

Optimization of Bandwidth and Energy Consumption in Wireless Local Area Networks

Marco Conti and Enrico Gregori

Consiglio Nazionale delle Ricerche
IIT Institute, Via G. Moruzzi, 1
56124 Pisa, Italy
{marco.conti, enrico.gregori}@iit.cnr.it

Abstract. In the recent years the proliferation of portable computers, handheld digital devices, and PDAs has led to a rapid growth in the use of wireless technologies for the Local Area Network (LAN) environment. Beyond supporting wireless connectivity for fixed, portable and moving stations within a local area, the wireless LAN (WLAN) technologies can provide a mobile and ubiquitous connection to the Internet information services. The design of WLANs has to concentrate on bandwidth consumption because wireless networks deliver much lower bandwidth than wired networks, e.g., 2-11 Mbps [1] versus 10-150 Mbps [2]. In addition, the finite battery power of mobile computers represents one of the greatest limitations to the utility of portable computers [3], [4]. Hence, a relevant performance-optimization problem is the balancing between the minimization of battery consumption, and the maximization of the channel utilization. In this paper, we study bandwidth and energy consumption of the IEEE 802.11 standard, i.e., the most mature technology for WLANs. Specifically, we derived analytical formulas that relate the protocol parameters to the maximum throughput and to the minimal energy consumption. These formulas are used to define an effective method for tuning at run time the protocol parameters.

1 Introduction

In IEEE802.11 WLANs [1], the Medium Access Control (MAC) protocol is the main element that determines the efficiency in sharing the limited resources of the wireless channel. The MAC protocol coordinates the transmissions of the network stations and, at the same time, manages the congestion situations that may occur inside the network. The congestion level in the network negatively affects both the channel utilization (i.e. the fraction of channel bandwidth used from successfully transmitted messages), and the energy consumed to successfully transmit a message. Specifically, each collision reduces the channel bandwidth and the battery capacity available for successful transmissions. To decrease the collision probability, the IEEE 802.11 protocol uses a CSMA/CA protocol based on a truncated binary exponential backoff scheme that doubles the backoff window after each collision [1], [2]. However, the

time spreading of the accesses that the standard backoff procedure accomplishes has a negative impact on both the channel utilization, and the energy consumption. Specifically, the time spreading of the accesses can introduce large delays in the message transmissions, and energy wastages due to the carrier sensing. Furthermore, the IEEE 802.11 policy has to pay the cost of collisions to increase the backoff time when the network is congested.

In [5], [6] and [7], given the binary exponential backoff scheme adopted by the standard, solutions have been proposed for a better uniform distribution of accesses. The most promising direction for improving the backoff protocols is to adopt feedback-based tuning algorithms that exploit the information retrieved from the observation of the channel status [8], [9], [10]. For the IEEE 802.11 MAC protocol, some authors have proposed an adaptive control of the network congestion by investigating the number of users in the system [11], [12], [13]. This investigation could result expensive, difficult to obtain, and subject to significant errors, especially in high contention situations [12]. Distributed (i.e. independently executed by each station) strategies for power saving have been proposed and investigated in [14], [15]. Specifically, in [14] the authors propose a power controlled wireless MAC protocol based on a fine-tuning of network interface transmitting power. [15] extends the algorithm presented in [7] with power saving features.

This paper presents and evaluates a distributed mechanism for the contention control in IEEE 802.11 WLANs that extends the standard access mechanism without requiring any additional hardware. Our mechanism dynamically adapts the backoff window size to the current network contention level, and guarantees that an IEEE 802.11 WLAN *asymptotically* achieves its optimal channel utilization and/or the minimum energy consumption. For this reason we named our mechanism Asymptotical Optimal Backoff (AOB). To tune the parameters of our mechanism we analytically studied the bandwidth and energy consumption of the IEEE 802.11 standard, and we derived closed formulas that relate the protocol backoff parameters to the maximum throughput and to the minimal energy consumption.

Our analytical study of IEEE 802.11 performance is based on a p -persistent model of the IEEE 802.11 protocol [11], [15]. This protocol model differs from the standard protocol only in the selection of the backoff interval. Instead of the binary exponential backoff used in the standard, the backoff interval of the p -persistent IEEE 802.11 protocol is sampled from a geometric distribution with parameter p . In [11], it was shown that a p -persistent IEEE 802.11 closely approximates the standard protocol.

In this paper, we use the p -persistent model to derive analytical formulas for the IEEE802.11 protocol capacity and energy consumption. From these formulas we compute the p value (i.e., the average backoff window size) corresponding to maximum channel utilization (*optimal capacity* p , also referred to as p_{opt}^C), and the p value corresponding to minimum energy consumption (*optimal energy* p , also referred to as p_{opt}^E). The properties of the optimal operating points (both from the efficiency and power saving standpoint) are deeply investigated. In addition, we also provide closed formulas for the optimal p values. These formulas are used by AOB to dynamically tune the WLAN backoff parameters either to maximize WLAN efficiency, or to minimize WLAN energy consumption.

Via simulation, we compared the performance figures of the IEEE 802.11 protocol with or without our backoff tuning algorithm. Results obtained indicate that the enhanced protocol significantly improves the standard protocol, and (in all the configurations analyzed) produces performance figures very close to the theoretical limits.

2. IEEE 802.11

The IEEE 802.11 standard defines a MAC layer and a Physical Layer for WLANs. The MAC layer provides both contention-based and contention-free access control on a variety of physical layers. The standard provides 2 Physical layer specifications for radio (Frequency Hopping Spread Spectrum, Direct Sequence Spread Spectrum), operating in the 2,400 - 2,483.5 MHz band (depending on local regulations), and one for infrared. The Physical Layer provides the basic rates of 1 Mbit/s and 2 Mbit/s. Two projects are currently ongoing to develop higher-speed PHY extensions to 802.11 operating in the 2.4 GHz band (project 802.11b, handled by TGb) and in the 5 GHz band (project 802.11a handled by TGa), see [16] and [17].

The basic access method in the IEEE 802.11 MAC protocol is the *Distributed Coordination Function* (DCF) which is a *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA) MAC protocol. In addition to the DCF, the IEEE 802.11 also incorporates an alternative access method known as the *Point Coordination Function* (PCF) - an access method that is similar to a polling system and uses a point coordinator to determine which station has the right to transmit. In this section we only present the aspects of the DCF access method relevant for the scope of this paper. For the detailed explanation of the IEEE 802.11 standard we address interested readers to [1], [2], [18].

The DCF access method is based on a CSMA/CA MAC protocol. This protocol requires that every station, before transmitting, performs a Carrier Sensing activity to determine the state of the channel (idle or busy). If the medium is found to be idle for an interval that exceeds the *Distributed InterFrame Space* (DIFS), the station continues with its transmission. If the medium is busy, the transmission is deferred until the ongoing transmission terminates, and a Collision Avoidance mechanism is adopted. The IEEE 802.11 Collision Avoidance mechanism is a *Binary Exponential Backoff* scheme [1], [19], [20], [21]. According to this mechanism, a station selects a random interval, named *backoff interval*, that is used to initialize a *backoff counter*.

The backoff counter is decreased as long as the channel is sensed idle, stopped when a transmission is detected on the channel, and reactivated when the channel is sensed idle again for more than a DIFS. A station transmits when its backoff counter reaches zero.

When, the channel is idle the time is measured in constant length units (*Slot_Time*) indicated as slots in the following. The backoff interval is an integer number of slots and its value is uniformly chosen in the interval $(0, CW_Size-1)$, where *CW_Size* is, in each station, a local parameter that defines the current station *Contention Window* size. Specifically, the backoff value is defined by the following expression [1]:

$$Backoff_Counter = INT(Rnd() \cdot CW_Size) ,$$

where $Rnd()$ is a function which returns pseudo-random numbers uniformly distributed in $[0..1]$.

The Binary Exponential Backoff is characterized by the expression that gives the dependency of the CW_Size parameter by the number of *unsuccessful transmission attempts* (N_A) already performed for a given frame. In [1] it is defined that the first transmission attempt for a given frame is performed adopting CW_Size equal to the minimum value CW_Size_min (assuming low contention). After each unsuccessful (re)transmission of the same frame, the station doubles CW_Size until it reaches the maximal value fixed by the standard, i.e. CW_Size_MAX , as follows:

$$CW_Size(N_A) = \min(CW_Size_MAX, CW_Size_min \cdot 2^{(N_A-1)}) .$$

Positive acknowledgements are employed to ascertain a successful transmission. This is accomplished by the receiver (immediately following the reception of the data frame) which initiates the transmission of an acknowledgement frame (ACK) after a time interval *Short Inter Frame Space* (SIFS), which is less than DIFS.

