Performance Analysis of the IEEE 802.11 Distributed Coordination Function

Giuseppe Bianchi
Universitá di Palermo, Dipartimento di Ingegneria Elettrica
Viale Delle Scienze, Parco D'Orleans, 90128 Palermo, Italy
Tel: +39-091-6566-247 Fax: +39-091-488452
email: bianchi@elet.polimi.it

Abstract— **IEEE has recently standardized the 802.11 protocol for Wireless Local Area Networks. The primary Medium Access Control (MAC) technique of 802.11 is called Distributed Coordination Function (DCF). DCF is a Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) scheme with binary slotted exponential backoff. This paper provides a simple, but nevertheless extremely accurate, analytical model to compute the 802.11 DCF throughput, in the assumption of finite number of terminals and ideal channel conditions. The proposed analysis applies to both the packet transmission schemes employed by DCF, namely the Basic Access and the RTS/CTS access mechanisms. In addition, it applies also to a combination of the two schemes, in which packets longer than a given threshold are transmitted according to the RTS/CTS mechanism. By means of the proposed model, in this paper we provide an extensive throughput performance evaluation of both access mechanisms of the 802.11 protocol.**

I. Introduction

In recent years, much interest has been involved in the design of wireless networks for local area communication [1], [2]. Study group 802.11 was formed under IEEE project 802 to recommend an international standard for Wireless Local Area Networks (WLANs). The final version of the standard has recently appeared [3], and provides detailed Medium Access Control (MAC) and Physical layer (PHY) specification for WLANs.

In the 802.11 protocol, the fundamental mechanism to access the medium is called *Distributed Coordination Function* (DCF). This is a random access scheme, based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol. Retransmission of collided packets is managed according to binary exponential backoff rules. The standard also defines an optional *Point Coordination Function* (PCF), which is a centralized MAC protocol able to support collision free and time bounded services. In this paper we limit our investigation to the DCF scheme.

DCF describes two techniques to employ for packet transmission. The default scheme is a two-way handshaking technique called Basic Access mechanism. This mechanism is characterized by the immediate transmission of a positive acknowledgement (ACK) by the destination station, upon successful reception of a packet transmitted by the sender station. Explicit transmission of an ACK is required since, in the wireless medium, a transmitter cannot determine if a packet is successfully received by listening to its own transmission.

In addition to the Basic Access, an optional four way handshaking technique, known as *Request-To-Send/Clear-To-Send* (RTS/CTS) mechanism has been standardized. Before transmitting a packet, a station operating in RTS/CTS mode "reserves" the channel by sending a special *Request-To-Send* short frame. The destination station acknowledges the receipt of an RTS frame by sending back a *Clear-To-Send* frame, after which normal packet transmission and ACK response occurs. Since collision may occur only on the RTS frame, and it is detected by the lack of CTS response, the RTS/CTS mechanism allows to increase the system performance by reducing the duration of a collision when long messages are transmitted. As an important side effect, the RTS/CTS scheme designed in the 802.11 protocol is suited to combat the so called problem of *Hidden Terminals* [4], which occurs when pairs of mobile stations result to be unable to hear each other. This problem has been specifically considered in [5] and in [6], which, in addition, studies the phenomenon of packet capture.

In this paper we concentrate on the performance evaluation of the DCF scheme, in the assumption of ideal channel conditions and finite number of terminals. In the literature, performance evaluation of 802.11 has been carried out either by means of simulation [7], [8] or by means of analytical models with simplified backoff rule assumptions. In particular, constant or geometrically distributed backoff window has been used in [5], [9], [10], while [11] has considered an exponential backoff limited to two stages (maximum window size equal to twice the minimum size) by employing a two dimensional Markov chain analysis.

In this paper, which revises and substantially extends [12], we succeed in providing an extremely simple model that accounts for all the exponential backoff protocol details, and allows to compute the saturation (asymptotic) throughput performance of DCF for both standardized access mechanisms (and also for any combination of the two methods). The key approximation that enables our model is the assumption of constant and independent collision probability of a packet transmitted by each station, regardless of the number of retransmissions already suffered. As proven by comparison with simulation, this assumption leads to extremely accurate (practically exact) results, especially when the number of stations in the wireless LAN is fairly large (say greater than 10).

The paper is outlined as follows. In section II we briefly review both Basic Access and RTS/CTS mechanisms of the DCF. In section III we define the concept of Saturation Throughput, and in section IV we provide an analytical technique to compute this performance figure. Section V validates the accuracy of the model by comparing the analytical results with that obtained by means of simulation. Additional considerations on the maximum throughput theoretically achievable are carried out in section VI. Finally, the performance evaluation of both DCF access schemes is carried out in section VII. Concluding remarks

This work has been partially supported by CNR and MURST, Italy

PHY	Slot Time (σ)	$CW_{\min}$	$CW_{\max}$
FHSS	50 μs	16	1024
DSSS	20 μs	32	1024
IR	8 μs	64	1024

TABLE I

Slot Time, *minimum, and maximum contention window values for the three PHY specified by the 802.11 standard: Frequency Hopping Spread Spectrum (FHSS), Direct Sequence Spread Spectrum (DSSS), Infrared (IR)*

are given in Section VIII.

II. 802.11 Distributed Coordination Function

This section briefly summarizes the Distributed Coordination Function (DCF) as standardized by the 802.11 protocol. For a more complete and detailed presentation, refer to the 802.11 standard [3].

A station with a new packet to transmit monitors the channel activity. If the channel is idle for a period of time equal to a Distributed InterFrame Space (DIFS), the station transmits. Otherwise, if the channel is sensed busy (either immediately or during the DIFS), the station persists to monitor the channel until it is measured idle for a DIFS. At this point, the station generates a random backoff interval before transmitting (this is the Collision Avoidance feature of the protocol), to minimize the probability of collision with packets being transmitted by other stations. In addition, to avoid channel capture, a station must wait a random backoff time between two consecutive new packet transmissions, even if the medium is sensed idle in the DIFS time[1].

For efficiency reasons, DCF employs a discrete-time backoff scale. The time immediately following an idle DIFS is slotted, and a station is allowed to transmit only at the beginning of each *Slot Time*. The Slot Time size, σ, is set equal to the time needed at any station to detect the transmission of a packet from any other station. As shown in table I, it depends on the physical layer, and it accounts for the propagation delay, for the time needed to switch from the receiving to the transmitting state (RX_TX_Turnaround_Time), and for the time to signal to the MAC layer the state of the channel (Busy Detect Time).

DCF adopts an exponential backoff scheme. At each packet transmission, the backoff time is uniformly chosen in the range $(0, w-1)$. The value w is called *Contention Window*, and depends on the number of transmissions failed for the packet. At the first transmission attempt, w is set equal to a value $CW_{\min}$ called minimum contention window. After each unsuccessful transmission, w is doubled, up to a maximum value $CW_{\max} = 2^m CW_{\min}$. The values $CW_{\min}$ and $CW_{\max}$ reported in the final version of the standard [3] are PHY-specific and are summarized in table I.

The backoff time counter is decremented as long as the channel is sensed idle, "frozen" when a transmission is detected on the channel, and reactivated when the channel is sensed idle

[1] As an exception to this rule, the protocol provides a fragmentation mechanism, which allows the MAC to split an MSDU (the packet delivered to the MAC by the higher layers) into more MPDUs (packets delivered by the MAC to the PHY layer), if the MSDU size exceeds the maximum MPDU payload size. The different fragments are then transmitted in sequence, with only a SIFS between them, so that only the first fragment must contend for the channel access.

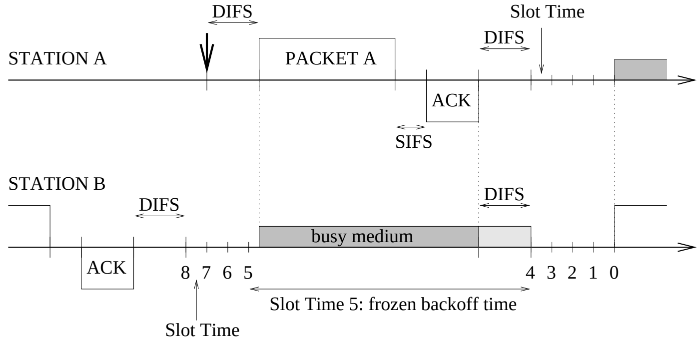

Fig. 1. *Example of Basic Access Mechanism*

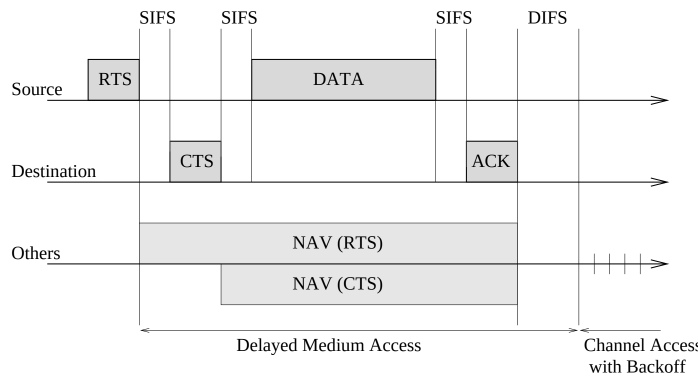

Fig. 2. *RTS/CTS Access Mechanism*

again for more than a DIFS. The station transmits when the backoff time reaches 0.

