Carrier Sense Multiple Access with Enhanced Collision Avoidance

A thesis presented
by

Jaume Barceló Vicens

to
The Department of Information and Communication Technologies
in partial fulfillment of the requirements
for the degree of
Philosophy Doctor
in the subject of

Digital Communications

Advisor: Miquel Oliver i Riera

Barcelona
January 2009

UNIVERSITAT
POMPEU FABRA

Per sa familia, que sempre m'ha recolzat.
I per na Geni.

The Thesis in Plain Words

If you are familiar with random medium access mechanisms, you can safely skip this page. However, if you are one of those that believes that IEEE 802.11 is a character in the Star Wars saga, this page is for you.

Your laptop can connect to the Internet wirelessly. This is possible because data is transmitted from your laptop to the access point (a device that can be usually recognized by its antennae and flashing lights) using radio waves. It is useful to think that the devices talk to each other.

If two or more devices talk simultaneously, the words are garbled and no one can understand what the others are saying. For this reason the devices listen before talking. If a device detects that another device is currently talking, it will postpone its transmission.

It is still possible that two devices begin to talk simultaneously. This thesis presents an approach that minimizes the chances that two devices simultaneously begin to transmit.

Summary

This thesis explores the limits of the CSMA/CA random multiple access protocol, which is used in WLANs. It is shown that the performance of the protocols that randomly choose the slot at which transmission occurs is bounded by a fundamental trade-off. If the contenders aggressively transmit, the probability of collision is high. Conversely, if the contenders use a low transmission probability (i.e. separate their transmission attempts by a large number of slots), the performance suffers because most of the slots remain empty. Although the transmission probability can be optimized, the resultant efficiency is still far from satisfactory. A conceptual change in the protocol is required to overcome the aforementioned fundamental bound.

Nevertheless, randomness is of paramount importance for resolving collisions. After a collision, it is desired that the implicated parts backoff for a different number of slots, in order to prevent that they collide in their next transmission attempt. Given the facts that random selection of the transmission slot limits the performance and that randomness is necessary to resolve collisions, a modification to CSMA/CA is proposed. It is suggested to use a deterministic backoff after successful collisions and a random backoff otherwise.

The immediate consequence is that, in saturation conditions, the stations that successfully transmitted in their last transmission attempt cannot collide in their next transmission attempt. Hence, the new protocol reduces the chances of collisions and thus it is named CSMA with Enhanced Collision Avoidance (CSMA/ECA). Moreover, if the number of contenders does not exceed the value of the deterministic backoff after successes, the systems converges to a collision free operation. After all the stations successfully consecutively transmit, collisions disappear. The suppression of collisions have a positive impact on the channel efficiency, which is the fraction of channel time devoted to successful transmissions. Actually, the performance of CSMA/ECA surpasses the efficiency upper bound associated to those protocols that randomly select the transmission slot.

A Markov Chain is proposed to model the convergence of the system. As expected, the system converges almost instantaneously when the number of contenders is low, but it takes a longer time as the number of contenders approaches the value of the deterministic backoff. Only probabilistic guarantees can be provided regarding the convergence time. As the system can be moved from the stationary operation to a new transitory due to a channel error or a new entrant, the recovery process back to the stationary operation is also studied.

By means of simulation, it is shown that it is possible that stations using legacy CSMA/CA and the proposed CSMA/ECA can smoothly coexist in the same network. Simulations are also used to assess the performance of CSMA/ECA in lossy channels, and the results indicate that CSMA/ECA also ouperforms CSMA/CA in challenging channel conditions. The protocols are also tested with a variety of traffic patterns: rigid flows, elastic flows, and mixed scenarios. The experiments are repeated with and without RTS/CTS. In all cases, the proposed protocol outperforms the existing one.

To gain insight in the operation of CSMA/ECA, a model that takes into account the queue occupation is proposed and validated. The model accurately predicts channel status probabilities and system throughput. The results also show that in saturation conditions, collisions are prevented in the stationary operation of the network.

Since in infrastructure WLANs the traffic is highly asymmetric, in the sense that the AP has to send data to all the stations and thus requires more channel time, the support for traffic differentiation in CSMA/ECA is developed. Two of the well-known techniques for traffic differentiation in CSMA/CA are also applicable to CSMA/ECA: namely, the transmission opportunity (TXOP) and variable contention windows (CW_{min} and CW_{max}).

Resum

Aquesta tesi explora els límits del protocol d'accés al medi CSMA/CA, que és utilitzat a les xarxes locals sense fils WLAN. Es mostra que l'eficiència dels protocols que elegeixen de manera aleatòria la ranura temporal en la que es produeix la transmissió està acotada per un límit fonamental. Si els nodes transmeten sovint, la probabilitat de col·lisió és alta. En canvi, si els nodes transmeten poc (és a dir, separen les seves transmissions per un nombre elevat de ranures), el rendiment pateix ja que la majoria de ranures romanen buides. Tot i que la probabilitat de transmissió es pot optimitzar, l'eficiència resultant encara no és del tot satisfactòria. Cal un canvi conceptual en el protocol per superar el límit fonamental abans mencionat.

De totes maneres, l'aleatorietat és molt important per a resoldre les col·lisions. Després d'una col·lisió, és desitjable que les parts implicades s'esperin durant un nombre diferent de ranures abans de reintentar transmetre. Donat que l'aleatorietat total limita l'eficiència i que la aleatorietat és necessària per evitar col·lisions, es proposa una modificació a CSMA/CA. Es suggereix que els nodes s'esperin un nombre de ranures determinista després de les transmissions exitoses i un nombre de ranures aleatori en qualsevol altre cas.

La conseqüència immediata és que, en condicions de saturació, les estacions que han transmès amb èxit en el seu darrer intent de transmissió no poden col·lisionar entre elles en el seu proper intent de transmissió. Per tant, el nou protocol redueix les possibilitats de col·lisió i s'anomena CSMA amb evitament de col·lisions millorat (CSMA with Enhanced Collision Avoidance). Encara més, si el número de nodes que estan competint pel canal no supera el valor determinista del compte enrere utilitzat després de les transmissions exitoses, el sistema convergeix a un mode d'operació sense col·lisions. Després de que totes les estacions transmetin amb èxit de manera consecutiva, les col·lisions desapareixen. La supressió de les col·lisions té un impacte positiu en l'eficiència de canal, que és la fracció del temps de canal que es dedica a transmissions exitoses. En realitat, l'eficiència de CSMA/ECA supera el límit teòric associat a aque-

lls protocols que sempre elegeixen de manera aleatòria la ranura de transmissió.

Es proposa una cadena de Markov per a modelar la convergència del sistema. Com era d'esperar, el sistema convergeix de manera gairebé instantània quan el nombre de competidors és baix, però triga més a mesura que el nombre de competidors augmenta. Únicament es poden donar garanties probabilístiques pel que fa al temps de convergència. El sistema pot passar de l'estacionari al transitori a causa d'un error de canal o bé de la incorporació d'un nou competidor. Per tant, s'estudia també el temps de recuperació per tornar un altre cop a l'estacionari.

Es mostra mitjançant simulacions que estacions que utilitzen CSMA/CA poden conviure en la mateixa xarxa que estacions que utilitzen CSMA/ECA. Les simulacions també s'utilitzen per avaluar l'eficiència de CSMA/ECA en canals que perden paquets, i els resultats indiquen que CSMA/ECA també supera a CSMA/CA en aquest tipus de canals. Ambdós protocols són provats enfront una varietat de patrons de tràfic: fluxes rígids, fluxes elàstics i escenaris mixts. Els experiments també es realitzen amb i sense RTS/CTS. En tots els casos, el protocol proposat supera a l'existent.

Per tal d'entendre millor com funciona CSMA/ECA, es proposa i valida un model que té en compte l'ocupació de les cues. El model prediu amb precisió quin és l'estat del canal i el rendiment. Els resultats mostren que, en saturació i durant l'estacionari, s'eviten les col·lisions.

Com que a les xarxes sense fils basades en infrastructura el tràfic és molt asimètric, en el sentit que un mateix punt d'accés ha de transmetre dades a totes les estacions i per tant necessita més temps de canal, es desenvolupa el suport per a diferenciació de tràfic en CSMA/ECA. Dues tècniques de diferenciació de tràfic ben conegudes de CSMA/CA també són aplicables a CSMA/ECA: l'oportunitat de transmissió (TXOP) i les finestres de contenció variables (CW_{min} i CW_{max}).

Preface

This thesis dissertation is presented as a compilation of the following articles:

J. Barcelo, B.Bellalta, C. Cano, M. Oliver,
“Dynamic P-Persistent Backoff for Higher Efficiency and Implicit Prioritization”,
In Proceedings JITEL’08.

J. Barcelo, B.Bellalta, C. Cano, M. Oliver,
“Learning-BEB: Avoiding Collisions in WLAN”,
In Proceedings EUNICE’08.

J. Barcelo, A. Lopez-Toledo, C. Cano, M. Oliver,
“CSMA/ECA: Carrier Sense Multiple Access with Enhanced Collision Avoidance”.

J. Barcelo, B.Bellalta, A. Sfairopoulou, C. Cano, M. Oliver,
“CSMA with Enhanced Collision Avoidance: a Performance Assessment”.
In Proceedings IEEE VTC Spring’09.

J. Barcelo, B.Bellalta, A. Sfairopoulou, C. Cano, M. Oliver,
“CSMA with Enhanced Collision Avoidance: a Performance Analysis”.

J. Barcelo, B.Bellalta, C. Cano, A. Sfairopoulou, M. Oliver, J. Zuidweg,
“Traffic Prioritization for Carrier Sense Multiple Access with Enhanced Collision Avoidance ”.
In Proceedings MACOM (ICC’09).

A complete discussion about the particularities of this thesis format can be found in [1]. The benefits of this approach are twofold. First, the young researcher is trained in the type of writing that will be used after receiving the doctorate. And second, it eases the dissemination of the pre-doctoral contributions to a wide audience of professional colleagues.

The articles selected for inclusion in this dissertation, represent only a fraction of the publications generated during the Ph.D. program. The following is a partial list of the remaining of the first-authored papers generated throughout the doctoral program. They are sorted in reverse-chronological order:

J. Barcelo, A. Sfairopoulou, B.Bellalta,
“ Wireless Open Metropolitan Access Networks”,
In ACM Mobile Computing and Communication Review. Tentative publication date: July 2009.

J. Barcelo, B.Bellalta, C. Cano, M. Oliver,
“ No Ack in IEEE 802.11e Single-Hop Ad-Hoc VoIP Networks”,
In book IFIP Volume 265, Advances in Ad Hoc Networking, Boston:Springer, pp. 157-166.

J. Barcelo, B.Bellalta, C. Cano, A. Sfairopoulou,
“ VoIP Packet Delay in Songle-Hop Ad-Hoc IEEE 802.11 Networks”,
In Proceedings IFIP/IEEE WONS’08.

J. Barcelo, B.Bellalta, C. Cano, A. Sfairopoulou,
“ Position Information for VoIP Emergency Calls Originating from a Wireless Metropolitan Access Network”,
In Proceedings INSTICC WEBIST’07.

J. Barcelo, C. Macian, P. Novell, E. Arago and M. Isidre,
“ Offering VoIP Services in a Wireless Neutral Operator Environment ”,
In Proceedings HOJ/IEEE MCWC’06.

J. Barcelo, M. Oliver and J. Infante,
“Adapting a Captive Portal to Enable SMS-Based Micropayment for Wireless Internet Access”,
Lecture Notes in Computer Science, Volume 4033, 2006, Pages 78 - 89. ISSN: 0302-9743.

J. Barcelo, C. Macian, J. Infante, M. Oliver and A. Sfairopoulou,
“Barcelona’s Open Access Network Testbed”,
In Proceedings IEEE TRIDENTCOM’06.

The Thesis in Context

Data communication networks are governed by protocols, which are sets of rules that orchestrate the information interchange. These protocols are organized in layers that conform the *protocol stack*. The Internet protocol stack comprehends four layers: namely, the application layer, the transport layer, the network layer, the link layer and the physical layer.

Each layer accomplishes a different function. The application layer is the one that provides services to the user, such as web browsing or e-mail. The transport layer is in charge of transporting the application layer messages between processes in different hosts. The network layer takes information from the source host and delivers it to the destination host, often travelling across multiple networks. The link layer is responsible for one one-hop communications and the physical layer converts the information into signals that can be propagated through a communication medium. This signals typically take the form of electrical, radio or light waveforms.

The link layer protocols vary from network to network, since they are closely related to the physical medium being used. Popular link layer protocols include ethernet for wired communications and WiFi for wireless communications. The specifications of these two link layer protocols can be found in the IEEE 802.3 and IEEE 802.11 standards, respectively. The IEEE specifications span to the physical layer, providing support for different media, modulations and data rates.

There are some networks that provide each terminal with a dedicated medium, such as current switched Ethernet networks. There are other networks in which a common channel is shared among various stations. This is the case of wireless local area networks (WLANs). When the medium is shared, there is one aspect of the link layer that becomes particularly critical, which is the Medium Access Control (MAC). A comprehensive study of MAC can be found in [2].

This dissertation presents a contribution in the field of Medium Access Control (MAC) in WLANs. There are different strategies in sharing the channel, and this thesis is focussed on random medium access channel, which has proven extremely successful since it is the one implemented in the pervasive WiFi networks. The proposed protocol is called Carrier Sense Multiple Access with Enhanced Collision Avoidance (CSMA/ECA), and the randomness is administered with great caution. Actually, a deterministic behaviour is used whenever the randomness is not strictly required. The result is that the system converges to a deterministic medium access mechanism attaining a performance that has no precedent in previous random protocols.

Thesis Outline

The first article of the thesis studies CSMA/CA, which is the protocol that is adopted in the IEEE 802.11 standard for WLANs. It presents an analysis of the maximum achievable channel efficiency and a mechanism to attain near-optimal performance in CSMA/CA. The performance limit on CSMA/CA has its origin in the fact that all the stations randomly decide when to transmit.

This subtlety becomes obvious in the second article, in which it is proposed that a deterministic backoff should be used after successful transmissions. By this simple modification, the system clearly outperforms the maximum theoretical limit of CSMA/CA. This article also explains that the benefits of using a deterministic backoff may not be apparent for the first transmissions attempts. It takes some time for the system to transition from a CSMA/CA operation to a collision-free operation.

There is an overlap between the first and second article. Specifically, the Sec. 3 of both articles present the same contribution. The reason is that the second article was submitted before the first one was accepted, and thus it was not possible to cite the first paper in the second one.

The third article further studies the idea presented in the second one. The new protocol receives its name (CSMA/ECA) and its behaviour is dissected. A model for the convergence process is introduced, by using a Markov Chain. The model reinforces the intuitive idea that the convergence is almost immediate for the most usual scenarios. The convergence process can also be avoided by using *smart entry*. Key advancements of this third article are the study of the feasibility of the coexistence of both CSMA/CA and CSMA/ECA in a same network and the evidence that CSMA/ECA outperforms CSMA/CA also in non-ideal channel conditions.

The fourth article addresses the performance of CSMA/ECA in the presence of rigid rigid flows. This kind of flows, as opposed to the elastic ones, transmit a fixed bandwidth and do not adjust to the network capacity or the network conditions. The results show that the behaviour of CSMA/CA and CSMA/ECA are exactly the same in non-saturated conditions. However, as the network approaches and reaches saturation, CSMA/ECA outperforms CSMA/CA.

The fifth article proposes a model that analytically reflects the findings of the previous article. It is required that the model takes into account the occupation of the queues, since the behaviour of CSMA/ECA is drastically different for empty and non-empty queues.

The sixth article, presents traffic prioritization for CSMA/ECA. For the new protocol to be widely adopted, it is required that it provides traffic differentiation which is no worse than the one already available in current network equipment. It is shown that CSMA/ECA can be easily adapted for traffic differentiation and, again, it beats the performance of CSMA/CA.

The six articles that constitute the core of this thesis gather the research results obtained so far. However, by working on this ideas and collecting feedback from colleagues, it becomes apparent that the idea of deterministic backoff after successes can be taken even further. The *final remarks* of this thesis review ongoing work and propose ideas and thoughts that might crystallize in the near future.

Motivation, Goals and Contribution

The link layer of WLANs suffers from efficiency problems since only a fraction of the channel time is devoted to successful transmissions. For this reason, it has received much attention from the research community in recent years and a myriad of protocols have been proposed. However, until now, there is not a strong candidate to be the successor of CSMA/CA. To succeed, a new MAC protocol should meet the following requirements:

- Significantly improve the efficiency by reducing the chances of collision.
- It has to be a distributed algorithm that does not require a central controlling entity.

- It has to be backward compatible with CSMA/CA, to allow a seamless transition in time from one protocol to the other.
- It should not introduce additional signaling or overhead.
- Support for traffic asymmetry, to prevent uplink/downlink unfairness.
- Suitability for bursty traffic.
- Generality, to be applicable to the whole IEEE 802.11 protocol family and other networks.
- It should not present additional requirements in terms of memory of computation power.
- Similarity to CSMA/CA is also a desired feature, in order to ease the transition path to standardization bodies and manufacturers.

The main contribution of this thesis is presenting a modification to the MAC protocol used in the pervasive IEEE 802.11 networks. The modified protocol, which is called CSMA/ECA, uses a deterministic backoff value after successful transmissions, as opposed to the ran dom value used in current implementations. To the best of our knowledge, CSMA/ECA is the first proposal that satisfies all the above mentioned requirements.

January 2009 Jaume Barcelo

References

1. Nell K. Nuke and Sarah W. Beck. Education Should Consider Alternative Formats for the Dissertation. *Educational Researcher*, 28(3):31–36, 1999.
2. R. Rom and M. Sidi. *Multiple access protocols: performance and analysis.* Springer-Verlag New York, Inc. New York, NY, USA, 1990.

Acknowledgements

First of all, I want to acknowledge Anna, for she is the one that had to share the office with me for a longer time (which might not be easy). Same thing with Cristina and Laura. Boris, Carlos, Jorge, Victor and Xavi and all of those that were around and came (or were called) to help me at some moment. Also a special thanks for Miquel, my advisor. The fact that I continued under his supervision after finishing my master thesis should suffice to prove how much I trust and appreciate him. I am very grateful for his continuous support and expert advise. Boris took the role of a second advisor, and he was the one that came up with the ideas for some of the papers included in this thesis. And thanks to Alberto and Domenico, for the inspiring discussions during the last steps of the thesis.

I would like to thank the secretaries (for handling the paperwork) and the helpdesk and communication services, for attending those unusual request coming from someone that likes to tinker with computers and data networks.

Since the thesis contains published material, I acknowledge the task of the anonymous reviewers that uninterestedly provided comments and suggestions.

I also feel indebted with the open-source community that provides great software that makes our daily work easier (and sometimes even fun). The algorithm suggested in this thesis can be seen as a payback, since making the wireless networks faster will make the nerd community happier.

Jorge Cham: thanks for the PhD (Piled Higher and Deeper) comics, who helped me to understand what a grad student life is about.

This thesis about collision avoidance can be seen as a perfect blend of my mother's mathematical background and my father's passion for conflict resolution.

