Design and Performance Evaluation of a Distributed Contention Control (DCC) Mechanism for IEEE 802.11 Wireless Local Area Networks

Luciano Bononi, Marco Conti(*), Lorenzo Donatiello

e-mail: {bononi, donat}@cs.unibo.it
Dipartimento di Scienze dell'Informazione
Università di Bologna
Mura Anteo Zamboni, 7
40127, Bologna, Italy

(*) e-mail: M.Conti@cnuce.cnr.it
CNR Istituto CNUCE
Via S.Maria, 36
56100, Pisa, Italy

Abstract

This paper focuses on the design and the performance evaluation of a new mechanism, named Distributed Contention Control (DCC), for the adaptive reduction of contention in LAN networks that utilize random access MAC protocols. Specifically, we consider the IEEE 802.11 wireless LAN (WLAN). The proposed mechanism could be executed on the top of a pre-existent access scheduling protocol, with no additional overhead introduced. The DCC mechanism requires a simple and rough estimate of any parameter easy to investigate, and directly connected with the amount of contention currently suffered by the shared channel. The main characteristics of the proposed mechanism are represented by its simplicity, integrability, complete distribution, absence of overheads, and adaptiveness. The adaptiveness is given in a fully automatic way, by means of a feedback effect connected to the estimated contention parameter. This contributes to enhance the system stability. Simulation experiments to evaluate the performance of an IEEE 802.11 WLAN, with and without the DCC mechanism, have been carried out. Results confirmed the effectiveness of the DCC mechanism in improving the performance, stability, and congestion reaction of the IEEE 802.11 access scheme.

1 Introduction

This paper proposes a Distributed Contention Control (DCC) mechanism for improving the performance of the IEEE 802.11 access scheme for Wireless LANs. The IEEE 802.11 access scheme is based on two access methods: Distributed Coordination Function (DCF) for asynchronous, contention-based accesses to the channel, and Point Coordination Function (PCF) for centralized, contention-free accesses.

We will concentrate our analysis over DCF, because it is in this case that the congestion problems could arise.

The DCF method adopts a distributed, contention-based access scheme belonging to the class of the CSMA/CA MAC protocols [4,7,11]. Collisions occur when two or more users try to contemporary utilize the resource. The contention reduction in accessing to the shared channel is obtained with a variable time-spreading of the users' accesses. Hence, in this system a resource wastage is caused both from collisions and from the resource idle periods introduced by the time-spreading of the accesses. As the reduction of the idle periods generally produces an increase in the number of collisions, to maximize the resource utilization the protocol should balance these two costs [3,8]. Since these costs change dynamically, depending on the network load, it is evident the need for a kind of adaptiveness of such CSMA protocols with respect to congestion variations in the system [4,8,12]. Such an adaptive behavior is currently obtained by means of backoff protocols, based on the local history regarding successes or collisions [12]. Specifically, each user is not assumed to have any kind of knowledge about other users' successes or collisions, or even about the number of users in the system. Each station accesses the channel within a random self-defined amount of time, whose mean size depends on the number of collisions already experienced by the station while transmitting a given frame. For these protocols, when the number of users grows, or even in case of bursty arrivals, due to high collision rates, a low resource utilization is often obtained. This fact seems difficult to avoid because the system's reaction to high contention conditions is based upon collision happenings. Examples of these channel allocation problems arise in the context of CSMA/CD multiple access protocols for LANs such as Ethernet and IEEE 802.3, or in the IEEE 802.11 MAC protocol for WLANs [7]. Specifically, it is known that for CSMA/CD MAC protocols, given a great number of users, the communication capability of the network could collapse [4,12].

Several authors have investigated the enhancement of the IEEE 802.11 DCF MAC protocol to increase its performance when it is utilized in WLANs. In [5,6], via a performance analysis, it is studied the tuning of the Standard parameters. In [17] solutions have been proposed for a better uniform distribution of accesses, given the Binary Exponential Backoff scheme [7, 10] adopted by the Standard.

Trying to extend backoff protocols, a great deal of work has been done in the analysis of models that assume the use of more information to define the access scheduling [9,13,16]. Some models try to approximate the knowledge about the number of users involved in the accesses by means of the analysis of the past history of the system. Examples of such works regarding the IEEE 802.11 DCF MAC protocol relates to the attempt to make the reduction of contention adaptive and optimal by investigating the number of users

---

[0] This work was partially supported by MURST and CNR funds.

WOWMOM 98 Dallas Texas USA

in the system [1,3]. It is worth observing how this investigation could result expensive, difficult to obtain and subject to some vagueness, especially in high contention situations.

Starting from these considerations, we propose a new contention-reduction mechanism, named DCC, simply based on a rough estimate of the contention level. Our mechanism is adaptive to the contention level of the system, exploiting the proposed contention level estimate. Hence, the contention reaction performed does not necessarily depend on collisions, but contributes to prevent collisions' occurrence. Moreover, the reduction of contention can be achieved together with the introduction of distributed priorities, zero cost and simple distributed management. This opportunity constitutes a step towards the obtainment of queue-emptying-like protocols, which contribute to a greater stability [12], also avoiding the risk that a single user dominates the resource. Then, it results possible to adopt a contention-based access scheme, without worring about the eventuality of a congested system to arise, and then taking full advantage from it. With more generality, by means of the DCC mechanism implemented on top, it results possible to extend even a simple 1-persistent scheduling protocol, whatever the backoff function adopted, obtaining a powerful adaptive contention control in the accesses to the channel.

The work is organized as follows: in Section 2 we give a characterization of the system IEEE 802.11 considered, and we describe some congestion reaction problems; in Section 3 we discuss the implementation of our mechanism; in Section 4 we study the stability aspects of the system obtained; in Section 5 we present simulation results of the proposed mechanism, implemented on top the IEEE 802.11 protocol; in Section 6 conclusions and future researches are given.

2 IEEE 802.11 Standard DCF for Wireless LANs

For the detailed explanation of the IEEE 802.11 Standard, and in particular of the 802.11 Standard DCF, we address interested readers to [5,6,7,14].