Figure 1 illustrates this operation. Two stations A and B share the same wireless channel. At the end of the packet transmission, station B waits for a DIFS and then chooses a backoff time equal to 8, before transmitting the next packet. We assume that the first packet of station A arrives at the time indicated with an arrow in the figure. After a DIFS, the packet is transmitted. Note that the transmission of packet A occurs in the middle of the Slot Time corresponding to a backoff value, for station B, equal to 5. As a consequence of the channel sensed busy, the backoff time is frozen to its value 5, and the backoff counter decrements again only when the channel is sensed idle for a DIFS.

Since the CSMA/CA does not rely on the capability of the stations to detect a collision by hearing their own transmission, a positive acknowledgement (ACK) is transmitted by the destination station to signal the successful packet reception. The ACK is immediately transmitted at the end of the packet, after a period of time called Short InterFrame Space (SIFS). As the SIFS (plus the propagation delay) is shorter than a DIFS, no other station is able to detect the channel idle for a DIFS until the end of the ACK. If the transmitting station does not receive the ACK within a specified ACK_Timeout, or it detects the transmission of a different packet on the channel, it reschedules the packet transmission according to the given backoff rules.

The above described two-way handshaking technique for the packet transmission is called *Basic Access* mechanism. DCF defines an additional four-way handshaking technique to be optionally used for a packet transmission. This mechanism, known

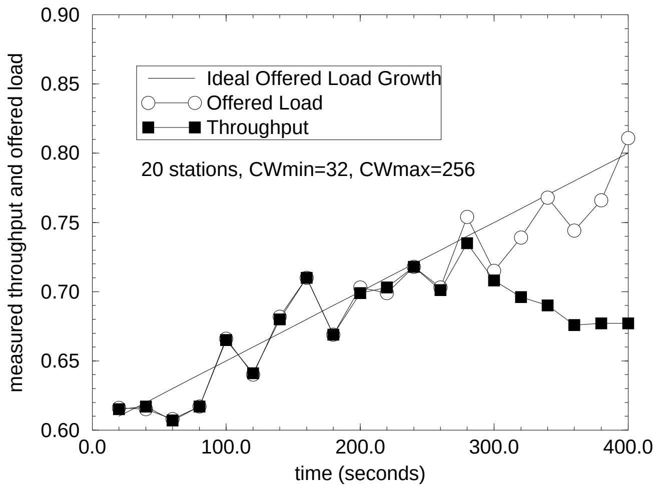

Fig. 3. *Measured Throughput with slowly increasing offered load*

with the name RTS/CTS, is shown in figure 2. A station that wants to transmit a packet, waits until the channel is sensed idle for a DIFS, follows the backoff rules explained above, and then, instead of the packet, preliminarily transmits a special short frame called *Request To Send* (RTS). When the receiving station detects an RTS frame, it responds, after a SIFS, with a *Clear To Send* (CTS) frame. The transmitting station is allowed to transmit its packet only if the CTS frame is correctly received.

The frames RTS and CTS carry the information of the length of the packet to be transmitted. This information can be read by any listening station, which is then able to update a *Network Allocation Vector* (NAV) containing the information of the period of time in which the channel will remain busy. Therefore, when a station is *hidden* from either the transmitting or the receiving station, by detecting just one frame among the RTS and CTS frames, it can suitably delay further transmission, and thus avoid collision.

The RTS/CTS mechanism is very effective in terms of system performance, especially when large packets are considered, as it reduces the length of the frames involved in the contention process. In fact, in the assumption of perfect channel sensing by every station, collision may occur only when two (or more) packets are transmitted within the same slot time. If both transmitting stations employ the RTS/CTS mechanism, collision occurs only on the RTS frames, and it is early detected by the transmitting stations by the lack of CTS responses. A quantitative analysis will be carried out in section VII.

III. Maximum and Saturation Throughput Performance

In this paper we concentrate on the “Saturation Throughput”. This is a fundamental performance figure defined as the limit reached by the system throughput as the offered load increases, and represents the maximum load that the system can carry *in stable conditions*.

It is well known that several random access schemes exhibit an unstable behavior. In particular, as the offered load increases, the throughput grows up to a maximum value, referred to as “Maximum Throughput”. However, further increases of the offered load lead to an eventually significant decrease in the system throughput. This results in the practical impossibility to operate the random access scheme at its maximum throughput for a “long” period of time, and thus in the practical meaningless of the maximum throughput as performance figure for the access scheme. The mathematical formulation and interpretation of this instability problem is the object of a wide and general discussion in [13].

Indeed, the 802.11 protocol is known to exhibits some form of instability (see for example [5], [11]. To visualize the unstable behaviour of 802.11, in figure 3 we have run simulations in which the offered load linearly increases with the simulation time. The general simulation model and parameters employed are summarized in section V. The results reported in the figure are obtained with 20 stations. The straight line represents the ideal offered load, normalized with respect of the channel capacity. The simulated offered load has been generated according to a Poisson arrival process of fixed size packets (payload equal to 8184 bits), where the arrival rate has been varied throughout the simulation to match the ideal offered load. The figure reports both offered load and system throughput measured over 20 seconds time intervals, and normalized with respect to the channel rate.

From the figure, we see that the measured throughput follows closely the measured offered load for the first 260 seconds of simulation, while it asynptotically drops to the value 0.68 in the second part of the simulation run. This asymptotic throughput value is referred to, in this paper, as Saturation Throughput, and represents the system throughput in overload conditions. Note than, during the simulation run, the instantaneous throughput temporarily increases over the saturation value (up to 0.74 in the example considered), but ultimately it decreases and stabilizes to the saturation value. Queue build-up is observed in such a condition.

IV. Throughput Analysis

The core contribution of this paper is the analytical evaluation of the saturation throughput, in the assumption of ideal channel conditions (i.e. no hidden terminals and capture [6]). In the analysis, we assume a fixed number of stations, each always having a packet available for transmission. In other words, we operate in *saturation* conditions, i.e. the transmission queue of each station is assumed to be always non-empty.

The analysis is divided into two distinct parts. First, we study the behavior of a single station with a Markov model, and we obtain the stationary probability τ that the station transmits a packet in a generic (i.e. randomly chosen) slot time. This probability does not depend on the access mechanism (i.e. Basic or RTS/CTS) employed. Then, by studying the events that can occur within a generic slot time, we express the throughput of both Basic and RTS/CTS access methods (as well as of a combination of the two) as function of the computed value τ.

A. Packet Transmission Probability

Consider a fixed number n of contending stations. In saturation conditions, each station has immediately a packet available for transmission, after the completion of each successful transmission. Moreover, being all packets “consecutive”, each packet needs to wait for a random backoff time before transmitting.

Let $b(t)$ be the stochastic process representing the backoff time counter for a given station. A discrete and integer time scale is adopted: t and $t+1$ correspond to the beginning of two consecutive slot times, and the backoff time counter of each station decrements at the beginning of each slot time. Note that this discrete time scale does not directly relates to the system time. In fact, as illustrated in figure 1, the backoff time decrement is stopped when the channel is sensed busy, and thus the time interval between two consecutive slot time beginnings may be much longer than the slot time size σ, as it may include a packet transmission. In what follows, unless ambiguity occurs, with the term slot time we will refer to either the (constant) value σ, and the (variable) time interval between two consecutive backoff time counter decrements.

Since the value of the backoff counter of each station depends also on its transmission history (e.g. how many retransmission the head-of-line packet has suffered), the stochastic process $b(t)$ is non markovian. However, define for convenience $W = CW_{\min}$. Let m, "maximum backoff stage", be the value such that $CW_{\max} = 2^m W$, and let us adopt the notation $W_i = 2^i W$, where $i \in (0,m)$ is called "backoff stage". Let $s(t)$ be the stochastic process representing the backoff stage $(0,\cdots,m)$ of the station at time t.

The key approximation in our model is that, at each transmission attempt, and regardless of the number of retransmissions suffered, each packet collides with constant and independent probability p. It is intuitive that this assumption results more accurate as long as W and n get larger. p will be referred to as *conditional collision probability*, meaning that this is the probability of a collision seen by a packet being transmitted on the channel.