Table of Contents

Carrier Sense Multiple Access with Enhanced Collision Avoidance

Dynamic P-Persistent Backoff for Higher Efficiency and Implicit Prioritization

J. Barcelo, B. Bellalta, C. Cano, M. Oliver

Universitat Pompeu Fabra

Abstract. This article studies the efficiency of backoff algorithms. The fraction of channel time devoted to successful transmissions is maximized when the stations choose the optimal transmission probability. The binary exponential backoff algorithm does not come close to optimal channel efficiency, thus a new backoff mechanism that attains near-optimal efficiency is proposed. This algorithm is called Dynamic-P-Persistent backoff and is based on the observation that, under optimal efficiency conditions, the fraction of channel slots busy with collisions is constant. The stations monitor the channel to estimate the fraction of collision slots and adjust their transmission probabilities consequently. As opposed to previous backoff proposals, DPP does not require any estimation of the number of concurrent active stations. Further, DPP offers implicit prioritization that reduces the delay of real time and interactive traffic while maintaining optimal throughput for background traffic.

1 Introduction

Wireless networks build upon the IEEE 802.11 [1] standard and its different flavors are growing and proliferating at universities, enterprises and homes. In each of these networks, the stations and access points share a common channel to transmit data. Being the air a broadcast channel, the participants in the network should avoid to transmit simultaneously. If two participants do transmit at the same time a collision occurs and the data of both senders might be lost. It is the duty of the Medium Access Control (MAC) layer to handle collisions and minimize their impact on performance.

This is not a new problem; it already appeared in early Aloha [5] and Ethernet [2] networks. There are two general techniques that effectively improve the efficiency of this kind of networks. The first one consists on sensing the channel before transmitting (Carrier Sense Multiple Access, CSMA [12]). If the channel is sensed busy, it means that there is an ongoing transmission and the other participants will refrain from transmitting to avoid a collision. Further, limiting the instants at which the participants can begin a new transmission, also reduces the number of collisions. The time is divided in slots and transmissions are allowed only at the beginning of each slot. There is a collision if two or more stations choose the same slot to transmit. To reduce the probability of a collision, it is necessary to randomize the selection of the time slot at which a given station transmits.

In P-persistent protocols, the stations involved in a collision retransmit in the following slot with probability P. With probability $1 - P$ the retransmission is postponed for the next slot. This operation repeats until the station finally retransmits. In a more sophisticated backoff algorithm, the stations involved in a collision draw a random number from a contention window (*e.g.* a number between 0 and 31) and then wait for that number of slots before re-attempting transmission. If the random values are selected from a contention window that doubles after each failed attempt, the mechanism is called Binary Exponential Backoff (BEB). A variant of this scheme called Truncated BEB (T-BEB) is the contention algorithm of choice for IEEE 802.11 networks.

IEEE 802.11 medium access comes in two different flavors. The most simple (Basic Access) consists on a two-way handshake in which the sender transmits a packet and waits for the receiver to explicitly acknowledge the correct reception with a short packet. When a collision occurs, a considerable amount of time is wasted since the senders cannot detect the collision while they are transmitting. This implies that the senders will not immediately interrupt transmission when a collision occurs. Conversely, the transmitters will send the whole packet and will only realize that a collision has happened because of the lack of acknowledgement.

To prevent collisions, RTS/CTS can be used. It is a more elaborated four-way handshaking mechanism in which the sender requests permission to send (Request-To-Send) and the receiver grants the permission (Clear-To-Send) effectively reserving the channel for the duration of the transmission and acknowledgement. This approach also solves the hidden terminal problem. The hidden terminal problem occurs when two terminals that can not hear to each other have a packet ready to transmit. If this is the case, the carrier sense mechanism will not work and both stations will transmit simultaneously. The problem arises when the receiver is in the hearing range of both transmitting stations and the collision occurs.

Due to the additional control messages, RTS/CTS access places an additional overhead on the channel that penalizes performance. For this reason, the rest of the article focuses on the Basic Access two-way handshaking mechanism. To simplify the analysis, it is considered that all the participating stations share a common broadcast channel, and each station can hear the transmissions of all the other stations.

After this first introductory section, the remaining of the paper is organized as follows. Sec. 2 reviews previous art and highlights the contribution of this paper. Sec. 3 describes T-BEB and proposes a general framework to assess the efficiency of backoff mechanisms in general. This framework is used to derive the optimum efficiency, which can be used as a benchmark to compare backoff schemes. It is observed that the maximum efficiency is a function of both the packet length and the number of contending stations. Further, it can be concluded that T-BEB performs less-than-optimal in most of the cases. The finding that the fraction of collision slots is constant when optimal transmission probability is used is crucial to derive a near-optimal backoff algorithm.

Sec. 4 introduces Dynamic-P-Persistent (DPP) backoff protocol. It is a variant of P-Persistent backoff that constantly monitors the number of collision slots and adjusts the transmission probability to attain optimal collision probability. Since the collission probability is independent of the number of active stations, this proposal delivers near-optimal performance for any number of competing stations. It is noticeable that the estimation of the number of backlogged stations is not required.

Sec. 5 presents simulations results to support the analysis of the previous sections. A first simulation shows how the stations adjust their transmission probability as the number of stations varies. This simulation offers an intuitive understanding of the behaviour of the mechanism in a dynamic environment. Then, extensive simulations assess the efficiency of DPP and show how close it is to the upper bound obtained in Sec. 3.

The proposed backoff scheme comes with advantageous implicit prioritizing features that are explored in Sec. 6. DPP benefit stations that generate real-time and interactive traffic and penalizes those that are permanently active sending background traffic.

Finally, Sec. 7 summarizes the paper and provides some concluding remarks.

2 Related Work

The Truncated Binary Exponential Backoff is a protocol to control multiple-access broadcast channels. It is a distributed access mechanism in the sense that each station independently executes the algorithm to decide whether to transmit or not in a given time slot. Each station selects a number from a contention window and waits for that number of slots before attempting transmission. The contention window doubles after each failed transmission attempt and resets to its minimum value after a successful transmission. It is called Truncated, because when reaching a maximum backoff stage (m) the contention window does not double any more. Additionally, a packet is dropped after reaching the maximum number of retransmission attempts (R). The properties of BEB and T-BEB have been extensively studied in [7, 11, 13] to cite a few.

CSMA and T-BEB are widely used in WLAN since they are at the core of the Distributed Coordinated Function (DCF) defined in IEEE 802.11. Any improvement in the backoff mechanisms would traduce in increased performance of the ubiquitous WiFi networks. Moreover, CSMA and T-BEB also appear as an ingredient of many MAC layer proposals supporting upcoming networks such as (Mobile) Ad-Hoc Networks [18], Mesh Networks and Personal Area Networks [3].

The studies are performed under saturation conditions, *i.e.* each station has always a packet to transmit. This is the maximum load that can be offered to the network and it is assumed that it is the maximum strain to which the network may be exposed. The properties of interest include fairness (both short-term and long-term), stability and efficiency. In this paper the focus is placed on efficiency (the fraction of channel time devoted to successful transmissions).

Given a data rate, this metric can be translated to throughput which is widely used in the literature.

The backoff protocols put the stations on hold thus diminishing the chances that a station attempts transmission in any given slot. The backoff effectively influences the frequency with which stations transmit. Another way to interpret the effect of the backoff is to understand that it tunes the transmission probability.

In [9], it was already stated that the optimal transmission probability is a function of the packet length (l) and the number of competing stations (n). A p-persistent backoff mechanism was also suggested to study the behaviour of T-BEB. The maximum efficiency of T-BEB was estimated by minimizing the average virtual transmission time. Similarly to our work, an algorithm to tune the transmission probability to improve the efficiency was proposed. The main difference lies in that the estimation of the number of competing stations is not required in our algorithm.

Previous efforts focused on inferring the number of stations from the number of empty, busy and collision slots. Specifically, [8] shows that the number of active stations can be expressed as a function of the collision probability encountered on the channel. Additionally, it proposes an extended Kalman filter coupled with a change detection mechanisms to estimate the number of contending stations n. A remarkable advancement was presented in [14] in which a bayesian approach was adopted to estimate the number of competing terminals.

Other works [6] assume that the number of contending stations is known (either using one of the estimation techniques cited above or assuming that the information is directly available at the AP) and then compute the optimal – fixed – contention window. A fixed (as opposed to T-BEB's exponentially-growing) optimal contention window increases performance both in terms of efficiency and fairness.

Another line of research consists on cross-layer techniques that combine BEB, Tree Algorithms [10] , and successive interference cancellation [17]. However, these studies maximize the number of successful slots while neglecting the fact that empty slots are much shorter than collision slots. In Sec. 3 it is explained that the different duration of the slots is of paramount importance in computing channel efficiency.

Finally, there is a game-theoretical approach presented in [19]. It is extended in [20] to include Virtual-CSMA, a technique that helps to estimate the conditional collision probability. This estimation is used to compute the number of contending stations (n) which, in turn, is used to obtain the minimum contention window as

$$CW_{min} = [n \cdot RAND(7, 8)]. \tag{1}$$

The contributions of this paper are as follows. First, it provides a general framework to study the efficiency of the backoff protocols. From this framework, the optimal transmission probability is derived and the optimal efficiency is compared to the efficiency obtained when using T-BEB. The comparison shows that there is room for improvement and that it is possible to design a backoff algorithm that performs better than T-BEB. It is observed that the fraction of slots containing a collision is independent of the number of contending stations when optimal transmission probability is used. Conversely, the fraction of slots containing collisions increases with the number of stations when T-BEB is used.

Inspired by this observation, a variant of the P-Persistent backoff algorithm is proposed. It is called Dynamic P-Persistent backoff (DPP) and dynamically adjusts the transmission probability to reach the optimal (constant) target fraction of collision slots. Thus the problem of estimating the number of contending stations is suppressed and substituted by an easier one which is estimating the fraction of collision slots. This estimation is performed using an exponential moving average estimator based on direct channel observations.

In addition to being simpler than the other optimization proposals mentioned in this section, DPP also presents advantageous implicit prioritization properties. The behaviour of DPP reduces the delay suffered by real-time traffic and interactive traffic in the presence of background traffic, when compared to the other backoff solutions. While previous research focused on either optimization or prioritization, DPP presents simultaneous improvements in both fields.

3 Binary Exponential Backoff and Performance analysis

This section introduces T-BEB which is part of the popular suite of protocols IEEE 802.11. This protocol is an example of CSMA algorithm in which the stations transmit without any previous knowledge about other stations intentions to transmit. The second part of this section assesses the performance of T-BEB, and finds the theoretical efficiency upper bound for this sort of algorithms.

3.1 Binary Exponential Backoff

The MAC mechanism used in IEEE 802.11 networks is called Distributed Coordination Function (DCF). Although the standard considers also a centralized alternative (the Point Coordination Function), it has been sparsely implemented.

In T-BEB, when a station that has its MAC queue empty receives a packet from the upper layer, it is allowed to transmit the packet after sensing the channel empty [1]. Otherwise, when the MAC queue is not empty or a packet arrives to the Head-Of-Line (HOL) of the MAC queue after the previous packet is successfully transmitted, the station has to backoff.

The backoff consists on randomly selecting a value from a Contention Window (CW) and waiting for that number of slots before transmitting. For the first transmission attempt the minimum congestion window is used (CW_{min}). If there is a collision, the congestion window doubles ($CW = 2 \cdot CW_{min}$) and the station randomly chooses a new number and waits for that number of slots before re-attempting transmission. The CW doubles after each collision until it reaches a maximum value CW_{max}. After a successful transmission the value of CW is reset to its minimum. IEEE 802.11b takes the values 32 and 1024 for its minimum and maximum contention windows, respectively.

With the IEEE 802.11e [4] standard amendment for Quality of Service support, the values of CW_{min} and CW_{max} can vary. However, the essence of the T-BEB remains the same.

---

[1] The channel has to be sensed for a DIFS (Distributed-coordination-function Inter Frame Space).

Four our analysis we will consider traffic sources that are saturated, *i.e.* each active station has always a packet ready to transmit. Intuitively, if there is only one active station in the network, it is expected to transmit one slot in every 16 slots.

It is apparent that an efficiency problem exists, since only one of every 16 slots is used while the rest remain empty. Nevertheless the problem is not as acute as it may seem at a first glance, because an empty slot is much shorter than a busy slot. Actually, the duration of an empty slot is $20\mu s$ in IEEE 802.11b while the duration of a successful slot is in the order of ms. The exact value of the latter depends on the length of the data contained in the packet.

As the number of stations increases, there are chances that two or more stations transmit on the same slot and that the transmissions are lost due to collision. A collision slot is as long as the longest of the packets involved in the collision. Therefore it is critical to reduce the number of collisions.

T-BEB reacts to collisions by doubling the contention window, thus diminishing the transmission rate of the stations. This reaction reduces the load on the network and should decrease the collision probability. Note, however, that it is necessary that there is one collision for the algorithm to realize that the network is highly loaded. Since the value of CW is reset to CW_{min} after a successful transmission, the station has to learn about the network congestion conditions for every packet, and every time there has to be a collision for the station to adjust its CW value. This is a relatively high price to pay for adjusting the CW to its optimal value.

It is shown in [6] that small contention windows are desirable when the number of contending stations is low, to reduce the number of empty unused slots. Conversely, for a large number of stations, larger contention windows offer better performance because reduce the collision probability. The framework provided by IEEE 802.11e can be used to dynamically tune the values of CW_{min} and CW_{max} to adapt to the number of contending stations. However, as explained in the previous section, this strategy requires previous estimation of the number of active stations n [16].

This qualitative analysis of T-BEB can help to understand the trade-off in choosing the right CW. A quantitative analysis of the algorithm can be obtained using Markov Chains and the assumption that, regardless of the number of retransmissions, a packet collides with constant probability [7]. Using that model, it is possible to compute the probability that a given station attempts transmission in a given slot (τ). This probability can then be used to obtain the probability of an empty, successful and collision slot. With these values, the overall performance of T-BEB can be evaluated and compared to other mechanisms.

The backoff process pursues the random distribution of the transmission attempts among the slots. An important goal is to maximize the number of successful transmissions while minimizing the collision probability. It is also important to keep the number of empty slots relatively low. However, an empty slot is much more desirable than a collision since the duration of the empty slots is orders of magnitude lower than the duration of a collision.

3.2 Efficiency of CSMA Algorithms

In CSMA algorithms, the stations autonomously decide whether to transmit or not. The probability that a station transmits (τ) is the key parameter to compute the probability of empty (P_e), successful (P_s) or collision [2] (P_c) slot. For a given number of contending stations n:

$$P_e = (1-\tau)^n, \tag{2}$$
$$P_s = n\tau(1-\tau)^{n-1}, \tag{3}$$
$$P_c = 1 - P_e - P_s. \tag{4}$$

The probability that a station transmits τ can be derived from [7] and is:

$$\tau = \frac{2(1-2p_{cc})}{(1-2p_{cc})(CW_{min}-1) + p_{cc}CW_{min}(1-(2p_{cc})^m)},$$
$$p_{cc} = 1-(1-\tau)^{n-1}. \tag{5}$$

---

[2] The notation P_c is used in this paper to denote the probability that a slot is busy with collision. This is different to the conditional collision probability (p or p_c in many papers) which is the probability that a collision occurs conditioned to the event that a tagged station attempts transmission.

where p_{cc} is the conditional collision probability, *i.e.* the probability that a collision occurs given that one tagged station is attempting transmission. CW_{min} is the minimum congestion window and m the maximum backoff stage.

We define the efficiency as the fraction of time that the channel is used for successful transmissions. It is understood that the time that the channel remains empty or busy with collisions is wasted.

$$\phi = \frac{T_s P_s}{T_e P_e + T_s P_s + T_c P_c}. \tag{6}$$

In (6) we can observe that the duration of empty, successful and collision slots also affect the observed efficiency. While T_e is constant and defined in the standard, T_s and T_c are a function of the length of the frames. Under the assumption of fixed packet length, the duration of successful and collision slots are similar. Thus the duration of a collision can be approximated to the duration of a successful slot $T_c \approx T_s$. Using the approximation and substituting (2) - (4) into (6) we obtain:

$$\phi = \frac{n\tau(1-\tau)^{n-1}}{1 - \frac{T_s - T_e}{T_s}(1-\tau)^n} \tag{7}$$

From (7) it can be observed that the efficiency increases when large frames are used. Given a number of contending stations n and a successful slot duration T_s, the optimal transmission probability τ that maximizes efficiency satisfies:

$$\begin{aligned} &\frac{d\phi}{d\tau} = \\ &\frac{(1-\tau)^{n-1} + (n-1)\tau(1-\tau)^{n-2}}{1 - \frac{T_s - T_e}{T_s}(1-\tau)^n} - \\ &\frac{\frac{T_s - T_e}{T_s} n\tau(1-\tau)^{2(n-1)}}{(1 - \frac{T_s - T_e}{T_s}(1-\tau)^n)^2} = 0 \end{aligned} \tag{8}$$

In Fig. 1, the efficiency using optimal values of τ is plotted. Fig. 2 shows that when using an optimal transmission probability, the collision probability is (almost) independent of the number of active

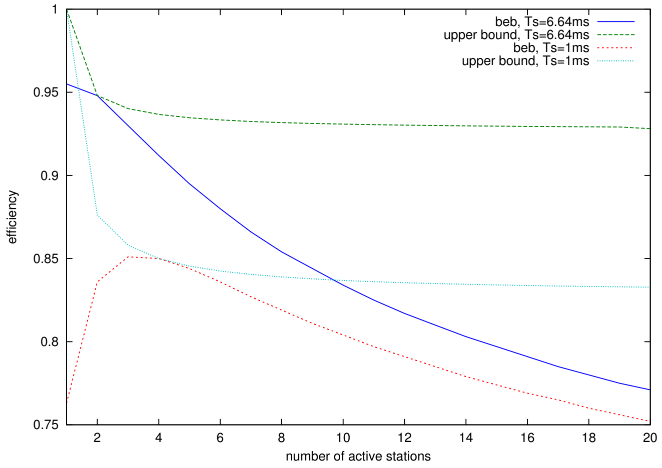


Fig. 1. This figure compares the performance of BEB to the theoretical maximum for different values of successful slot duration T_s.

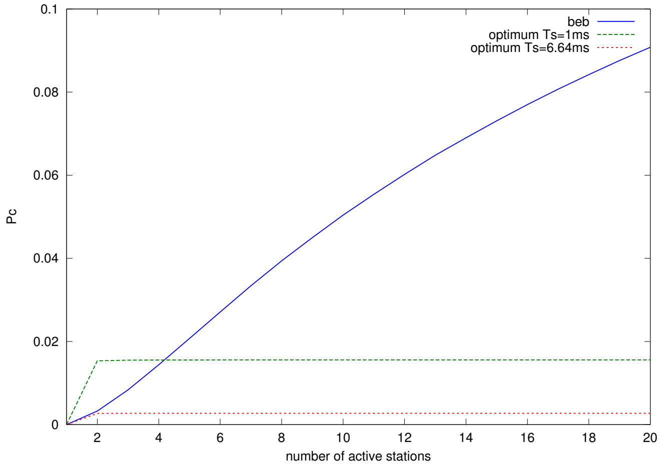


Fig. 2. This figure compares the collision probability obtained when using BEB with one that would be obtained when using optimal transmission probability.

stations. This interesting property can be used to derive a near-optimal contention algorithm based on a variant of the P-Persistent mechanism explained in the introduction.

