Modeling and Analysis of Slow CW Decrease for IEEE 802.11 WLAN

Qiang Ni[1], Imad Aad[2], Chadi Barakat[1], and Thierry Turletti[1]

[1]Planete Group
INRIA Sophia Antipolis
Sophia Antipolis, France
{qni, cbarakat, turletti}@sophia.inria.fr

[2]Planete Group
INRIA Rhône-Alpes
Grenoble, France
imad.aad@inrialpes.fr

Abstract—**The IEEE 802.11 Medium Access Control (MAC) protocol provides a contention-based distributed channel access mechanism for mobile stations to share the wireless medium, which may introduce a lot of collisions in case of overloaded active stations. Slow Contention Window (CW) decrease scheme is a simple and efficient solution for this problem. In this paper, we use an analytical model to compare the slow CW decrease scheme to the IEEE 802.11 MAC protocol. Several parameters are investigated such as the number of stations, the initial CW size, the decrease factor value, the maximum backoff stage and the coexistence with the RequestToSend and ClearToSend (RTS/CTS) mechanism. The results show that the slow CW decrease scheme can efficiently improve the throughput of IEEE 802.11, and that the throughput gain is higher when the decrease factor is larger. Moreover, the initial CW size and maximum backoff stage also affect the performance of slow CW decrease scheme.**

Keywords - IEEE 802.11; DCF; slow CW decrease scheme; RTS/CTS

I. Introduction

In recent years, IEEE 802.11 wireless LAN (WLAN) [1] has emerged as one of the most deployed wireless access technologies all over the world. This technology provides people with a ubiquitous environment in offices, hospitals, campuses, factories, airports and stock markets. The IEEE 802.11 standard provides both Medium Access Control (MAC) layer and the physical (PHY) layer specification for WLAN. IEEE 802.11 MAC has defined two medium access coordination functions: the contention-based *Distributed Coordination Function* (DCF) and the contention-free based *Point Coordination Function* (PCF) [1]. 802.11 can operate both in DCF mode and PCF mode. Every 802.11 station should implement DCF mode, which is based on the *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA) protocol [1]. Unlike DCF, the implementation of PCF is not mandatory in the standard. In this paper, we limit our investigation to the DCF and corresponding enhanced schemes.

In the DCF scheme, all stations compete for the resources and channel with the same priorities. The number of collisions increases with the number of stations. Throughput degradation and high delays are caused by the increasing time needed by contending stations to access the channel. Although the RequestToSend and ClearToSend (RTS/CTS) scheme is known to provide better performance than basic access scheme in some cases [2], it induces a considerable overhead when packet size is small. Recently, IEEE 802.11 Task Group e (TGe) has been working on a new mechanism, the *Enhanced Distributed Coordination Function* (EDCF), to enhance the performance of 802.11 DCF [4]. However, latest research works [8, 9] have shown that EDCF only reduces the internal collisions within a station, and external collisions between stations remain high in ad-hoc networks. This motivates the research on the slow Contention Window (CW) decrease scheme [9].

To analyze the performance of 802.11 DCF, [2] proposes an analytical model for the computation of 802.11 saturation throughput. This model makes the following assumptions: Ideal channel conditions (i.e. no hidden terminals and capture), a fixed number of stations and each mobile station always have packets to send. [3] extends this model further to consider the case of dynamic number of stations. The active stations are modeled with a Continuous Time Markov Chain Single Server Queue (CTMC-SSQ) process. [5] extends the model in [2] to consider the frame retansmission limits, which is specified in the 802.11 standard. [6] analyzes the throughput and fairness issues of the DCF function concerning the effect of hidden terminals and capture. [7] uses a *p-persistent* protocol to study the maximum protocol capacity of 802.11. The authors in [7] claim that this method gives very close approximation of the 802.11 standard protocol if the average backoff interval is always the same. Unlike 802.11, they propose to compute the optimized contention window size that maximizes the channel utilization. But this scheme requires the knowledge of the number of active stations, which is difficult to obtain in real implementations. The slow CW decrease scheme in [9] is simpler than the one in [7], since it only requires multiplying the previous CW by a constant decrease factor to compute the new CW after successful transmission. Given that there are no analytical models to analyze the performance of slow CW decrease scheme, we present in this paper a Markov chain

0-7803-7822-9/03/$17.00 ©2003 IEEE.

model that allows this analysis. Our analytical model is based on the one proposed in [2], which has already proven good performance results.