If the transmission generates a collision¹, the CW_Size parameter is doubled for the new scheduling of the retransmission attempt thus obtaining a further reduction of contention.

The increase of the CW_Size parameter value after a collision is the reaction that the 802.11 standard DCF provides to make the access mechanism adaptive to channel conditions.

2.1 IEEE 802.11 Congestion Reaction

Fig. 1 shows simulation data regarding the channel utilization of a standard 802.11 system running in DCF mode, with respect to the contention level, i.e. the number of active stations with continuous transmission requirements. The parameters adopted in the simulation, presented in Table 1, refer to the Frequency Hopping Spread Spectrum implementation [1].

Fig. 1 plots the channel utilization versus the number of active stations obtained assuming asymptotic conditions, i.e. all the stations have always a frame to transmit. By analyzing the behavior of the 802.11 DCF mechanism some problems could be identified. Specifically, the results presented in the figure show that the channel utilization is negatively affected by the increase in the contention level.

These results can be explained as, in the IEEE 802.11 backoff algorithm, a station selects the initial size of the Contention Window by assuming a low level of congestion in the system. This choice avoids long access delays when the load is light. Unfortunately, this choice causes efficiency problems in bursty arrival scenarios, and in

¹ A collision is assumed whenever the ACK from the receiver is missing

congested systems, because it concentrates the accesses in a reduced time window, and hence it may cause a high collision probability.

Table 1: System's physical parameters

parameter	value
Number of Stations (M)	variable from 2 to 200
CW_Size_min	16
CW_Size_MAX	1024
Channel transmission rate	2 Mb/s
Payload size	Geometric distribution (parameter q)
Acknowledgement size	200 μ sec (50 Bytes)
Header size	136 μ sec (34 Bytes)
SlotTime (t_{slot})	50 μ sec
SIFS	28 μ sec
DIFS	128 μ sec
Propagation time	< 1 μ sec

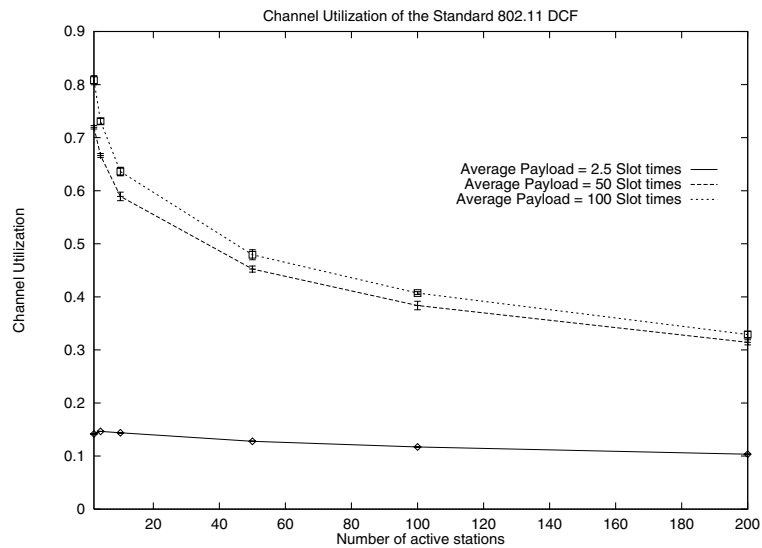


Fig. 1.: Channel utilization of the IEEE 802.11 DCF access scheme

In high-congestion conditions each station reacts to the contention on the basis of the collisions so far experienced while transmitting a frame. Every station performs its attempts blindly, with a late collision reaction performed (increasing CW_Size). Each increase of the CW_Size is obtained paying the cost of a collision. It is worth noting that, as a collision detection mechanism is not implemented in the IEEE 802.11, a collision implies that the channel is not available for the time required to transmit the longest colliding packet. Furthermore, after a successful transmission the CW_Size is

set again to the minimum value without maintaining any knowledge of the current contention level. To summarize the IEEE 802.11 backoff mechanism has two main drawbacks: *i*) the increase of the CW_Size is obtained paying the cost of a collision, and *ii*) after a successful transmission no state information indicating the actual contention level is maintained.

3. Low-Cost Dynamic Tuning of the Backoff Window Size

The drawbacks of the IEEE 802.11 backoff algorithm, explained in the previous section, indicate the direction for improving the performance of a random access scheme, by exploiting the information on the current network congestion level that is already available at the MAC level. Specifically, the utilization rate of the slots (*Slot Utilization*) observed on the channel by each station is used as a simple and effective estimate of the channel congestion level. The estimate of the Slot Utilization must be frequently updated. For this reason in [7] it was proposed an estimate that has to be updated by each station in every *Backoff interval*, i.e., the defer phase that precedes a transmission attempt.

A simple and intuitive definition of the slot utilization estimate is then given by:

$$Slot_Utilization = \frac{Num_Busy_Slots}{Num_Available_Slots} ,$$

where Num_Busy_Slots is the number of slots, in the backoff interval, in which a transmission attempt starts, hereafter referred as *busy slots*. A transmission attempt can be either a successful transmission, or a collision; and $Num_Available_Slots$ is the total number of slots available for transmission in the backoff interval, i.e. the sum of idle and busy slots.

In the 802.11 standard mechanism every station performs a Carrier Sensing activity and thus the proposed slot utilization (S_U) estimate is simple to obtain. The information required to estimate S_U is already available to an IEEE 802.11 station, and no additional hardware is required.

The current S_U estimate can be utilized by each station to evaluate (before trying a “blind” transmission) the opportunity to perform, or to defer, its scheduled transmission attempt. If a station knows that the probability of a successful transmission is low, it should defer its transmission attempt. Such a behavior can be achieved in an IEEE 802.11 network by exploiting the DCC mechanism proposed in [7]. According to DCC, each IEEE 802.11 station performs an additional control (beyond carrier sensing and backoff algorithm) before any transmission attempt. This control is based on a new parameter named *Probability of Transmission* $P_T(\dots)$ whose value is dependent on the current contention level of the channel, i.e., S_U . The heuristic formula proposed in [7] for $P_T(\dots)$ is:

$$P_T(S_U, N_A) = 1 - S_U^{N_A} ,$$

where, by definition, S_U assumes values in the interval $[0,1]$, and N_A is the number of attempts already performed by the station for the transmission of the current frame.

The N_A parameter is used to partition the set of active stations in such a way that each subset of stations is associated with a different level of privilege to access the channel. Stations that have performed several unsuccessful attempts have the highest transmission privilege [7].

The P_T parameter allows to filter the transmission accesses. When, according to the standard protocol, a station is authorized to transmit (i.e., backoff counter is zero and channel is idle) in the protocol extended with the Probability of Transmission, a station will perform a real transmission with probability P_T , otherwise (i.e. with probability $1-P_T$) the transmission is re-scheduled as a collision would have occurred, i.e. a new backoff interval is sampled.

To better understand the relationship between the P_T definition and the network congestion level, we can observe the Fig. 2.

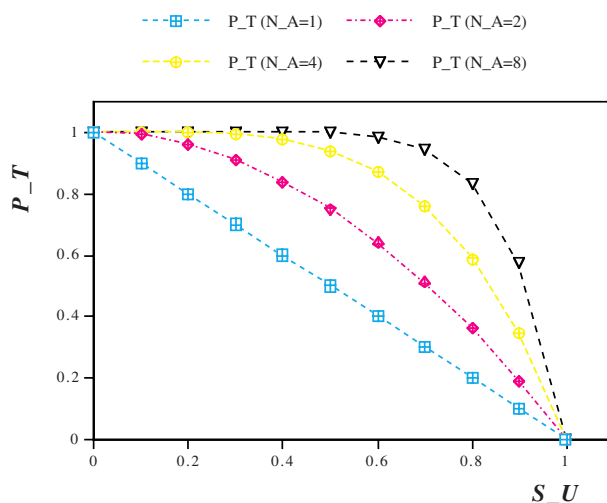


Fig. 2: DCC Probability of Transmission

In the figure we show the P_T curves (for users with different N_A) with respect to the estimated S_U values. Assuming a slot utilization near to zero, we can observe that each station, independently by its number of performed attempts, obtains a Probability of Transmission near to 1. This means that the proposed mechanism has no effect on the system, and each user performs its accesses just like in the standard access scheme, without any additional contention control. This point is significant as it implies the absence of overhead introduced in low-load conditions. The differences in the users behavior as a function of their levels of privilege (related to the value of the N_A parameter) appear when the slot utilization grows. For example, assuming a slot utilization near to 1, say 0.8, we observe that the stations with the highest N_A value

obtains a Probability of Transmission close to 1 while stations at the first transmission attempt transmit with a probability equal to 0.2.