4 DP-Persistent CSMA

The observation that the collision probability is almost constant when the transmission probability τ is optimal can be exploited to increase the efficiency to values closer to the theoretical optimum.

The proposal consists on observing the channel to estimate the collision probability. Then the stations adapt the transmission probability τ to adjust the collision probability to the target (optimal) collision probability.

Algorithm 1 explains how the transmission probability is distributedly adjusted to attain the optimal collision probability. $\hat{P}_c$ is the estimated collision probability and is computed as an Exponential Moving Average (EMA) based on the observation of the channel. Then, the estimated collision probability ($\hat{P}_c$) is compared to the target collision probability (P_c^T).

If $\hat{P}_c > P_c^T$, the transmission probability (τ) is decremented. Otherwise, the transmission probability is increased. We adopt an Additive Increase Multiplicative Decrease (AIMD) approach for the tuning of τ. The reason for this choice is that it provides long-term fairness among competing flows, even when they begin with different values of τ.

It can be observed that Algorithm 1 includes a number of parameters (P_c^T, τ_0, $\hat{P}_{c0}$, ϵ, α, μ, τ_{max}). Each of this parameters conditions the overall performance of the backoff mechanism, and the selection of these parameters also involve some kind of trade-off. In the following, we summarize and discuss the values of these parameters.

P_c^T is the target collision probability, *i.e.* the collision probability that delivers optimal performance. Unfortunately, P_c^T is a function of the duration of a successful transmission (T_s). Assuming a data rate of 11Mbps, T_s takes values from 0.6 ms (when the frame carries no data) to 9.9 ms (when the payload is maximum, 2304 bytes). The actual packet size distribution in WLAN [15] is trimodal, being most of the packets smaller than 100 bytes or larger than 1470 bytes, with a lower fraction around 600 bytes. Since the duration of a collision is approximately equal to the duration of the longest packet involved in the transmission, the conservative decision of assuming a payload size of 1500 bytes is adopted.

```
/* $\tau$ is the transmission probability                 */
/* $\hat{P}_c$ is the estimated collision probability     */
/* $P_c^T$ is the target collision probability            */
/*                                                         */
/* $\tau$ and $\hat{P}_c$ are initialized                 */
$\tau \leftarrow \tau_0$
$\hat{P}_c \leftarrow \hat{P}_{c0}$
while There are packets ready to transmit do
    Sense the channel /* EMA is used to update $\hat{P}_c$ */
    if Collision then
        $\hat{P}_c \leftarrow \epsilon + (1-\epsilon) \cdot \hat{P}_c$
    else
        $\hat{P}_c \leftarrow (1-\epsilon) \cdot \hat{P}_c$
    end
    /* $\tau$ is updated using AIMD                        */
    if $\hat{P}_c < P_c^T$ then
        $\tau \leftarrow MIN\left[\tau + \alpha(P_c^T - \hat{P}_c), \tau_{max}\right]$
    else
        $\tau \leftarrow \frac{\tau}{1+\mu(\hat{P}_c - P_c^T)}$
    end
end
```

Algorithm 1: Transmission probability adaptation

If the payload size is 1500 bytes, the duration of a slot containing a successful transmission is 6.64ms and the optimal collision probability (as described in Sec. 3) is 0.0027. Therefore, the target collision probability P_c^T is set to 0.0027.

Since the minimum contention window in IEEE 802.11b is 32 (the stations would transmit every 16 slots on average if there were no collisions), a value of 1/16 have been chosen as initial transmission probability τ_0. The initial estimated collision probability $\hat{P}_{c0}$ is set to the target collision probability P_c^T. As the station senses the channel, it will obtain a finer value of $\hat{P}_c$ that can be used to adapt τ and take it closer to the optimal value.

The EMA estimator uses the parameter ϵ. It must take values between 0 and 1. A high value of ϵ gives more weight to what has happened in recent slots and makes the estimation to react faster to new conditions (*i.e.* addition or suppression of a contending station or changes in transmission probability τ). However, since collisions happen seldom, a high value of ϵ can easily lead to excessive oscil-

Table 1. Parameter Values

P_c^T	τ_0	$\hat{P}_{c0}$	ϵ	α	μ	τ_{max}
0.0027	1/16	0.0027	0.001	0.01	0.05	1/8

lations that would set τ far from its optimal value. Thus a value of 0.001was chosen for ϵ.

The parameters α and μ represent the Additive Increase and Multiplicative Decrease of τ respectively. As happens with ϵ, a higher value offers prompt reactions but also increases the risk of larger oscillations that penalize performance. Their values $\alpha = 0.01$ and $\mu = 0.05$ were choosen empirically, after observing their impact in simulation results.

Finally, there is a need to limit the maximum transmit probability τ_{max}. The purpose of τ_{max} is to prevent τ to grow to 1 in the special case in which there is only one active station. A transmission probability of 1 would boost the efficiency to 100% but would hamper the entry of a new contender. A value $\tau_{max} = 1/8$ is a good compromise to guarantee high efficiency when there is only one station while leaving 7 out of 8 slots free for the new contender to successfully transmit.

Table 1 summarizes the parameters and its values.

5 Simulation Results

Using the algorithm and parameters described in the previous section, simulations[3] can be used to observe the results obtained using the proposed alternative backoff algorithm. First we present a toy scenario in which the number of stations is increased from two to eleven. The increments happen every 4000 slots. The case with only one station is omitted in the figures because it presents results so different from the other cases that obfuscate the resultant plots. When there is only one station the collision probability is equal to zero, and the transmission probability tends to τ_{max}.

The following plots show the actual collision probability compared to the target collision probability (Fig. 3), the actual trans-

---

[3] The simulations and the numerical computations were performed using octave. All the scripts are available upon request to the corresponding author.

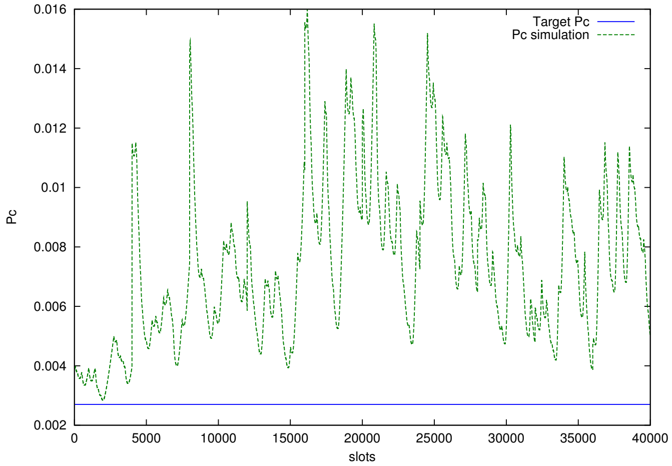


Fig. 3. The actual P_c is compared to the target P_c^T. The number of active stations is increased from 2 to 11. A station is added every 4000 slots

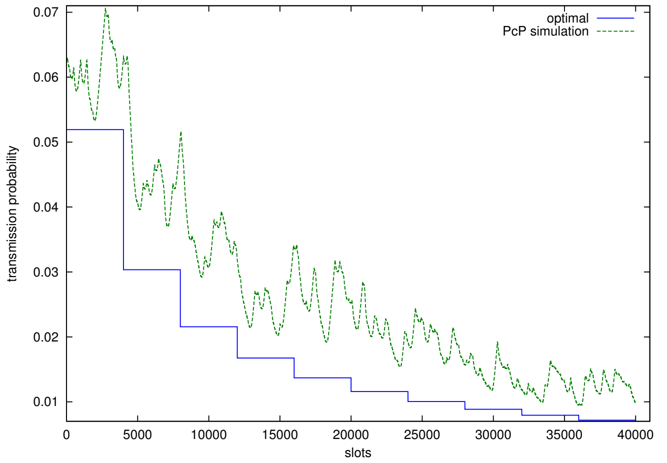


Fig. 4. The actual transmission probability τ is compared to the optimal transmission probability τ^{opt}. The number of active stations is increased from 2 to 11. A station is added every 4000 slots

mission probability compared to the optimal transmission probability (Fig. 4) and the actual efficiency compared to the theoretical maximum (Fig. 5).

In Fig. 3 it can be observed that that the backoff algorithm tries to keep the collision probability close to the (constant) target col-

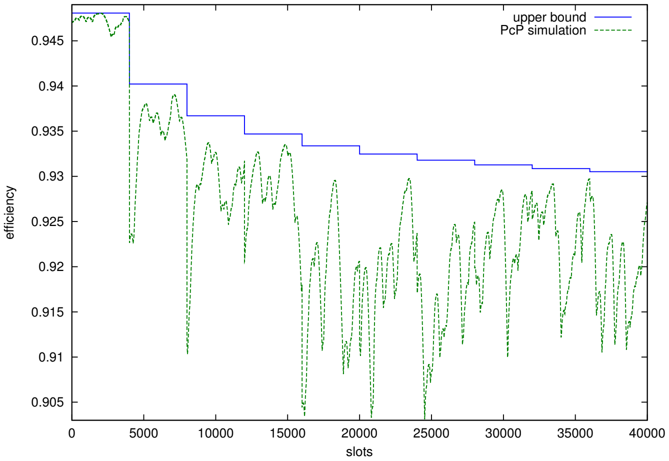


Fig. 5. The actual efficiency ϕ is compared to the optimal efficiency ϕ^{opt}. The number of active stations is increased from 2 to 11. A station is added every 4000 slots

lision probability for any number of stations. When the number of stations increases (at slot 4000, 8000, etc.) a spike appears in the actual collision probability. It takes some time for the stations to detect the increased number of collisions and reduce the transmission probability and thus adjust the collision probability to a value closer to to the desired one. A careful observer would notice that the actual collision probability (P_c) is larger than the target collision probability (P_c^T). There are two causes for this misadjustment: the estimator fails to capture the instant collision probability and the τ parameter tuning is a slow iterative process. Nevertheless, P_c is close enough to P_c^T to offer excellent efficiency.

Fig. 4 shows the transmission probability observed in the simulations compared to the optimum transmission probability. Again, it can be observed that the stations require some time to adapt to a scenario change. However, in the long term, the actual transmission probability approximately follows the optimal transmission probability.

Finally, in Fig. 5, we can observe the benefits of the proposed backoff scheme. The obtained efficiency closely sticks to the optimal efficiency for any number of stations.

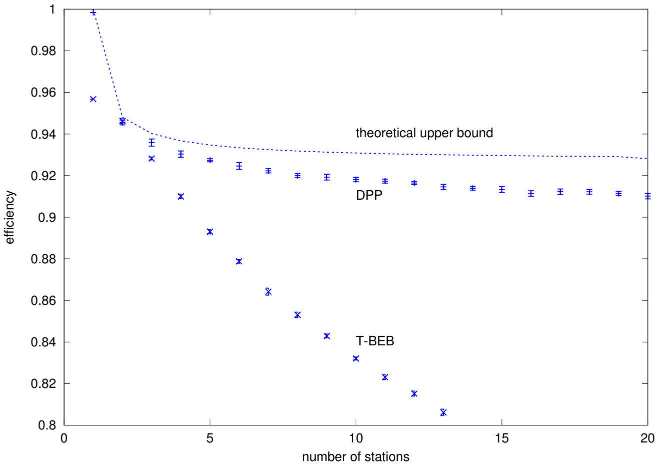


Fig. 6. Theoretical maximum (dashed line) compared to simulations results of DP-Persistent CSMA and T-BEB. The 95% confidence intervals are plotted.

In the previous example and figures, the dynamic behaviour of the algorithm has been explained by observing a simulation in which the number of active stations is variable and the control loop implemented in the backoff algorithm actuates to adjust the probability of a collision slot to a fixed (optimal) value.

In order to asses with greater accuracy the performance delivered by DPP, simulations for a fixed number of stations have been performed. Each simulation comprises 80,000 slots and has been repeated 10 times with different random seeds. Fig. 6 shows the results and compares them to the theoretical maximum computed in Sec. 3 and depicted in Fig. 1. It can be observed that DPP performs close to the theoretical maximum in steady-state operation.

6 Implicit Prioritization

Current data networks carry heterogeneous traffic. Internet traffic can be classified in background, interactive and real-time traffic. Background traffic transfers large amounts of data with no stringent delay constraints. This traffic is carried by long-lived TCP flows that are permanently active. A good example of background traffic is peer-to-peer file sharing. This data is transferred without the active participation of any human being.

Interactive traffic is originated and consumed by users. It consists in small data burst such as a request for a webpage and the consequent response from the server. This are short-lived TCP interactions in which a relatively small amount of data needs to be transmitted in a reasonable amount of time. Reasonable is a lax definition and depends on the expectations from the users, and is probably in the order of one second. Users would prefer a shorter reaction time; therefore, for this kind of traffic, delay does matter.

The last kind of traffic is real-time traffic. Very small quantities of data are sent periodically to maintain a voice or video flow. For real-time flows delay is critical, and those packets that suffer excessive delay are useless at reception and are discarded.

It is a desired property of a network that allows the harmonious coexistence of different kinds of traffic. Ideally, real-time traffic would traverse the networks with the highest priority to reach the destination in tens of milliseconds. Interactive traffic comes second in the priority row, since there is a user waiting for an answer and that waiting time should be minimized. When neither real-time nor interactive traffic is transmitted, the network can be used to transmit background traffic.

From the previous argumentation it can be concluded that the priority of a data transfer maintains an inverse relationship with its duration. In the following, it will be explained that this is exactly the treatment that stations deserve under the DPP backoff mechanism.

It has to be noticed that every station enters the playground with a initial transmission probability $\tau_0 = 1/16$. In its commitment to lower the number of collisions to achieve the maximum efficiency, DPP lowers the transmission probability. The result is a large fraction of empty slots (about 90%) and transmission probabilities lower than τ_0 for a number of stations equal or larger than 3. With this scenario, a station becoming active after an inactivity period enjoys priority for a limited initial period of time.

Due to the slow nature of the EMA average and the τ adjustment mechanism explained in Sec. 4, it takes some time for the newcomer to lower its own transmission probability from the initial value τ_0 to the optimal value τ_{opt}. This time can be used to transmit with higher priority than the other stations that have been active for a long time. A station transmitting a burst of data will observe that the first

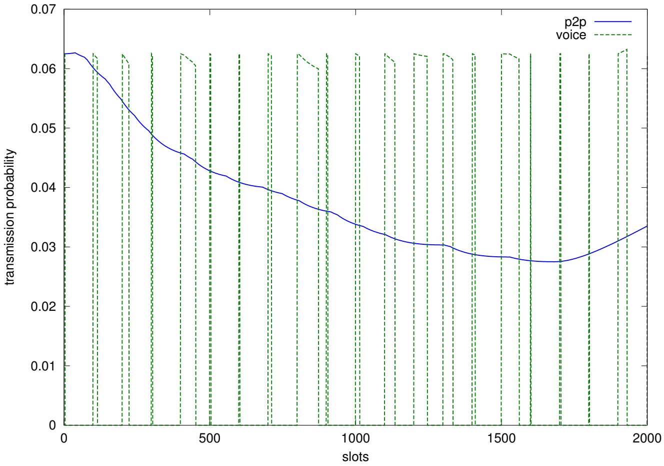


Fig. 7. A single station generating voice traffic competes against five peer-to-peer stations for the channel. The voice station periodically enters the contention with transmission probability τ_0 and leaves the contention once the voice packet has been transmitted.

packets of the burst enjoy priority, but that priority vanishes as times passes and its own transmission probability is slowly decreased. The result is that shorter burst will be transmitted with higher priority than longer bursts.

The behaviour of DPP can be summarized as assigning priority to stations that become active after an inactivity period. This priority fades away as the station continues active for a longer period. Fig. 7 shows a single station generating voice traffic competing against five peer-to-peer saturated stations. The voice station has a new packet to send one in every 100 slots, it competes for the channel until it has sent that packet and then leaves the contention. When the voice station rejoins the contention to send a new packet, it uses the initial transmission probability τ_0. The peer-to-peer stations are constantly contending for the channel and do not have the chance to reset their transmission probability to τ_0.

Even though DPP exhibits convenient prioritizing properties, it does not completely solve priority issues. There are two aspects in which DPP falls short of solving the problem. The first one involves uplink/downlink unfairness in infrastructure scenarios. All the stations transmit to the access point and the access point transmits to all stations. The latter easily becomes the bottleneck of the network and requires higher priority.

DPP does not solve the issue of stations transmitting heterogeneous traffic. A station that sends both real-time and background traffic would be continuously active and would not benefit from the early priority commented in this section.

Nevertheless, DPP offers advantageous implicit prioritizing properties when compared with IEEE 802.11.

7 Conclusion

This paper studies the performance of backoff mechanisms in terms of efficiency, *i.e.* the fraction of time that is devoted to successful transmissions compared to the time wasted in empty slots and collisions. Optimal efficiency can be obtained by adjusting the transmission probability τ of the stations. It is shown that the optimal transmission probability τ_{opt} depends on the packet length and the number of active stations. It is also observed that the fraction of slots containing a collision P_c is almost constant when optimal transmission probability is used.

The efficiency of T-BEB is compared to the optimum to show that there is room for improvement. Then an algorithm called DPP is proposed. This algorithm dynamically adjusts the transmission probability τ to achieve optimal collision probabiliy P_c which is known and constant. As opposed to backoff mechanisms proposed in previous art, DPP does not need to estimate the number of contending stations. Additionally, DPP outperforms BEB and achieves near-optimal efficiency.

DPP is a completely distributed backoff scheme in which the stations monitor the channel to estimate the collision probability and dynamically adjust their transmission probability in the quest for optimal efficiency. Both the estimation and the parameter adjustment takes some time. This results in stations awaking from an inactivity period having higher priority than those that have been active for a longer period of time. This proves beneficial since reduces the delay of real-time and interactive applications while maintains near-optimal throughput for background traffic.

Acknowledgment

We would like to acknowledge the anonymous reviewers for their insightful comments.

References

1. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999 Edition (Revised 2003).
2. Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD), Access Method and Physical Layer Specifications, 2000 Edition.
3. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications for low-rate wireless personal area networks, 2003 Edition (Revised 2006).
4. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications; Amendment: Medium Access Control(MAC) Quality of Service Enhancements. (IEEE Std 802.11e), 2005.
5. N. Abramson. The ALOHA System–Another Alternative for Computer Communications. *Cluster Computing*, 5:187–201, 1970.
6. H. Anouar and C. Bonnet. Optimal Constant-Window Backoff Scheme for IEEE 802.11 DCF in Single-Hop Wireless Networks Under Finite Load Conditions. *Wireless Personal Communications*, 43(4):1583–1602, Dec. 2007.
7. G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination-function. 18(3):535–547, 2000.
8. G. Bianchi and I. Tinnirello. Kalman filter estimation of the number of competing terminals in an IEEE 802.11 network. *IEEE INFOCOM*, 2, 2003.
9. F. Cali, M. Conti, E. Gregori, and P. Aleph. Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit. *IEEE/ACM Transactions on Networking*, 8(6):785–799, 2000.
10. J. Capetanakis. Tree algorithms for packet broadcast channels. *IEEE Transactions on Information Theory*, 25(5):505–515, 1979.
11. J. Goodman, A.G. Greenberg, N. Madras, and P. March. Stability of binary exponential backoff. *Journal of the ACM (JACM)*, 35(3):579–602, 1988.
12. J. Jubin and JD Tornow. The DARPA packet radio network protocols. *Proceedings of the IEEE*, 75(1):21–32, 1987.
13. B.J. Kwak, N.O. Song, and L.E. Miller. Analysis of the Stability and Performance of Exponential Backoff. *In IEEE WCNC*, pages 1754–1761, 2003.
14. A. Lopez-Toledo, T. Vercauteren, and X. Wang. Adaptive Optimization of IEEE 802.11 DCF Based on Bayesian Estimation of the Number of Competing Terminals. 5(9):1283, 2006.
15. C. Na, J.K. Chen, and T.S. Rappaport. Measured traffic statistics and throughput of IEEE 802.11 b public WLAN hotspots with three different applications. *IEEE Trans. Wireless Commun*, 5(11):3296–3305, 2006.
16. T. Vercauteren, A. Lopez-Toledo, and X. Wang. Batch and Sequential Bayesian Estimators of the Number of Active Terminals in an IEEE 802.11 Network. 55(2):437–450, 2007.
17. X. Wang, Y. Yu, and G.B. Giannakis. Design and analysis of cross-layer tree algorithms for wireless random access. *IEEE Transactions on Wireless Communications*, 7(3):909–919, March 2008.