The rest of this paper is organized as follows. Section II describes the slow CW decrease scheme briefly. Section III derives our analytical model. Section IV shows the numerical results of the model and analyzes the performance of the slow decrease scheme. Section V concludes the paper.

II. Slow CW decrease scheme

In a distributed 802.11 DCF mode, a mobile station has no knowledge of the number of other contending terminals. Thus, the MAC layer adapts its CW to the current congestion level by doubling its CW upon each collision, and by resetting it upon each successful transmission. Doubling the CW assumes that each unsuccessful transmission indicates a high congestion level. On the contrary, when a node succeeds to transmit a packet, it assumes the congestion level decreasing and resets its CW size to its minimal value: CW_{min}.

However, when a transmission succeeds at a given CW, this does not correspond to a congestion level decrease, but to a convenient CW value. Therefore the CW value should be kept the same as long as the congestion level remains the same. Normally, congestion level is not likely to drop sharply. By resetting the CW to CW_{min}, a node takes the risk of experiencing collisions and retransmissions until it reaches the high CW value again, wasting time and channel bandwidth. Although a "post" backoff, i.e. DIFS plus backoff after a successful transmission, is used in the standard to help slow-start after each successful transmission [1], this is not enough to avoid collision. Slow CW decrease scheme provides a solution to this problem. The main advantage of slow CW decrease scheme is more collision avoidance during congestion, which results in less collisions and retransmissions, and hence in a better throughput. The disadvantage is keeping high CW values when congestion level sharply drops, increasing the overhead and maybe decreasing the throughput. The slow CW decrease scheme induces then a tradeoff between wasting some backoff time and risking a collision following a packet transmission.

[9] proposes three different slow CW decrease schemes: multiplicative CW decrease scheme, linear CW decrease scheme and adaptive CW decrease scheme. In this paper, we propose a Markov model to analyze the performance of the multiplicative slow CW decrease scheme and we denote this scheme as SD scheme. Let δ be the constant slow decrease factor in the range of (0,1). The SD scheme studied in this paper is defined as follows:

$CW_{new} = max\ (CW_{min},\ \delta\ CW_{old})$, after each successful transmission.

$CW_{new} = 2\ CW_{old}$, after each unsuccessful transmission.

III. Analytical model of slow CW decrease (SD) scheme

Our analysis is divided into two parts: First, we study the behavior of a single mobile station with a SD Markov model, and we compute the stationary probability τ that the station transmits a packet in a randomly chosen slot time. This probability does not depend on the access mechanisms (with or without RTS/CTS scheme). Second, by studying the events occurring within a slot time, we express the channel throughput as a function of τ with and without RTS/CTS scheme. We get then a system of two equations that we solve for the channel throughput by getting rid of τ.

A. Analysis of packet transmission probability

We make the same assumptions as [2]. A fixed number n of contending stations is considered and the transmission queue of each station is always nonempty. Each packet has to wait for a random backoff time decrement to zero before being transmitted. The slot time is defined as σ, and p denotes the probability that a packet collides. A slot time is equal to real PHY slot time if no packets are transmitted. If a packet is transmitted, σ is equal to the busy period until the channel is idle again. We define two stochastic processes to model the protocol behavior, see Fig. 1. First, $b(t)$ represents the backoff counter of the time a station has to wait before it can transmit. This process has the range from 0 to the current CW size. Another stochastic process $s(t)$ is defined as the backoff stage at a different CW level. $s(t)$ scales from 0 to m, with m being the maximum CW stage.