The 802.11 DCF is based on a CSMA/CA MAC protocol which requires every station to perform a Carrier Sensing activity to determine the current state of the channel (idle or busy), and to evaluate the opportunity to start a transmission without interfering with any other on going transmission. If the medium is found to be busy, the station defers keeping to listen the channel, until it becomes idle for at least a Distributed Interframe Space time interval (DIFS), then the station starts its Basic Access mechanism. Obviously, to avoid collisions as soon as an idle DIFS is sensed on the channel, a Collision Avoidance mechanism is needed. The Collision Avoidance mechanism adopted in the 802.11 Standard is based on a Binary Exponential Backoff scheme, associated with a slottization of time during which accesses are allowed. The objectives of the backoff scheme are: a distribution the most uniform as possible of scheduled transmission attempts over a variable sized time window, and a limitation of the mean access delay. The Binary Exponential Backoff scheme is implemented by each station by means of a parameter, named Backoff Counter, which maintains the number of empty slots the tagged station must observe on the channel before performing its own transmission attempt. At the time a tagged station needs to schedule a new transmission, it selects a particular slot among those of its Contention Window, whose size is maintained in a local parameter CW_Size. Specifically, the Backoff value is defined by the following expression [7]:

$$Backoff\_Counter(CW\_Size) = int(Rnd() * CW\_Size)$$

where Rnd() is a function which returns pseudo-random numbers uniformly distributed in [0..1]. The Backoff_Counter is decreased as long as a slot time is sensed as idle, it is frozen when a transmission is detected, and reactivated after the channel is sensed as idle for at least a further DIFS. As soon as the Backoff Counter reaches the value Zero the station transmits its own frame.

If the transmission generates a collision[1], it results opportune to double the CW_Size parameter for the new scheduling of the retransmission attempt, in order to obtain a further reduction of contention.

The Binary Exponential Backoff is then characterized by means of the expression which gives the dependency of the CW_Size parameter by the number of unsuccessful transmission attempts (Num_Att) already performed for a given frame. In [7] it is defined that the first transmission attempt of a given frame is to be performed with 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:

$$CW\_Size(Num\_Att) = \\ \min(CW\_Size\_MAX, CW\_Size\_min * 2^{Num\_Att-1})$$

The increasing of the CW_Size, on behalf of every station whose transmission attempt resulted in a collision, is the reaction that the 802.11 Standard DCF provides to react to a congestion condition, and to make the access mechanism adaptive to channel conditions.

2.1 Congestion reaction analysis

Analyzing the behavior of the 802.11 DCF mechanism, some problems could be identified. Figure 2 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 are presented in Table 1 of Section 5. The results show how the channel utilization is negatively influenced by the contention level growth.

Following the 802.11 Standard indications, a station selects the initial size of the Contention Window for a frame transmission by assuming a low level of congestion in the system, because this avoids a great access delay when the load is light. This represents a problem in bursty arrival scenarios, and in congested systems, because it introduces a concentration of accesses in a reduced time window, and hence it causes a high collision probability. The reaction in high congestion conditions is performed by each station only on the basis of the experienced collisions in previous unsuccessful attempts for each given frame. Every station performs its attempts blindly, with a late collision reaction performed (increasing CW_Size), but with the need for successive retransmission attempts, leading to a further worsening of the congestion. Moreover, in wireless channels, Collision Detection (CD) results not easy to implement, and each collision implies a cost proportional to the maximum size among those of the colliding frames.

---

[1] A collision is assumed whenever the Ack from the receiver is missing [7]

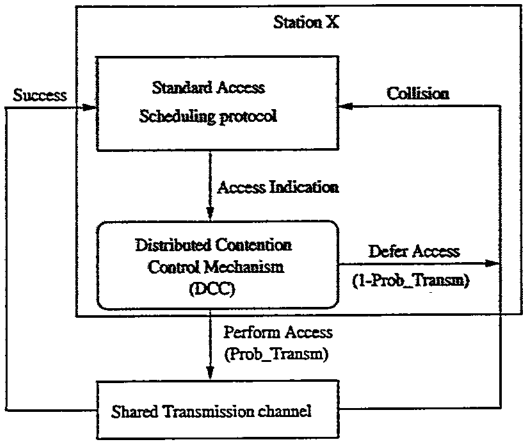

Figure 1: Architectural collocation of DCC

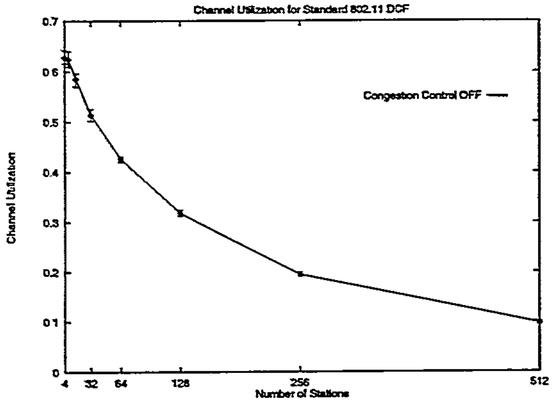

Figure 2: Channel utilization of Standard 802.11 DCF

The information used in the 802.11 Standard has two negative aspects: it costs a collision for each successive approximation towards a greater level of significance, and it does not provide an accurate indication of the actual contention level, due to stochastic variability in the slot selection[2], and due to the initially limited number of transmission attempts performed.

For the reasons explained above, we studied the possibility to obtain early and meaningful information concerning the actual state of congestion of the channel. Our idea involves an estimate of the channel's congestion level that is given by the utilization rate (Slot Utilization) of the slots observed on the channel by each station with a frame scheduled for transmission. It is simple to realize that the Slot Utilization could assume the role of an early Contention Level indication.