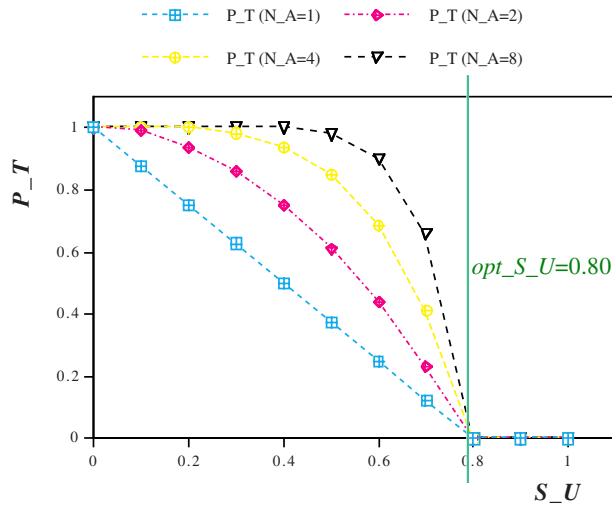


Fig. 3: Generalized Probability of Transmission

It is worth noting a property of the DCC mechanism: the slot utilization of the channel never reaches the value 1. Assuming S_U near or equal to 1, the DCC mechanism reduces the Probabilities of Transmission for all stations close to zero thus reducing the network contention level. This effect was due to the P_T definition, and in particular to the explicit presence of the upper bound 1 for the slot utilization estimate. The DCC choice to use 1 as the asymptotical limit for the S_U is heuristic, and it is induced by the lack of the knowledge of which is the optimal upper bound for the S_U value (opt_S_U) to guarantee the maximum channel utilization. It is worth noting that, if opt_S_U is known, the P_T mechanism can be easily tuned to guarantee that the maximum channel utilization is achieved. Intuitively, if the slot-utilization boundary value (i.e. the value one for DCC) would be replaced by the opt_S_U value, we reduce all the probabilities of transmission to zero in correspondence of slot utilization values greater or equal to the opt_S_U . This can be achieved by generalizing the definition for the Probability of Transmission:

$$P_T(opt_S_U, S_U, N_A) = 1 - \min\left(1, \frac{S_U}{opt_S_U}\right)^{N_A} \quad (1)$$

Specifically, by applying this definition of the transmission probability we obtain the P_T curves shown in Fig. 3. These curves have been obtained by applying the generalized P_T definition with $opt_S_U=0.80$. As expected the curves indicate the effectiveness of the generalized P_T definition in providing a S_U bounding that changes with the opt_S_U value.

The generalized Probability of Transmission provides an effective tool for controlling, in an optimal way, the congestion inside a IEEE 802.11 WLAN provided that

the opt_S_U value is known. In the next section we will show how this optimal value can be computed. The computation is based on the exploitation of the results derived in [11].

4. Protocol Model

Consider a system with M active stations accessing a slotted multiple-access channel. The random access protocols for controlling this channel can be either a S-Aloha or a p -persistent CSMA algorithm. In the first case (Aloha) the stations transmit constant-length messages with length l that exactly fits in a slot. In the latter case (CSMA), the message length is a random variable L , with average l . Hereafter, to simplify the presentation we will assume that L values always correspond to an integer number of slots. In both cases, when a transmission attempt is completed (successfully or with a collision), each network station with packets ready for transmission (hereafter *backlogged station*) will start a transmission attempt with probability p .

To derive a closed formula of the channel utilization, say ρ , for the p -persistent CSMA protocols we observe the channel between two consecutive successful transmissions. Let us denote with t_i the time between the $(i-1)$ th successful transmission and the i th successful transmission, and with s_i the duration of the i th successful transmission. Hence, the channel utilization is simply obtained:

$$\rho = \lim_{n \rightarrow \infty} \frac{s_1 + s_2 + \dots + s_n}{t_1 + t_2 + \dots + t_n}. \quad (2)$$

By dividing both numerator and denominator of Equation (3) by n , after some algebraic manipulations, it follows that:

$$\rho = \frac{E[S]}{E[T]}, \quad (3)$$

where $E[S]$ is the average duration of a successful transmission, and $E[T]$ is the average time between two successful transmissions. The $E[T]$ formula is straightforwardly obtained by considering the behavior of a CSMA protocol. Specifically, before successfully transmitting a message, a station will experience $E[N_c]$ collisions, on average. Furthermore, each transmission is preceded by an idle time, say $E[Idle]$, during which a station listens to the channel. Therefore, we can write:

$$E[T] = (E[N_c] + 1) \cdot E[Idle] + E[N_c] \cdot E[Coll | Coll] + E[S], \quad (4)$$

where $E[Coll | Coll]$ is the average duration of a collision, given that a collision occurs. Finally, the unknown quantities in Equation(4) are derived in Lemma 1 under the following assumptions: *i*) the channel idle time is divided in fixed time lengths, say t_{slot} ; *ii*) all the stations adopt a p -persistent CSMA algorithm to access the channel; *iii*) all the stations operate in saturation conditions, i.e., they have always a mes-

sage waiting to be transmitted; *iv*) the message lengths, say l_i , are random variables identically and independently distributed.

Lemma 1. In a network with M stations operating according to assumptions i) to iv) above, by denoting with $\Pr\{L \leq h\}$ the probability that the time occupied by the transmission of the message L is less or equal to $h \cdot t_{slot}$:

$$E[Idle] = \frac{(1-p)^M}{1-(1-p)^M} \quad , \quad (5)$$

$$E[N_c] = \frac{1-(1-p)^M}{Mp(1-p)^{M-1}} - 1 \quad , \quad (6)$$

$$E[Coll | Coll] = \quad (7)$$

$$= \frac{t_{slot}}{1 - [(1-p)^M + Mp(1-p)^{M-1}]} \cdot \sum_{h=1}^{\infty} h \left\{ [1 - (1 - \Pr\{L \leq h\})p]^M - [1 - (1 - \Pr\{L < h\})p]^M - M[\Pr\{L \leq h\} - \Pr\{L < h\}]p(1-p)^{M-1} \right\} \quad ,$$

Proof. The proof is obtained with standard probabilistic considerations (see [11]).

◇

$E[S]$ is a constant independent of the p value, but dependent only on the message length distribution. As it appears from Equation (4) and Lemma 1, the channel utilization is a function of the protocol parameter p , the number M of active stations and the message length distribution. The protocol capacity, say ρ_{MAX} , is obtained by finding the p value, say p_{opt}^C , that maximizes Equation (3). Since $E[S]$ is a constant, the utilization-maximization problem is equivalent to the following minimization problem:

$$\min_{p \in [0,1]} \left\{ (E[N_c] + 1) \cdot E[Idle] + E[N_c] \cdot E[Coll | Coll] \right\} \quad , \quad (8)$$

For instance, for the Slotted-ALOHA access scheme the p_{opt}^C value is calculated by considering in Equation (8) constant messages with length 1 t_{slot} . By solving Equation (8) we obtain that $p_{opt}^C = 1/M$ and $\rho_{MAX} \xrightarrow{M \gg 1} e^{-1}$. However, from Equation (8),

it is not possible to derive a general exact closed formula for the p_{opt}^C value in the general message-length distribution case.

Due to the choice of $l=1$, collisions and successful transmissions always last a single t_{slot} . However, for a *p-persistent* CSMA protocol with a general message-length distribution, the protocol capacity depends on the message length, hence, the p_{opt}^C also depends on the traffic characteristics. The p_{opt}^C value can be computed for any network configuration by numerically maximizing Equation (3). Hence, if a station has an exact knowledge of the network configuration it is possible to tune its backoff algorithm to achieve a channel utilization close to the protocol capacity. Unfortunately, in a real case a station does not have this knowledge but it can only estimate it. Using this estimate, the station can in principle compute at run time the p_{opt}^C value. However, the p_{opt}^C computation obtained by the numerical maximization of Equation (3) is too complex to be executed at run-time. Furthermore, the numerical maximization of Equation (3) does not provide a closed formula for the optimal working conditions, and hence it does not provide any insight on the optimal protocol behavior. Hence, the Equation (3) can be adopted to derive the optimal capacity state in an off-line analysis, but it would be convenient to derive a simpler relationship to provide an approximation of the p_{opt}^C value to guarantee a quasi-optimal capacity state.