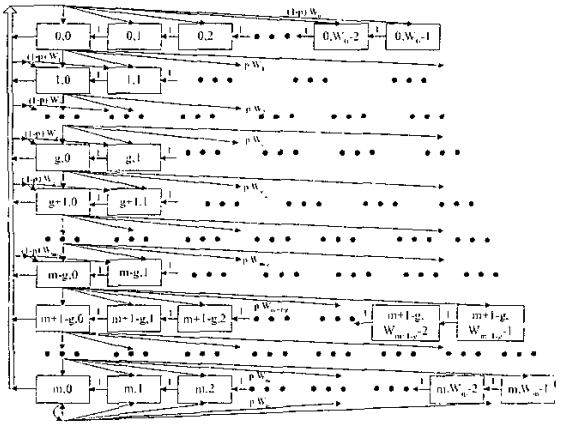

Fig.1 Markov chain model for the SD scheme

With these assumptions, the bi-dimensional stochastic process $\{s(t), b(t)\}$ fulfills the properties of a homogenous discrete Markov chain. The Markovian property does not hold for the process $b(t)$ alone, which is dependent on the backoff stage history. For simplicity, we write W_i instead of CW_i and W_0 instead of CW_{min}. Since the contention window doubles after each collision, we can write $W_i = 2^i W_0$, where $0 \le i \le m$. The maximum backoff stage m is the value such that $CW_{max}=2^m W_0$. We suppose that the constant decrease factor δ has a power of two form $\delta = 1/(2^s)$, where the constant factor

g is a positive integer with $g>0$. This choice of δ limits the number of states of the Markov chain and simplifies the analysis, without impacting the results. Another reason for choosing δ as power of two comes from implementation requirements. Current IEEE 802.11 contention window updating algorithms are implemented in hardware, where power of two multiplicative factors can be easily supported. Thus, the new CW value when a packet is correctly transmitted will be:

$$CW_{new} = max\,(W_0,\ \delta W_i) = max(W_0,\ 2^{i-g} W_0) = max(W_0,\ W_{i-g}).$$

Consider the transitions of the SD scheme between slot times. Fig. 1 explains the behavior of the Markov chain. The only non-null one-step transition probabilities are:

$$\begin{aligned}
&P\{i,\ k \mid i,\ k+1\} = 1, && k \in [0, W_i - 2],\ i \in [0, m]\\
&P\{0,\ k \mid i,\ 0\} = (1-p)/\ W_0, && k \in [0, W_0 - 1],\ i \in [0,\ g-1]\\
&P\{i\text{-}g,\ k \mid i,\ 0\} = (1-p)/\ W_{i-g}, && k \in [0, W_{i-g} - 1],\ i \in [g,\ m]\\
&P\{i,\ k \mid i-1,\ 0\} = p/\ W_i, && k \in [0,\ W_i - 1],\ i \in [1,\ m]\\
&P\{m,\ k \mid m,\ 0\} = p/\ W_m, && k \in [0,\ W_m - 1]. \qquad (1)
\end{aligned}$$

The first equation in (1) accounts for the fact that the backoff timer has not yet reached 0 and that it is decremented by 1 at the beginning of each slot time. The second and third equations are specific to the SD scheme. The second equation accounts for the fact that when δW_i is smaller than W_0, we reset W_i to W_0, and a new backoff is uniformly chosen in the range $[0,\ W_0 - 1]$. The third equation accounts for the fact that when δW_i is larger than W_0, we decrease W_i slowly to the new value W_{i-g} and we choose the new backoff counter randomly in the range $[0,\ W_{i-g}]$. The fourth and the fifth equations correspond to the cases where a collision occurs.

Let $\pi_{i,k} = \lim_{t\to\infty} P\{s(t) = i, b(t) = k\},\ i \in [0,\ m],\ k \in [0,\ W_i - 1]$, be the stationary distribution of the chain. As the Markov Chain is ergodic, this distribution exists and is unique. First, we express all $\pi_{i,k}$ as a function of $\pi_{0,0}$, then we use the normalization equation to solve for $\pi_{0,0}$, and hence for all $\pi_{i,k}$.

From the Markov chain above, we can see that the incoming traffic to stage i from either stage $i+g$ after a successful transmission, or from stage i-1 after a collision, is uniformly distributed over all possible backoff values at this stage. Afterwards, the counter is decremented by one and finally reaches state $(i,0)$. So, the stationary probability $\pi_{i,0}$ is given by:

$$\begin{aligned}
\pi_{0,0} &= (1-p) \sum_{j=0}^{g} \pi_{j,0}\\
\pi_{i,0} &= p\,\pi_{i-1,0} + (1-p)\,\pi_{i+g,0}, \quad 0 < i \le m-g \qquad (2)\\
\pi_{i,0} &= p\,\pi_{i-1,0}, \quad m-g < i < m\\
p\,\pi_{m-1,0} &= (1-p)\,\pi_{m,0} \rightarrow \pi_{m,0} = \frac{p}{1-p}\pi_{m-1,0}, \quad i = m.
\end{aligned}$$

The first equation in (2) accounts for the fact that stage 0 can only be reached from stage j ($j \le g$) in the SD scheme, the stage i ($j > g$) can not directly decrease to stage 0. The second equation in (2) says that when $0 < i \le m - g$, there are two different inputs: from the previous stage with collision probability p and from stage $i + g$ after a successful transmission with probability $1-p$. For i larger than $m-g$, there will be no input from stage $i + g$, because $i + g$ is bigger than the maximum stage number m. For $i = m$, we fall into a special case, since after a collision the contention window remains at this stage.