18. S.L. Wu, Y.C. Tseng, C.Y. Lin, and J.P. Sheu. A Multi-channel MAC Protocol with Power Control for Multi-hop Mobile Ad Hoc Networks. *The Computer Journal*, 45(1):101–110, 2002.
19. L. Zhao, J. Zhang, K. Yang, and H. Zhang. Using Incompletely Cooperative Game Theory in Mobile Ad Hoc Networks. *In IEEE ICC'07*, pages 3401–3406, 2007.
20. L. Zhao, J. Zhang, and H. Zhang. Using Incompletely Cooperative Game Theory in Wireless Mesh Networks. *IEEE Network*, 22(1):39–44, 2008.

Learning-BEB: Avoiding Collisions in WLANs

J. Barcelo, B. Bellalta, C. Cano, M. Oliver

Universitat Pompeu Fabra

Abstract. Random access protocols have been the mechanism of choice for most WLANs, thanks to their simplicity and distributed nature. Nevertheless, these advantages come at the price of sub-optimal channel utilization because of empty slots and collisions. In previous random access protocols, the stations transmit on the channel without any clue of other stations' intentions to transmit. In this article we provide a framework to study the efficiency of channel access protocols. This framework is used to analyze the efficiency of the Binary Exponential Backoff mechanism and the maximum achievable efficiency that can be obtained from any completely random access protocol. Then we propose Learning-BEB (L-BEB).

L-BEB is exactly the same as legacy BEB, with one exception: L-BEB chooses a deterministic backoff value after a successful transmission. We call this value the virtual frame size (V). This subtle modification significantly reduces the number of collisions. It can be observed that, as the system runs, the number of collisions is progressively reduced. Thus we conclude that the system learns. Further, if the number of contending stations is equal or lower than V and all stations consecutively successfully transmit, collisions disappear. This collision-free operation is maintained until a new station is activated and joins the contention.

L-BEB pushes the system performance beyond the upper bound inherent to completely-random access mechanisms. Moreover, L-BEB does not introduce any additional complexity to the algorithms currently in use in WLANs. All the claims in the paper are supported by extensive simulation results.

1 Introduction

The radio channel is a broadcast medium and nodes which are in each other interference range should take turns in transmitting. Simultaneous transmissions are called collisions. As a result of a collision, the messages being transmitted might be lost.

The Medium Access Control (MAC) is the function that arbitrates the access to the channel. In wireless networks, the MAC protocols play a key role in maximizing the channel utilization.

Carrier Sense Multiple Access (CSMA) and Time Division Multiple Access (TDMA) are two well known medium access mechanisms for shared medium communication systems. The former relies on stations sensing the medium before transmitting whereas the latter reserves time slices for each active station.

CSMA is simpler to operate. No tight time synchronization is required and the stations simply transmit when they have data ready to be transmitted. Thus they benefit from statistical multiplexing, supporting a larger number of bursty traffic sources. The problems arrive when two stations decide to transmit simultaneously. It may happen that the data from both transmissions is lost. Then a collision resolution mechanism must be activated. To avoid collisions, the stations distributedly execute backoff algorithms that randomly delay the transmission. Because of the random nature of the selection of transmission times, only a fraction of the time is devoted to successful transmissions, while the rest is wasted (either in the form of empty channel, or busy with collisions).

TDMA, on the other side, requires tight time synchronization among the participating stations. Additionally, a prior set-up is required to assign a time-slice (or slot) to the active stations. This set-up causes extra signaling overhead and often requires the presence of a central decision point. After the time slices are assigned, those slices are reserved for a given station. If that station has no data ready to transmit, the channel time is wasted. Conversely, if the station has a large amount of data to transmit, it can only transmit the fraction that fits in the reserved time slice. The rest of the data is buffered for later transmission. The great advantage of TDMA is that it avoids collisions and may achieve high channel efficiency.

Both CSMA and TDMA have advantages and disadvantages. The combination of the advantages of both mechanisms has been a long sought after goal. Sec. 2 describes related work in the subject. The prior art is characterized by its complexity, which have prevented widespread implementation of the ideas.

Sec. 3 briefly describes the CSMA mechanisms in IEEE 802.11 [1] which strongly rely on the Binary Exponential Backoff (BEB). This section also provides a framework to compute the performance of a backoff algorithm. It is demonstrated that an upper efficiency limit exists, under the assumption that the stations are unaware of other stations' intentions to transmit.

In Sec. 4 we show that, by a simple modification of BEB, the channel access medium is converted from pure CSMA to a hybrid CSMA-TDMA. We call the new mechanism Learning-BEB (L-BEB) because it progressively learns from both successful and unsuccessful transmission attempts, in order to migrate to TDMA-like operation. L-BEB is even simpler than legacy BEB and does not require any additional signaling. In the worst case, the performance delivered by L-BEB is the same as the performance that is currently obtained from legacy BEB.

The argumentation of Sec. 4 is supported by the simulation results in Sec. 5. Specifically, it is shown that L-BEB outperforms legacy BEB for any number of stations by reducing collisions and increasing the number of successful slots. Further, cumulative collision plots are used to show how the system learns from previous transmission attempts and the number of collisions is reduced as the simulation progresses.

Finally, the paper is concluded in Sec. 6.

2 Related Work

The Aloha [3] protocol laid the foundations for many random access protocols to come. In random access protocols, the nodes optimistically send their packets. In Aloha, a node with data ready to send, sends it immediately. The nodes involved in a collision wait a random period of time before attempting retransmission. In CSMA [11], the nodes are smarter and listen before talking, thus reducing the chances of collisions.

Reservation-Aloha (R-Aloha), presented in [8] and further analyzed in [14], already proposed a combination of random access and TDMA. The time is divided in slots which are grouped in frames. The duration of the slots is fixed and the duration of frames is chosen to be longer than the propagation delay of the broadcast channel. When station X successfully transmits in slot Y of a frame, it implicitly reserves slot Y for the next frame. The reservation can be released either explicitly, using a special flag in the last packet transmission, or implicitly by not sending a packet in the reserved slot.

R-Aloha presents several disadvantages when compared to protocols that are currently in use, such as IEEE 802.11. First, the fixed length of slots implies that a high fraction of the channel time is wasted due to empty slots (In IEEE 802.11, the empty slots are orders of magnitude shorter than busy slots). Second, R-Aloha requires time synchronization among terminals. And third, the number of slots in a frame effectively limits the maximum number of active terminals. As the frame becomes full, new entrants do not have any chance to transmit. If the frame size is variable, additional signaling is required to inform all the stations about the current frame size.

Packet Reservation Multiple Access (PRMA) [10] and Centralized-PRMA [6] further extended the idea of R-Aloha to support heterogeneous (real-time and bursty) traffic. However, the improvements came at the price of higher complexity and signaling requirements.

A CSMA-TDMA hybrid MAC protocol was also explored in [9]. It is called Probabilistic TDMA (PTDMA). As in TDMA, the time is divided in time slices called slots which are grouped in frames. A station can own a slot in the frame. If this is the case, a station can transmit in that slot with probability a. Otherwise, if the station does not own the slot, it can also transmit with probability b. a and b satisfy the following equation:

$$a + (n - 1)b = 1 \tag{1}$$

where n is the number of senders. For low values of n, the behaviour of PTDMA is closer to CSMA. As the number of stations increase, the probability that a station transmits in a non-owned slot is reduced and the behaviour of PTDMA is biased towards TDMA.

In the context of wireless sensor networks, there have been recent research efforts in the field of CSMA-TDMA hybrids. Z-MAC [13] aims to combine the advantages of CSMA and TDMA in a single protocol. From CSMA, it takes high channel utilization and low latency under low contention; as TDMA, it offers high channel utilization and a limited number of collisions under high contention. Differently from our proposal, it specifically addresses multi-hop networks. The downside of Z-MAC is its increased complexity, which include neighbor discovery, slots assignment, local frame exchange and global time synchronization.

As opposed to the related work described in this section, the protocol proposed in this paper is based on an extremely simple modification to the protocol currently in use. This modification can be even considered a simplification. Another key differentiation aspect is that our proposal supports different slot durations (as IEEE 802.11 does), allowing the empty frames to be shorter than transmission frames. This option dramatically boosts the performance by reducing the time that the channel remains empty. All the previous work cited above assumes fixed slot duration.

A separate line of research consists on squeezing the maximum efficiency out of BEB by tuning its operation parameters, without making any CSMA-TDMA hybridization attempts. This avenue of research has its origins in the finding that the optimal transmission probability in BEB is a function of the packet length (l) and the number of competing stations (n) [15].

It is natural to attempt to estimate the number of contending stations to optimize the performance of BEB. The fast and accurate estimation of n is not a trivial task and advanced filtering techniques are required. An extended Kalman filter is used in [7] while [12] further improves the estimation by means of a bayesian approach.

Nevertheless, even if perfect estimation of the number of contending stations is achieved, the obtained efficiency never surpasses the upper bound for BEB, which is further detailed in the next section.

Our proposal easily breaks the upper bound for BEB and neither requires the estimation of n nor the dynamic adjustment of the operation parameters.

3 Binary Exponential Backoff and Performance analysis

This section introduces Binary Exponential Backoff (BEB) which is part of the popular suite of protocols IEEE 802.11. This protocol is an example of a CSMA algorithm in which the stations transmit without any previous knowledge about other stations' intentions to transmit. The second part of this section assesses the performance of BEB, and finds the theoretical efficiency upper bound for this sort of algorithms.

Throughout the analysis, a number of usual assumptions are adopted. These include the supposition that all the stations are in the transmission range of one another, *i.e.* there is no hidden terminal effect [16]. The time is divided in slots, and the stations are synchronized to those slots. Transmission attempts can occur only at the beginning of a slot. Additionally, an ideal channel is assumed and frame losses are caused only by collisions. To simplify the analysis, all the stations transmit using the same data rate. The frame length is also the same for all stations.

3.1 Binary Exponential Backoff

The Medium Access Control (MAC) mechanism used in IEEE 802.11 networks is called Distributed Coordination Function (DCF). Although the standard considers also a centralized alternative (the Point Coordination Function) it has been sparsely implemented.

DCF uses a truncated Binary Exponential Backoff strategy. When a station that has its MAC queue empty receives a packet from the upper layer, it is allowed to transmit the packet after sensing the channel empty [1]. Otherwise, when the MAC queue is not empty and a packet arrives to the head-of-line of the MAC queue after the previous packet is successfully transmitted (or discarded), the station has to backoff.

The backoff consists on drawing a number from a Contention Window $[0, CW)$ and waiting for that number of slots before transmitting. For the first transmission attempt the minimum contention window is used (CW_{min}). If there is a collision, the contention window doubles ($CW = 2 \cdot CW_{min}$) and the station randomly chooses a new number and waits for that number of slots before re-attempting transmission. The CW doubles after each collision until it reaches a maximum value CW_{max}. After a successful transmission, the value of CW is reset to its minimum. IEEE 802.11 takes the values 32 and 1024 for its minimum and maximum contention windows, respectively.

---

[1] The channel has to be sensed idle for a DIFS (DCF Inter Frame Space) period of time.

With the IEEE 802.11e [2] standard amendment for Quality of Service support, the values of CW_{min} and CW_{max} can vary. However, the essence of the BEB remains the same.

For our analysis we will consider traffic sources that are saturated, *i.e.* each active station has always a packet ready to transmit. Intuitively, if there is only one active station in the network, this station is expected to transmit one slot in every 16 slots. The reason is that the actual number of empty slots between transmissions will uniformly vary from 0 to 31.

It is apparent that an efficiency problem exists, since only one of every 16 slots is used while the rest remain empty. Nevertheless the problem is not as acute at it may seem at a first glance, because an empty slot is much shorter than a busy slot. Actually, the duration of an empty slot is $20\mu s$ in IEEE 802.11b while the duration of a successful slot is in the order of ms. The exact value of the latter depends on the length of the data contained in the packet.

As the number of stations increases, there are higher chances that two or more stations transmit on the same slot and that the packets are lost due to a collision. The length of a collision slot is equal to the longest of the transmissions involved in the collision. Therefore it is critical to reduce the number of collisions.

The BEB reacts to collisions by doubling the contention window, thus diminishing the transmission rate of the stations. This reaction reduces the load on the network and should decrease the collision probability. Note, however, that it is necessary that there is one collision for the algorithm to realize that the network is highly loaded. Since the value of CW is reset to CW_{min} after a successful transmission, the station has to learn about the network congestion conditions for every packet, and every time there has to be a collision for the station to adjust its CW value. This is a relatively high price to pay for adjusting the CW to its optimal value.

Studies in [4] show that small contention windows are desirable when the number of contending stations is low, since a small contention window reduces the number of empty unused slots. Conversely, for a large number of stations, larger contention windows offer better performance because they reduce the collision probability. The framework provided by IEEE 802.11e can be used to dynamically tune the values of CW_{min} and CW_{max} to adapt to the number of contending stations. However, as explained in the previous section, this strategy requires previous estimation of the number of active stations n [12].

The qualitative analysis of BEB presented above describes the trade-off incurred in choosing the right CW. A quantitative analysis of BEB can be obtained using Markov Chains and the assumption that, regardless of the number of retransmissions, a packet collides with constant probability [5]. Using that model, it is possible to compute the probability that a given station attempts transmission in a given slot (τ). This probability can then be used to obtain the probability of an empty, a successful and a collision slot. With these values, the overall performance of BEB can be evaluated and compared to other mechanisms.

The backoff process pursues the random distribution of the transmission attempts among the slots. An important goal is to maximize the number of successful transmissions while minimizing the collision probability. It is also important to keep the number of empty slots relatively low. However, an empty slot is much more desirable than a collision since the duration of the empty slots is orders of magnitude lower than the duration of a collision.

3.2 Efficiency of CSMA Algorithms

In CSMA algorithms the stations autonomously decide whether to transmit or not. The transmission probability (τ) is the key parameter that determines the probability of empty, successful or collision slot (P_e, P_s and P_c [2] respectively). For a given number of contending stations n:

$$P_e = (1-\tau)^n, \tag{2}$$

$$P_s = n\tau(1-\tau)^{n-1}, \tag{3}$$

$$P_c = 1 - P_e - P_s. \tag{4}$$

---

[2] The notation P_c is used in this paper to denote the probability that a slot is busy with collision. This is different to the conditional collision probability (p or p_c in many papers) which is the probability that a collision occurs conditioned to the event that a tagged station attempts transmission.

The transmission probability τ for BEB can be derived from [5] and is:

$$\tau = \frac{2(1-2p_{cc})}{(1-2p_{cc})(CW_{min}-1)+p_{cc}CW_{min}(1-(2p_{cc})^m)},$$
$$p_{cc} = 1-(1-\tau)^{n-1}. \tag{5}$$

p_{cc} is the conditional collision probability; the probability that a collision occurs given that one tagged station is attempting transmission. CW_{min} is the minimum contention window and m the maximum backoff stage:

$$m = log_2\left[\frac{CW_{max}}{CW_{min}}\right]. \tag{6}$$

We define the efficiency (ϕ) as the fraction of time that the channel is used for successful transmissions. It is understood that the time that the channel remains empty or busy with collisions is wasted. The efficiency is a function of the probabilities described in (2) - (4) and the duration of an empty, successful and collision slot (T_e, T_s and T_c respectively).

$$\phi = \frac{T_sP_s}{T_eP_e+T_sP_s+T_cP_c}. \tag{7}$$

In (7) we can observe that the duration of empty, successful and collision slots also affect the observed efficiency. While T_e is constant and defined in the standard, T_s and T_c are a function of the length of the frames. The duration of successful and collision slots are similar, thus the duration of a collision can be approximated to the duration of a successful slot $T_c \approx T_s$. Using the approximation and substituting (2) - (4) into (7) we obtain:

$$\phi = \frac{n\tau(1-\tau)^{n-1}}{1-\frac{T_s-T_e}{T_s}(1-\tau)^n} \tag{8}$$

From (8) it can be observed that the efficiency increases when using large frames. Given a number of contending stations n and a successful slot duration T_s, the optimal transmission probability τ that maximizes efficiency satisfies:

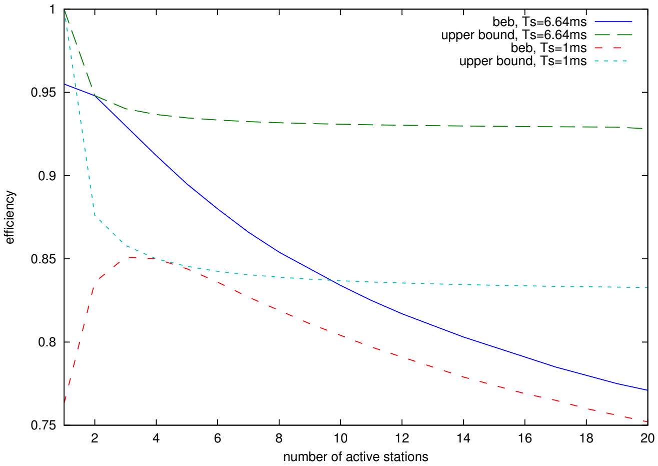


Fig. 1. This figure compares the efficiency of BEB to the efficiency that would be obtained if optimal transmission probability was used. The results are presented for different values of successful slot duration T_s.

$$\frac{d\phi}{d\tau} = \frac{(1-\tau)^{n-1} + (n-1)\tau(1-\tau)^{n-2}}{1 - \frac{T_s - T_e}{T_s}(1-\tau)^n} - \frac{\frac{T_s - T_e}{T_s} n\tau(1-\tau)^{2(n-1)}}{(1 - \frac{T_s - T_e}{T_s}(1-\tau)^n)^2} = 0 \qquad (9)$$

In Figure 1, the efficiency using optimal values of τ (derived from (9)) is plotted. This values are compared to the ones that are actually obtained when using the values of τ provided by legacy BEB (which are derived from (5)).