4.1 Protocol Capacity: An Approximate Analysis

Equation (8) can be adopted to derive the optimal capacity state in an off-line analysis, but it is important to derive a simpler relationship to provide an approximation of the p value corresponding to optimal capacity, i.e. p_{opt}^C . In [22], the network operating point in which the time wasted on idle periods is equal to the time spent on collisions was identified as the condition to obtain the maximum protocol capacity in ALOHA and CSMA protocols where all messages require one time slot for transmission. In [12] the same network operating point was proposed to determine a quasi-optimal capacity state in the *p-persistent* IEEE 802.11 protocol where the message length was sampled from a geometric distribution. This condition can be expressed with the following relationship:

$$E[Idle_p] = E[Coll | N_{tr} \geq 1] \quad , \quad (9)$$

where $E[Coll | N_{tr} \geq 1]$ is the average collision length given that at least a transmission occurs.

Equation (9) was proposed in previous papers using heuristic considerations. Specifically, it is straightforward to observe that $E[Idle_p]$ is a decreasing function of the p value, whereas $E[Coll | N_{tr} \geq 1]$ is an increasing function of the p value. Therefore, Equation (9) suggests that a quasi-optimal capacity state is achieved when each station behaves in such a way to balance these two conflicting costs.

Hereafter, first, we give an analytical justification and a numerical validation of the above heuristic. Second, we provide closed formulas for the p_{opt}^C , and we show that the optimal capacity state, given the message length distribution, is characterized by an invariant figure: the $M \cdot p_{opt}^C$ product.

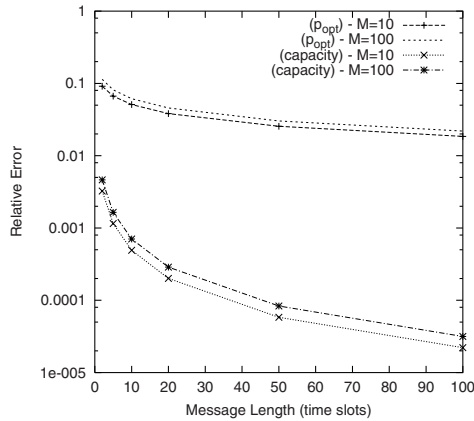


Fig. 4: Relative errors for deterministic message-length distribution

4.2 A Balancing Equation to Derive the Optimal Capacity State

Equation (9) presents a heuristic that was numerically validated. In this section we analytically investigate its validity. To make tractable the problem defined by Equation (5) in the general message-length distribution case, we assume that $E[Coll | Coll] = \bar{C} = \max\{l_1, l_2\}^2$. According to this assumption, substituting formulas (5) and (6) in Equation (8), after some algebraic manipulations, it follows that the p_{opt}^C value is calculated by solving:

$$\min_{p \in [0,1]} \left\{ \frac{\bar{C} - (1-p)^M (\bar{C} - 1)}{Mp(1-p)^{M-1}} \right\} = \min_{p \in [0,1]} \{F(p, M, \bar{C})\} \quad (10)$$

As shown by the following lemma, the solution of Equation (10) is often equivalent to the solution of Equation (9).

² This approximation is motivated by considering that: *i*) the collision probability is low when the network is close to its optimal capacity state, and *ii*) given a collision, the probability that more than two stations collide is negligible (as shown in [11] for the *p*-persistent IEEE 802.11 protocol).

Lemma 2. For $M \gg 1$, the p value that satisfies Equation (10) can be obtained by solving the following equation:

$$E[Idle_p] = E[Coll | N_{tr} \geq 1] \quad , \quad (11)$$

Proof. The proof is reported in Appendix A. ◇

Lemma 2 shows that, asymptotically ($M \gg 1$), in p -persistent CSMA protocols the optimal capacity state is characterized by the balancing between collisions' duration and idle times. To verify the existence of this relationship for small medium M values, we numerically solved both Equation (4) and Equation (9) for a wide range of M values and several message length distributions. Specifically, in Fig. 4 and 5 we show, for several the average message length (l), the relative error³ between the optimal p (curves tagged capacity), and the p value that solve Equation (7), curves tagged capacity p_{opt} .

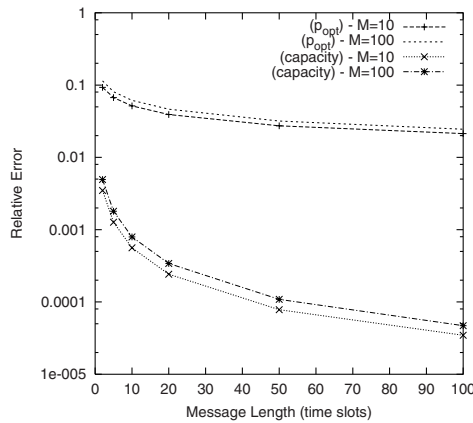


Fig. 5: Relative errors for geometric message-length distribution

4.3 Protocol Capacity: A Closed Formula

The Equation (9) allows us to evaluate a quasi-optimal p value by simply determining the p value that balance $E[Idle_p]$ and $E[Coll | N_{tr} \geq 1]$ costs. By exploiting this approach it is possible to afford, at run-time, the channel-utilization maximization. Indeed, each station by exploiting the carrier sensing mechanism is able to distinguish the idle periods by collisions and by successful transmissions. Furthermore, as shown in this subsection, from Equation (10) it is easy to derive an approximated closed

³ The relative error is defined as the error between the exact value and its approximation, normalized to the exact value.

formula for the p_{opt}^C value. Note that this closed formula generalizes results already known for S-ALOHA networks [23].

Lemma 3. In an M -station network that adopts a p-persistent CSMA access scheme, in which the message lengths $\{L_i\}$, normalized to t_{slot} , are a sequence of i.i.d. random variables, if the stations operate in asymptotic conditions, under the condition $Mp \ll 1$, the p_{opt}^C value is⁴

$$p_{opt}^C \cong \frac{\sqrt{1 + 2(\bar{C} - 1)\frac{M-1}{M}} - 1}{(M-1)(\bar{C}-1)}, \quad (12)$$

where $\bar{C} = E[\max\{L_1, L_2\}]$.

PROOF. The proof is reported in Appendix B.

Equation (12) is important because it allows to easily compute the p_{opt}^C value within a considered scenario (i.e. fixed the number M of stations in the network) and hence it can be used for the optimal tuning of a network if the network and traffic parameters are known, i.e. M and the message-length distribution. However, in a real networking environment, the number of active stations frequently and sharply changes. Hence, M is not a-priori known and must be estimated at run-time. Unfortunately, the M estimation could result expensive, difficult to obtain and subject to significant errors. However as shown in Proposition 1 it is possible to characterize the network optimal behavior without an explicit knowledge of M . Specifically, (given a message length distribution) the optimal state corresponds to a quasi-constant value of the $M \cdot p_{opt}^C$ product. This interesting result requires two assumption: $Mp \ll 1$ and $M \gg 1$. These two conditions are usually met by current WLAN traffic, i.e., long transmitted messages and several active stations. In other words, this means that for any number of active stations the optimal operating point has the same average number of stations that transmit in a slot. The following proposition provides an analytical explanation of this behavior.

Proposition 1. In an M -station network that adopts a p-persistent CSMA access scheme, in which the message lengths $\{L_i\}$, normalized to t_{slot} , are a sequence of i.i.d. random variables, under the condition $Mp \ll 1$ and $M \gg 1$ the $M \cdot p_{opt}^C$ value is

$$M \cdot p_{opt}^C \cong \frac{\sqrt{1 + 2(\bar{C} - 1)} - 1}{(\bar{C} - 1)}, \quad (13)$$

⁴ This assumption is as more correct as more long is the average message length. See, for example, the numerical results reported in Figure 4 and 5.

PROOF. The proof of this proposition is straightforward. Under the condition $M \gg 1$, Equation (12) can be rewritten as (13) by noting that $(M - 1) \approx M$.

◇

Remark. In an M -station network that adopts a p -persistent CSMA access scheme, in which the message lengths $\{L_i\}$, normalized to t_{slot} , are a sequence of i.i.d. random variables, under the condition $\bar{C} \gg 1$ the p_{opt}^C value is

$$p_{opt}^C \cong \frac{1}{M \cdot \sqrt{\bar{C}/2}} \quad (14)$$

Results presented in this section indicate that the $M \cdot p_{opt}^C$ product can be easily computed by estimating only the average collision length. The average collision length estimates are easy to obtain and reliable if the average collision length does not change frequently and sharply. In addition, these results indicate that the estimation of the number of active stations is not necessary for driving the system to its optimal capacity state.

5. Energy Consumption in P-Persistent CSMA Access Schemes

As explained in the introduction, the finite battery power represents the other greatest limitation to the utility of portable computers [25], [26], [27]. The network interface is one of the main system components from the battery consumption standpoint [28]. Hence, the energy used by the network interface to successfully transmit a message is the other important figure for a mobile computer. Thus the target for our environment would be both to maximize the network capacity and to minimize the network-interface energy consumption. It may appear that these two targets cannot be achieved together. It seems that to maximize the network capacity the users must be greedy, i.e. transmitting as much as possible. On the other hand, minimization of the network-interface energy seems to indicate that the network should be lightly loaded, i.e. sporadic accesses to the network. In this paper we will show that, for *p-persistent* CSMA access schemes, the main performance figures (capacity and energy consumption) are not orthogonal, i.e. the network state that optimizes one index is not far from being optimal also for the other one. To show this result, firstly we need to study our system from an energy consumption standpoint, i.e. to identify the relationships to characterize the p value that minimizes the energy consumption.