Once independence is assumed, and p is supposed to be a constant value, it is possible to model the bidimensional process $\{s(t), b(t)\}$ with the discrete-time Markov chain depicted in figure 4. In this Markov chain, the only non null one-step transition probabilities are[2]:

$$\begin{cases} P\{i,k|i,k+1\} = 1 & k \in (0,W_i-2) \quad i \in (0,m) \\ P\{0,k|i,0\} = (1-p)/W_0 & k \in (0,W_0-1) \quad i \in (0,m) \\ P\{i,k|i-1,0\} = p/W_i & k \in (0,W_i-1) \quad i \in (1,m) \\ P\{m,k|m,0\} = p/W_m & k \in (0,W_m-1) \end{cases} \tag{1}$$

The first equation in (1) account for the fact that, at the beginning of each slot time, the backoff time is decremented. The second equation accounts for the fact that a new packet following a successful packet transmission starts with backoff stage 0, and thus the backoff is initially uniformly chosen in the range $(0, W_0 - 1)$. The other cases model the system after an unsuccessful transmission. In particular, as considered in the third equation of (1), when an unsuccessful transmission occurs at backoff stage $i-1$, the backoff stage increases, and the new initial backoff value is uniformly chosen in the range $(0, W_i)$. Finally, the fourth case models the fact that once the backoff stage reaches the value m, it is not increased in subsequent packet transmissions.

[2]we adopt the short notation:
$P\{i_1,k_1|i_0,k_0\} = P\{s(t+1) = i_1, b(t+1) = k_1 | s(t) = i_0, b(t) = k_0\}$.

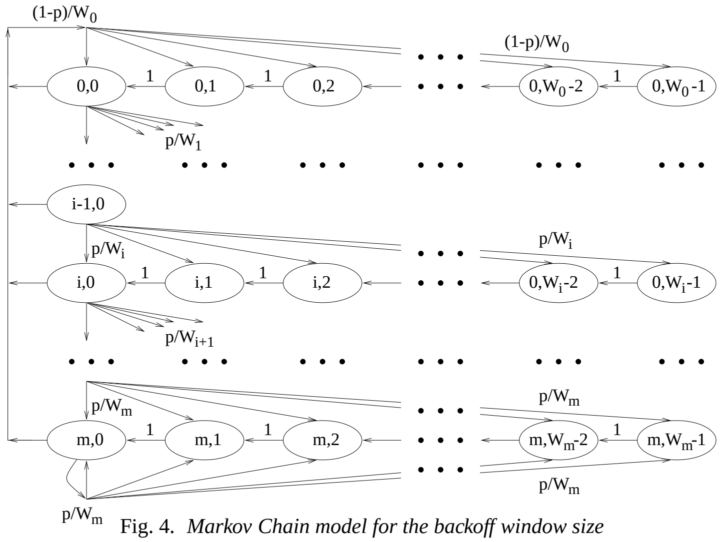

Fig. 4. *Markov Chain model for the backoff window size*

Let $b_{i,k} = \lim_{t\to\infty} P\{s(t) = i, b(t) = k\}$, $i \in (0,m)$, $k \in (0, W_i - 1)$ be the stationary distribution of the chain. We now show that it is easy to obtain a closed-form solution for this Markov chain. First, note that

$$\begin{array}{lll} b_{i-1,0}\cdot p = b_{i,0} & \rightarrow & b_{i,0} = p^i b_{0,0} \quad 0 < i < m \\ b_{m-1,0}\cdot p = (1-p) b_{m,0} & \rightarrow & b_{m,0} = \frac{p^m}{1-p} b_{0,0} \end{array} \tag{2}$$

Owing to the chain regularities, for each $k \in (1, W_i - 1)$, it is:

$$b_{i,k} = \frac{W_i - k}{W_i} \cdot \begin{cases} (1-p)\sum_{j=0}^{m} b_{j,0} & i = 0 \\ p\cdot b_{i-1,0} & 0 < i < m \\ p\cdot (b_{m-1,0} + b_{m,0}) & i = m \end{cases} \tag{3}$$

By means of relations (2), and making use of the fact that $\sum_{i=0}^{m} b_{i,0} = b_{0,0}/(1-p)$, equation 3 rewrites as:

$$b_{i,k} = \frac{W_i - k}{W_i} b_{i,0} \quad i \in (0,m), k \in (0, W_i - 1) \tag{4}$$

Thus, by relations (2) and (4), all the values $b_{i,k}$ are expressed as function of the value $b_{0,0}$ and of the conditional collision probability p. $b_{0,0}$ is finally determined by imposing the normalization condition, that simplifies as follows:

$$1 = \sum_{i=0}^{m}\sum_{k=0}^{W_i-1} b_{i,k} = \sum_{i=0}^{m} b_{i,0} \sum_{k=0}^{W_i-1} \frac{W_i - k}{W_i} = \sum_{i=0}^{m} b_{i,0}\frac{W_i+1}{2} =$$

$$= \frac{b_{0,0}}{2}\left[W\left(\sum_{i=0}^{m-1}(2p)^i + \frac{(2p)^m}{1-p}\right) + \frac{1}{1-p}\right] \tag{5}$$

from which:

$$b_{0,0} = \frac{2(1-2p)(1-p)}{(1-2p)(W+1) + pW(1-(2p)^m)} \tag{6}$$

We can now express the probability τ that a station transmits in a randomly chosen slot time. As any transmission occurs when the backoff time counter is equal to 0, regardless of the backoff stage, it is:

$$\tau = \sum_{i=0}^{m} b_{i,0} = \frac{b_{0,0}}{1-p} = \frac{2(1-2p)}{(1-2p)(W+1) + pW(1-(2p)^m)} \tag{7}$$

As a side note, it is interesting to highlight that, when $m = 0$, i.e. no exponential backoff is considered, the probability τ results to be independent of p, and equation (7) becomes the much simpler one independently found in [9] for the constant backoff window problem:

$$\tau = \frac{2}{W+1} \tag{8}$$

However, in general, τ depends on the conditional collision probability p, which is still unknown. To find the value of p it is sufficient to note that the probability p that a transmitted packet encounters a collision, is the probability that, in a time slot, at least one of the $n-1$ remaining stations transmit. The fundamental independence assumption given above implies that each transmission "sees" the system in the same state, i.e. in steady state. At steady state, each remaining station transmits a packet with probability τ. This yields:

$$p = 1 - (1-\tau)^{n-1} \tag{9}$$

Equations (7) and (9) represent a non linear system in the two unknowns τ and p, which can be solved using numerical techniques. It is easy to prove that this system has a unique solution. In fact, inverting (9), we obtain $\tau^*(p) = 1 - (1-p)^{1/(n-1)}$. This is a continuous and monotone increasing function in the range $p \in (0,1)$, that starts from $\tau^*(0) = 0$ and grows up to $\tau^*(1) = 1$. Equation $\tau(p)$ defined by (7) is also continuous in the range $p \in (0,1)$: continuity in correspondence of the critical value $p = 1/2$ is simply proven by noting that $\tau(p)$ can be alternatively written as:

$$\tau(p) = \frac{2}{1 + W + pW\sum_{i=0}^{m-1}(2p)^i}$$

and therefore $\tau(1/2) = 2/(1+W+mW/2)$. Moreover, $\tau(p)$ is trivially shown to be a monotone decreasing function that starts from $\tau(0) = 2/(W+1)$ and reduces up to $\tau(1) = 2/(1+2^m W)$. Uniqueness of the solution is now proven noting that $\tau(0) > \tau^*(0)$ and $\tau(1) < \tau^*(1)$.

B. Throughput

Let S be the normalized system throughput, defined as the fraction of time the channel is used to successfully transmit payload bits. To compute S, let us analyze what can happen in a randomly chosen slot time. Let P_{tr} be the probability that there is at least one transmission in the considered slot time. Since n stations contend on the channel, and each transmits with probability τ,

$$P_{tr} = 1 - (1-\tau)^n \tag{10}$$

The probability P_s that a transmission occurring on the channel is successful is given by the probability that exactly one station

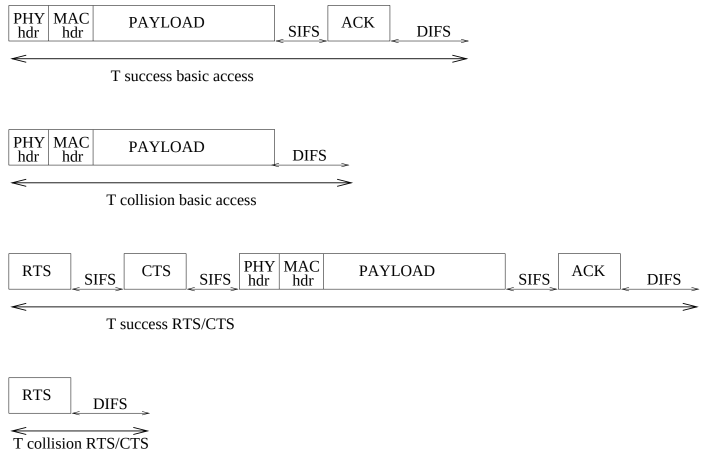

Fig. 5. *T_s and T_c for Basic Access and RTS/CTS mechanisms*

transmits on the channel, conditioned on the fact that at least one station transmits, i.e.:

$$P_s = \frac{n\tau(1-\tau)^{n-1}}{P_{tr}} = \frac{n\tau(1-\tau)^{n-1}}{1-(1-\tau)^n} \tag{11}$$