The curves in Figure 1 for optimal transmission probability represent an upper bound for BEB, and for those protocols that simply tune the parameters of BEB in response to the number of competing terminals. To surpass that upper bound, it is not sufficient to adjust the size of the CW_{min}. Conversely, it is required that the stations gain some kind of knowledge about the other stations' future intentions to transmit. This can be achieved by setting the stations' backoff to a deterministic value after a successful transmission.

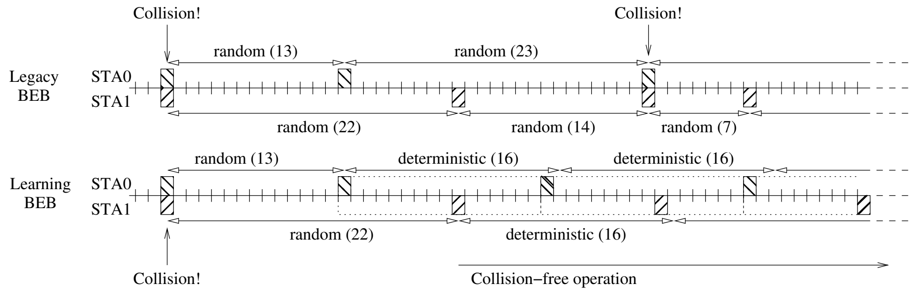


Fig. 2. The operation of BEB and L-BEB is compared. While the former computes a random backoff after a successful transmission, the latter always chooses 16. In L-BEB, after all the participating stations successfully consecutively transmit, the system enters in collision-free operation.

4 Learning-BEB

In BEB, the stations randomly access the channel, without any attempt to collect any feedback from previous transmission attempts. This means that, if two saturated stations compete for the channel for a long time, the collision probability perceived by the stations remains constant. After a transmission attempt, a station samples a random backoff number from CW_{min} if the transmission attempt was successful. Otherwise, the current contention window is doubled before drawing the backoff number.

It is easy to modify the protocol to allow the stations to learn from previous transmission attempts and decrease the number of collisions. Consider the same example of two stations competing for the channel. In this case, the stations use a constant backoff value ($V = 16$) after a successful transmission. At the beginning, the two stations randomly transmit without any knowledge about the other station's intention to transmit. However, as soon as the two stations successfully consecutively transmit, each of the stations periodically transmit every $V = 16$ slots. Since the selection of the transmission slot is deterministic, the chances of suffering collisions disappear, and the stations will orderly transmit in a TDMA fashion.

Figure 2 shows a graphical example. It represents two time lines divided in slots. Even though the actual duration of empty, successful and collision slots is different, in the figure they are all represented equal for simplicity.

In the upper time line, the stations operate using legacy BEB. The two stations collide in their first transmission attempt. The stations double their contention window and draw backoff values, specifically 13 and 22. After 13 empty slots, *STA 0* transmits, and obtains a new backoff value equal to 23. After 8 empty slots, *STA 1* successfully transmits and draws a backoff value of 14. The result, in this example, is that both stations will collide in their following transmission attempt.

In the lower time line of Figure 2 the stations use Learning-BEB (L-BEB), which is the backoff mechanism proposed in this paper. The beginning of the example is similar: the two stations collide and draw different backoff numbers. At this point, the behaviour of the system will become completely deterministic. *STA 0* successfully transmits and thus sets its backoff counter to $V = 16$. After eight empty slots *STA 1* successfully transmit and sets its backoff counter to $V = 16$. 6 empty slots later, *STA 0* successfully transmits again. And after eight empty slots it is *STA 1*'s turn. Both stations continue to transmit in turns occupying slots 0 and 9 of a virtual 16-slot TDMA frame. The suppression of collisions should be warmly welcomed because implies more efficient channel utilization.

We introduce the concept of *virtual frame* to highlight the similarities with TDMA. The virtual frame consists on V slots. Throughout this article we consider $V = 16$ for similarity with legacy BEB. In legacy BEB a station uniformly draws a random number between 0 and 31 after a successful transmission. Thus we choose L-BEB to wait for 16 slots after a successful transmission. The value of V can be tuned to adjust the behaviour of L-BEB. Although we provide some insights about the implications of tuning V by the end of this section, an exhaustive study is considered out of the scope of this work.

In the example in Figure 2, the virtual frames appear as a dotted line. After a successful transmission, a station will retransmit in the same slot position in the next virtual frame. In Figure 2, if we number the frame's position from 0 to 15, *STA 0* and *STA 1* transmit in positions 0 and 9 respectively.

The frame is virtual because there is neither explicit signaling nor configuration to assign a slot to a station. Additionally, the virtual frame only applies to those stations that successfully transmit, because the rest operate as in legacy BEB by selecting random backoff numbers. Moreover, a station that deterministically selects its next transmission slot does not have any kind of reservation for that slot.

For exemplifying purposes we have considered the simplest case of two contending stations. Nevertheless, the same conclusions apply for an increasing number of stations up to V. There is an initial transitory phase in which collisions occur and the stations try to find their place in the virtual TDMA frame. A station that successfully transmits in a given slot of the virtual frame, will keep transmitting in the same slot until a collision occurs. If a collision occurs, the station should draw a random backoff number from a doubled contention window. Eventually, all the stations will sequencially successfully transmit. At this point, each station has found its slot in the virtual frame and collisions will vanish.

Obviously, the higher the number of contending stations, the longer it will be the transitory operation of the protocol, since it is more difficult that all stations choose a different slot. In a naive approximation, we consider the probability that n stations choose different slots from a V-slot virtual frame.

$$\prod_{i=1}^{n-1}\left(1-\frac{i}{V}\right) \quad ; \quad 1 < n \leq V \tag{10}$$

If the value of n is low, all the stations will probably choose a slot different from the others. Oppositely, when the value of n is larger, it is more probable that some of the stations successfully transmit while others collide. Continuing with the approximation of the virtual frame, if n_s ($n_s < n$) stations successfully transmitted in the previous virtual frame, the probability that the rest of stations choose a slot that does not result in collision is:

$$\prod_{i=n_s}^{n-1}\left(1-\frac{i}{V}\right) \quad ; \quad 1 \leq n_s < n \leq V \tag{11}$$

The intuition is that the higher the value of n_s, the closer we are to the TDMA operation of the system. In the next section, simulations will be used to find out how long it takes to reach the stationary condition, depending on the number of active stations.

Special attention deserves the case in which the number of contending stations n is greater thant the size of the virtual frame (V).

It is not possible to fit more than V stations in a frame containing V slots. Thus, the system will not reach stability and the collisions will not completely disappear, no matter how long the system is running. Nevertheless, the system still performs as a CSMA/TDMA hybrid and, therefore, outperforms pure CSMA. There will be some stations that successfully transmit and deterministically choose their backoff, while the others collide and operate as CSMA stations.

4.1 Limitations of L-BEB

L-BEB shows its full potential after a short period of learning process. Ideally, after each station has found its place in the virtual frame, the system operates without collisions, until a perturbation moves the system back to the transitory phase. This perturbation could appear in the form of a new station entering the contention. It might happen that the new entrant successfully transmits in its first transmission attempt. If this is the case, no collision occurs and the systems continues its TDMA-like operation. Otherwise, when a collision occurs, there will be two stations selecting a random backoff algorithm before re-attempting transmission. This two stations might, in turn, generate new collisions initiating a chain reaction that brings the system to its transitory CSMA-like operation. Therefore, in a scenario with a high number of new entrants (in the order of multiple new incorporations per second), the medium access mechanism will be closer to CSMA and the advantages of using L-BEB will not be so obvious.

The transitory operation of the protocol can be shortened by increasing V. It can be observed that a higher value of V leads to higher success probabilities in (10) and (11). Moreover, a higher value of V allows for more terminals to operate in a collision-free fashion since collision-free operation is only possible when $n \leq V$. However, increasing the virtual frame size V has the side effect of lowering the efficiency when the number of contending stations is low. One could argue that V should be chosen as a function of the number of contending stations. However, the estimation of the number of contending stations is not trivial. For this reason, we opt for a static configuration of V for the paper and leave the dynamic selection of V for further study.

In a realistic scenario, a packet might be lost due to bad channel conditions. A station losing a packet would not be able to differentiate whether the packet was lost due to a collision or because of poor channel conditions. In any case, the station will double the contention window and draw a random backoff number. This action will also endanger the stability the same way a new entrant does.

Finally, the argumentation in Figure 2 is valid only if all the stations share the same vision of the channel, i.e. if all the stations can listen to all the successful and collision slots. If there is a station that cannot listen to another station transmission, the slot count would be different for different stations and the system performance would be the same as the one obtained in BEB. This last problem can be alleviated by using request-to-send and clear-to-send packets (RTS/CTS).

RTS/CTS signaling packets are transmitted before the actual data transmission and include a Network Allocation Vector (NAV) that describe the channel occupation intentions. RTS and CTS are sent by the sender and the receiver, respectively. Therefore, the channel occupation information reaches all the stations that can hear either the sender or the receiver. The RTS/CTS also limits the impact of collisions, since collisions can only occur in signaling (short) packets. Nevertheless, the RTS/CTS mechanism adds extra signaling overhead thus reducing the overall efficiency of the channel.

5 Simulation Results

The goal of the simulations[3] is to show that a performance improvement can be obtained by substituting BEB for L-BEB. The performance is a function of the number of empty, successful and collision slots. It is desirable to maximize the number of successful slots while minimizing the number of collisions. Empty slots play a minor role in the performance evaluations, because they are much shorter than successful transmissions and collisions.

By counting the number of successful and collision slots, the performance of a backoff algorithm can be evaluated. Nevertheless, while

---

[3] The simulations were performed in Octave. All the scripts are available upon request to the corresponding author.

the performance of BEB is maintained along the simulation, L-BEB learns and delivers better performance by the end of the simulation. However, we will postpone the analysis of the evolution of the performance of L-BEB. For the first round of simulations, the average number of empty, successful and collision slots in the first 1000 slots are studied.

The value of 1000 is arbitrary, choosing a higher value would highlight the advantage of L-BEB whereas a lower value would bring the curves in the plots closer. Note that even though all the simulations contain the same number of slots, the simulated time of the different simulations is not equal. The reason is that the duration of empty, successful and collision slots is different. Assuming a successful transmission time $T_s = 6.64$ ms and that halve of the slots are empty, the duration of 1000 slots would be 3.33 seconds.

The number of active stations in the simulations ranges from 2 to 20. Figure 3 compares the number of collisions in the first 1000 slots when using BEB and L-BEB. Each simulation is repeated 100 times and both the average and the 95% confidence interval are computed. It can be observed that by employing L-BEB instead of BEB, the number of collisions is reduced for any number of stations from 2 to 20. Even when the number of stations is greater than the size (in slots) of the virtual frame (V), L-BEB consistently achieves a lower number of collisions than BEB.

Figure 4 shows the number of successful slots. The first observation is that the number of successful slots is much higher when using L-BEB. This is a direct consequence of the lower number of collisions. Remember that, after a collision, the stations double their contention window and therefore reduce their transmission rate. L-BEB, reduces the number of collisions and allows the stations to keep a higher sending rate. Further, thanks to the CSMA-TDMA hybridization, L-BEB permits that the higher transmission rate does not translate to a higher number of collisions. This is true even when the number of contending stations is higher than the number of slots in the virtual frame.

The values in Figure 4 are those obtained in the first 1000 (transitory) slots. In steady state (collision-free) operation, the fraction of successful slots is n/V for $n \leq V$.

The original goal of the article was to increase the efficiency of the MAC access protocol by reducing the number of collisions. When the stations orderly and deterministically transmit, it is possible to outperform legacy BEB. It is also possible to cross the upper limit associated to random transmission. This is shown in Figure 5. The values of the axes replicate those of Figure 1 to ease comparison. The efficiency values presented in Figure 5 are those obtained during the first 1000 slots. As the system keeps learning, the efficiency further increases. The efficiency is a also a function of the frame length; for this reason, the efficiency is plotted for two values of T_s (the same values that were used in Figure 1).

Figure 6 evaluates the duration of the transitory in L-BEB. The transitory is characterized by collisions, while in steady-state conditions collisions theoretically disappear for $n \leq V$. To evaluate how the number of collisions fluctuate along the first 1000 slots, we plot the cumulative number of collisions. The cumulative number of collisions steadily grows at the beginning of simulations and becomes flat as the simulation advances and collisions disappear. The results presented in the plot are the average of the 100 simulations.

It can be observed that when the number of stations is 8, the steady-state condition is reached in about 200 slots. If the number of active stations is increased to 12, the steady-state condition is not reached within the simulation, since the curve does not become completely flat. Nevertheless, a reduction in the number of collision can be appreciated as the simulation progresses. Finally, for the case of $n = V$, the number of collisions is high even by the end of the simulation. Even though it is theoretically feasible to reach a steady-state condition without collisions for $n = V$, the probabilities are so small that it is not something that we can expect to happen in simulated or real scenarios.

6 Conclusion

This article addresses MAC protocols for wireless local area networks. In the extensively used Binary Exponential Backoff, the stations randomly select backoff (waiting) values to separate transmission attempts. Prior art struggled to optimize the parameters of BEB to improve its efficiency. Nevertheless, even if optimal transmission

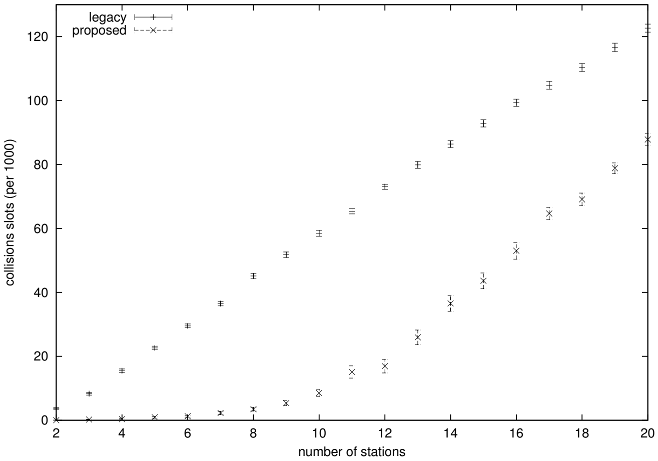


Fig. 3. The number of collision slots out of the total 1000 slots are plotted in this figure. The number of active stations range from 2 to 20. The results are presented with 95% confidence intervals

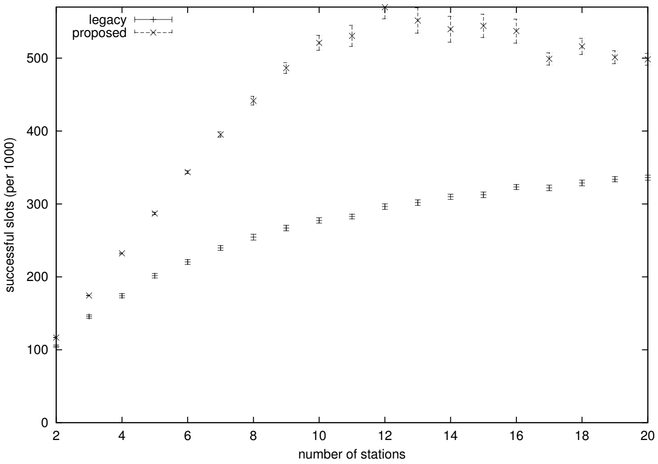


Fig. 4. The number of successful slots out of the total 1000 slots are plotted in this figure. The number of active stations range from 2 to 20. The results are presented with 95% confidence intervals

probability is used, the efficiency of the channel utilization is far from 100%. The explanation is that the stations blindly transmit, unaware of other stations's intentions to transmit.

We propose a framework to compute the efficiency of a MAC mechanism and we apply it to analyze BEB. We also derive the

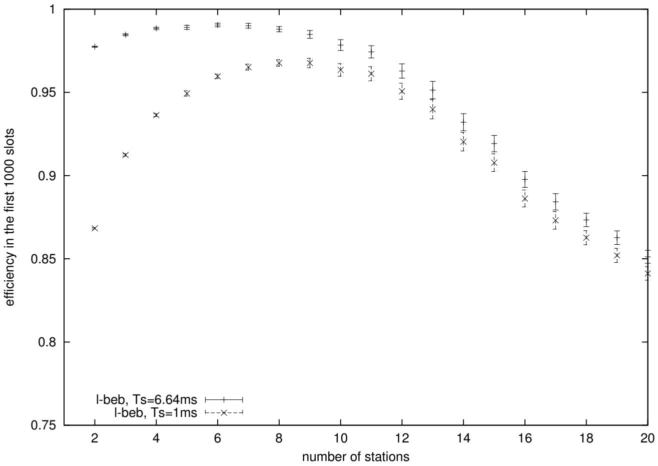


Fig. 5. The efficiency of L-BEB during the first 1000 slots as obtained from the simulations. The number of contending stations range from 2 to 20. The 95% confidence intervals are also plotted.

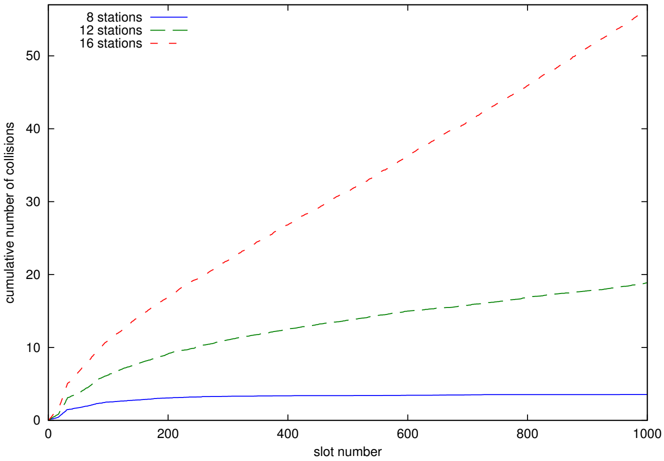


Fig. 6. The cumulative number of collisions in each of the first 1000 slots of the simulation of L-BEB. As the system reaches steady-state, collisions disappear and the curves become flat.

maximum efficiency that can be obtained from non-learning backoff schemes. Then we suggest a minor change to BEB: choosing a deterministic backoff value after a successful transmission, instead of a random one. By this simple modification, we allow that the stations learn from both collisions and successful transmission, thus reducing the chances of future collision. The system initially performs as legacy BEB, but after a few transmissions, the benefits of learning become clear and the collisions diminish. When the number of stations is lower than V, the collisions eventually vanish and the system operates in a collision-free fashion.

We call the new MAC algorithm Learning-BEB, since its performance improves over time, until a new station is activated. If this occurs, L-BEB has to learn again to adapt to the new scenario. Simulations have been used to show that L-BEB, in addition to reduce the number of collisions, also increases the number of successful transmissions. The combination of both effects positively impacts the efficiency, pushing it higher than the upper bound for non-learning algorithms.

The simulations are also used to understand the learning curve of L-BEB by analyzing the cumulative number of collisions over the first 1000 slots of the simulations. The conclusion is that, when the number of contending stations is low (lower than 8), the system quickly enters in collision-free operation. However, when the number of contending stations is higher, the learning pace is slower and the systems spends a long time in transitory operation. During the transitory, collisions still occur. Nevertheless, the number of collisions is lesser than in legacy BEB. If the number of contending stations is greater than the virtual frame size V, the system never reaches a steady-state (collision-free) condition. However, even in this extreme situation, L-BEB still outperforms BEB.