5.1 An Analytical Model for the Energy Consumption

In this section we develop an analytical model to investigate the energy consumption. From an energy consumption standpoint, the network interface alternates between two

different phases: the transmitting phase, during which it consumes power to transmit the message to the physical channel, and the receiving phase, during which it consumes power to listen to the physical channel. Hereafter, we denote with PTX and PRX the power consumption (expressed in mW) of the network interface during the transmitting and receiving phase, respectively.

We assume that the message lengths are i.i.d. random variables. An analytical model for the energy consumption in p -persistent IEEE 802.11 MAC protocol was developed in [15].

Following the same line of reasoning of Section 4 and focusing on the successful transmission of a tagged station, we derive an analytical expression for the energy drained by a tagged-station network-interface to successfully transmit a packet. Hence, the system efficiency, from the energy standpoint, can be expressed as:

$$\rho_{energy} = \frac{PTX \cdot l \cdot t_{slot}}{E[Energy_{virtual_transmission_time}]}, \quad (15)$$

where $E[Energy_{virtual_transmission_time}]$ is the energy consumed by the tagged station in a tagged_station virtual transmission time (i.e., the time between two tagged station successful transmission) and l is the average message transmission time, normalized to the t_{slot} .

5.2 Energy Consumption: An Approximate Analysis

As for the throughput maximization, it is desirable to have a simpler relationship than Equation (15) to provide an approximation for the p_{opt}^E value. To this end, in the following we will investigate the role of the various terms of Equation (15) in determining the Energy Consumption. Specifically, in the energy consumption formula we separate the terms that are increasing function of the p value from the terms that are decreasing function of the p value.

To achieve this, it is useful to introduce the following proposition.

Proposition 2. In an M -station network that adopts a p -persistent CSMA access scheme, if the stations operate in asymptotic conditions, ρ_{energy} is equal to.

$$\rho_{energy} = PTX \cdot l \cdot t_{slot} \left\{ E[N_{ta}] E[Energy_{idle_p}] + PTX \cdot l \cdot t_{slot} + PRX \cdot (M-1) \cdot l \cdot t_{slot} + (E[N_{ta}] - M) \cdot E[Energy_{Coll} | N_{tr} \geq 1] \right\}, \quad (16)$$

PROOF. The proof requires some algebraic manipulations of Equation (15). We first observe that the number of successful transmissions in a virtual transmission time is M . Furthermore, in the virtual transmission time there is exactly one successful transmission of the tagged station, and in average there are $(M-1)$ successful transmissions of the other stations. It is also straightforward to derive that the average number of collisions that we observe within a virtual transmission time is

$(E[N_{ta}] - M)$ i.e., the total number of transmission attempts in a virtual transmission time less the average number of successful transmissions, i.e.. According to these assessments, with standard probabilistic considerations Equation (16) is derived.

◇

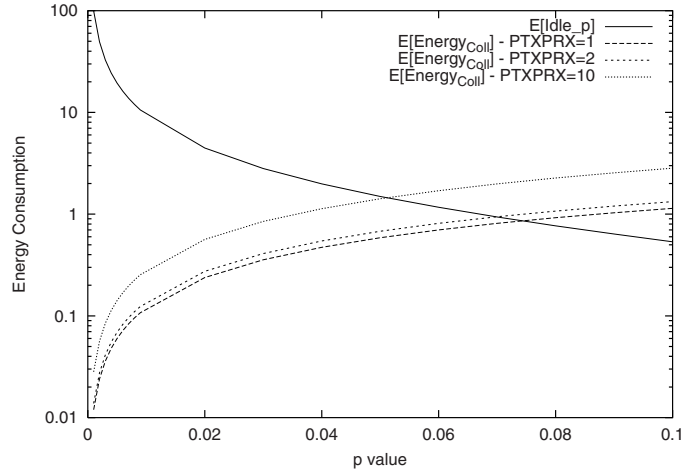


Fig. 6: p_{opt}^E approximation with $M = 10$ and $l = 2 t_{slot}$.

From Proposition 2 it follows that p_{opt}^E corresponds to the p value that minimizes the denominator of Equation (16). It is also worth noting that the second and third terms of this denominator do not depend on the p value and hence they play no role in the minimization process. Now our problem reduces to find

$$\min_p \left\{ E[N_{ta}] \cdot E[Energy_{Idle_p}] + (E[N_{ta}] - M) \cdot E[Energy_{Coll} | N_{tr} \geq 1] \right\} . \quad (17)$$

Now let us note that the first term in Equation (17), i.e. $E[N_{ta}] \cdot E[Energy_{Idle_p}]$, after some algebraic manipulations, reduces to $[PRX \cdot t_{slot} \cdot (1 - p)]/p$, and thus this first term is a decreasing function of p . On the other hand, the second term in Equation (21) is an increasing function of p . Hence, following the same arguments that drive to propose the protocol-capacity approximation (see Equation (9)), we propose to approximate the p_{opt}^E value with the p value that balances the increasing and decreasing costs of p :

$$E[N_{ta}] \cdot E[Energy_{Idle_p}] = (E[N_{ta}] - M) \cdot E[Energy_{Coll} | N_{tr} \geq 1] . \quad (18)$$

Equation (22) can be re-written as:

$$E[Energy_{Idle_p}] = E[Energy_{Coll} | N_{tr} \geq 1] \cdot P_{Coll | N_{tr} \geq 1} . \quad (19)$$

To ease the computation of $E[Energy_{Coll} | N_{tr} \geq 1]$ we subdivide the collisions in two subsets depending on whether they involve the tagged station. Specifically, Equation (19) can be re-written as:

$$E[Energy_{Idle_p}] = E[Energy_{tag_Coll} | tag_Coll] \cdot P_{tag_Coll | N_{tr} \geq 1} + E[Energy_{not_tag_Coll} | not_tag_Coll] \cdot P_{not_tag_Coll | N_{tr} \geq 1} \quad (20)$$

Equation (20) defines a simple but approximate relationship to characterize p_{opt}^E . Specifically, in Fig. 6 we have plotted $E[Energy_{Idle_p}]$ and $E[Energy_{Coll} | N_{tr} \geq 1]$ versus the p value, for various PTX/PRX values. $E[Energy_{Idle_p}]$ is equal to $E[Idle_p]$ due to the assumption that $PRX=1$. The p value that corresponds to the intersection point of the $E[Energy_{Idle_p}]$ and $E[Energy_{Coll} | N_{tr} \geq 1]$ curves is the approximation of the p_{opt}^E value, as Equation (16) indicates. As the $E[Energy_{Coll} | N_{tr} \geq 1]$ related to $PTX/PRX = 1$ is equal to the average length of a collision given a transmission attempt, i.e. $E[Coll | N_{tr} \geq 1]$, the p value that corresponds to the intersection point of the $E[Idle_p]$ and $E[Coll | N_{tr} \geq 1]$ curves provides a good approximation of the p_{opt}^C value, as Equation (6) indicates. We note that by increasing the PTX value also $E[Energy_{Coll} | N_{tr} \geq 1]$ grows due to the rise in the energy consumption of tagged-station collisions. However, $E[Energy_{Idle_p}]$ does not depend on the PTX value, hence, only a decrease in the p_{opt}^E value can balance the increase in $E[Energy_{Coll} | N_{tr} \geq 1]$.

5.3 Energy Consumption: A p_{opt}^E Closed Formula

Similarly to Section 4, we conclude the characterization of our system, from the Energy Consumption standpoint, by providing closed (approximate) formulas to identify the network state that minimizes the Energy Consumption.

Lemma 4. In an M -station network in which the message lengths $\{L_i\}$, normalized to t_{slot} , are a sequence of i.i.d. random variables, under the condition $Mp \ll 1$ the p_{opt}^E value is

$$p_{opt}^E \cong \frac{\sqrt{1 + 2 \frac{M-1}{M} \left[\bar{C} \cdot \frac{(M-2)}{M} + \frac{\bar{E}_{CT}}{PRX} \cdot \frac{1}{M} - 1 \right]} - 1}{(M-1) \left[\bar{C} \cdot \frac{(M-2)}{M} + \frac{\bar{E}_{CT}}{PRX} \cdot \frac{1}{M} - 1 \right]}, \quad (21)$$

where $\bar{C} = E[\max\{L_1, L_2\}]$ and $\bar{E}_{CT} = E[Energy_{tag_Coll} | tag_Coll, N_{tr} = 2]$.