We are now able to express S as the ratio:

$$S = \frac{E\,[\text{payload information transmitted in a slot time}]}{E\,[\text{length of a slot time}]} \tag{12}$$

Being $E[P]$ the average packet payload size, the average amount of payload information successfully transmitted in a slot time is $P_{tr}P_sE[P]$, since a successful transmission occurs in a slot time with probability $P_{tr}P_s$. The average length of a slot time is readily obtained considering that, with probability $1 - P_{tr}$, the slot time is empty; with probability $P_{tr}P_s$ it contains a successful transmission, and with probability $P_{tr}(1-P_s)$ it contains a collision. Hence, (12) becomes:

$$S = \frac{P_s P_{tr} E[P]}{(1-P_{tr})\sigma + P_{tr}P_sT_s + P_{tr}(1-P_s)T_c} \tag{13}$$

Here, T_s is the average time the channel is sensed busy (i.e. the slot time lasts) because of a successful transmission, and T_c is the average time the channel is sensed busy by each station during a collision. σ is the duration of an empty slot time. Of course, the values $E[P]$, T_s, T_c, and σ must be expressed with the same unit.

Note that the throughput expression (13) has been obtained without the need to specify the access mechanism employed. To specifically compute the throughput for a given DCF access mechanism it is now necessary only to specify the corresponding values T_s and T_c.

Let us first consider a system completely managed via the Basic Access mechanism. Let $H = \mathrm{PHY_{hdr}} + \mathrm{MAC_{hdr}}$ be the packet header, and δ be the propagation delay. As shown in figure 5, in the Basic Access case we obtain:

$$\begin{cases} T_s^{bas} = & H + E[P] + \mathrm{SIFS} + \delta + \mathrm{ACK} + \mathrm{DIFS} + \delta \\ T_c^{bas} = & H + E[P^*] + \mathrm{DIFS} + \delta \end{cases} \tag{14}$$

where $E[P^*]$ is the the average length of the longest packet payload involved in a collision.

In the case all packets have the same fixed size, $E[P^*] = E[P] = P$. In the general case, the payload size of each collided packet is an independent random variable P_i. It is thus necessary to assume a suitable probability distribution function $F(.)$ for the packet's payload size. Let P_{max} be the maximum payload size. Taking the conditional expectation on the number k of colliding packets, $E[P^*]$ writes as follows:

$$E[P^*] = E\left[E\left[\max(P_1, \cdots, P_k)|k\right]\right] =$$

$$= \frac{\sum_{k=2}^{n} \binom{n}{k} \tau^k (1-\tau)^{n-k} \int_0^{P_{max}} (1-F(x)^k)dx}{1-(1-\tau)^n - n\tau(1-\tau)^{n-1}} \quad (15)$$

When the probability of three or more packets simultaneously colliding is neglected, expression (15) simplifies to:

$$E[P^*] = \int_0^{P_{max}} (1-F(x)^2)dx \quad (16)$$

T_c is the period of time during which the channel is sensed busy by the *non colliding stations*. We neglect the fact that the two or more colliding stations, before sensing the channel again, need to wait an ACK Timeout, and thus the T_c for these colliding stations is greater than that considered here (the same approximation holds in the following RTS/CTS case, with a CTS Timeout instead of the ACK timeout).

Let us now consider a system in which each packet is transmitted by means of the RTS/CTS Access mechanism. As, in such a case, collision can occur only on RTS frames, it is (see figure 5):

$$\begin{cases} T_s^{rts} = & \mathrm{RTS} + \mathrm{SIFS} + \delta + \mathrm{CTS} + \mathrm{SIFS} + \delta + H + \\ & +E[P] + \mathrm{SIFS} + \delta + \mathrm{ACK} + \mathrm{DIFS} + \delta \\ & \\ T_c^{rts} = & \mathrm{RTS} + \mathrm{DIFS} + \delta \end{cases} \quad (17)$$

and the throughput expression depends on the packet size distribution only through its mean.

Finally, formula (13) can be also adopted to express the throughput of an "Hybrid" system in which, as suggested in the standard [3], packets are transmitted by means of the RTS/CTS mechanism only if they exceed a given predetermined threshold $\bar{P}$ on the packet's payload size. More specifically, being, again, $F(.)$ the probability distribution function of the packet size, $F(\bar{P})$ is the probability that a packet is transmitted according to the Basic Access mechanism (i.e. the packet size is lower than $\bar{P}$), while $1-F(\bar{P})$ is the probability that a packet is transmitted via the RTS/CTS mechanism. For convenience, let us indicate with

$$O_{rts} = T_s^{rts} - T_s^{bas} = \mathrm{RTS} + \mathrm{SIFS} + \delta + \mathrm{CTS} + \mathrm{SIFS} + \delta \quad (18)$$

the RTS/CTS overhead for a successful packet transmission. It is easy to recognize that, for the described hybrid access scheme, it is:

$$T_s = T_s(\bar{P}) = T_s^{bas} F(\bar{P}) + T_s^{rts}(1-F(\bar{P})) =$$

$$= T_s^{bas} + O_{rts}(1-F(\bar{P})) \quad (19)$$

To compute $T_c = T_c(\bar{P})$ in the case of the Hybrid Access scheme, we rely on the simplifying assumption that the probability of a collision of more than two packets in the same slot time is negligible. Hence, three possible collision cases may occur: (i) collision between two RTS frames, with probability $(1-F(\bar{P}))^2$; (ii) collision between two packets transmitted via Basic Access, with probability $F(\bar{P})^2$, and (iii) collision between a basic access packet and an RTS frame. Hence, indicating with $T_c^{rts/rts}$, $T_c^{bas/bas}$ and $T_c^{bas/rts}$ the respective average collision durations, we obtain:

$$T_c(\bar{P}) = (1-F(\bar{P}))^2 T_c^{rts/rts} +$$

$$+2F(\bar{P})(1-F(\bar{P}))T_c^{rts/bas} + F^2(\bar{P})T_c^{bas/bas} \quad (20)$$

The average collision durations adopted in equation (20) detail as follows. Let $O_h = (T_c^{bas} - P - T_c^{rts}) = (H - \mathrm{RTS})$ be the extra length of the packet header with respect of the RTS frame, and let $\alpha = H + \mathrm{DIFS} + \delta$. The value $T_c^{rts/rts}$ has been already computed in the case T_c^{rts} of (17), and can be rewritten with new notation as:

$$T_c^{rts/rts} = \mathrm{RTS} + \mathrm{DIFS} + \delta = \alpha - O_h \quad (21)$$

To compute the average length of a collision between an RTS frame and a Basic Access packet, let us note that, according to the numerical values provided by the standard [3], the length of an RTS frame is always lower than the packet header size, or, in other words, the value O_h defined above is strictly positive. Thus the average length of such a collision is given by the average amount of time the channel is kept busy by the unsuccessful transmission of the Basic Access Packet. Since $F(x)/F(\bar{P})$, $x \in (0, \bar{P})$ is the conditional probability distribution function of the payload size of the packets transmitted according to the Basic Access mechanism, we readily obtain:

$$T_c^{rts/bas} = \alpha + \int_0^{\bar{P}} \left(1 - \frac{F(x)}{F(\bar{P})}\right) dx \quad (22)$$

Finally, noting that in the case of collision between two Basic Access packets, the probability distribution function of the length of the longest packet payload involved in a collision is the square of the conditional probability distribution function of the packet size distribution,

$$T_c^{bas/bas} = \alpha + \int_0^{\bar{P}} \left(1 - \frac{F^2(x)}{F^2(\bar{P})}\right) dx \quad (23)$$

By substituting (21), (22) and (23) in equation (20), we finally obtain:

$$T_c(\bar{P}) = \alpha - (1-F(\bar{P}))^2 O_h +$$

$$+2F(\bar{P})(1-F(\bar{P})) \int_0^{\bar{P}} \left(1 - \frac{F(x)}{F(\bar{P})}\right) dx +$$

$$+F^2(\bar{P}) \int_0^{\bar{P}} \left(1 - \frac{F^2(x)}{F^2(\bar{P})}\right) dx \quad (24)$$

For simplicity, in the rest of this paper, we restrict our numerical investigation to the case of fixed packet size, and therefore

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
ACK_Timeout	300 μs
CTS_Timeout	300 μs

TABLE II

FHSS system parameters and additional parameters used to obtain numerical results