The main contribution of this paper is suggesting a minor change in the Binary Exponential Backoff that reduces the complexity and dramatically boosts efficiency.

Acknowledgment

The authors would like to thank the anonymous reviewers for their insightful comments.

References

1. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999 Edition (Revised 2003).

2. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications; Amendment: Medium Access Control(MAC) Quality of Service Enhancements. (IEEE Std 802.11e), 2005.
3. N. Abramson. The ALOHA System–Another Alternative for Computer Communications. *Cluster Computing*, 5:187–201, 1970.
4. H. Anouar and C. Bonnet. Optimal Constant-Window Backoff Scheme for IEEE 802.11 DCF in Single-Hop Wireless Networks Under Finite Load Conditions. *Wireless Personal Communications*, pages 1–20.
5. G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE J. Sel. Areas Commun.*, 18(3):535–547, 2000.
6. G. Bianchi, F. Borgonovo, L. Fratta, L. Musumeci, and M. Zorzi. C-PRMA: a centralized packet reservation multiple access for localwireless communications. *IEEE Transactions on Vehicular Technology*, 46(2):422–436, 1997.
7. G. Bianchi and I. Tinnirello. Kalman Filter Estimation of the Number of Competing Terminals in an IEEE 802.11 Network. *IEEE INFOCOM*, 2.
8. W. Crowther, R. Rettberg, D. Walden, S. Ornstein, and F. Heart. A System for Broadcast Communication: Reservation-ALOHA. *Proc. 6th Hawaii Int. Conf. Syst. Sci*, pages 596–603, 1973.
9. A. Ephremides and O.A. Mowafi. Analysis of a Hybrid Access Scheme for Buffered Users- Probabilistic Time Division. *IEEE Transactions on Software Engineering*, 8(1):52–60, 1982.
10. DJ Goodman, RA Valenzuela, KT Gayliard, and B. Ramamurthi. Packet Reservation Multiple Access for Local Wireless Communications. *IEEE Trans. Commun.*, 37(8):885–890, 1989.
11. L. Kleinrock and F. Tobagi. Packet Switching in Radio Channels: Part I–Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics. *IEEE Transactions on Communications [legacy, pre-1988]*, 23(12):1400–1416, 1975.
12. A. Lopez-Toledo, T. Vercauteren, and X. Wang. Adaptive Optimization of IEEE 802.11 DCF Based on Bayesian Estimation of the Number of Competing Terminals. *IEEE Trans. Mobile Comput.*, 5(9):1283, 2006.
13. I. Rhee, A. Warrier, M. Aia, and J. Min. Z-MAC: a hybrid MAC for wireless sensor networks. In *Embedded Networked Sensor Systems*, pages 90–101. ACM Press New York, NY, USA, 2005.
14. S. Tasaka. Stability and Performance of the R-ALOHA Packet Broadcast System. *IEEE Trans. Comput.*, C-32:717–726, Aug. 1983.
15. YC Tay and KC Chua. A Capacity Analysis for the IEEE 802.11 MAC Protocol. *Wireless Networks*, 7(2):159–171, 2001.
16. A. Zahedi and K. Pahlavan. Natural hidden terminal and the performance of the wireless LANs. *IEEE Universal Personal Communications Record*, 2, 1997.

CSMA/ECA: Carrier Sense Multiple Access with Enhanced Collision Avoidance

J. Barcelo[1], A. Lopez-Toledo[2], C. Cano[1], M. Oliver[1]

[1] Universitat Pompeu Fabra
[2] Telefonica Research Labs

Abstract. This paper presents CSMA/ECA, which combines the efficiency of reservation-based protocols and the simplicity of random access mechanisms. CSMA/ECA stations fairly coexist with legacy CSMA/CA and increase the portion of time that is devoted to successful transmissions while decreasing the number of collisions and empty slots. The system initially behaves as a CSMA/CA network, but it progressively converges to a collision-free deterministic operation. The convergence process can be modelled as a Markov Chain to assess the duration of the transitory phase. We show that the proposed mechanism outperforms the upper theoretical limit of CSMA/CA with optimal parameter adjustment.

1 Introduction

In many communications systems, a broadcast channel is shared by a set of stations. There are different strategies to arrange the sharing, which are called multiple access mechanisms. One option is to divide the resources (time, frequency, carriers or codes) among the different participating nodes. The nodes can also take turns in transmitting, and explicitly signal the end of each turn. Those alternatives prevent that two stations simultaneously transmit.

A popular medium access technique in local area networks is Carrier Sense Multiple Access (CSMA) [17]. The key property of CSMA networks is that the stations listen before transmitting. A station with data ready to transmit senses the channel for a given amount of time and, if the channel is detected idle, the station transmits.

It is still possible that collisions occur in CSMA because the propagation of the communication signals is not instantaneous, and real communication systems require a certain amount of time to switch from a listening mode to a transmitting mode.

In CSMA with Collision Avoidance (CSMA/CA), the stations defer their transmission a random number of slots. The efforts to reduce the number of collisions are motivated by the fact that collisions represent a significant waste of resources in wireless networks, since it is not feasible to immediately detect a collision and interrupt the transmission. The stations either transmit or receive, and cannot collect any feedback from the radio channel while they are transmitting.

CSMA/CA combined with truncated Binary Exponential Backoff (BEB) is at the core of the Medium Access Control (MAC) specification in the suite of protocols IEEE 802.11 [1]. These protocols are widely used in Wireless Local Area Networks (WLANs) and, for this reason, they have been the subject of extensive research with the goal of reducing collisions and improving performance.

In spite of the possibility of collisions, CSMA/CA is still an appealing protocol for WLAN. It is lightweight, it takes advantage of statistical multiplexing to accommodate bursty traffic and it can be executed in a distributed fashion. CSMA/CA is especially fitted for networks with a large number of stations that sporadically send one packet. However, CSMA/CA was not designed to benefit from the fact that some stations have multiple-packet messages [8, 18], *i.e.* stations that store several packets in their transmission queues.

When stations send multiple consecutive packets, it is possible to use the feedback obtained from previous transmissions attempts to adequately schedule future transmissions. For this reason, we suggest a modification to the CSMA/CA protocol that further reduces the number of collisions while maintaining all its versatility and power. We call the new protocol CSMA with Enhanced Collision Avoidance (CSMA/ECA).

The main features of the presented CSMA/ECA protocol are the following:

- It outperforms the theoretical upper bound efficiency of CSMA/CA with optimal parameter adjustment.
- It provides a collision-free medium access after a transitory phase.
- It fairly coexists with legacy CSMA/CA.
- It works in a distributed fashion.
- It does not require additional computational efforts and can be easily implemented.

- It is robust against channel errors.

The rest of the paper is organized as follows: Section 2 defines the CSMA/ECA algorithm, then in Section 3 a Markov Chain model to predict the length of the transitory phase is described. Implementation issues and the performance evaluation results are discussed in Section 4 while Section 5 presents an overview of the related work in the area. Finally, some conclusions are given.

2 Enhanced Collision Avoidance

In CSMA/CA, whenever there are backlogged stations with a packet ready to be transmitted, the channel time is implicitly divided into slots. Three different kinds of slots are differentiated: empty, successful and collision. A slot is empty when no station attempts transmission; successful if one (and only one) station transmits; and collision if more than one station simultaneously transmit. The channel time spent in empty slots or collision slots is wasted.

Whenever a station has to defer its transmission, it chooses a random backoff value B from a contention window.

$$B \sim \mathcal{U}[0, CW-1], \tag{1}$$

where $\mathcal{U}$ is the uniform distribution and CW is the contention window.

We consider that the stations are saturated (*i.e.* the stations always have a packet ready to transmit). As a consequence, the stations are either transmitting, receiving or backing off, but they are never idle. After each transmission attempt, the stations choose a backoff value.

The stations have to backoff both after collisions and successful transmissions. For the first case, the backoff has to be necessarily random to prevent a new collision in the retransmission attempt. However, for the second case, the backoff value can be deterministically selected.

2.1 Deterministic Backoff After Successful Transmissions

By choosing a deterministic backoff after a successful transmission and a random backoff otherwise, the system converges to a collision-free operation when the number of active stations is not greater

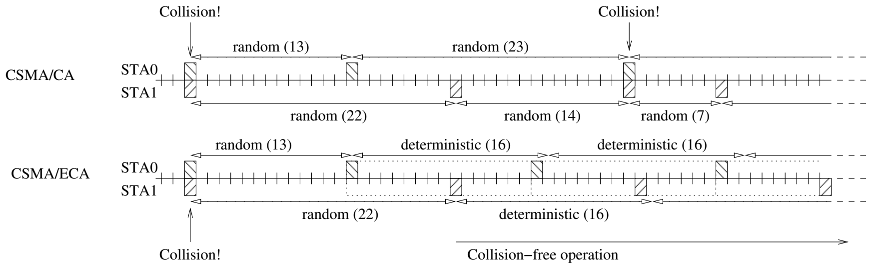


Fig. 1. CSMA/CA is compared to CSMA/ECA in an example in which two saturated stations contend for the channel. When CSMA/ECA is used, after both stations have successfully transmitted, the behaviour of the stations is deterministic and no more collisions occur.

than the value of the deterministic backoff. In the case of a successful transmission, the deterministic behaviour stabilizes the system (hopefully leading to another success). Conversely, if there is a collision, the randomness of the backoff provides a change that would (desirably) avoid more collisions. The system exploits the information gathered from previous transmission attempts to further reduce the collisions, thus we call it Enhanced Collision Avoidance (ECA). The terminals perform a random search to find free slots, until collisions disappear.

It has to be clear that a station keeps using a deterministic backoff while it successfully transmits. As soon as it suffers a collision, it moves back to the random behaviour. That collision will always be caused by a station that randomly selected its transmission slot, since collisions among stations that behave deterministically are not possible.

This principle can be better understood by an example. Consider the simplest case of two stations (*STA 0* and *STA 1*) contending for a channel, as shown in Fig. 1. The channel time advances from left to right and it is divided in slots.

Even though the actual duration of empty, successful and collision slots differ, all the slots are equally represented in Fig. 1 for simplicity reasons. The upper channel time line corresponds to legacy CSMA/CA, while the lower one incorporates the modifications we have proposed for CSMA/ECA. The constant backoff after successes is a value that depends on the 802.11 flavor, as will be explained in subsection 4.1.

The figure shows the slots at which each of the stations transmits. It also shows the backoff value chosen by each station (between brackets). The label of the backoff value also indicates whether it has been chosen randomly or deterministically.

In the example, the two CSMA/CA stations collide, then successfully transmit and, finally, collide again. When CSMA/ECA is used, collisions disappear after all stations have successfully transmitted, because the backoff is selected deterministically. It is useful to imagine a virtual frame[3] of V slots (represented with a dotted line in the figure) and observe that, after collisions disappear, the stations transmit in fixed slot positions within the virtual frame, similarly to a TDMA operation.

Algorithm 2 represents the protocol that is distributedly executed in each of the contending stations. The meaning of each of the variables is as folows:

- b is the backoff counter.
- CW_{min} is the minimum contention window.
- CW_{max} is the maximum contention window.
- a is the number of transmission attempts.
- A is the maximum number of transmission attempts.
- V is the deterministic backoff value after successful transmissions.

Let us define the channel efficiency (ϕ) as the fraction of channel time that is devoted to successful transmissions,

$$\phi = \frac{P_s T_s}{P_e T_e + P_s T_s + P_c T_c}, \tag{2}$$

where P_e, P_s and P_c are the empty, success and collision probabilities, respectively. And T_e, T_s and T_c are the duration of an empty, successful and collision slot, respectively.

Then, for a number of contending stations (ς) not greater than the size of the virtual frame, the efficiency that can be obtained from

---

[3] Some works refer to data-link layer PDUs as frames. In this article, a frame is a group of slots. Data-link layer PDUs are called packets.

```
/* Initialize $b$.                                              */
$b \leftarrow \mathcal{U}[0, CW_{min} - 1]$;
while there is a packet to transmit do
    /* Initialize $a$.                                          */
    $a \leftarrow 0$ ;
    while $a < A$ do
        /* First, backoff.                                      */
        while $b > 0$ do
            wait 1 slot ;
            $b \leftarrow b - 1$ ;
        end
        Attempt transmission ;
        if success then
            /* Deterministic backoff.                           */
            $b \leftarrow V$;
            break ;
        else
            /* If transmission fails.                           */
            $a \leftarrow a + 1$ ;
            /* Random backoff value.                            */
            $b \leftarrow \mathcal{U}[0, \min(CW_{min} * 2^a, CW_{max}) - 1]$;
        end
    end
end
```

Algorithm 2: CSMA/ECA

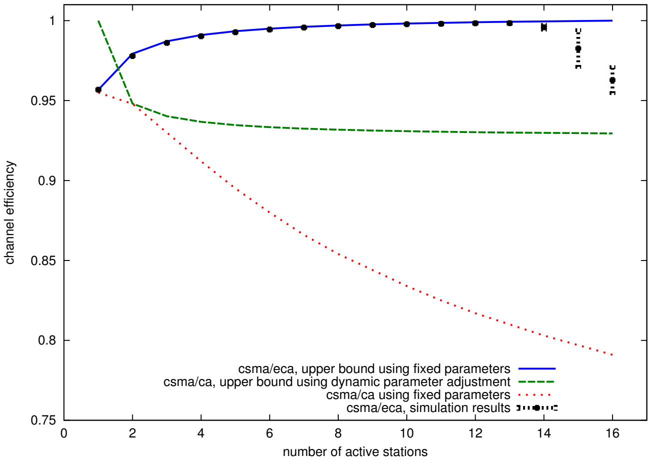


Fig. 2. The performance of CSMA/ECA with fixed parameters is compared to CSMA/CA with fixed and dynamic parameters. Simulations results are provided for CSMA/ECA.

CSMA/ECA in steady-state collision-free operation is :

$$\phi = \frac{\varsigma \cdot T_s}{\varsigma \cdot T_s + (V - \varsigma) \cdot T_e} \; ; \; \varsigma \leq V. \tag{3}$$

The channel efficiency as presented in (3) is plotted for the typical number of simultaneously active stations[4] and a backoff value $V = 16$ in Fig. 2. It is compared to legacy CSMA/CA with and without dynamic parameter adjustment. T_e and T_s are taken from the 11 Mbps IEEE 802.11b specification, considering a data load of 1500 bytes. The performance of CSMA/CA is derived from [6]. The curve for the upper bound of CSMA/CA is computed as proposed in [4]. Simulation [5] results are presented for CSMA/ECA with 95% confidence intervals.

Before reaching the steady-state and obtaining the efficiency as presented in (3), the system goes through a transitory operation. The efficiency obtained in the transitory operation is a value between the

---

[4] A station is active if it has a packet ready to transmit and it is competing for the channel. The number of active stations is usually only a fraction of the total number of stations registered to the network.

[5] A custom simulator of the medium access sharing mechanism has been used. It is programmed in Octave and the source code is available upon request.

efficiency delivered by CSMA/CA and the efficiency in (3), because only a fraction of the collisions is avoided. During this transitory phase, the number of stations that successfully transmit (and thus use a deterministic backoff) is a random variable. In the next section, the evolution of this number is modelled as a Markov Chain in order to draw additional conclusions about the transition process.

3 A Dissection of the Convergence Process

Consider a scenario with ς saturated stations and a virtual frame size of V slots, $2 \leq \varsigma \leq V$. We will assume that the transition process occurs in a frame-by-frame basis. Let X_n be the random variable that represents the number of stations that successfully transmitted in the frame n. Then we can model the transition process as a time-homogeneous Markov Chain and the state space is

$$\mathbf{S} = \{S_i | 0 \leq i \leq \varsigma\} \tag{4}$$

As the system runs, it transitions from an initial state S_0 to a (stable) state S_ς.

We are interested in computing the transition probability matrix $\mathbf{P}$ which is the matrix of one step transition probabilities $p_{i,j}$ defined by [6]

$$p_{i,j} = Pr(X_{n+1} = j | X_n = i) \; ; \; 0 \leq i, j \leq \varsigma. \tag{5}$$

Before dealing with the general computation of $p_{i,j}$, we will analyze some results that immediately arise from the definition of the problem and provide some insights about the behaviour of the model. Note that the following properties apply only to the model, and not necessarily to the system that is being modelled. However, they are helpful in computing the transition matrix for the model.

Claim. The system is stable when $X_n = \varsigma$, i.e. state S_ς is absorbing.

$$Pr(X_{n+1} = \varsigma | X_n = \varsigma) = 1. \tag{6}$$

---

[6] Note that we index the rows of the matrix from 0 to ς. This is for coherence with the numbering of the states of the Markov Chain.

Proof. $X_n = \varsigma$ implies that all the stations successfully transmitted in virtual frame n. Therefore, all the stations will deterministically choose the transmission slot in virtual frame $n+1$, specifically they will transmit exactly in the same position in the frame as they did in virtual frame n. As there were no collisions in frame n, there will be no collisions in frame $n+1$.

Claim. It is not possible that there is one and only one station that randomly selects the transmission slot in a given virtual frame.

$$Pr(X_n = \varsigma - 1) = 0\ ;\ n > 0. \tag{7}$$

Proof. Seeking a contradiction we assume that there is only one station that randomly selects the transmission slot in virtual frame n. This implies that this station suffered a collision in the previous frame $n-1$. Since a collision occurs when a minimum of two stations transmit in the same slot, there are at least two stations that will randomly select the transmission slot in virtual frame n. This contradicts our assumption.

3.1 Computing the Transition Probability Matrix

After these preliminary results, we face the general problem of computing $p_{i,j}$, i.e. the probability that we have j successful transmissions in the current virtual frame given that there were i successes in the previous frame. There are i stations that deterministically transmit in i different slots, while the rest of the stations $(\varsigma - i)$ randomly transmit in any of the V slots.

Note that for the special case $i = 0$, the problem is reduced to the computation of the number of successes that are obtained when ς stations transmit in V slots and can be solved using the model suggested in [15]. For any other value of i $(i \neq 0)$, the approach in [15] is no longer applicable, since it assumes that there are slots reserved for the stations that successfully transmitted in the previous frame. Hence we are interested in finding another scheme that can be used for any value of i.

For large values of V, a brute force approach that sweeps all the different combinations to obtain the transition probability matrix $\mathbf{P}$ is computationally impractical. To compute the first row of the

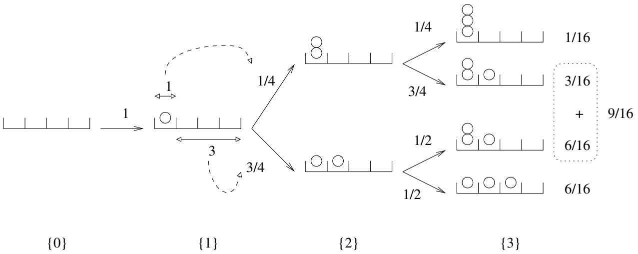


Fig. 3. A tree is used to evaluate the different outcomes that are possible in a system with $\varsigma = 3$ and $V = 4$.

transition matrix, it would be necessary to consider ς stations that could transmit in any of the V available slots, which would account for V^{ς} possibilities.

Nevertheless, certain shortcuts are possible to accelerate the computation of $\mathbf{P}$. The reason is that we are interested only in the number of successful slots in a virtual frame, but not in which are those successful slots. In other words, the slots are interchangeable. Similarly, we are not interested in which are the stations that successfully transmitted; all the stations are equivalent from our point of view.