2.2 The Slot Utilization estimate

The Slot Utilization estimate we propose to utilize is simple and with zero costs and overheads. All the informations needed are already available in an 802.11 station and no additional hardware is required. The estimate of the Slot Utilization has to be maintained and updated by each station, during the defer phase that precedes a transmission attempt, also indicated as Backoff interval. For the use we considered to make with, the Slot Utilization estimate in the proposed DCC mechanism has to satisfy only two conditions:

- values included in [0..1]: the Zero value should indicate that no slots observed in the Backoff interval resulted as busy, while the value One should indicate that every observed slot resulted as busy;
- intermediate values should be distributed in the [0..1] interval, proportional to the contention level, and indicate the rate of busy slots compared to the total number of slots observed.

The 802.11 Standard mechanism assumes that every station is able to perform a Carrier Sensing activity such that the evolution of the state of the channel is known to each station, see Figure 3. This means that it is possible to obtain a Slot Utilization estimate following our assumptions. Specifically, every station during its Backoff period, counts the number of transmission attempts it observes on the channel (Num_Busy_Slots), and then divides this number for the initial value of its Backoff Counter (Init_Backoff), which also represents the number of slots available for transmission, observed on the channel. Hence, a simple and intuitive definition of the Slot Utilization estimate is given by:

$$Slot\_Utilization = \frac{Num\_Busy\_Slots}{Init\_Backoff} \quad .$$

It is interesting to observe that the Slot Utilization constitutes a lower bound to the effective number of stations that have tried to access the channel during the last observed Backoff interval. This implies that if the value of the Slot Utilization is high (i.e. near to One), we observed, during the last backoff interval, a high level of contention on the channel. As we can assume a time locality of contention, such an information is significant about the status of the channel in immediately successive slots. We cannot say the opposite, that is, if the value of Slot Utilization is low, then we are dealing with a low level of congestion, because of stochastic variability in slot selection[3]. Anyway, this fact is statistically true in most cases. The reactions that the proposed mechanism for congestion control realizes, are based on the above considerations. The Slot Utilization estimate is an useful indicator of the state of the channel and it guarantees that the DCC mechanism activates only when a real congestion condition occurs. So, the stochastic variability of the Slot Utilization estimate can only attenuate the mechanism's effects, and never amplify them in cases where no congestion occurs. This is an improvement with respect to the earlier mechanism, because no overheads are introduced in low load cases. The Slot Utilization estimate constitutes the reference that every station in the system will utilize to evaluate the kind of reaction to the level of congestion in the channel.

To summarize, the positive aspects of the Slot Utilization estimate are: it constitutes a contention indication before the access (without the need to experience collisions), it is completely distributed, it provides a significant indication of the actual contention level, and it is obtained with no costs and overheads introduced. Furthermore, it is obtainable with a full integration with the standard.

---

[2] Collisions could occur even with few stations, so the contention indication obtained could be overestimated

[3] As said before, the Slot_Utilization is only a lower bound of the contention level, in fact a busy slot observed could have been selected for transmission by an arbitrary number of colliding stations

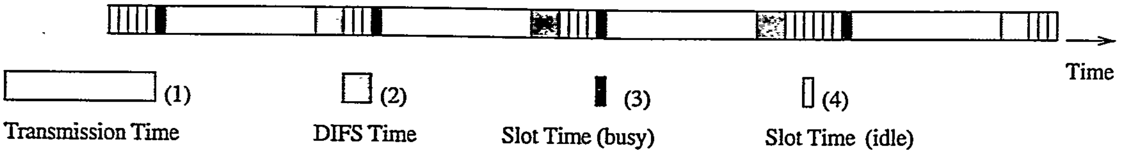

Figure 3: A sample of the channel state observed by each station.

The use of the Slot Utilization estimate presents some problems, due to the possible zero assignment for the Backoff Counter, in the 802.11 Standard. A station which obtains a zero value in the scheduling definition is authorized to perform its transmission attempts immediately after the first idle DIFS interval it senses on the channel. This implies that such a station will perform an access to the channel without obtaining any information about the congestion level of the channel. Another problem is given with respect to the Slot Utilization definition. In particular, given some stations with initial Backoff equal to zero, it results possible to obtain values of Slot Utilization greater than one, because the number of transmission attempts a station can observe during its Backoff interval it is no longer upper bounded by the Initial Backoff value. So we modified the Slot utilization estimate, as

$$Slot\_Utilization = \frac{Num\_Busy\_Slots\_After\_Idle\_One}{Init\_Backoff}$$

where Num_Busy_Slots_After_Idle_One represents the number of busy slots belonging to the Backoff interval of the given station, and following at least one idle Slot. This simply means that we don't consider at all stations with Initial Backoff equal to zero, in the Slot Utilization estimate. As the number of stations with Initial Backoff equal to zero is a small percentage of the number of active stations, the proposed mechanism results in any case effective.

3 The DCC mechanism

In this section we introduce a new parameter, local to each station, which constitutes the core of the proposed mechanism. It allows to realize a filtering of access indications between the standard access scheme adopted by the system, and the physical layer, as shown in Figure 1. We start from the consideration that each station should evaluate its opportunity to perform or to defer the scheduled transmission attempt. This decision is based on the information obtained before to try a blind transmission attempt, according to the standard. In few words, the only reasonable behavior of a station that knows there are few possibilities for a successful transmission, is to defer its transmission attempt. For this reason, according to our mechanism, each station should base its transmission attempt on the value of a new parameter named Probability of Transmission. The value of this parameter is dependent both on the contention level of the channel, and on the number of unsuccessful transmission attempts already performed by the station. The latter parameter is used in our mechanism as an indicator of the dynamic priority achieved by a station. The lowest priority is given to stations performing the first transmission attempt for a given frame, while the priority level linearly increases with the number of collisions experienced. The reduction of contention is then obtained on a statistical basis, because each station, after the Slot Utilization estimate, computes the Probability of Transmission value, and adopts it to evaluate its opportunity to perform a true access to the shared channel. If the station decides to defer the transmission, it reschedules a new attempt, as in the case of a collision occurred (Figure 1). The probability to defer an access should obviously increase with the decrease of the users priority. Hence, this mechanism realizes the expected effective reduction of contention, facilitates the accesses by high priority stations, and provides a low probability of collision, with respect to the standard. This fact allows a greater drain by older transmission requests, and could introduce a progressive reduction of stations waiting for transmission, leading to a lower probability of collision, then to a lower mean access delay.