Now, according to the Markov chain regularities, for each $k \in [1, W_i-1]$, $\pi_{i,k}$ can be written as:

$$\pi_{i,k} = \frac{W_i - k}{W_i} \begin{cases} (1-p)\sum_{j=0}^{g} \pi_{j,0} & i = 0\\ p\,\pi_{i-1,0} + (1-p)\,\pi_{i+g,0} & 0 < i \le m-g\\ p\,\pi_{i-1,0} & m-g < i < m\\ p\,(\pi_{m-1,0} + \pi_{m,0}) & i = m. \end{cases} \qquad (3)$$

The ratio before the parenthesis accounts for the distribution of probabilities for each state in a stage. When we move in a stage to the right, the state probability decreases by $1/W_i$, since we do not get the input of the previous state in the same stage. From there, we can obtain the relation between $\pi_{i,k}$ and $\pi_{i,0}$: $\pi_{i,k} = [(W_i - k)/\ W_i]\ \pi_{i,0}$. Using (2), we obtain the term on the right-hand side of the parenthesis in (3). By combining (2) and (3), one can compute all stationary probabilities as a function of $\pi_{0,0}$ and p. In opposite to [2], obtaining closed-form expressions does not seem possible, so we proceed by solving the system numerically with Matlab: first we solve formulas in (2) to obtain $\pi_{i,0}$ that are only dependent on $\pi_{0,0}$ and p. Then we plug them into (3) to obtain $\pi_{i,k}$ that are only dependent on $\pi_{0,0}$ and p. $\pi_{0,0}$ is finally computed by using the normalization condition:

$$\sum_{i=0}^{m} \sum_{k=0}^{W_i-1} \pi_{i,k} = 1. \qquad (4)$$

Now we compute τ, the probability that a station transmits in a slot time. This probability is simply the sum of probabilities of all $(i,0)$ states,

$$\tau = \sum_{i=0}^{m} \pi_{i,0} = f(p, W_0, g, m). \qquad (5)$$

This expression of τ is a function of p, which is unknown. Let us assume independence of all stations sharing the medium, i.e. the probability that a station encounters a contention is independent of the status of the other stations. All stations transmit packets in a slot time with the same probability τ. Consider that a station transmits a packet in a slot time. p is then the probability that at least one other station transmits a packet in the same slot time:

$$p = 1 - (1-\tau)^{(n-1)}. \qquad (6)$$

We obtain therefore a non-linear system of two equations (5) and (6), that we can solve for p and τ. This system certainly has a solution, since the expression of p as a function of τ is continuously increasing with τ, with $p = 0$ for $\tau = 0$ and $p = 1$

for $\tau = 1$. A sufficient condition for this solution to be unique is that the expression of τ as a function of p given in (5) is continuously decreasing. Our numerical results in section IV show that a unique solution for our model always exists.

B. Throughput

Denote by S the normalized system throughput, which is defined as the fraction of time the channel is used to transmit payloads successfully. Consider a random slot time, let P_{tr} be the probability that there is at least one transmission in this slot time, and let P_s be the probability of one successful transmission given that there is at least one transmission. Note that $P_{tr}=1-(1-\tau)^n$ and $P_s = \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^n}$. Hence,

$$S = \frac{P_{tr}P_s\mathrm{E[P]}}{(1-P_{tr})\sigma + P_{tr}P_sT_s + P_{tr}(1-P_s)T_c}, \qquad (7)$$

where T_s is the average time the channel is sensed busy because of a successful transmission, and T_c is the average time the channel is sensed busy by each station during a collision. We use in our analysis the values of T_s and T_c computed in [2]. Note that the throughput expression (7) does not specify the access mechanism employed. To account for whether RTS/CTS scheme is used or not, we only need to specify the corresponding values T_s and T_c [2].