Assume that the previous state is S_0 and we want to compute the probabilities $p_{0,j}$ for all values $0 \leq j \leq \varsigma$. Now consider a transmission in the current frame. This transmission can be in any of the V (for now, empty) slots. Since all these slots are empty, the V possible outcomes are equivalent for our analysis. Each of the V outcomes consists of a slot with one transmission and $V - 1$ empty slots. Following the same reasoning, for a second transmission in the same virtual frame, there are only two possible outcomes: *a)* that the transmission slot is the same as the one as the first transmission (which occurs with probability $1/V$) or *b)* the two transmissions are in different slots (which occurs with probability $(V - 1)/V$). The same rationale can be used to build a tree to obtain all the possible outcomes of interest and the probabilities associated with each outcome. A graphical example is presented in Fig. 3.

In Fig. 3 we show an example for $\varsigma = 3$ and $V = 4$. It is a tree with $\varsigma + 1$ levels. The root represents the $V = 4$ empty slots, and in every level, a new transmission (represented as a ball) is included. The levels are labeled as $\{0\}$, $\{1\}$, $\{2\}$ and $\{3\}$. The edges of the tree are labeled with probability values. At the first level, there is only one node, since the only possible situation (with only one transmission) is one success and three empty slots. Therefore, the edge from the root to the node at the first level is labeled with probability 1. In the transition from level {1} to level {2} there are two possible options: *a)* that the two transmissions occur in the same slot (with probability 1/4) and *b)* that the transmissions occur in different slots (with probability 3/4). This process is iterated until all the transmissions are included, and 4 leafs are obtained. By following the path from the root to the leaf, the probability of each leaf is computed. The probability that no station successfully transmits can be obtained from the first leaf: From the tree it can be observed that the transition probability from state S_0 to state S_0 is

$$p_{0,0} = Pr(X_{n+1} = 0|X_n = 0) = \frac{1}{16}. \tag{8}$$

The probability that there is only one success is $p_{0,1} = \frac{3}{16} + \frac{6}{16} = \frac{9}{16}$. The probability of two successes is zero $p_{0,2} = 0$ and the probability of three successes is $p_{0,3} = \frac{6}{16}$. With these values, we have already completed the first row of the transition matrix $\mathbf{P}$. To obtain the values for the second row, one has to assume that there was a successful collision in the previous virtual frame. Therefore, we consider only a subtree of the tree represented in Fig. 3, particularly the one with the root at the node of level {1}. To compute the third row of the matrix we use as a root the lower node of level {2}. The last row is computed using only one node, which is the lowest leaf. The transition matrix which is obtained[7] for this example is:

$$\mathbf{P}_{\varsigma=3,V=4} = \begin{pmatrix} \frac{1}{16} & \frac{9}{16} & 0 & \frac{6}{16} \\ \frac{1}{16} & \frac{9}{16} & 0 & \frac{6}{16} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{9}$$

It is not a coincidence that the first two rows of $\mathbf{P}_{\varsigma=3,V=4}$ are the same. Actually, since in level {0} all the slots are empty and thus equivalent, there is only one way to place the first ball. As a

---

[7] A script in Octave to compute the transition matrix for any value of V and ς is available upon request.

consequence, there is always only one node in level $\{1\}$, and the edge from $\{0\}$ to $\{1\}$ takes the value 1.

Claim. The first two rows of the transition probability matrix $\mathbf{P}$ are equal.

$$p_{0,j} = p_{1,j} \ ; \ 0 \leq j \leq \varsigma \tag{10}$$

Proof. Consider a tree as the one exemplified in Fig. 3. Then take the subtree with the node of level $\{1\}$ as a root. From this tree, we can obtain the values of the second row (indexed as 1) of the transition matrix $p_{1,j}$. Now, to obtain the first row (indexed as 0), we observe that we use exactly the same tree, but with an additional edge with value 1 and an additional node as a root. Then we can obtain the values of the first row by multiplying the values of the second row by one.

We are interested in evaluating how long does it take for the system to leave the transitory phase and begin the collision-free operation. We consider an initial state S_0 in which all the stations randomly choose their transmission slot and then we use the transition matrix $\mathbf{P}$ to evaluate the marginal distributions in subsequent frames. Let

$$\boldsymbol{\pi}_n = \{Pr(X_n = i), 0 \leq i \leq \varsigma\} \tag{11}$$

be the vector of the marginal probabilities at stage n, and $\boldsymbol{\pi}_0 = [1, 0, ..., 0]$ the initial vector. Then the vector $\boldsymbol{\pi_n}$ can be obtained by:

$$\boldsymbol{\pi}_n = \boldsymbol{\pi}_0 \mathbf{P}^n. \tag{12}$$

The last term of the vector, $\boldsymbol{\pi}_n(\varsigma)$, is precisely the value of interest for our study $Pr(X_n = \varsigma)$, which is the probability that the system has reached the stable collision-free state. One particularity of our evaluation of the transition curve is that we have considered that the transition step contains $2 * V$ slots *i.e.* two virtual frames. This is an approximation of the expected backoff of those stations that suffered a collision.

3.2 Validation by Simulation

The model presented above is based on two approximations with respect to the actual CSMA/CA operation. The first one is that, in the

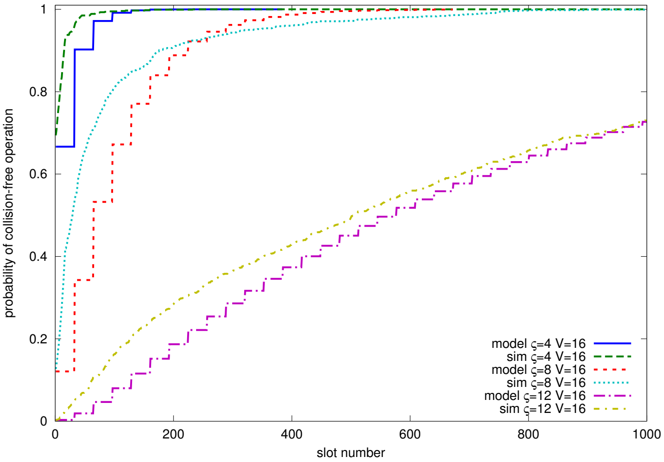


Fig. 4. The transition curves obtained using the model and simulation are compared for a value of $V = 16$ and various values of ς.

model, the convergence process occurs in a frame-by-frame basis. On the contrary, the CSMA/ECA algorithm allows that the same station re-attempts transmission (and eventually succeeds) in the same virtual frame. Actually, the virtual frame concept is not intrinsic of CSMA/ECA and it is an abstraction we have used for the analysis. The second concession to simplicity is that the exponential growing of the contention window has been neglected in our model. As a consequence of these two concessions (frame-by-evolution and static contention window), our model provides only an approximation to the expected behaviour of CSMA/ECA.

The probabilities of reaching the stationary operation have also been obtained by means of simulation. In Fig. 4, the probability that the system reaches the collision-free operation in a given slot is plotted, making it possible to compare the analytic and simulation results. It can be observed that, as the number of active stations ς increases, the transition process becomes slower.

As we can see in Fig. 3, the analytical model follows closely the simulation results. The small mismatch at the beginning of the convergence process is due to the aforementioned approximations, however the length of the transitory period is accurately predicted.

3.3 Disruption of the Stationary Operation

Although the system is expected to run in the collision-free mode of operation for most of the time, there are two events that can disrupt the stationary operation: a channel error and a new entrant. The model can be used to assess the recovery curves associated with these events. It is necessary to force the initial state to $S_{\varsigma-1}$. Regardless of the fact that the system will never transition to $S_{\varsigma-1}$, it is possible to use it as an initial state. It precisely reflects the fact that all stations but one are using a deterministic backoff. The initial vector under consideration is: $\boldsymbol{\pi}_0^D = [0, ..., 0, 1, 0]$.

And the marginal probabilities of subsequent steps:

$$\boldsymbol{\pi}_n^D = \boldsymbol{\pi}_0^D \mathbf{P}^n. \tag{13}$$

Provided that current state is S_i, we use the maximum number of collisions (worst case) in the previous step as an approximation of the actual number of collisions in the previous step:

$$\kappa_i \approx \lfloor \frac{\varsigma - i}{2} \rfloor. \tag{14}$$

where $\lfloor \cdot \rfloor$ is the floor operator. Then, using the approximation $T_c \approx T_s$, the efficiency of the system in the step $n-1$ is:

$$\phi_{n-1} \approx \sum_{i=0}^{\varsigma} \frac{2 \cdot i \cdot T_s}{(2 \cdot i + \kappa_i) \cdot T_s + (2 \cdot V - 2 \cdot i - \kappa_i) \cdot T_e} \pi_n^D(i), \tag{15}$$

where the expectation of the backoff of those stations that suffer collisions is considered to be twice as much as V.

Fig. 5 shows the recovery curves obtained from (13)-(15). The transitory phase associated with new incorporations to the contention can be avoided by means of Smart Entry, which will be described in Subsection 4.2.

4 Implementation Issues

In this section we address the coexistence of CSMA/ECA with the legacy protocol. We also study the impact of releasing assumptions such as the fixed number of contenders, saturated stations, fixed deterministic backoff value and ideal channel conditions.

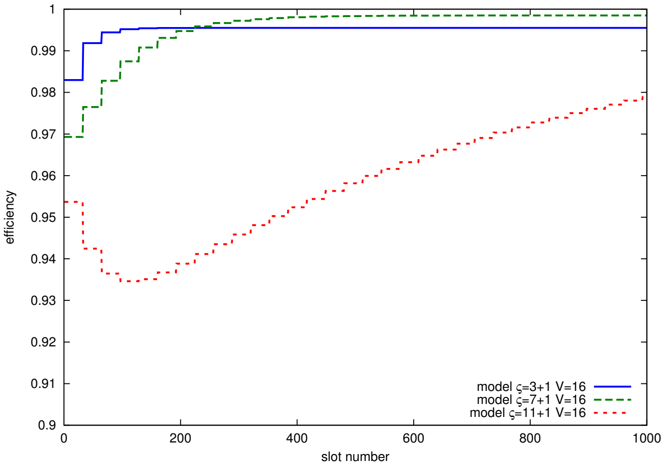


Fig. 5. Recovery curves after a channel error or new entrant.

4.1 Coexistence with legacy CSMA/CA

A promising field of application of the proposed CSMA/ECA is the successful protocol suite IEEE 802.11. Nevertheless, given the large number of deployed networks and terminals, any new version of the medium access control algorithm should be backward compatible with the already existing equipment. Further, to guarantee the smooth coexistence of new and legacy stations, those stations running CSMA/ECA should consume a fair amount of the available bandwith.

The only difference between CSMA/CA and CSMA/ECA as presented in Algorithm 2 can be found in line 11. CSMA/CA randomly chooses the backoff value from the minimum contention window ($b \leftarrow \mathcal{U}[0, CW_{min} - 1]$), while CSMA/ECA deterministically chooses as a value the size of the virtual frame ($b \leftarrow V$). In order to fairly compete with legacy stations, it is desired that

$$V = \lceil E\left[\mathcal{U}[0, CW_{min} - 1]\right]\rceil, \tag{16}$$

where $E\left[\cdot\right]$ represents the expectation operator and $\lceil \cdot \rceil$ is the ceiling operator. This selection of the virtual frame size guarantees that the expected number of slots that a station waits after a successful transmissions is approximately the same, for both CSMA/CA and CSMA/ECA.

To validate this idea, we performed simulations for a scenario in which half of of the stations run CSMA/CA while the other half use CSMA/ECA. The values chosen for the MAC parameters are $CW_{min} = 32$ and $V = 16$. The rest of the parameters are taken from the IEEE 802.11b specification. The efficiency obtained by each group of stations (CSMA/ECA and CSMA/CA) is computed separately. Each simulation runs for 10000 slots and each scenario is repeated ten times. The number of competing stations range from two to forty (only even values are considered). When a value of 40 stations is indicated, it actually means 20 CSMA/ECA stations plus 20 CSMA/CA stations.

The results are presented in Fig. 6. The plot also shows the aggregated channel efficiency, which is the sum of the efficiencies obtained by the two groups of stations.

It can be observed that CSMA/ECA flows obtain higher channel utilization thanks to the reduced collision probability. This small advantage can be seen as an incentive for legacy networks to shift to CSMA/ECA for the greater benefit of the network. The Jain's fairness [14] index has been computed as:

$$fairness = \frac{(\phi_{csma/eca} + \phi_{csma/ca})^2}{2(\phi^2_{csma/eca} + \phi^2_{csma/ca})}. \tag{17}$$

The possible outcomes range from 0.5 (worst case) to 1 (best case). We obtained results higher than 0.98 when comparing the efficiency of CSMA/ECA and CSMA/CA in a mixed scenario.

The benefits of using CSMA/ECA are greatly diminished in the presence of legacy stations since the collision-free operation is never reached. Nevertheless, a network running a mixture of CSMA/CA and CSMA/ECA stations will offer equal or better performance that a pure CSMA/CA network, since some of the collisions will be avoided.

To assess the benefits of using CSMA/ECA, we have repeated the simulations described above for a pure CSMA/ECA scenario and a pure CSMA/CA scenario. The results are shown in Fig. 7, which also includes the hybrid scenario. It can be observed that, thanks to the enhanced collision avoidance mechanism, a larger fraction of the channel time is devoted to successful transmissions when

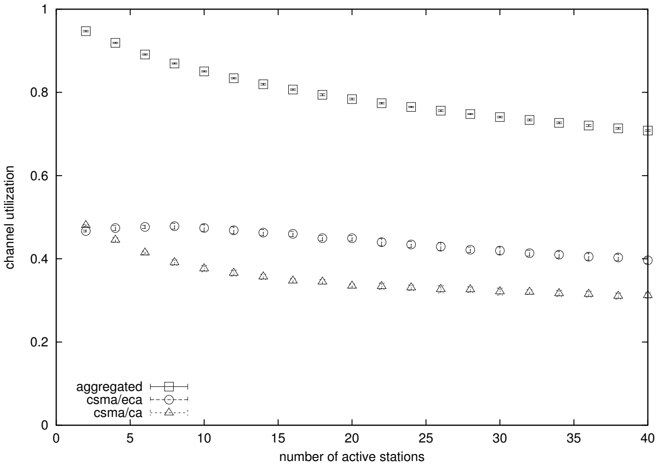


Fig. 6. Half of the stations run CSMA/ECA, while the other half run CSMA/CA. The figure shows the channel utilization achieved by each group.

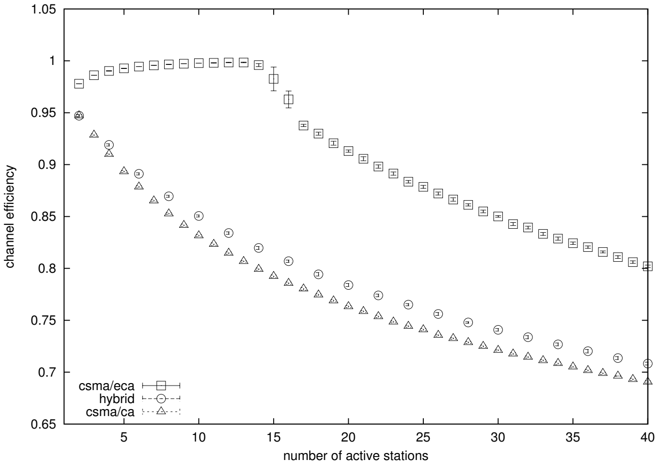


Fig. 7. The channel efficiency obtained for pure CSMA/ECA and pure CSMA/CA scenarios. The efficiency in a hybrid (mixed) scenario is also included.

CSMA/ECA is used. For a number of active stations up to the size of the virtual frame size V, the efficiency is almost 1.

It is noteworthy that, while CSMA/ECA delivers the best results for a number of contenders lower than the size of the virtual frame ($\varsigma < V$), it still clearly outperforms CSMA/CA when the number of active stations is larger.

4.2 Smart Entry

So far we have assumed that the number of contenders is fixed. Nevertheless, in a real network, the stations join an leave the contention depending on the load that they receive from the upper layers of the protocol stack.

Ideally, the system will run in the collision-free stable mode of operation. At this point, if a station that joins the contention selects the first transmission slot randomly, it poses the collision-free mode of operation of the system at risk: it may provoke a collision and move the system back to its transitory (collision-prone) mode of operation. To avoid this situation, the stations that are not actively contending for the channel should keep track of the empty slots in each virtual frame. When one of those stations receives a packet from the upper layer, it already knows which slots are expected to be empty, and can schedule the first transmission accordingly.

If Smart Entry is to be used, the first line of Algorithm 2 has to be substituted by Algorithm 3. It includes an array called `slotNumber[]` to keep track of the status of each slot of the frame. The size of this array is precisely the size of the virtual frame V. With the modification presented in Algorithm 3, a station joining the contention transmits in the first empty slot.

Note that while the station is delaying the first transmission attempt, it marks the positions in the array as free. This behaviour prevents a deadlock in the case in which all the slots are busy. If there are no free slots, the station will delay its transmission attempt V slots, and then deliberately prompt a collision in order to free some slots for a future transmission attempt.

4.3 Dynamic Parameter Adjustment

For the sake of completeness, we will also consider the dynamic adjustment of the parameter V for CSMA/ECA. However, one should be aware of the difficulties associated with the implementation of this approach: *i)* backward compatibility with legacy networks is compromised (since a modification of the beacon is needed to distribute the value of V), and *ii)* since it requires a central entity, it is not suited for ad-hoc deployments.

```
/* Initialize slotNumber[]                                          */
for $i \leftarrow$ 0 to $V - 1$ do
    slotNumber[$i$] $\leftarrow$ unknown ;
end
i $\leftarrow$ 0 ;
/* Scan the channel while waiting for a packet from the upper
   layers.                                                          */
while True do
    if there is a packet ready to transmit then
        if slotNumber[i] is free then
            transmit ;
            /* Leave Smart Entry and move to normal CSMA/ECA
               operation.                                           */
            break ;
        else
            wait 1 slot ;
            slotNumber[$i$] $\leftarrow$ free ;
        end
    else
        wait 1 slot ;
        if channel sensed busy then
            slotNumber[$i$] $\leftarrow$ busy ;
        else
            slotNumber[$i$] $\leftarrow$ free ;
        end
    end
    $i \leftarrow (i + 1) \pmod{V}$ ;
end
```

Algorithm 3: Smart Entry

If the dynamic adjustment of V could be implemented, the central entity would broadcast a value of V equal to the number of active stations ς. Combining the dynamic adjustment with Smart Entry, the maximum efficiency 1 can be achieved. This value is obtained by evaluating (3) for $V = \varsigma$. The virtual frame size would be equal to the number of contending stations and each station would transmit in one of the slots, thus each slot would contain a successful transmission. There would be no collisions nor empty slots.

4.4 Non-Saturated Stations

Throughout the article, we have considered that the stations are saturated. Nevertheless there exist possible scenarios in which the saturation assumption does not hold. This is the case of WLANs devoted to voice communications [5,13]. In this kind of network, the stations periodically join and leave the network to send one single packet. Thus, the deterministic backoff after successful transmissions does not apply: after a success, there is not a second packet to send.