Now we present a heuristic expression to evaluate the Probability of Transmission (P_T) adopted in the proposed mechanism. To control a congested system, it results opportune to reduce the number of accesses, independently by the priority level of each user. This can be easily obtained with a definition of the Probability of Transmission (P_T) parameter which depends from the estimate of the Slot_Utilization as:

$$P\_T(Slot\_Utilization) = (1 - Slot\_Utilization) \quad .$$

The Slot_Utilization parameter, whose values are in [0..1] by definition, is interpreted here as the inhibition power of the mechanism on the accesses, depending on the contention level it represents. However, we can observe that such a definition would conduce the system to fluctuate between two states, with Slot Utilization zero and one, i.e. with no utilization of the channel. If the Slot_Utilization is high, then every user would obtain a low P_T value, implying a future low Slot_Utilization. This will cause every user to obtain a high P_T value, and again a future high Slot_Utilization. To avoid this harmful fluctuating behavior, the given definition is now extended introducing a second parameter connected with the priority level reached by the stations, i.e. the number of attempts already performed (Num_Att):

$$\begin{aligned} &P\_T(Slot\_Utilization, Num\_Att) = \\ &\qquad (1 - Slot\_Utilization^{Num\_Att}) \end{aligned}$$

The effect of the Num_Att parameter represents then a reduction of the inhibition power of the mechanism, for users with higher right of access (i.e. higher values of Num_Att), still maintaining the general effect to reduce the accesses to the network as a function of the Slot_Utilization. Note also that the mechanism's effect is entrusted to the effect of a probability value and not to a threshold value. A threshold value would be less smoothed and could introduce the risk of fluctuations. To understand what we obtain, we can observe the Figure 4. In this figure are shown the Probability of Transmission curves for users with different numbers of attempts performed, and with respect to the Slot_Utilization values estimated. Assuming a Slot_Utilization near to zero, we can observe how each station, independently by its number of attempts performed, obtains a Probability of Transmission near to one. This means that the proposed mech-

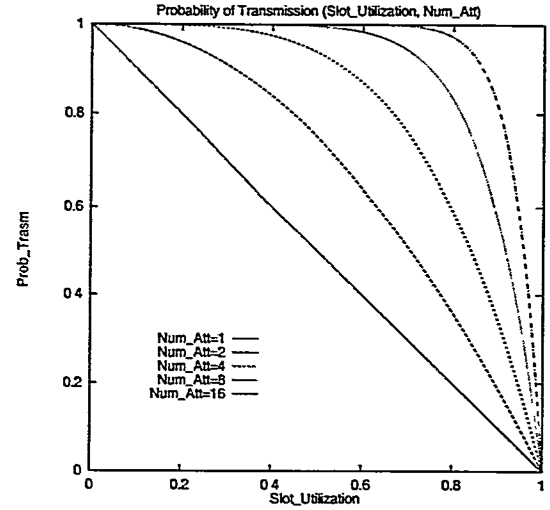

Figure 4: The Probability of Transmission functions.

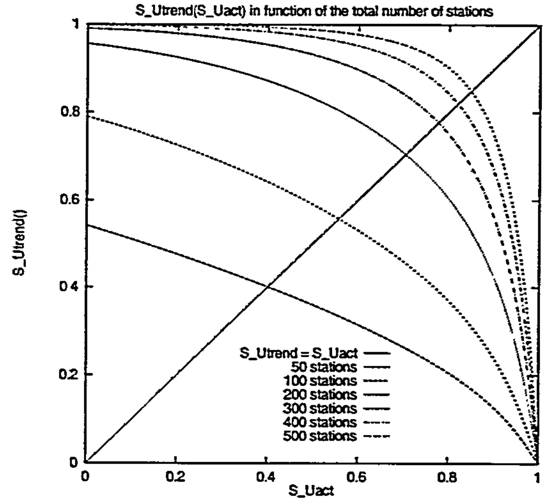

Figure 5: System's stability

anism has no effect on the system, and each user performs its accesses just like in the standard access scheme, without any contention control. This point is significant considering our target of absence of overhead introduced in low load cases. The difference between users with different levels of priority connected to the Num_Att parameter becomes clear when the Slot_Utilization grows. Assuming a Slot_Utilization near to one, e.g. 0.8, we see how a station with a high Num_Att value obtains a Probability of Transmission greater than others. It is simple to verify how the priority order relationship connected to the local parameter Num_Att maintains its invariance while it is mapped on the priority order relationship connected to the Probability of Transmission parameter. This is exactly what we expected to obtain to realize a reduction of contention based upon a priority principle, and limited to congested systems. Note that this mechanism is completely distributed, and each station does not have to worry about the Num_Att values obtained by other users. It is immediate to verify also that the Slot_Utilization of the system, under this mechanism, never reaches the value one. In this case the inhibition mechanism would perform asymptotically, by reducing the Probability of Transmission for each user, see figure 4. This also means that we could obtain as asymptotic limit for the resource utilization a given fixed limit, by means of the opportune definition of the expression for the Probability of Transmission (P_T).

The system behavior is then governed by the Probability of Transmission values, and the fastest reactions to variations in the Slot Utilization estimate are given for the highest priority stations. This fact means that in each case, statistically, the channel is utilized by the stations with the highest priority level in the system. Note that the deferring of the transmission attempts by low priority users is interpreted as a collision happening. Thus, it results in an increase of their priority level (by increasing Num_Att), and hence it maintains a gap with respect to newer access requirements.