IV. Numerical Analysis

We use the Matlab tool to solve our model for the throughput of the channel. The 802.11 WLAN system parameters used in the model are reported in Table 1. We study the performance impact of the SD scheme on 802.11 throughput for several system parameters, such as with or without RTS/CTS mode, the number of stations, the $CW_{\min}$ value, the maximum backoff stage number m, and the value of SD factor g. Note that g=1 means $CW_{new} = 0.5\ CW_{old}$, which is the slowest decrease scheme we consider in this paper. Our numerical results show that in all cases, g=1 achieves the best performance in terms of throughput. We validate this result with *ns* simulations and obtain a channel throughput very close to what is predicted by our model. The results of the simulations are not included in this paper for lack of space.

TABLE 1 SIMULATION PARAMETERS

Packet payload	8184 bits
MAC header	272 bits
PHY header	128 bits
ACK	112 bits+PHY header
RTS	160 bits+PHY header
CTS	112 bits+PHY header
Channel bit rate	1 Mbit/s
Propagation Delay	1 μs
Slot time	50 μs
SIFS	28 μs
DIFS	128 μs

A. Without RTS/CTS mechanism

Fig. 2 shows the saturation throughput for standard 802.11 and for the SD scheme. The figure reports six different values for the number of stations n: 5, 10, 15, 20, 30 and 50. We clearly see how the throughput decreases when n increases (more contention) and how the total throughput of the SD scheme is always higher than that of the basic 802.11 access scheme, especially for the smallest value of g (g=1). For example, when n =50, the throughput gain of the SD scheme over standard 802.11 is about 28% for g = 1, about 13% for g=2, about 6% for g=3, and about 1% for g=5.

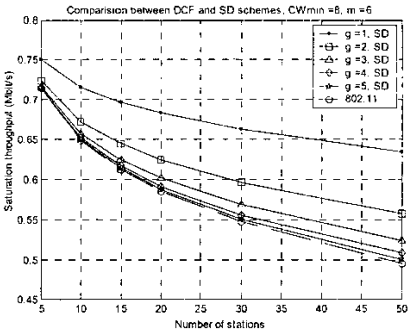

Fig. 2 Saturation throughput for SD and 802.11

Fig. 3 describes the impact of the initial CW size (W_0) on the SD scheme for different values of g. We set the maximum number of backoff stages to 6, i.e. $W_m=2^6\ W_0$. The initial CW size strongly affects the SD gain. For example, when n =50, a high throughput gain (28%) is obtained with a small initial CW (W_0=8), and the gain decreases to 4% with a large initial CW size (W_0=128). A large initial CW reduces the number of collisions, which makes the SD scheme less effective than the case when a small initial CW is used and the number of collisions is high.

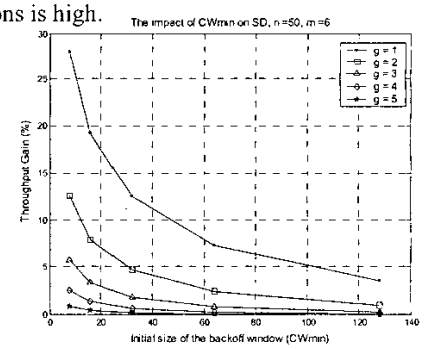

Fig. 3 Throughput gain vs. initial CW size

To better understand the above results, we study the following two measures:

i). The average number of idle slot times per successful transmission, which can be expressed as: $(1-P_{tr})/(P_{tr}P_s)$;

ii). The average channel time wasted in collisions per successful transmission, which is expressed as: $T_c(\frac{1}{P_s}-1)$.

Fig. 4 and Fig. 5 show the idle time and the collision time versus the number of stations, for the SD scheme with 5 different values of g and for the 802.11 scheme when W_0=8. We observe that the SD scheme slightly increases the idle

time but significantly decreases the collision probability. For example, when n=15 and g=1, the idle channel time for the SD scheme is 0.6 slot times longer than 802.11, and the time wasted in collision for the SD scheme is about 38 slot times shorter than 802.11. As mentioned in Section II, the SD scheme involves a tradeoff between wasting some backoff time and risking a collision followed by the retransmission.