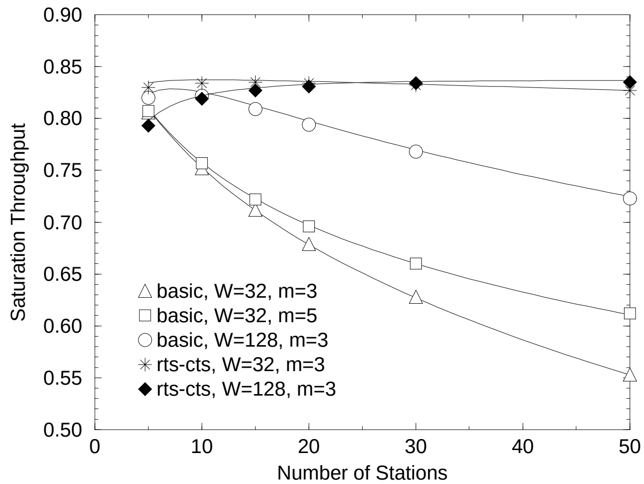

Fig. 6. *Saturation Throughput: analysis versus simulation*

	analysis	simulation
n=2, BAS	0.8473	0.846 $\pm$ 0.001
n=2, RTS	0.8198	0.817 $\pm$ 0.001
n=3, BAS	0.8368	0.835 $\pm$ 0.001
n=3, RTS	0.8279	0.823 $\pm$ 0.001

TABLE III

Analysis versus simulation: comparison for a very low number of stations - $W = 32, m = 3$

we will evaluate the performance of systems in which all stations operate either according to the Basic Access Mechanism or according to the RTS/CTS mechanism (i.e. never operating in the hybrid mode[3]).

V. MODEL VALIDATION

To validate the model, we have compared its results with that obtained with the 802.11 DCF simulator used in [9]. Ours is an event-driven custom simulation program, written in the C++ programming language, that closely follows all the 802.11 protocol details for each independently transmitting station. In particular, the simulation program attempts to emulate as closely as possible the real operation of each station, including propagation times, turnaround times, etc.

The values of the parameters used to obtain numerical results, for both the analytical model and the simulation runs, are summarized in table II. The system values are those specified for the FHSS (Frequency Hopping Spread Spectrum) PHY layer [3]. The channel bit rate has been assumed equal to 1 Mbit/s. The frame sizes are those defined by the 802.11 MAC specifications, and the PHY header is that defined for the FHSS PHY. The values of the ACK_Timeout and CTS_Timeout reported in table II, and used in the simulation runs only (our analysis neglects the effect of these timeouts) are not specified in the standard, and they have been set equal to 300 μs. This numerical value has been chosen as it is sufficiently long to contain a SIFS, the ACK transmission and a round trip delay.

Unless otherwise specified, we have used in the simulation runs a constant packet payload size of 8184 bits, which is about one fourth of the maximum MPDU size specified for the FHSS PHY, while it is the maximum MPDU size for the DSSS PHY.

Figure 6 shows that the analytical model is extremely accurate: analytical results (lines) practically coincide with the simulation results (symbols), in both Basic Access and RTS/CTS cases. All simulation results in the plot are obtained with a 95% confidence interval lower than 0.002. Negligible differences, well below 1%, are noted only for a small number of stations (results for the extreme case of as low as 2 and 3 stations are tabulated in table III).

[3] A detailed performance analysis of the hybrid mode requires to assume one or more suitable probability distribution functions for the packet's payload size, and also to determine the sensitivity of the throughput on the assumed distributions. Such a straightforward, but lengthy, study is out of the scopes of the present work.

VI. MAXIMUM SATURATION THROUGHPUT

The analytical model given above is very convenient to determine the maximum achievable saturation throughput. Let us rearrange (13) to obtain:

$$S = \frac{E[P]}{T_s - T_c + \frac{\sigma(1-P_{tr})/P_{tr} + T_c}{Ps}} \tag{25}$$

As T_s, T_c, $E[P]$, and σ, are constants, the throughput S is maximized when the following expression is maximized:

$$\frac{Ps}{(1-P_{tr})/P_{tr} + T_c/\sigma} = \frac{n\tau(1-\tau)^{n-1}}{T_c^* - (1-\tau)^n(T_c^* - 1)} \tag{26}$$

where $T_c^* = T_c/\sigma$ is the duration of a collision measured in slot time units σ. Taking the derivative of (26) with respect to τ, and imposing it equal to 0, we obtain, after some simplifications, the following equation:

$$(1-\tau)^n - T_c^*\{n\tau - [1-(1-\tau)^n]\} = 0 \tag{27}$$

Under the condition $\tau << 1$,

$$(1-\tau)^n \approx 1 - n\tau + \frac{n(n-1)}{2}\tau^2$$

BASIC ACCESS		
n	Max Throughput	Max Throughput Approx.
5	0.832827 (τ=0.022869)	0.832662 (τ=0.021426)
10	0.828279 (τ=0.010848)	0.828272 (τ=0.010713)
20	0.826111 (τ=0.005294)	0.826105 (τ=0.005357)
50	0.824841 (τ=0.002089)	0.824814 (τ=0.002143)
∞	0.823957	
RTS/CTS ACCESS		
n	Max Throughput	Max Throughput Approx.
5	0.838511 (τ=0.090399)	0.838436 (τ=0.097940)
10	0.837281 (τ=0.043712)	0.837129 (τ=0.048970)
20	0.836686 (τ=0.021520)	0.836490 (τ=0.024485)
50	0.836335 (τ=0.008532)	0.836110 (τ=0.009794)
∞	0.835859	

TABLE IV

Comparison between maximum throughput and throghput resulting from approximate solution (28) - the case $n = \infty$ is obtained from equation (31)

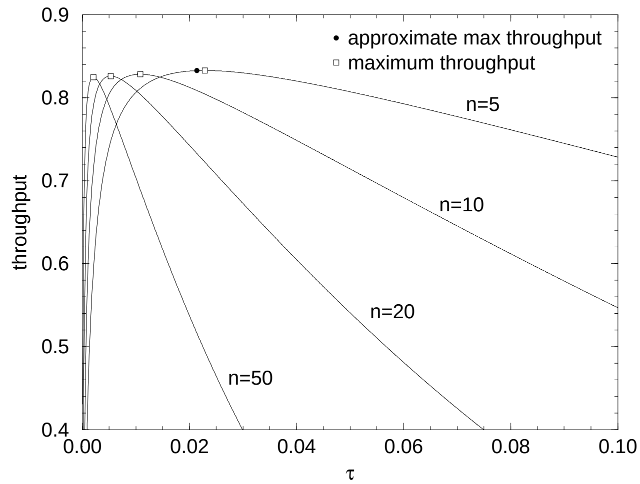

Fig. 7. *Throughput versus the transmission probability τ for the Basic Access method*

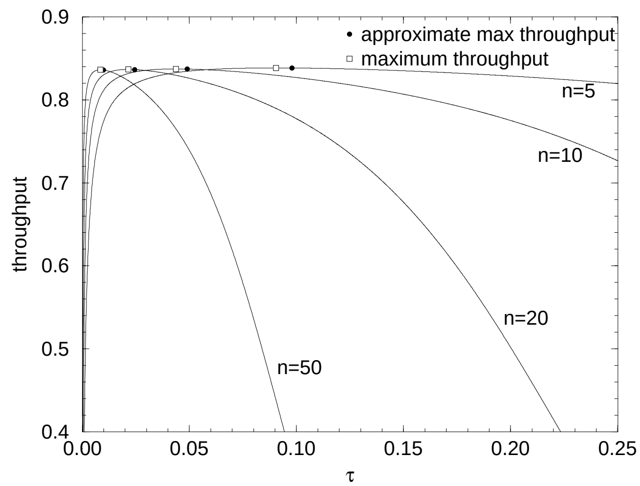

Fig. 8. *Throughput versus the transmission probability τ for the RTS/CTS mechanism*

holds, and yields the following approximate solution:

$$\tau = \frac{\sqrt{[n + 2(n-1)(T_c^* - 1)]/n} - 1}{(n-1)(T_c^* - 1)} \approx \frac{1}{n\sqrt{T_c^*/2}} \qquad (28)$$

Equation (27) and its approximate solution (28) are of fundamental theoretical importance. In fact, they allow to explicitly compute the optimal transmission probability τ that each station should adopt in order to achieve maximum throughput performance within a considered network scenario (i.e. number of stations n). In other words, they show that (within a PHY and an access mechanism, which determine the constant value T_c^*) maximum performance can be, in principle, achieved for every network scenario, through a suitable sizing of the transmission probability τ in relation to the network size.

However, equations (7) and (9) show that τ depends only on the network size and on the system parameters m and W. As n is not a directly controllable variable, the only way to achieve optimal performance is to employ adaptive techniques to tune the values m and W (and consequently τ) on the basis of the estimated value of n.

This problem has been specifically considered in [9] for the case of fixed backoff window size (i.e. $m = 0$). In such a case, τ is given by (8), and therefore the backoff window that maximizes the system throughput is readily found as

$$W_{\text{opt}} \approx n\sqrt{2T_c^*}$$

Refer to [9] for a large discussion related to the problem of estimating the value n.