On the basis of the Basic Access Mechanism of the Standard 802.11 DCF, we simply need to insert a new condition before the access indication reaches the hardware. The algorithm adopted by each station becomes the following:

```
...
if (the scheduled slot for transmission is reached)
   then calculate the Slot Utilization estimate
        evaluate the Probability of Transmission,
            P_T(Slot_Utilization, Num_Att)
        if Rnd() < P_T()
            then {transmit in the scheduled slot}
            else {defer}
        if (the station deferred) or (a collision occurred) then
            Schedule a new retransmission following
                standard indications.
```

It has to be noted how this algorithm can be executed on top of the standard basic access method, because it interferes only when a station performs an access to the channel (Fig. 1). Also, It has to be noted how it is possible now to obtain the indication about the channel congestion status, and then about the opportunity to double the CW_Size parameter, without necessarily performing a blind access with high collision risk. This allows to tune the CW_Size parameter even equal to CW_Size_MAX, but without performing any access to the channel, and so without contributing to collisions' generation. Moreover, this behavior is particularly desirable considering the possibility of bursty transmission requirements in the system.

To sum up, we realized the abstraction of a distributed statistical access authorization mechanism, based upon a priority principle, and devoted to obtain a reduction of contention, when it is needed. Our mechanism wishes to lead to a capture of the resource by those users with the highest priority. The mechanism is also adaptive by means of a feedback effect connected to the Slot_Utilization estimate, and it realizes an automatic tuning of the system behavior, without overheads introduced, and with optimal stability characteristics, as we can see in next section.

The station priority level, depends in our mechanism on the value of the Num_Att parameter. A good management of such a parameter is thus necessary. To this end it is important to accurately define how a station's priority increases. Simulation studies have indicated that a fast increment of Num_Att implies a fast flattening in the priority mechanism. To maintain the effectiveness of the priority mechanism, it is useful to delay the priority effect of the Num_Att parameter at least after the CW_Size reaches the CW_Size_MAX value, and only after each true attempt is performed. Initially, every deferring in the transmission attempt, based upon the Slot Utilization estimate, has to be connected only to the increasing of the CW_size parameter. Once the CW_Size reached the maximum value, and the transmission is still on going, then each unsuccessful transmission attempt implies the increment of the priority level connected to Num_Att. No increment of Num_Att is given when a station decides to defer a transmission attempt, for the previous reason.

It is also interesting to note that it is possible to extend our priority mechanism by introducing several classes of traffic in the network, without negotiations required. This extension could be realized with a little change in the Probability of Transmission definition, simply adding a local parameter Priority Level (Prior_Lev), with values greater or equal to 1 (default = 1):

$$P\_T(Slot\_Utilization, Num\_Att, Prior\_Lev) = (1 - Slot\_Utilization^{Prior\_Lev * Num\_Att})$$

The effect of the Prior_Lev parameter introduced is to provide a fast increment in the priority of the station, with respect to the number of transmission attempts performed. Simulation studies of such a mechanism have shown its effectiveness.

4 Stability of the system

In this section we analize, from a stability standpoint, a system constituted by a set of peer stations with continuous transmission requirements, implementing our congestion control mechanism. Our study is based on some basic principles of Equilibrium Point Analysis (EPA) [15]. The study focuses upon the observation of the Slot_Utilization of the channel, because it results the main variable which influences and indicates the system behavior. Our stability analysis is based on the analytical expression of the future Slot_Utilization that will be generated on the channel (S_U_{trend}) as a function of the actual Slot_Utilization level (S_U_{act}), and of the system's offered load. To perform this study, we consider a set of M peer stations with continuous transmission requirements. We define an equivalence relationship (≡) between stations such that:

$$\text{stationA} \equiv \text{stationB} \Leftrightarrow \text{stationA.Num\_Att} = \text{stationB.Num\_Att}$$

and we assume there are H classes of equivalence. This partition gives the distribution of stations among classes related with variable levels of the right to transmit. Assuming that, at a given time, each station knows the actual S_U_{act} value of the channel, with less or no distortion, we can argue the immediately induced $S\_U_{trend}(S\_U_{act})$ level on the channel. This is done by calculating the probability of at least one transmission in the immediately successive slot on the channel. So, let M be the cardinality of the set of stations, and let H be the number of classes related to the number of transmission attempts already performed. We can define a vector $\vec{h}$ such that h_i, $i \in [1..H]$, represents the number of stations trying to perform their i-th transmission attempt at the given time, and $\sum_{i=1}^{H} h_i = M$. We can also define a vector $\vec{cw}$, such that cw_i, $i \in [1..H]$, maintains the Contention Window size of a station belonging to class i, i.e.

$$cw_i = min(CW\_Size\_MAX, CW\_Size\_min * 2^{i-1}) \quad .$$

Given Num_Att the number of transmission attempts performed, and S_U_{act} the Slot Utilization estimate obtained, the probability for a given station to access the successive slot, say $P_{acc\_next}(S\_U_{act}, Num\_Att)$, is the product of the following probabilities:
$P_{sel}(Num\_Att) = 1/cw_{Num\_Att}$ = probability to select a given slot among those belonging to the station's Contention Window, and
$P\_T(S\_U_{act}, Num\_Att) = (1 - S\_U_{act}{}^{Num\_Att})$ = probability of transmission in a given slot.
The obtained expression for $P_{acc\_next}(S\_U_{act}, Num\_Att)$ is:

$$P_{acc\_next}(S\_U_{act}, Num\_Att) = P\_T(S\_U_{act}, Num\_Att) * P_{sel}(Num\_Att)$$

So, we obtain the probability of *at least one transmission attempt in the next slot interval*, i.e. the $S\_U_{trend}(S\_U_{act})$ expression for the whole system, like

$$S\_U_{trend}(S\_U_{act}) = (1 - \prod_{i=1}^{H}(1 - P_{acc\_next}(S\_U_{act}, i))^{h_i})$$

Now assuming that one or more equilibrium points exists for the system behavior, from the Slot Utilization point of view, and trying to investigate them, we impose the actual level of Slot Utilization be equal to the future level, i.e.:

$$S\_U_{trend}(S\_U_{act}) = S\_U_{act} \quad .$$

The number of possible values of S_U_{act} satisfying the given condition, gives also the indication of the number of possible equilibrium points of the system. We want to analyze their number and their characteristics, in order to evaluate every possible evolution of the system. Our results are obtained by plotting the points satisfying the equilibrium condition, i.e. the points of intersection between the line representing the system's trend of the Slot Utilization level and the load line representing the actual Slot Utilization level on the channel.

