDCA performance analysis under the saturated traffic load was proposed. Compared with the existing analytical models of EDCA, the proposed model incorporated more features of EDCA into the analysis and eliminated their limitations. Both the effects of using different AIFSs and the backoff suspension caused by transmission from other stations are considered. Based on the proposed model, saturated throughput of EDCA was analyzed. Simulation study using OPNET was performed, which demonstrated that theoretical results obtained from the proposed model can closely match simulation results, and the proposed model has better accuracy than that in the literature.

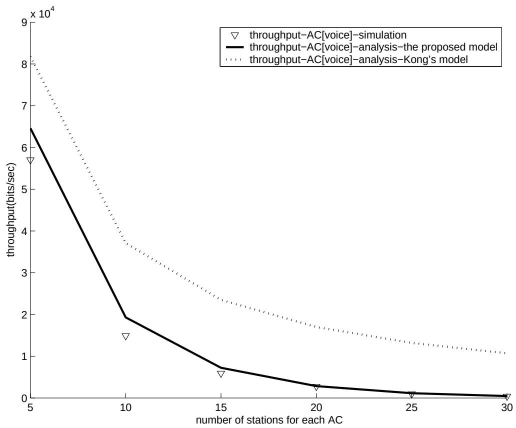
Despite the improvement, the analysis presented in this paper was based on the saturated throughput assumption. In a real network, traffic from a station is more likely to be non-saturated. Therefore a more interesting scenario will be throughput in non-saturated conditions. Moreover, wireless channel is characterized by the relatively higher bit error rate due to noise and interference. The effect of noise on EDCA performance should also be considered. These problems shall be addressed in our future research. These problems shall be addressed in our future research.

REFERENCES

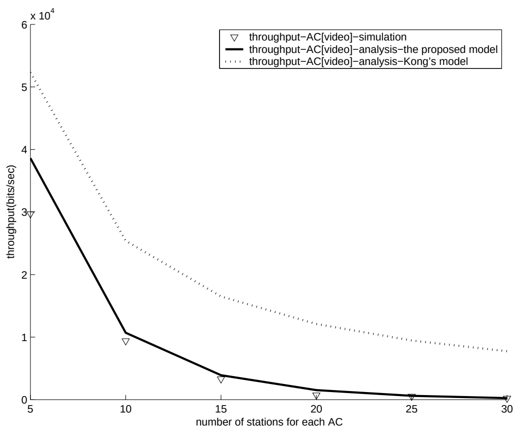
[1] "Information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements. Part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications," 1999.

[2] "IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements," 2005.

[3] Y. Chen, Q.-A. Zeng, and D. Agrawal, "Performance analysis of IEEE 802.11e enhanced distributed coordination function," in *Networks, 2003. ICON2003. The 11th IEEE International Conference on*, 2003, pp. 573–578.

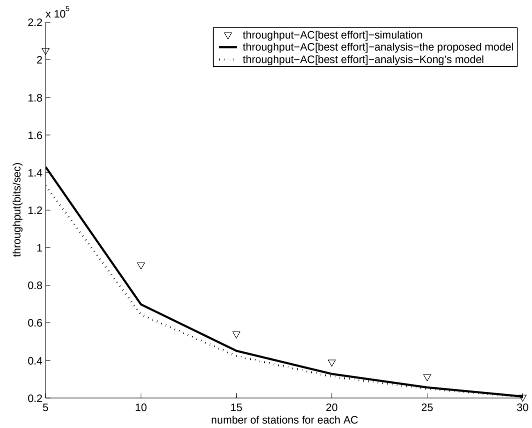


(a) vocie

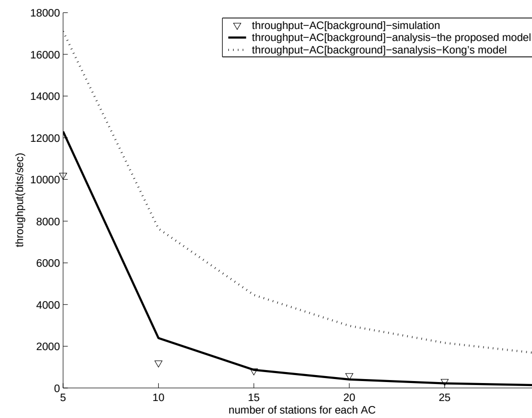


(b) video

Fig. 7. Comparison results for AC[voice] and AC[video]



(a) best effort



(b) background

Fig. 8. Comparison results for AC[best effort] and AC[background]

[4] Y.-L. Kuo, C.-H. Lu, E. H.-K. Wu, G.-H. Chen, and Y.-H. Tseng, "Performance analysis of the enhanced distributed coordination function in the IEEE 802.11e," in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, vol. 5, 2003, pp. 3488–3492 Vol.5.

[5] J. Robinson and T. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 5, pp. 917–928, 2004.

[6] K. Xu, Q. Wang, and H. Hassanein, "Performance analysis of differentiated QoS supported by IEEE 802.11e enhanced distributed coordination function (EDCF) in WLAN," in *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 2, 2003, pp. 1048–1053 Vol.2.

[7] A. Banchs and L. Vollero, "A delay model for IEEE 802.11e EDCA," *Communications Letters, IEEE*, vol. 9, no. 6, pp. 508–510, 2005.

[8] X. Chen, H. Zhai, and Y. Fang, "Enhancing the IEEE 802.11e in QoS Support: Analysis and Mechanisms," in *Quality of Service in Heterogeneous Wired/Wireless Networks, 2005. Second International Conference on*, 2005, p. 23.

[9] H. Zhu and I. Chlamtac, "Performance analysis for IEEE 802.11e EDCF service differentiation," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1779–1788, 2005.

[10] J. Hui and M. Devetsikiotis, "A Unified Model for the Performance Analysis of IEEE 802.11e EDCA," *Communications, IEEE Transactions on*, vol. 53, no. 9, pp. 1498–1510, 2005.

[11] Y. Xiao, "Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1506–1515, 2005.

[12] P. Engelstad and O. Osterbo, "Delay and Throughput Analysis of IEEE 802.11e EDCA with Starvation Prediction," in *Local Computer Networks, 2005. 30th Anniversary. The IEEE Conference on*, 2005, pp. 647–655.

[13] Z.-n. Kong, D. Tsang, B. Bensaou, and D. Gao, "Performance analysis of IEEE 802.11e contention-based channel access," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 10, pp. 2095–2106, 2004.

[14] J. Tantra, C. H. Foh, and A. Mnaouer, "Throughput and delay analysis of the IEEE 802.11e EDCA saturation," in *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, vol. 5, 2005, pp. 3450–3454 Vol. 5.

[15] J. Zhao, Z. Guo, Q. Zhang, and W. Zhu, "Performance study of MAC for service differentiation in IEEE 802.11," in *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, vol. 1, 2002, pp. 778–782 vol.1.

[16] G. Bianchi and I. Tinnirello, "Remarks on IEEE 802.11 DCF performance analysis," *Communications Letters, IEEE*, vol. 9, no. 8, pp. 765–767, 2005.

[17] "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," pp. i–513, 2003.

[18] G. Bianchi, I. Tinnirello, and L. Scalia, "Understanding 802.11e contention-based prioritization mechanisms and their coexistence with legacy 802.11 stations," *Network, IEEE*, vol. 19, no. 4, pp. 28–34, 2005.

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

[20] "OPNET University Program: http://www.opnet.com/services/university/."