PROOF. The proof of this Lemma can be found in [24]

As we have done in Proposition 1, the following proposition provides an analytical investigation of the $M \cdot p_{opt}^E$ for a large network-size population. This investigation is useful because it shows how for a large network size population the p_{opt}^E value tends to the p_{opt}^C value.

Proposition 3. In an network with a large number of active stations ($M \gg 1$) in which the message lengths $\{L_i\}$, normalized to t_{slot} , are a sequence of i.i.d. random variables, the optimal $M \cdot p_{opt}^E$ value is:

$$M \cdot p_{opt}^E \cong \frac{\sqrt{1 + 2\left(\bar{C} + \frac{\bar{E}_{CT}}{PRX} \frac{1}{M} - 1\right)} - 1}{\left(\bar{C} + \frac{\bar{E}_{CT}}{PRX} \frac{1}{M} - 1\right)} \approx \frac{\sqrt{1 + 2(\bar{C} - 1)} - 1}{(\bar{C} - 1)} \quad (22)$$

PROOF. The proof of this proposition is straightforward. Under the condition ($M \gg 1$), Equation (21) can be rewritten as (22) by noting that $(M - 1) \approx M$ and $(M - 2) \approx M$.

◇

Equation (22) presents a tight analogy with Equation (13). For a large network-size populations it is straightforward to observe that $Mp_{opt}^E \approx Mp_{opt}^C$ (as Equation (22) and Equation (13) show). Since \bar{E}_{CT} is divided by M , the energy consumption during tagged-station collisions decreases as the M value increases. Therefore, the PTX value has no significant impact on the $M \cdot p_{opt}^E$ computation for a large network-size population as it contributes only to \bar{E}_{CT} . Obviously $p_{opt}^E = p_{opt}^C$ when $PTX = PRX$. However, the comparison between the structure of Equation (13) and Equation (22) show also that the correspondence between the optimal p values continues to hold.

6. Effectiveness of the AOB Mechanism

In the remaining part of the paper, by means of the discrete event simulation, we extensively investigate the performance of the IEEE 802.11 protocol enhanced with the AOB mechanism. This mechanism utilize the current S_U estimate to evaluate, the opportunity to perform or to defer a transmission attempt authorized by the standard protocol (see Section 3). Specifically AOB uses the probability of transmission defined by Equation (1). As discussed before, Equation (1) requires the knowledge of the opt_S_U parameter. Below we show how results derived in Sections 4 and 5 can be

used to estimate the opt_S_U value. Specifically, in the previous sections we show that the optimal capacity and energy state are characterized by invariant figures: $M \cdot p_{opt}^C$ and $M \cdot p_{opt}^E$. Hereafter we investigate the relationship between S_U and $M \cdot p_{opt}^x$ ($x \in \{C, E\}$).

We denote with N_{tr} the number of stations that make a transmission attempt in a slot. Hence, $P\{N_{tr} = i\}$ is the probability that exactly i stations transmit in a slot, and $P\{N_{tr} = 0\}$ is the probability that a slot remains empty. Let us now observe that $M \cdot p_{opt}^x$ is the average number of stations which transmits in a slot:

$$M \cdot p_{opt}^x = \sum_{i=1}^M i \cdot P\{N_{tr} = i\} \geq \sum_{i=1}^M P\{N_{tr} = i\} = 1 - P\{N_{tr} = 0\} = S_U \quad .$$

The above formula indicates that $M \cdot p_{opt}^x \geq S_U$. By exploiting this result we derive the opt_S_U parameter value used by AOB: $opt_S_U \approx M \cdot p_{opt}^x$.

6.1 Steady-State Analysis

In this section we analyze the AOB behavior when the network operates under steady-state conditions. The protocol analysis in transient conditions, and the protocol robustness are studied in the subsequent sections. The physical characteristics and parameter values of the investigated system are reported in Table 1.

The target of a MAC protocol is to share resources efficiently among several users. This efficiency can be expressed in terms of *capacity* [25], [29]. However from the user standpoint other performance figures are needed to measure the Quality of Service (QoS) that can be relied on. The most widely used performance measure is the delay, which can be defined in several forms, depending on the time instants considered during its measurement (access delay, queueing delay, propagation delay, etc.). Hereafter, we will focus on the MAC delay. The MAC delay of a station in a LAN is the time between the instant at which a packet comes to the head of the station transmission queue and the end of the packet transmission [29].

To study the protocol performance we run a set of simulative experiments with different M of values. Active stations are assumed to operate in asymptotic conditions (i.e., with continuous transmission requirements). We use a maximum number of 200 active stations because the number of stations expected in the future for such a system could raise the order of hundreds [30]. For example, let us think to a conference room in which the participants use mobile devices with a wireless interface.

The effectiveness of the proposed AOB mechanism is shown in Fig. 7. This figure shows the channel utilization level achieved by adopting the AOB system and compares this index with the analytically defined optimal utilization levels (OPT curves in the figure). The results show that the AOB mechanism drives an IEEE 802.11 network very close to its optimal behavior at least from the channel utilization viewpoint. Only a little overhead is introduced when only few stations are active. It is

worth noting that, with the AOB mechanism, the channel utilization remains close to its optimal value even in high-contention situations. In such cases, AOB almost doubles the channel utilization with respect to the standard protocol.

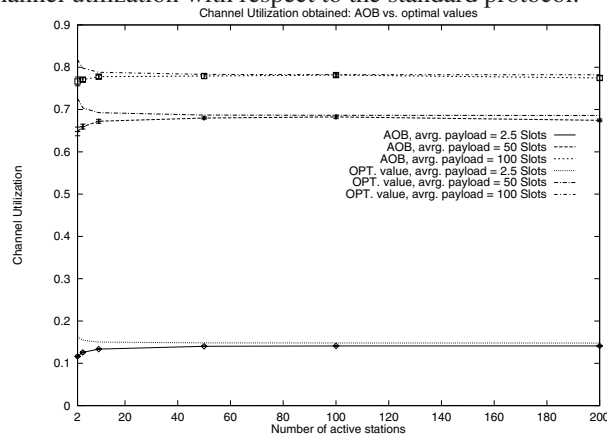


Fig. 7: Channel utilization of the IEEE 802.11 protocol with the AOB mechanism vs. optimal value

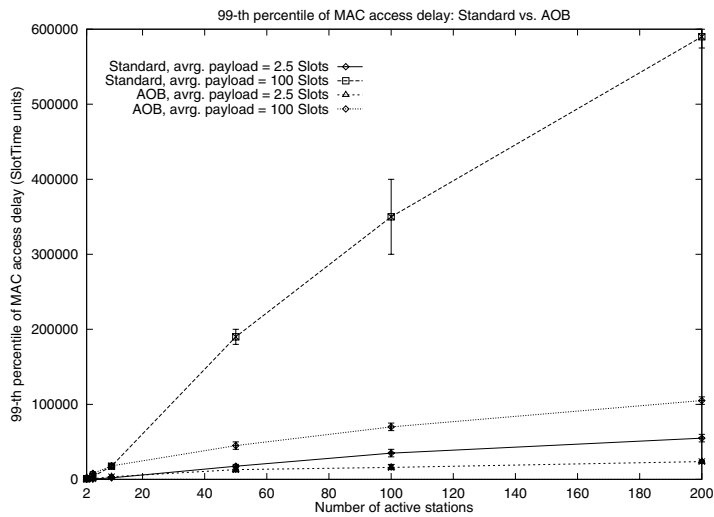


Fig. 8: 99-th percentile of MAC delay

In Fig. 8 we report the 99-th percentile of the MAC delay vs. contention level (i.e. number of active stations) for various average sizes of the transmitted frames. Simulative results show that the AOB mechanism leads to a great reduction of the tail of the MAC delay distribution with respect to the standard access scheme alone. By noting that when the network operates in asymptotic conditions the average MAC delay is the inverse of the station throughput, we can verify that AOB is really effective in

reducing the tail of the MAC Delay. For example, with 100-slot average payload, the ratio between the 99-th percentile of the MAC Delay with or without the AOB mechanism is about 6 while the ratio between the average MAC Delay is only about 2.