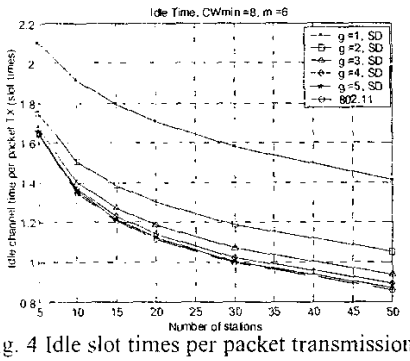

Fig. 4 Idle slot times per packet transmission (W_0=8)

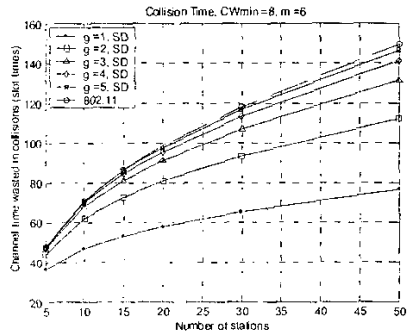

Fig. 5 Channel time wasted in collision (W_0=8)

B. *With RTS/CTS mechanism*

Fig. 6 compares the SD throughput gain obtained with and without the use of the RTS/CTS mechanism. The gain without RTS/CTS is much higher than when RTS/CTS is used. This means that the SD scheme is more useful when the RTS/CTS is not used. The reason is that RTS/CTS reduces the collision time to a small value, which makes the use of SD less effective since the collision time is already small.

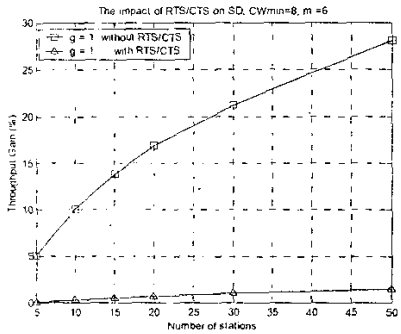

Fig. 6 SD throughput gain with and without RTS/CTS

V. Conclusion

This paper presents an analytical model for the slow CW decrease scheme, which has been proposed to improve the performance of the basic IEEE 802.11 MAC. Our model takes into account the different parameters that affect the channel throughput, such as the number of mobile stations, the initial CW size, the decrease factor value, the maximum number of backoff stages and the use of RTS/CTS. The numerical results we obtained show that the Slow CW Decrease (SD) scheme improves the throughput of IEEE 802.11 in all cases, especially when the number of stations is large. Another finding is that the SD scheme significantly increases the throughput of basic CSMA/CA mode when using a large decrease factor (e.g. δ=0.5), while it is not very helpful when the RTS/CTS mode is used since the collision time is small with RTS/CTS. In addition, the initial CW size and the maximum backoff stage also affect the performance of the SD scheme and the gain in throughput. Future work will include the modeling analysis of the SD scheme with the effect of hidden terminals, and the impact of the SD scheme on fairness issues.

Acknowledgment

The authors would like to thank Romain Zanolla and Marco Schmitt for their earlier works on Matlab codes.

References

[1] IEEE Std 802.11-1999, Part 11:Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, 1999

[2] G. Bianchi. "Performance analysis of the IEEE 802.11 distributed coordination function". IEEE Journal on Selected Areas In Communications, Vol. 18, No. 3, March 2000

[3] C. H. Foh and M. Zukerman. "Performance analysis of the IEEE 802.11 mac protocol". Proceedings of the EW 2002 Conference, Florence, Italy, pp. 184-190, February 2002

[4] IEEE 802.11 WG, Draft Supplement to Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), IEEE 802.11e/Draft 4.1, Februray 2003

[5] H. Wu, Y. Peng, K. Long, S. Cheng and J. Ma. "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement". IEEE Infocom 02, New York, June 2002

[6] H.S. Chhaya and S. Gupta. "Performance modeling of the asynchronous data transfer methods of IEEE 802.11 MAC protocol". Wireless Networks, No. 3, pp. 217-234, 1997

[7] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit". IEEE/ACM Transaction on Networking, Vol. 8, No. 6, December 2000

[8] L. Romdhani, Q. Ni and T. Turletti, "Adaptive EDCF: enhanced service differentiation for IEEE 802.11 wireless ad hoc networks". IEEE WCNC'03, New Orleans, Louisiana, March 16-20, 2003

[9] I. Aad, Q. Ni, C. Castelluccia, and T. Turletti, "Enhancing IEEE 802.11 performance with slow CW decrease". IEEE 802.11e working group document 802.11-02/674r0, November 11, 2002