The Figure 5 shows a set of different decreasing lines which represent the values of $S\_U_{trend}()$ in function of S_U_{act} on the channel. Each line should refer to a system with a specified number of stations M, while H and the distribution of stations over H classes are specifically defined as $H = 10$ and the M stations equally distributed over the H classes. The straight line $S\_U_{trend}() = S\_U_{act}$ represents graphically the equilibrium condition. Thus, given the actual S_U_{act} value on the channel, the interpretation of the state of a given system is the following:

- if the corresponding value of $S\_U_{trend}(S\_U_{act})$ is over the value of S_U_{act} (i.e. over the straight line), then the expected Slot Utilization in the near future will grow, and the system will be pushed to work towards the S_U_{act} value corresponding to the point of intersection between the lines.
- if the value of $S\_U_{trend}(S\_U_{act})$ is under the value of S_U_{act} (i.e. under the straight line), then the expected Slot Utilization in the near future will decrease, and again the system will be pushed to work towards the S_U_{act} value corresponding to the point of intersection between the lines.

We can observe how, for each given system, and under the Slot Utilization point of view, there is only a single stable equilibrium point of the system. This fact has been proved by demonstrating that, for each system considered, the $S\_U_{trend}(S\_U_{act})$ expression is decreasing monotonic, and so there is a single point of intersection with the line representing the equilibrium condition. This point constitutes the level of Slot Utilization that the system would produce in the near future. Note that what we have to consider is not the value itself, which depends on the evolution of the system, and could change in time, but the fact that, in every condition, only a single equilibrium point is present. This single equilibrium point, gives good indications about the congestion reaction and control of the system. It assures that there is no possibility for the system behavior to be deceived by bad or useless convergence points. In particular, it has to be noted that a level of Slot Utilization near to one is determined in greater measure by high priority stations (i.e. by their accesses), which also obtain a lower level of contention due to the probability of transmission influence.

5 Simulation results

Now we present a set of results, obtained via discrete event simulation, we derived to analyze the system's behavior under the Standard 802.11 DCF and under the same Standard with the proposed DCC mechanism on top. To perform each simulation, it has been developed a simulation model for the Research Queueing Package Version 3 (RESQ) simulation tool. The physical parameters of the system investigated have been taken by literature [1,3,6], and are specified in Table 1. In our simulation experiments the number of active stations changes from 4 to 512, in order to represent different levels of congestion. It has to be intended as the number of active stations, with continuous transmission requirements. Each time a station performs a successful transmission, it resets the mechanism and immediately reschedules a new transmission. In our work we don't consider the Hidden Terminals problem [4,5,7,17], and we assume that the only cause of collision between two or more transmissions is given by their common slot selection. By means of this assumption we focus our attention to the problem of the reduction of the Probability of Collision, and so we can study the MAC protocol behavior strictly about the congestion control and congestion reactions point of view. Results presented in this work have been obtained with a constant frame size. On going studies with a variable frame size confirm and strengthen the results presented in this paper. All data plotted show their confidence interval with a confidence level of 95%.

5.1 Number of collisions for each successful transmission

The first evaluated parameter of the system is given by the mean number of collisions suffered for each given frame before a successful transmission attempt. This observation gives a meaningful indication of the improvement in the probability of collision. As shown in Figure 6, the proposed DCC Mechanism allows a significant reduction in the number of collisions. This reduction increases as the contention level increases, as well. Moreover, it has to be noted how the congestion reaction of the proposed mechanism maintains its linear functionality even if the number of stations in the system exceeds the reaction capability of the Binary Exponential Backoff mechanism, i.e. the number of stations exceeds CW_Size_MAX. On the contrary, when the number of stations becomes greater than CW_Size_MAX (256), it is possible to observe a knee of the investigated parameter in the Standard system.

5.2 Mean Access Delay

The Figure 7 shows the Mean Access Delay in a system with different levels of congestion. It is simple to verify how the proposed mechanism significantly improves this performance index, with respect to Standard values. It has also to be noted that no overhead is introduced, (i.e. the mean access delay does not increase) even in cases with few stations (Figure 8). This fact is really important, because it confirms our first intuition about the possibility to reduce the Mean Access Delay of the system by introducing a delay in the transmission attempts to be performed, when the contention level is high. Even if this seems a contradiction, it results useful, because of the double combination between reduction of the contention obtained, and priority mechanism which forces a queue-emptying behavior. Anyway, the great advantage of the mechanism remains the reduction of contention obtained.

The figure 9 shows the effectiveness of the priority mechanism, in terms of the Mean Access Delay obtained by three identical sets of stations with different self-determined traffic priority levels (Prior_Lev = 1,2,3).

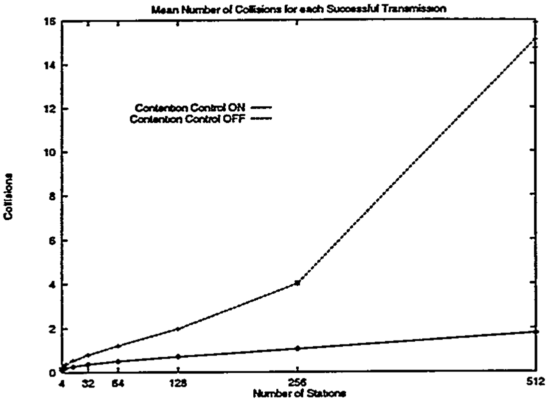

Figure 6: Collisions vs successful transmissions rate.