Unfortunately, in the 802.11 standard, the values W and m are hardwired in the PHY layer details (see table I for the standardized values), and thus they cannot be made dependent on n. As a consequence of this lack of flexibility, the throughput in some network scenarios can be significantly lower than the maximum achievable.

Figures 7 and 8 show the maximum throughput theoretically achievable by the DCF protocol in both the cases of Basic Access and RTS/CTS mechanisms. The values reported in these figures have been obtained assuming the system parameters reported in table II. The figure reports also the different throughput values obtained in the case of exact and approximate solution for τ. As the maximum is very smooth, even a non negligible difference in the estimate of the optimal value τ leads to similar throughput values. The accuracy of the throughput obtained by the approximate solution is better testified by the numerical values reported in table IV. Note that the agreement is greater in the Basic Access case, as T_c is greater.

A surprising result is that the maximum throughput achievable by the Basic Access mechanism is very close to that achievable by the RTS/CTS mechanism. Moreover, the maximum throughput is practically independent of the number of stations in the wireless network. This is easily justified by noting that the throughput formula can be approximated as follows. Let $K = \sqrt{T_c^*/2}$, and let us use the approximate solution $\tau = 1/(nK)$. For n sufficiently large,

$$P_{tr} = 1 - (1-\tau)^n = 1 - \left(1 - \frac{1}{nK}\right)^n \approx 1 - e^{-1/K} \qquad (29)$$

	bits or μs	slot time units (σ=50 μs)
Packet Payload	8184	163.68
T_s^{bas}	8982	179.64
T_c^{bas}	8713	174.26
T_s^{rts}	9568	191.36
T_c^{rts}	417	8.34

TABLE V

Values T_s and T_c measured in bits and in 50 μs slot time units, for the considered system parameters, for both Basic and RTS/CTS access methods

$$P_s = \frac{n\tau(1-\tau)^{n-1}}{P_{tr}} \approx \frac{n}{(nK-1)(e^{1/K}-1)} \approx \frac{1}{K(e^{1/K}-1)} \tag{30}$$

The maximum achievable throughput S_{max} can thus be approximated as:

$$S_{max} = \frac{E[P]}{T_s + \sigma K + T_c\left(K(e^{1/K}-1)-1\right)} \tag{31}$$

which results to be independent of n. Using the numerical values of table V, we obtain $K = 9.334$ for the Basic Access mechanism, and $K = 2.042$ for the RTS/CTS mechanism. The resulting maximum throughput approximation values are reported in table IV under the label $n = \infty$.

An advantage of the RTS/CTS scheme is that the throughput is less sensitive on the transmission probability τ. In fact, we see from figures 7 and 8 (note the different x-axis scale) that a small variation in the optimal value of τ leads to a greater decrease in the throughput for the Basic Access case than for the RTS/CTS case. Hence, we expect (see quantitative results in the following section VII) a much lower dependence of the RTS/CTS throughput on the system engineering parameters with respect of the Basic Access throughput.

VII. Performance Evaluation

Unless otherwise specified, the following results have been obtained assuming the parameters reported in table II and, in particular, assuming a constant payload size $P = 8184$ bits.

Figure 6 shows that the throughput for the Basic Access scheme strongly depends on the number of stations in the network. In particular, the figure shows that, in most cases, the greater is the network size, the lower is the throughput. The only partial exception is the case $W = 128$. For such an initial contention window size, the throughput is comparable in networks with 5 to 10 stations, although it smoothly decreases as the network size increases. The same figure shows that performance impairment does not occur for the RTS/CTS mechanism when n increases. In fact, the throughput is practically constant for $W = 32$, and even increases with the number of mobile stations when $W = 128$.

To investigate the dependency of the throughput from the initial contention window size, W, we have reported in figures 9 and 10 the saturation throughput versus the value W for, respectively, the Basic Access and the RTS/CTS mechanisms. In both figures, we have assumed a number of backoff stages equal to 6, i.e. $CW_{\max} = 2^6 W$. The figures report four different network sizes, i.e. number of stations n equal to 5, 10, 20 and 50.

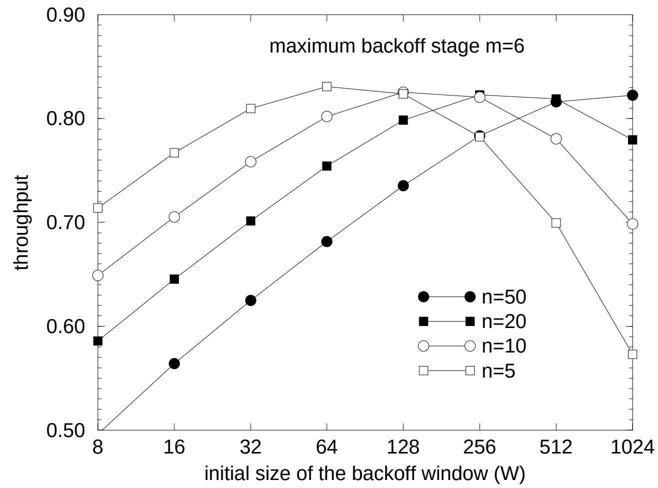

Fig. 9. *Saturation Throughput versus initial contention window size for the Basic Access mechanism*

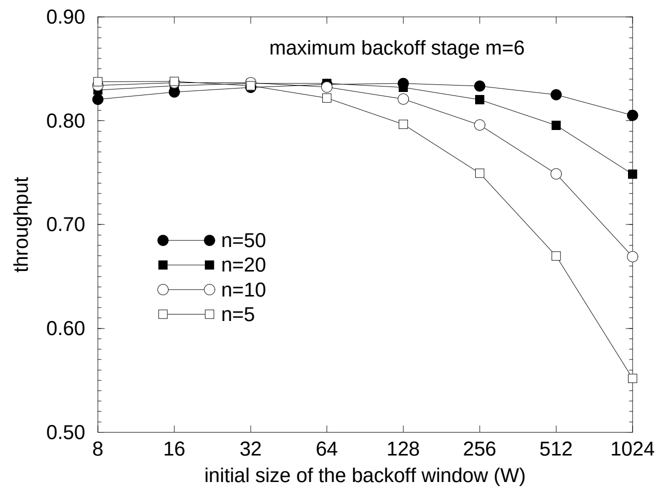

Fig. 10. *Saturation Throughput versus initial contention window size for the RTS/CTS mechanism*

Figure 9 shows that the throughput of the Basic Access mechanism highly depends on W, and the optimal value of W depends on the number of terminals in the network. For example, an high value of W (e.g. 1024) gives excellent throughput performance in the case of 50 contending stations, while it drastically penalizes the throughput in the case of small number (e.g. 5) of contending stations. This behavior is seen also in figure 10, where the RTS/CTS mechanism is employed. Large values of W may, in fact, limit the throughput of a single station, which, when alone in the channel is bounded by:

$$\frac{E[P]}{T_s + \sigma(W-1)/2} \tag{32}$$

where $E[P]$ and T_s are the average packet payload and the average channel holding time in case of successful transmission. Equation (32) is directly obtained from equation (13) of section IV-B by observing that, as there are no other stations which can collide with the considered one, the probability of success P_s is equal to 1. In addition, the probability P_{tr} that a transmission occurs on the channel is equal to the probability τ that the station transmits. Being the conditional collision probability p equal to 0, $P_{tr} = \tau$ is given by formula (8).

Of more practical interest is the case of small values of W, and particularly in correspondence of the values $W = 16$, 32, and 64 (i.e. those standardized for the three PHY - see table I). Figures 9 and 10 show that the two access mechanisms achieve a significantly different operation. In the case of the Basic Access mechanism, reported in figure 9, the system throughput increases as long as W gets closer to 64. Moreover, the throughput significantly decreases as the number of stations increases. On the contrary, figure 10 shows that the throughput obtained with the RTS/CTS mechanism is almost independent of the value $W \leq 64$, and, in this range, it is furthermore almost insensitive on the network size.

This surprising independence is quantitatively explained as follows. Dividing numerator and denominator of (13) by $P_{tr}P_s$, we obtain:

$$S = E[P] \Big/ \left[T_s + \sigma \frac{1 - P_{tr}}{P_s P_{tr}} + T_c \left(\frac{1}{P_s} - 1 \right) \right] \qquad (33)$$

The denominator of formula 33 expresses the average amount of time spent on the channel in order to observe the successful transmission of a packet payload. This time is further decomposed into three components.

T_s is the time spent in order to successfully transmit a packet. Table V reports the numerical values for T_s^{bas} and T_s^{rts}, computed according to equations (14) and (17), in the assumption of system and channel parameters of table II. The difference between T_s^{rts} and T_s^{bas} (586 bits) is the additional overhead introduced by the RTS/CTS mechanism.