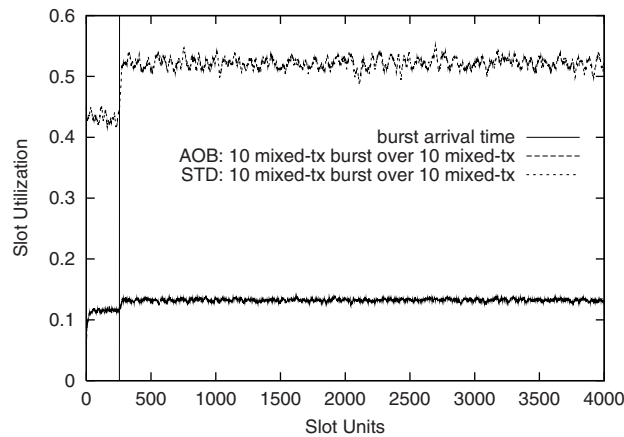


Fig. 9: A burst of 10 new stations activates when the network is operating in steady-state conditions with 10 active stations (Slot utilization)

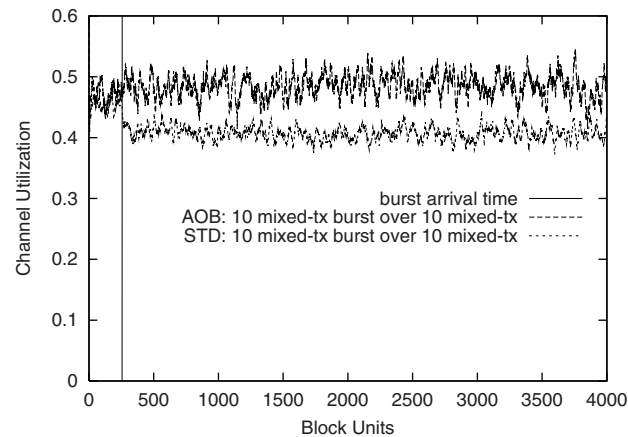


Fig. 10: A burst of 10 new stations activates when the network is operating in steady-state conditions with 10 active stations (Channel utilization)

Previous results seem to indicate that AOB, by reducing the average MAC delay, has also a positive effect on the power consumed by an IEEE 802.11 network interface. This intuition is further confirmed by the results presented in [15] showing that, by

adopting AOB with current IEEE802.11 cards, the optimal capacity state closely approximates the minimum energy consumption state.

6.2 AOB Behavior in Transient Situations

In this section we analyze the protocol promptness to re-tune when the network state sharply changes. Specifically, we investigate the effectiveness the *AOB* when there is an upsurge in the number of active stations. Specifically, we analyze a network operating in steady-state conditions with 10 active stations. After 256 block units⁵ (highlighted by the vertical bar “burst arrival time”), additional 10 stations become active. All stations transmit mixed traffic composed by 50% of long message (1500 bytes) and 50% of short messages (40 Bytes). Fig. 9 shows the effectiveness of the AOB mechanism. In the AOB case, the sharp increase in the number of active stations produces a negligible effect both in the slot utilization and in the channel utilization. On the other hand, the standard is negatively affected by this change: the slot utilization sharply increases (i.e., the network congestion level increases) and as a consequence the channel utilization decreases.

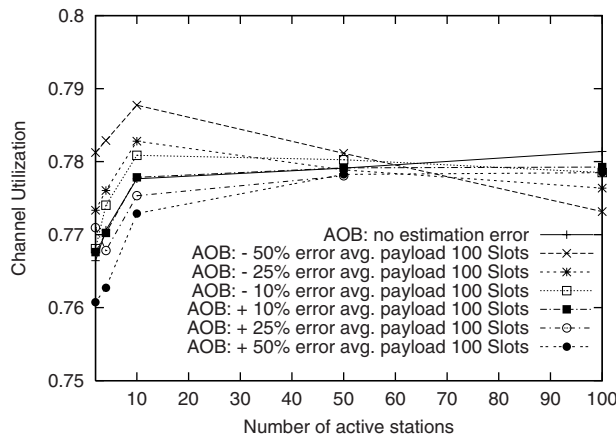


Fig. 11: Sensitiveness to errors in the estimation of the average message length (Slot utilization)

6.3 Protocol Robustness

The AOB mechanism, to tune the backoff algorithm, requires the knowledge of the network status that is identified by two parameters: the average message length (or

⁵ A block unit corresponds to 512 slots. The block unit is introduced to smooth the trace. The smoothing was introduced to reduce the fluctuations and thus increasing the figure readability.

equivalently the q parameter) and the slot utilization. As the values of these parameters are obtained through estimations some errors may occur. Hereafter, we discuss the sensitiveness of AOB to these possible errors. To this end we compare the channel utilization in the ideal case (no estimation errors) with the channel utilization when an error is added/subtracted to the correct q and S_U estimate. Specifically, in the following figures the curve tagged with $+x\%$ ($-x\%$) error is obtained by multiplying by $1+x/100$ ($1-x/100$) the real estimate of the parameter. Results obtained are summarized in Fig. 10 (errors on q by assuming an average message length equal to 100 slots, i.e., $q=0.99$) and in Fig. 12 (errors on S_U). These results indicate that the AOB channel utilization is scarcely affected by estimation errors. For example, assuming constant errors of 50%, the channel utilization fluctuates in a small interval (2-3%) around the no-error value. It is worth noting that due to the way AOB is defined: *i*) for large M errors always have a negative impact (AOB is tuned to optimize asymptotic performance), *ii*) for few active stations, underestimation errors generate a channel utilization which is higher than that obtained in the no-error case. The latter behavior was expected because the ACL value is too conservative (i.e., it excessively limits the transmission rate) when there are few active stations. The parameters underestimation produces the opposite effect thus resulting in an increased channel utilization.

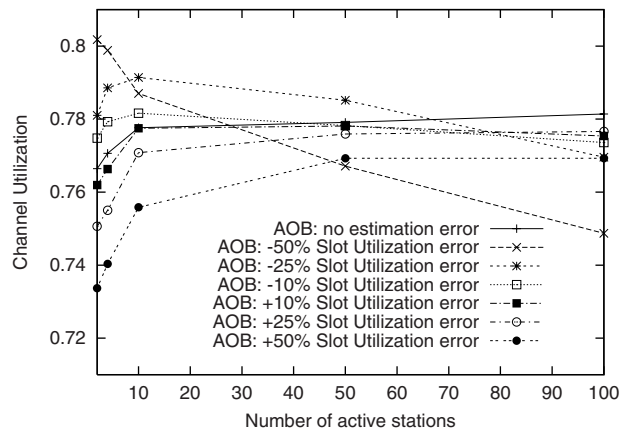


Fig. 12: Sensitiveness to errors in the estimation of the Slot Utilization

Acknowledgements

This work was partially supported by the NATO CLG.977405 project “Wireless Access to Internet exploiting the IEEE802.11 Technology” and by the the WILMA project “Wireless Internet and Location Management Architecture” funded by the Province of Trento, Italy.

References

1. ANSI/IEEE Standard 802.11, "Part 11: Wireless LAN- Medium Access Control (MAC) and Physical Layer (PHY) Specification", August 1999
2. Stallings W., *Local & Metropolitan Area Networks*, Fifth Edition, Prentice Hall 1996, pp. 356-383.
3. H. Woesner, J.P. Ebert, M. Schlager, A. Wolisz, "Power-saving mechanisms in emerging standards for wireless LANs: The MAC level perspective", *IEEE Personal Comm*, 1998, pp. 40-48.
4. N. Bambos, "Toward power-sensitive network architectures in wireless communications: Concepts, issues and design aspects", *IEEE Personal Comm*, 1998, pp. 50-59.
5. Weinmiller J., Woesner H., Ebert J.P., Wolisz A., "Analyzing and tuning the Distributed Coordination Function in the IEEE 802.11 DFWMAC Draft Standard", Proc. Int. Workshop on Modeling, MASCOT 96, San Jose, CA.
6. J. Weinmiller, M. Schläger, 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
7. L. Bononi, M. Conti, L. Donatiello, "Design and Performance Evaluation of a Distributed Contention Control (DCC) Mechanism for IEEE 802.11 Wireless Local Area Networks", *Journal of Parallel and Distributed Computing*, Accademic Press Vol.60 N.4 di Aprile 2000.
8. M. Gerla, L. Kleinrock, "Closed loop stability control for S-Aloha satellite communications", *Proc. Fifth Data Communications Symp., Sept. 1977*, pp. 2.10-2.19.
9. B. Hajek, T. Van Loon, "Decentralized dynamic control of a multiaccess broadcast channel", *IEEE Trans Automat. Control*, Vol.27, 1982, pp. 559-569.
10. F. Kelly, "Stochastic Models of computer communications systems", *J. Royal Statist. Soc., Series B*, Vol. 47, 1985, pp. 379-395.
11. F. Cali', Conti M., E. Gregori, "Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit", *IEEE/ACM Transactions on Networking*, Volume 8 , No. 6 (Dec. 2000), pp. 785 - 799.
12. F. Cali', Conti M., E. Gregori, "Dynamic IEEE 802.11: design, modeling and performance evaluation", *IEEE Journal on Selected Areas in Communications*, 18(9), September 2000. pp. 1774-1786.
13. Bianchi G., Fratta L., Olivieri M., "Performance Evaluation and Enhancement of the CSMA/CA MAC protocol for 802.11 Wireless LANs", proceedings of PIMRC 1996, 10/1996, Taipei, Taiwan, pp. 392-396.
14. J.P. Monks, V. Bharghavan, W.W. Hwu, "A Power Controlled Multiple Access Protocol for Wireless Packet Networks", in *Proc Infocom'01*, Anchorage, Alaska (Apr. 2001
15. L. Bononi, M. Conti, L. Donatiello, "A Distributed Mechanism for Power Saving in IEEE 802.11 Wireless LANs", *ACM/Kluwer Mobile Networks and Applic. Journal*, Vol. 6, N. 3 (2001), pp. 211-222.
16. <http://grouper.ieee.org/groups/802/11/main.html>
17. K. Bieseker, "The Promise of Broadband Wireless", *IT Pro* November/December 2000, pp. 31-39.
18. R. Bruno, M. Conti, E. Gregori, "WLAN technologies for mobile ad-hoc networks", Proc. HICSS-34, Maui, Hawaii, January 3-6, 2001. An extended version can be found in the Chapter 4 of *Handbook of Wireless Networks and Mobile Computing* (I. Stojmenovic Editor), John Wiley & Sons, New York, 2001.