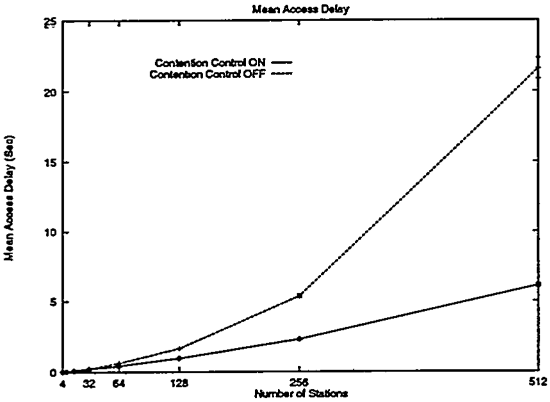

Figure 7: Mean Access Delay

5.3 99-th percentile of Access Delay

A further meaningful parameter to be investigated regarding the access delay is the 99-th percentile of Access Delay for each station in the system (Figure 10). This parameter confirms the generalized reduction in the access delay, assuring that the priority mechanism connected to the number of attempts performed results effective. This allow us to exclude that the reduction in the Mean Access Delay parameter was obtained only by means of a worsening in the access delay on behalf of a subset of unlucky stations. This is also a good indication about the reduction in the risk of starvation achieved with the given congestion control mechanism.

NUMBER OF STATIONS	(VARIABLE) 4..512	CW_Size_min	8 Slots
Channel Rate	2 Mb/s	CW_Size_MAX	256 Slots
Frame Payload	(CONSTANT) 1024 B	Ack size	30 B
Header size	34 B	Slot Time	20 µSec
SIFS	20 µSec	DIFS	50 µSec

Table 1: Simulation parameters for the considered system

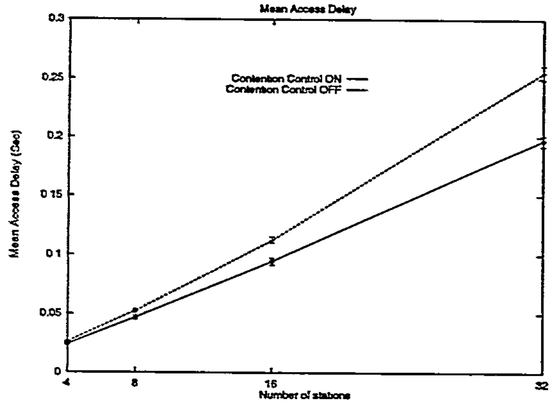

Figure 8: Mean Access Delay (zoomed)

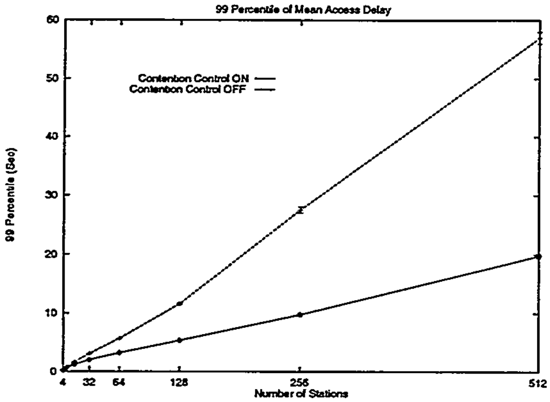

Figure 10: 99-th percentile of Access Delay

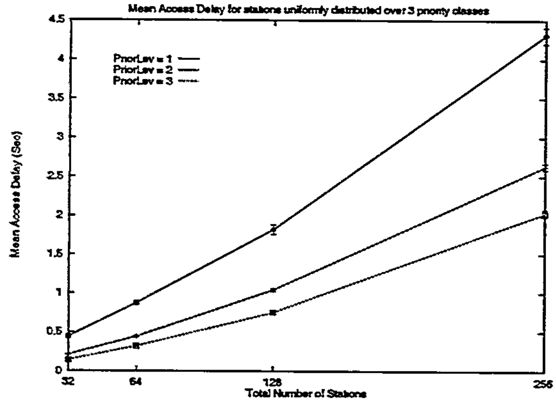

Figure 9: Mean access delay with Priority levels

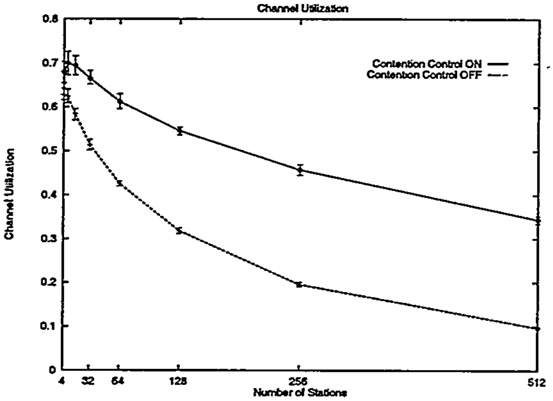

Figure 11: Channel Utilization

5.4 Channel utilization

Another important indication about the effectiveness of the proposed mechanism in the reduction of contention introduced in congested systems is given by the channel utilization level (Figure 11). It has to be noted the great improvement obtained in congested systems, and the absence of overheads (intended as reduction of the available channel bandwidth) introduced when there are few active stations. These observations are significant in terms of congestion prevention and recovery, considered the greater possibility to keep moving the system, with respect to the standard definition.

5.5 DCC behavior with bursty arrivals

Figures 12 and 13 show the analysis of the system behavior when the congestion level drastically changes. Specifically, in Figure 12, we considered a system with 128 steady-state active stations with continuos transmission requirements. At a given time instant (after about 6 seconds on the figure) additional 128 stations become active, each one with a single frame to transmit. In our experiments we measured the time the system takes to satisfy all the burst constituted by the 128 additional transmission requirements. In Figure 13 we report the same experiment's results considering an initially empty system, and a burst of 256 transmission requirements.