The second term at the denominator of (33) does not depend on the access mechanism employed, and represents the amount of time the channel is idle, per successful packet transmission. In fact, $1/(P_{tr}P_s)$ is the average number of slot times spent on the channel in order to have a successful transmission. Of those slot times, a fraction $(1 - P_{tr})$ is empty, and each empty slot time lasts σ. The average number of idle slot times per packet transmission, i.e. $(1 - P_{tr})/(P_{tr}P_s)$, is plotted in figure 11 versus the network size, for three different values of the initial contention window W. We see that, for $W = 16$ and $W = 64$, the amount of idle slot times per packet transmission is very low, particularly when compared with the values T_s given in table V. This value becomes significant only when W gets greater (the case $W = 256$ is reported in the figure) and the number of stations in the network is small.

Finally, the third term at the denominator of (33) represents the time wasted on the channel because of collisions, per successful packet transmission. In fact, $1/P_s - 1$ is the average number of collided transmissions per each successful transmission, which is multiplied by T_c, i.e. the amount of time the channel is held by a collision. Table V shows that the the RTS/CTS mechanism significantly reduces the time spent during a collision, with respect to the Basic Access mechanism. This reduction is extremely effective when the system parameter W and the network size n lead to a large collision probability. This fact is graphically shown in figure 12. This figure reports the average amount of time spent in collisions, per successful packet transmission, normalized with respect to the value σ. It shows

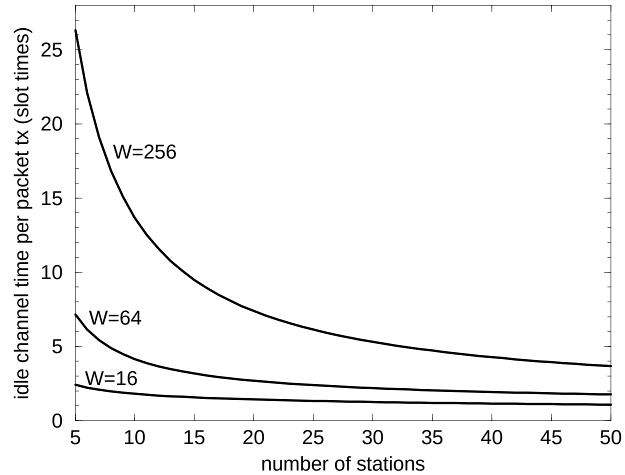

Fig. 11. *Average number of idle slot times per successful packet transmission*

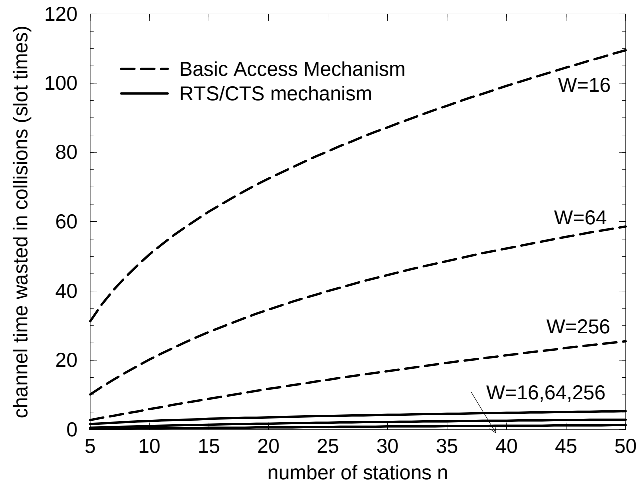

Fig. 12. *Average number of slot time units wasted on the channel because of packet collision, per successful packet transmission*

that, for the Basic Access mechanism, the amount of channel time wasted in collisions is extremely large for a small value W and a large number of stations in the network. Conversely, the additional amount of time wasted in collisions is negligible for the RTS/CTS mechanism, regardless of the values n and W. This explains the surprising constant RTS/CTS throughput in any practical system and network operation conditions.

Figure 13 shows that the dependence of the throughput from the maximum number m of backoff stages is marginal. The figure reports the cases of both Basic and RTS/CTS access schemes, with $W = 32$ (similar behaviour is observed for other values of the parameter W) and $n = 10, 50$. The points in the box indicate the throughput achieved when $m = 5$, i.e. in correspondence of the standardized engineering parameters of the DSSS PHY (table I). We see that the choice of m does not practically affect the system throughput, as long as m is greater than 4 or 5. The only case in which the throughput still grows, for m relatively large, is the Basic Access mechanism with a large network size.

Our model allows to obtain other measures of interest. The

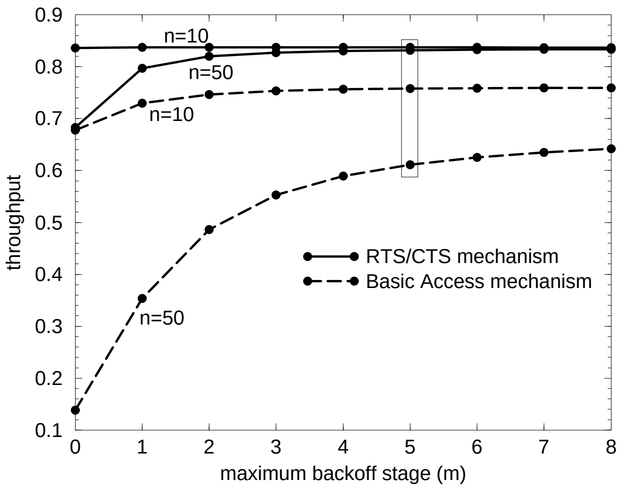

Fig. 13. *Throughput versus the maximum number of backoff stages:* $W = 32$

	slot time size (μs)				
	5	8	20	50	100
FH bas,10	.7105	.7101	.7088	**.7055**	.7000
rts,10	.8437	.8432	.8413	**.8367**	.8290
bas,50	.5658	.5657	.5652	**.5639**	.5618
rts,50	.8318	.8315	.8305	**.8278**	.8233
DS bas,10	.7659	.7654	**.7632**	.7577	.7488
rts,10	.8468	.8462	**.8435**	.8368	.8259
bas,50	.6134	.6133	**.6126**	.6108	.6079
rts,50	.8363	.8360	**.8347**	.8314	.8261
IR bas,10	.8171	**.8160**	.8120	.8021	.7862
rts,10	.8490	**.8479**	.8436	.8329	.8158
bas,50	.6696	**.6693**	.6683	.6656	.6613
rts,50	.8408	**.8404**	.8387	.8345	.8277

TABLE VI

Dependence of the saturation throughput on the slot time

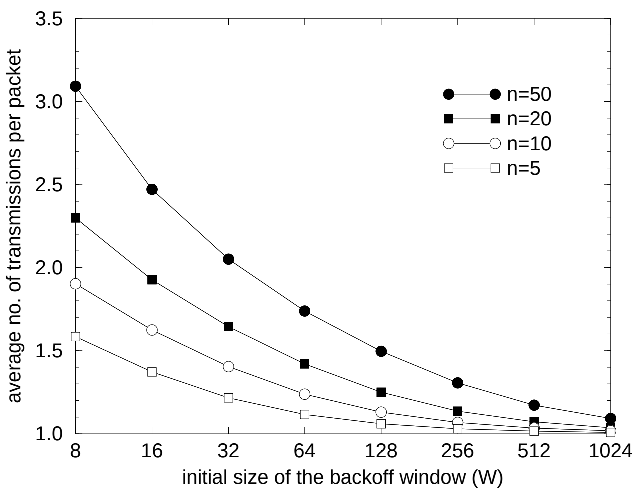

Fig. 14. *Average number of transmissions per packet*

conditional collision probability p is the probability, seen by the station, that its transmitted packet collides. Owing to the model's key assumption of independence at each retransmission, the average number of transmissions that each station must perform in order to successfully complete a packet transmission is given by $\frac{1}{1-p}$. This value is reported in figure 14, obtained with the same system parameters of figures 9 and 10. Figure 14 shows that the number of transmissions per packet significantly increases as the initial backoff window W reduces, and as the network size n increases.

At a first glance, it might seem that the throughput performance of the 802.11 protocol strongly depends on the slot time size σ. In particular, the lower is σ, the better is the expected performance. Instead, we note that, as far as saturation throughput performance is concerned, its dependence on the slot time size is only marginal. Table VI reports results for three different system configurations corresponding to the different PHYs. Results are obtained for both Basic Access and RTS/CTS mechanism, and for two different network sizes of 10 and 50 stations. Columns in boldface type correspond to the standardized slot time length for the related PHY. The marginal dependence of the throughput on the slot time size σ is related to the fact, commented above by means of equation (33) and figure 11, that the number of idle slot times per packet transmission is extremely small. A change of σ has the only effect to multiply by a constant value the amount of idle channel time per packet transmission. However, for any practical value of σ and W, the amount of idle channel time remains marginal with respect to the time spent in transmission and collision. This result is of fundamental importance for the future development of higher bit rate physical layer recommendations, as the slot time size is difficultly scalable.