19. Goodman J., Greenberg A.G., Madras N., March P., "Stability of Binary Exponential Backoff", app. in the Proc. of the 17-th Annual ACM Symp. on Theory of Comp., Providence, May 1985.
20. Hammond J.L., O'Reilly P.J.P., Performance Analysis of Local Computer Networks, Addison-Wesley 1988.
21. Hastad J., Leighton T., Rogoff B., "Analysis of Backoff Protocols for Multiple Access Channels", Siam J. Computing vol. 25, No. 4, 8/1996, pp. 740-774.
22. Gallager R.G., "A perspective on multiaccess channels", IEEE Trans. Information Theory, vol. IT-31, No.2, 3/1985, pp. 124-142.
23. D. Bertsekas, R. Gallager, "Data Networks" *Prentice Hall*, 1992.
24. R. Bruno, M. Conti, E. Gregori, "Optimization of Efficiency and Energy Consumption in p-persistent CSMA-based Wireless LANs", *IEEE Transactions on Mobile Computing*, Vol. 1 N.1, January 2002.
25. A. Chandra V. Gumalla, J.O. Limb, "Wireless Medium Access Control Protocols", IEEE Communications Surveys Second Quarter 2000.
26. G.H. Forman, J. Zahorjan, "The challenges of mobile computing", *IEEE Computer*, April 94, pp.38-47.
27. T. Imielinsky, B.R. Badrinath, "Mobile Computing: Solutions and Challenges in Data Management", *Communications of ACM*, Oct. 94.
28. R. Kravets, P. Krishnan, "Power Management Techniques for Mobile Communication", Proceedings of The Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'98).
29. Conti M., Gregori E., Lenzini L., "Metropolitan Area Networks", Springer Verlag, London, 1997.
30. Chen K.C., "Medium Access Control of Wireless LANs for Mobile Computing", IEEE Networks, 9-10/1994.
31. W.R.Stevens. TCP/IP Illustrated, Volume 1: The Protocols, Addison-Wesley, Reading, MA, 1994.
32. M. Stemm, R.H. Katz, "Measuring and Reducing Energy Consumption of Network Interfaces in Hand-Held Devices", Proc. 3rd International workshop on Mobile Multimedia Communications (MoMuC-3), Princeton, NJ, September 1996.

Appendix: A

Lemma 2. For $M \gg 1$, the p value that satisfies (6) can be obtained by solving the following equation: $E[Idle_p] = E[Coll | N_r \geq 1]$

where $E[Coll | N_r \geq 1]$ is the average duration of a collision given that at least a transmission occurs.

PROOF. Taking the derivative of $F(p, M, \bar{C})$ with respect p , and imposing it equal to 0, we obtain the following equation:

$$\begin{aligned} \left\{ (1-p)^M + p(1-p)^{M-1} \right\} &= & \text{(A.1)} \\ = \bar{C} \left\{ p(1-p)^{M-1} - \left[1 - (1-p)^M \right] + (M-1) \frac{p}{1-p} \right\} . \end{aligned}$$

The p_{opt} value is the solution of Equation (A.1).

First we analyze the l.h.s of Equation (A.1). It is easy to observe that the l.h.s is equal to $(1-p)^{M-1}$ that $M \gg 1$ tends to $(1-p)^M$, and $E[Idle] \cdot P_{N_r \geq 1} = (1-p)^M$, where $P_{N_r \geq 1} = 1 - (1-p)^M$, i.e., the probability that at least a station is transmitting. Under the condition $M \cdot p < 1$,⁶ the r.h.s of Equation (A.1) can be expressed as:

$$\left\{ p(1-p)^{M-1} - [1 - (1-p)^M] + (M-1) \frac{p}{1-p} \right\} = \frac{(M+2)(M-1)}{2} p^2 + O((Mp)^3) \quad (\text{A.2})$$

By indicating with $P_{Coll|N_r \geq 1}$ the collision probability conditioned to have at least a transmitting station, it also holds that:

$$P_{Coll|N_r \geq 1} \cdot P_{N_r \geq 1} = 1 - [(1-p)^M + Mp(1-p)^{M-1}] = \frac{M(M-1)}{2} p^2 + O((Mp)^3) \quad (\text{A.3})$$

It is worth noting the similarity between the r.h.s of Equation (A.3) and the r.h.s of Equation (A.2). Specifically the r.h.s of Equation (A.2) can be written as:

$$\left[\frac{(M-1)}{2} p^2 \right] \cdot (M+1) + O((Mp)^3) \xrightarrow{M \gg 1} \left[\frac{(M-1)}{2} p^2 \right] \cdot M + O((Mp)^3) = P_{Coll|N_r \geq 1} \cdot P_{N_r \geq 1} \quad (\text{A.4})$$

Hence, Equation (A.1) can be rewritten as:

$$E[Idle] \cdot P_{N_r \geq 1} = \bar{C} \cdot P_{Coll|N_r \geq 1} \cdot P_{N_r \geq 1}. \quad (\text{A.5})$$

By dividing all the terms in Equation (A.5) by $P_{N_r \geq 1}$, and substituting the \bar{C} approximation with $E[Coll|Coll]$, Equation (A.5) becomes:

$$E[Idle_{-p}] = E[Coll|N_r \geq 1] \quad (\text{A.6})$$

This concludes the proof. \diamond

Appendix: B

Lemma 3. In an M -station network that adopts a p -persistent CSMA access scheme, in which the message lengths $\{L_i\}$, normalized to t_{slot} , are a sequence of i.i.d. random

⁶ This assumption is always true in a stable system as the average number of transmitting stations in an empty slot must be less than one.

variables, if the stations operate in asymptotic conditions, under the condition $Mp \ll 1$, the p_{opt}^C value is

$$p_{opt}^C \cong \frac{\sqrt{1 + 2(\bar{C} - 1)\frac{M-1}{M}} - 1}{(M-1)(\bar{C}-1)}, \quad (\text{B.1})$$

where $\bar{C} = E[\max\{L_1, L_2\}]$.

PROOF.

In [24] it is shown that if the network stations are operating close to the p_{opt}^C value and if the stations operate in asymptotic conditions $E[Coll | Collision] \approx \bar{C}$. The $E[Coll | Collision] \approx \bar{C}$ assumption indicates that $E[Coll | Collision]$ depends only on the message-length distribution and Equation (11) can be rewritten as:

$$(1-p)^M - \bar{C} \cdot \left\{ 1 - \left[(1-p)^M + Mp(1-p)^{M-1} \right] \right\} = 0. \quad (\text{B.2})$$

Under the condition $Mp \ll 1$,

$$(1-p)^M \approx 1 - Mp + \frac{M(M-1)}{2} p^2 - O((Mp)^3), \quad (\text{B.3})$$

$$1 - \left[(1-p)^M + Mp(1-p)^{M-1} \right] \approx \frac{M(M-1)p^2}{2} - O((Mp)^3). \quad (\text{B.4})$$

By substituting (B.3) and (B.4) in (B.2) the following equation is obtained:

$$\frac{M(M-1)}{2} \cdot (\bar{C}-1)p^2 + Mp - 1 = 0. \quad (\text{B.5})$$

By solving Equation (B.5) we obtain the formula (B.1) and this concludes the proof of the Lemma.

◇