In both scenarios, the system which adopts the DCC Mechanism realizes a better and faster congestion reaction, with respect to the standard system. This indicates a greater capability to contrast the congestion growth, and so a better congestion prevention. Note that the greater velocity in satisfying the transmission requirements is mainly due to the greater channel utilization obtained. Note also the almost complete absence of tail in the DCC mechanism's curve (Figure 12), which demonstrates again the effectiveness of the priority mechanism introduced, based upon the Num_Att parameter, and then the achieved queue-emptying behavior.

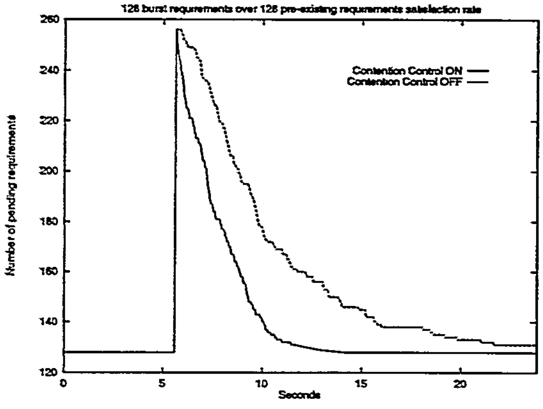

Figure 12: 128 bursty arrivals (congested system)

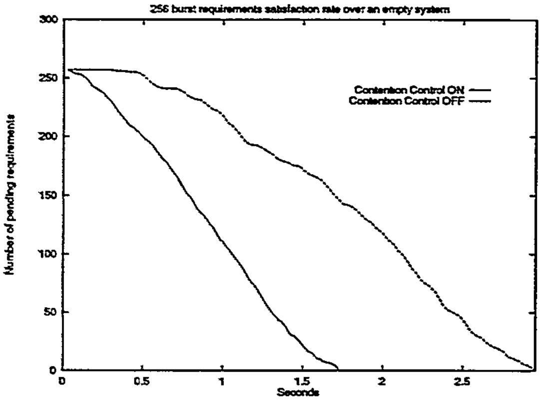

Figure 13: 256 bursty arrivals (empty system)

6 Conclusions and Future Researches

The proposed DCC mechanism results very effective in terms of contention control, congestion avoidance and congestion recovery, and it is completely distributed. It is able to individuate a congestion condition and it behaves in a fully adaptive way, leading to a system's behavior characterized by high stability, and with no overheads introduced even in low contention conditions. The analysis of the system under the proposed DCC mechanism shows the absence of harmful fluctuations, either during or after the transient phases. No further hardware, communication, power consumption or significant time is required with respect to the Standard requirements. The DCC mechanism can be used on top the 802.11 Standard DCF, with no significant modifications required, and moreover it allows to realize a better Contention Window sizing for the implemented Binary Exponential Backoff mechanism, with less or no impact on collision generation. Among the advantages of the proposed mechanism, we can observe an improvement in the channel utilization, mean access delay and 99-th percentile of access delay. Furthermore, our mechanism guarantees a better response of the system to transient changes in the contention level, and it reduces the risk of starvation. These advantages increase with the increase in the system's congestion level. It also results possible to introduce a distributed priority mechanism, simple and effective, which allows each station to self-determine its priority level without negotiations.

We are now investigating the evolution of the DCC mechanism to capture the relationship between the traffic characteristics (mainly the average message length) and the maximum channel utilization level [2]. This evolution is based both on an analytical study of the MAC protocol capacity similar to the one proposed in [3], and on a low cost estimate of the size of the transmitted frames. Future research on DCC mechanism will include the adoption of smoothing functions to estimate the Slot Utilization, the design of different scheduling functions, and the exploitation of the priority level mechanism for RTS/CTS packets [7], in order to privilege their transmission.

It is worth noting that the effectiveness of DCC could be exploited in other contexts of distributed systems involving resource sharing, provided that they adopt some kind of random access control, and that it is possible to obtain some indications about the contention level of resource accesses.

References

[1] 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.

[2] Bononi L., Conti M., Gregori E., "IEEE 802.11 Wireless LANs: Performance Analysis and Distributed Contention Control (DCC) Enhancement", Technical Report, 1998.

[3] Cali' F., Conti M., Gregori E., "IEEE 802.11 Wireless LAN: Capacity Analysis and Protocol Enhancement", Proc. INFOCOM'98, San Francisco, CA, March 29 - April 2, 1998, pp. 142-149.

[4] Chen K.C., "Medium Access Control of Wireless LANs for Mobile Computing", IEEE Networks, 9-10/1994.

[5] Chhaya H.S., "Performance Evaluation of the IEEE 802.11 MAC Protocol for Wireless LANs", Master Thesis, Graduate School of Illinois Institute of Technology, May 1996.

[6] Crow B.P., "Performance Evaluation of the IEEE 802.11 Wireless Local Area Network Protocol", Master thesis. University of Arizona, 1996.

[7] P802.11, IEEE Draft Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Dec. 1995.

[8] Gallagher R.G., "A perspective on multiaccess channels", IEEE Trans. Information Theory, vol. IT-31, No.2, 3/1985, pp. 124-142.

[9] Georgiadis L., Papantoni-Kazakos P., "Limited Feedback Sensing Algorithms for the Packet Broadcast channel", IEEE Trans. on Info. Theory, Vol IT-31, No.2, 3/1985, pp. 280-294.

[10] 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.

[11] Hammond J.L., O'Reilly P.J.P., Performance Analysis of Local Computer Networks, Addison-Wesley 1988.

[12] 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.

[13] Rivest R.L., "Network Control by Bayesian Broadcast", IEEE Trans. on Info. Theory, Vol IT-33, No.3, 5/1987, pp. 323-328.

[14] Stallings W., Local & Metropolitan Area Networks, Fifth Edition, Prentice Hall 1996, pp. 356-383

[15] Tasaka S., Performance Analysis of Multiple Access Protocols, MIT Press 1986.

[16] Tsitsiklis J.N., "Analysis on a multiaccess Control Scheme", IEEE Trans. on Auto. Contr., Vol AC-32, No. 11, 11/1987, pp. 1017-1020.

[17] 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.