Finally, let us add some considerations regarding the dependence of the access method on the packet length. It is often qualitatively stated that the RTS/CTS mechanism is effective when the packet size increases. This is justified in figure 15. This figure reports the system throughput for both Basic Access and RTS/CTS cases, for two different network sizes ($n = 10$ and $n = 50$), and for three different configuration parameters, referred to as FH, DS and IR, corresponding to the three PHY's reference values $CW_{\min}$, $CW_{\max}$ and Slot time size σ reported in table I. It is no more a surprise that the RTS/CTS mechanism achieves very similar performance in all the considered cases. This is due to the fact that the throughput performance marginally depends on the slot time, as shown in table VI, and on the fact that the RTS/CST scheme is negligibly dependent on the network size and on the minimum contention window size.

In the assumption of fixed packet payload size, it is very easy to quantify the threshold value for the packet size over which it is convenient to switch to the RTS/CTS mechanism. In fact, let us indicate with S^{bas} and S^{rts} the throughput achieved respectively by the Basic Access and RTS/CTS mechanism in the same system parameters and network size conditions. From equation (33), the inequality

$$S^{rts} > S^{bas}$$

implies that

$$T_s^{rts} - T_s^{bas} < \left(T_c^{bas} - T_c^{rts}\right)\left(\frac{1}{P_s} - 1\right) \tag{34}$$

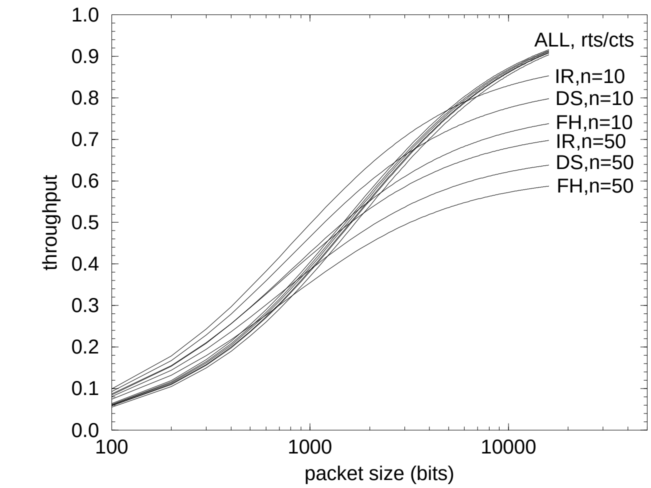

Fig. 15. *Throughput versus packet size for the standardized configuration parameters*

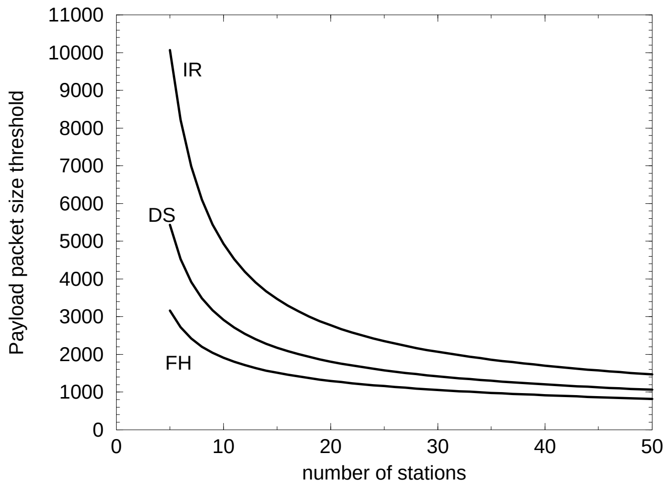

Fig. 16. *Packet payload threshold over which the RTS/CTS mechanism is advantageous*

Let now $O_{rts} = T_s^{rts} - T_s^{bas}$ the overhead introduced by the RTS/CTS mechanism, and let $O_h = H - RTS$ be the extra length of the packet header with respect of the RTS frame size (according to the values of table II, $O_{rts} = 586$ bits, and $O_h = 112$ bits). Indicating the packet payload with the variable P, condition (34) yields:

$$P > \frac{P_s O_{rts}}{1 - P_s} - O_h \tag{35}$$

The threshold value over which it is convenient to switch to the RTS/CTS scheme is plotted versus the network size in figure 16, for the three possible sets of parameters specified for the different PHYs. This figure shows that the threshold is highly dependent on the PHY employed. This is not a consequence of the different slot time size σ, which does not affect formula (35). Instead, it is a direct consequence of the different initial contention window sizes W adopted (see table I). The lower the value W, the greater is the performance impairment of the Basic Access scheme (see figure 9), and the greater (and thus for more packet size cases, as shown in figure 15) is the advantage of the RTS/CTS scheme.

Moreover, figure 16 runs counter to the “known” fact that the RTS/CTS mechanism should be employed when the packet size exceeds a given (meaning *fixed*) threshold. Instead, it shows that such a threshold strongly depends on the network size, and particularly it significantly decreases when the number of stations in the network increases. For example, in the case of 50 stations, the threshold is equal to about 1470 bits for the Infrared PHY, while it is as low as 820 bits for the Frequency Hopping PHY. The same threshold raises, respectively, to about 10065 bits and 3160 bits when the network is composed by 5 stations only.

VIII. Conclusions

In this paper, we have presented a simple analytical model to compute the saturation throughput performance of the 802.11 Distributed Coordination Function. Our model assumes a finite number of terminals and ideal channel conditions. The model is suited for any access scheme employed, i.e. for both Basic Access and RTS/CTS Access mechanisms, as well as for a combination of the two. Comparison with simulation results shows that the model is extremely accurate in predicting the system throughput.

Using the proposed model, we have evaluated the 802.11 throughput performance. We have shown that performance of the Basic Access method strongly depends on the system parameters, mainly minimum contention window and number of stations in the wireless network. Conversely, performance is only marginally dependent on the system parameters when the RTS/CTS mechanism is considered.

The RTS/CTS mechanism has proven its superiority in most of the cases. Notable is the advantage of the RTS/CTS scheme in large network scenarios, even with fairly limited packet sizes. When the capability of the RTS/CTS scheme to cope with hidden terminals is accounted, we conclude that this access method should be used in the majority of the practical cases.

Acknowledgements

The author wishes to thank the anonymous refereers for their helpful comments that have significantly improved the quality of the presentation.

References

[1] K. Pahlavan, A. H. Levesque, “Wireless Data Communications”, Proc. of the IEEE, Vol. 82, No. 9, September 1994, pp. 1398-1430.

[2] A. De Simone, S. Nanda, “Wireless Data: Systems, Standards, Services”, Journal of Wireless Networks, Vol. 1, No. 3, February 1996, pp. 241-254.

[3] P802.11, IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, November 1997.

[4] L. Kleinrock, F. Tobagi, “Packet Switching in radio channels, part II - the Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy Tone Solution”, IEEE Trans. Comm., Vol. COM-23, No. 12, December 1975, pp. 1417-1433.

[5] H. S. Chhaya, S. Gupta, “Performance Modeling of Asynchronous Data Transfer Methods of IEEE 802.11 MAC Protocol”, Wireless Networks, Vol. 3, 1997, pp. 217-234.

[6] K. C. Huang, K. C. Chen, “Interference Analysis of Nonpersistent CSMA with Hidden Terminals in Multicell Wireless Data Networks”, Proc. of IEEE PIMRC ’95, Toronto, Canada, September 1995, pp. 907-911.

[7] B. P. Crow, “Performance evaluation of the IEEE 802.11 Wireless Local Area Network Protocol”, Master thesis, Dept. of Electrical and Computer Engineering, University of Arizona, 1996.

[8] J. Weinmiller, M. Schlager, A. Festag, A. Wolisz, “Performance Study of Access Control in Wireless LANs IEEE 802.11 DFWMAC and ETSI RES 10 HIPERLAN”, Mobile Networks and Applications, Vol. 2, 1997, pp. 55-67.

[9] G. Bianchi, L. Fratta, M. Oliveri, “Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs”, Proc. PIMRC 1996, October 1996, Taipei, Taiwan, pp. 392-396.

[10] F. Cali’, M. Conti, E. Gregori, “IEEE 802.11 Wireless LAN: Capacity Analysis and Protocol Enhancement”, Proc. of INFOCOM ’98, March 1998, S. Francisco.

[11] T. S. Ho, K. C. Chen, “Performance Analysys of IEEE 802.11 CSMA/CA Medium Access Control Protocol”, Proc. PIMRC 1996, October 1996, Taipei, Taiwan, pp. 407-411.

[12] G. Bianchi “IEEE 802.11 - Saturation Throughput Analysis”, IEEE Communications Letters, Volume 2, No. 12, December 1998, pp. 318-320.

[13] D. Bertsekas, R. Gallager, *Data Networks*, Prentice Hall, Inc., 1987.