The result is that, for lightly loaded networks, CSMA/CA and CSMA/ECA behave exactly the same. However, as the load of the network increases (e.g. more stations join the contention) and the network approaches congestion, the MAC queues build up. At this point, when each queue has multiple packets to send, the stations are saturated and the enhanced collision avoidance increases the channel efficiency. Therefore, CSMA/ECA networks can accept higher loads than CSMA/CA before loosing packets due to MAC layer queue overflow.

4.5 Releasing the Ideal Channel assumption

So far we have considered an ideal channel that introduced no errors. Now we will assess the performance of CSMA/ECA when the channel is unreliable. Note that the behaviour of the protocol is the same for transmission errors and collisions. We want to stress the proposed protocol by introducing packet errors with probability of 10^{-2}. This 1% threshold was used as a standard measure of robustness by the IEEE 802.11 committee [12].

The simulations in Fig. 8 are performed in the presence of imperfect channel conditions. The packet errors are treated as collisions by

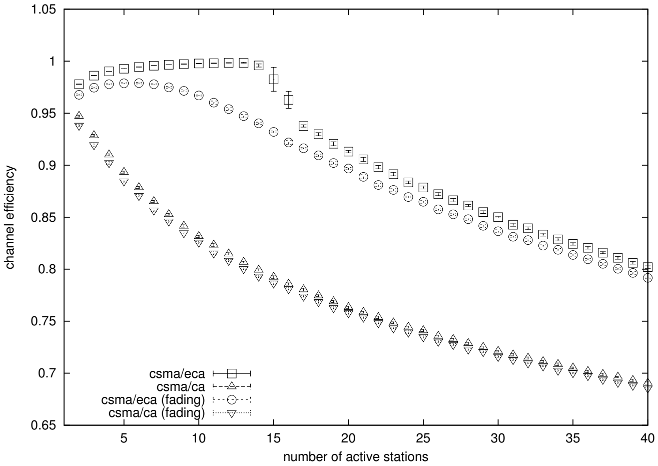


Fig. 8. The channel efficiency delivered by CSMA/ECA and CSMA/CA in an unreliable channel.

the stations and hence, interfere in the enhanced collision avoidance mechanism.

This is specially true when the value of the number of active stations (ς) is close to the size of the virtual frame (V). Nevertheless, CSMA/ECA still outperforms CSMA/CA for any number of competing terminals.

5 Related Work

The performance of CSMA/CA is particularly critical for the IEEE 802.11 family of protocols in its different physical flavors. The Distributed Coordination Function (DCF) is the algorithm that arbitrates the access to the channel. DCF combines CSMA/CA and truncated BEB, and its performance is modelled in [6]. The standard amendment for quality of service IEEE 802.11e [2] introduced an Enhanced Distributed Channel Access (EDCA). EDCA allowed the modification of some key MAC parameters and opened a new research avenue to improve the performance of WLAN.

The more successful line of research until now to maximize the performance of IEEE 802.11 CSMA/CA networks implies the estimation of the number of active terminals [7, 16] and the use of this value to tune the MAC parameters to optimize the throughput [3].

The estimation involves a relationship between the collision probability and the number of contenders that holds true only in ideal channel conditions.

Even if the number of active stations could be accurately estimated, a theoretical limit exists (See Fig. 2) for the maximum efficiency of DCF and EDCA [4,9]. There is a performance loss associated with the random selection of the transmission slot, because of the unavoidable collisions and the empty slots. The idea of choosing a deterministic backoff after successful transmissions to reduce the number of collisions and boost the channel utilization was already suggested in [4] and it has been further explored in the present article.

CSMA/ECA is closely related to the Reservation-Aloha [11, 18] protocol. Even though CSMA/ECA does not contemplate the possibility of reserving slots, the steady state operation of Reservation-Aloha and CSMA/ECA is very similar. In the Reservation Aloha protocol, the channel time is divided into slots with duration equal to the transmission time of one packet. The slots are grouped in frames, and the duration of the frame is chosen to be greater than the propagation delay of the channel. This detail is important because Reservation Aloha was initially designed for satellite communication, and users needed to be aware of the usage status of the slots in the previous frame.

In Reservation-Aloha, if a station successfully transmits in a slot in a given frame, it implicitly reserves the same slot for the following frame. Those slots that are not used in the current frame are not reserved and are available for contention in the following frame. Collisions can only occur in the frames that are free for contention.

Compared to Reservation-Aloha, the protocol presented in this paper has three advantages. First, CSMA/ECA can operate with a number of active stations greater than the size of the virtual frame V, while in Reservation-Aloha, the number of active contenders is strictly limited to the size of the frame. Second, CSMA/ECA uses variable duration slots *i.e.* empty slots are shorter than busy slots, which is a key factor for efficient channel utilization. And third, CSMA/ECA can fairly coexist with currently deployed devices.

In [10] an enhancement to the IEEE 802.11 DCF protocol called EBA was proposed. It consists on a distributed reservation mechanism which, as our proposal, reduces the chances of collision. In EBA, two new fields are added to the MAC headers. In these fields, a station announces its next backoff value. The stations keep track of other stations' intentions to transmit and adequately select backoff values that will not lead to collisions. The additional signaling to communicate future backoff intentions is included in the packet header and can be obtained by the rest of the stations only when that frame is successfully transmitted. If a frame from a station *STA 0* is lost due to bad channel conditions or collisions, the rest of the stations remain unaware of the backoff intentions of *STA 0*.

EBA attains a notable performance improvement by avoiding collisions. In our opinion, the main disadvantage of EBA is the fact that it requires the modification of the header's fields, thus complicating the coexistence of EBA stations with legacy stations. Further, in EBA, all the stations have to keep track of channel reservation and there is an specific algorithm to update that information and select advantageous backoff values. This translates in some increased complexity and memory requirements for the network cards. Nevertheless, these additional requirements should not be the limiting factor to the adoption of EBA.

We have shown that it is possible to attain performance results similar to those obtained by EBA without requiring the modification of the MAC headers and maintaining the same computation and memory requirements of DCF. EBA explicitly communicates its next backoff value by means of new fields in the header. Our proposal, CSMA/ECA, assumes that a constant and known bakoff value is used after a successful transmission. Note that the result obtained by both approaches is equivalent: all stations listening to the channel know the backoff intentions of the station that just successfully transmitted. However, CSMA/ECA does not place additional signaling requirements and uses the same headers as legacy networks, thus guaranteeing smooth coexistence.

6 Conclusions

In this article we address the problem of collisions in CSMA networks. Our finding is that, instead of using a random backoff after all transmission attempts, it is better to use a random one after collisions and a deterministic one after successes. It reduces the chances of collisions as soon as two or more stations successfully transmit. As the system runs, it progressively converges to a collision-free operation that considerably improves the channel efficiency.

The proposed protocol outperforms CSMA/CA and, in the most typical scenarios, it even surpasses the theoretical upper bound associated with CSMA/CA networks that allow for dynamic parameter adjustment. Further, CSMA/ECA does not add any additional complexity to the implementation, it can fairly coexist with already deployed networks and it is robust against unreliable channel conditions.

References

1. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, 1999 Edition (Revised 2003).
2. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications; Amendment: Medium Access Control(MAC) Quality of Service Enhancements, 2005.
3. A. Banchs and L. Vollero. Throughput Analysis and Optimal Configuration of 802.11e EDCA. *Computer Networks*, 50(11):1749–1768, 2006.
4. J. Barcelo, B. Bellalta, C. Cano, and M. Oliver. Learning-BEB: Avoiding Collisions in WLAN. In *Eunice Summer School*, 2008. Available at http://www.dtic.upf.edu/∼jbarcelo/papers/barcelo2008lba.pdf.
5. J. Barcelo, B. Bellalta, A. Sfairopoulou, and M. Oliver. *Advances in Ad Hoc Networks*, volume 265, chapter No Ack in IEEE 802.11e Single-Hop Ad-Hoc VoIP Networks, pages 157–166. Boston:Springer.
6. G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE J. Sel. Areas Commun.*, 18(3):535–547, 2000.
7. G. Bianchi and I. Tinnirello. Kalman Filter Estimation of the Number of Competing Terminals in an IEEE 802.11 Network. In *IEEE Infocom*, volume 2, pages 844–852, 2003.
8. F. Borgonovo and L. Fratta. A New Technique for Satellite Broadcast Channel Communication. In *Symposium on Data Communications*, pages 2.1–2.4, 1977.
9. F. Cali, M. Conti, E. Gregori, and P. Aleph. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Trans. Netw.*, 8(6):785–799, 2000.
10. J. Choi, J. Yoo, S. Choi, and C. Kim. EBA: An Enhancement of the IEEE 802. 11 DCF via Distributed Reservation. *IEEE Trans. Mobile Comput.*, 4(4):378–390, 2005.
11. W. Crowther, R. Rettberg, D. Walden, S. Ornstein, and F. Heart. A System for Broadcast Communication: Reservation-ALOHA. In *Hawaii Int. Conf. Syst. Sci*, pages 596–603, 1973.

12. C. Heegard, JT Coffey, S. Gummadi, PA Murphy, R. Provencio, EJ Rossin, S. Schrum, MB Shoemake, and T.I. Inc. High Performance Wireless Ethernet. *Communications Magazine, IEEE*, 39(11):64–73, 2001.
13. DP Hole and FA Tobagi. Capacity of an IEEE 802.11b Wireless LAN Supporting VoIP. In *IEEE ICC*, volume 1, pages 196–201, 2004.
14. R. Jain. *The Art of Computer Systems Performance Analysis*. John Wiley & Sons New York, 1991.
15. T.K. Liu, JA Silvester, and A. Polydoros. Performance Evaluation of R-ALOHA in Distributed Packet Radio Networks with Hard Real-Time Communications. In *IEEE VTC*, volume 2, pages 554–558, 1995.
16. A. Lopez-Toledo, T. Vercauteren, and X. Wang. Adaptive Optimization of IEEE 802.11 DCF Based on Bayesian Estimation of the Number of Competing Terminals. *IEEE Trans. Mobile Comput.*, 5(9):1283, 2006.
17. A.S. Tanenbaum. *Computer Networks*. Prentice Hall PTR, 2002.
18. S. Tasaka. Stability and Performance of the R-ALOHA Packet Broadcast System. *IEEE Trans. Comput.*, C-32:717–726, Aug. 1983.

CSMA with Enhanced Collision Avoidance: a Performance Assessment

J. Barcelo, B. Bellalta, A. Sfairopoulou, C. Cano, M. Oliver

Universitat Pompeu Fabra

Abstract. CSMA with Enhanced Collision Avoidance (CSMA/ECA) uses a deterministic backoff after successful transmissions to significantly reduce the number of collisions. This paper assesses by means of simulations the throughput and conditional collision probability obtained from a single-hop ad-hoc network using CSMA/ECA. A comparison with the legacy CSMA/CA reveals that the proposed protocol outperforms the legacy one in all considered scenarios. Specifically, it is shown that CSMA/ECA presents advantages for both rigid and elastic flows.

1 Introduction

The proliferation of IEEE 802.11 [1] networks makes the research associated to this family of protocols particularly relevant. After its success in Wireless Local Area Networks (WLANs), the IEEE 802.11 family is growing to cover other fields of applications, such as mesh and vehicular networks.

The Medium Access Control (MAC) is the mechanism that arbitrates the sharing of the channel among competing stations. In IEEE 802.11 networks, the MAC layer employs a combination of Carrier Sense Multiple Access and Collision Avoidance (CSMA/CA). The resultant protocol is called Distributed Coordination Function (DCF), and its behaviour significantly impacts the overall performance of the network.

The optimization of the performance of the MAC layer of IEEE 802.11 has deserved large research efforts. A simple model for the DCF is presented in [5] and its maximum throughput is derived in [7]. If the number of contending stations is known, the backoff mechanism can be tuned to attain the optimal performance of CSMA/CA. However, the estimation of the number of contending stations is not a trivial task. Under the assumption of ideal channel conditions and saturated stations, advanced filtering techniques ([6,10]) can be used to accurately estimate the number of contenders.

The estimation of the numbers of contenders is even more challenging when the saturation assumption is released. Unsaturated flows join the contention to transmit only one packet and hence, in most of the occasions, they manifest themselves only at the instant they stop competing for the channel.

Another line of research goes beyond the modification of the parameters of DCF and proposes a change in the protocol. In [11], it is proposed to modify the way that contention windows grow and shrink. Another approach is not to choose the backoff randomly, as proposed in [9]: if the stations are aware of the backoff values of the other contenders, collisions can be effectively avoided.

The aforementioned approaches exhibit one or more of the following weaknesses: *a)* they rely on the saturation assumption or on the ideal channel assumption, *b)* they require a modification of the packet headers or *c)* they cannot fairly coexist with legacy DCF.

In [2] it is shown that, by using a deterministic (and equal for all stations) backoff after successful transmissions, the collisions in the WLAN are significantly reduced, and even disappear. The reason is that collisions cannot occur among those stations that successfully transmitted and chose the same deterministic backoff value. We use the name CSMA with Enhanced Collision Avoidance (CSMA/ECA) to refer to the new protocol that uses a deterministic backoff after successes.

This last solution surpasses the maximum theoretical performance of DCF while maintaining the same packet headers and guaranteeing fair coexistence with legacy networks. In [2], the focus is placed on the channel efficiency (*i.e.* the fraction of channel time devoted to successful transmissions) under saturation conditions. The present paper completes that work by assessing the performance metrics as perceived by the stations (throughput and conditional collision probability), both for elastic and rigid flows.

In this paper, CSMA/ECA has been incorporated to an IEEE 802.11 simulator in order to evaluate the validity of the new protocol in a variety of scenarios. These are the main contributions:

- First assessment of the performance parameters of CSMA/ECA as perceived by the stations: throughput and conditional collision probability. The conditional collision probability is defined as the probability that a station suffers a collision conditioned to the fact that it is attempting a transmission.
- Evidence that CSMA/ECA outperforms CSMA/CA when the traffic is offered in the form of elastic flows, rigid flows or a combination of both.
- Comparison of CSMA/ECA and CSMA/CA for both the two-way-handshake and four-way handshake variants of IEEE 802.11.

The remainder of this paper is organized as follows: Section 2 describes the features of CSMA/CA that are relevant to the paper and also briefly introduces CSMA/ECA. Section 3 describes the scenario that has been used to assess the performance of CSMA/ECA. Section 4 presents simulation results that show that CSMA/ECA outperforms CSMA/CA in all scenarios under consideration. Final conclusions are drawn in Section 5.

2 The Medium Access Control

The stations running CSMA/CA sense the channel for ongoing transmissions before sending a packet. A station is allowed to transmit only if it senses the channel idle. It may happen that two or more stations begin a transmission (almost) simultaneously and a collision occurs. In order to reduce the chances of collision, the channel time is divided in slots and the transmissions are deferred a random number of slots.

The backoff values (B) are chosen from a contention window:

$$B \sim \mathcal{U}[0, \min(CW_{min} \cdot 2^{a}, CW_{max}) - 1], \tag{1}$$

where $\mathcal{U}$ represents the uniform distribution. CW_{min} and CW_{max} are the minimum and maximum contention windows, respectively. The number of transmission attempts for the current packet is denoted as a (It equals 0 for the first transmission attempt).

The contention window uses a minimum value CW_{min} for the first transmission attempt and doubles after each failed transmission attempt, up to a maximum value of CW_{max}. This binary exponential growth reduces the number of transmission attempts in a congested scenario.

It is common to use a simple two-way handshake mechanism in which the data is transmitted in one packet and acknowledged by the receiver in a second packet. This modality is called Basic Access (BA).

There is an optional four-way-handshake floor-reservation mechanism to minimize the channel time waste due to collisions and prevent the hidden terminal impairment [12]. Request-To-Send and Clear-To-Send (RTS/CTS) packets are used before the actual data transmission in order to reserve the channel. When RTS/CTS is in use, collisions can only occur among control (short) packets, thus the amount of channel time wasted in collisions is reduced. However, the additional control packets penalize the overall efficiency of the network.

2.1 CSMA with Enhanced Collision Avoidance

CSMA with Enhanced Collision Avoidance (CSMA/ECA), behaves exactly the same as the CSMA/CA protocol with the exception that a deterministic backoff is chosen after successful transmissions. To guarantee a fair coexistence with legacy CSMA/CA stations, the value of the deterministic backoff has to be:

$$V = \lceil E\left[\mathcal{U}[0, CW_{min} - 1]\right]\rceil = \lceil (CW_{min} - 1)/2 \rceil, \tag{2}$$

where $\lceil \cdot \rceil$ is the ceiling operator and $E[\cdot]$ is the expectation operator. The deterministic backoff after successes is a key parameter of the system, since it is also the maximum number of stations that can be accommodated in the collision-free mode of operation of CSMA/ECA. This parameter can also be adjusted to attain prioritization properties or to accommodate more contenders. More details on the adjustment of V can be found in [3].

For the first packet transmission and for transmission attempts following an unsuccessful transmission, the random backoff B as defined in (1) is used. The backoff behaviour of CSMA/ECA can be summarized as:

$$\text{backoff} = \begin{cases} V \text{ after a successful transmission;} \\ B \text{ otherwise.} \end{cases} \tag{3}$$

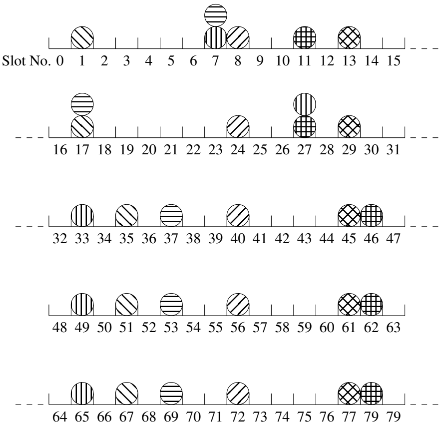


Fig. 1. A ball represents a transmission attempt in a given slot. Different filling patterns have been used to differentiate the transmissions of different stations. In CSMA/ECA the stations that successfully transmit use a deterministic backoff value.

Fig. 1 shows an example in which six CSMA/ECA saturated stations contend for the channel. The channel time is divided in numbered slots and the transmissions are represented as balls on that slots. The balls are filled with different patterns, each pattern corresponding to a different station.

If there is only one ball in a slot, it is a successful slot. In contrast, if there are two balls in the same slot, a collision has occurred. The two stations involved in the collision will randomly choose a backoff value. It can be observed that there is a collision in slot number 7. The two stations choose backoff values 10 and 20, leading to two new collisions in slots 17 and 27, respectively.

A station that successfully transmits, backoffs for $V = 16$ slots. As an example, the station that successfully transmits in slot number 13 in Fig. 1, also transmits in slots 29, 45, 61 and 77. It is useful to define the columns in the figure. A column is a set of slots whose numbers are equal modulo V (*e.g.* slots 0, 16, 32, 48 and 64 belong to the same column). Then, it can be observed that those stations that successfully transmit, use the same column in their next transmission attempt.

After all stations have successfully transmitted, in the slots numbered from 32 to 47, the behaviour of the system becomes deterministic and collisions disappear.

3 Evaluation Scenario

A single-hop ad-hoc network is considered where each station transmits one traffic flow to a randomly selected neighbor. It is assumed that all the stations are in the transmission range of one another, thus there is no hidden terminal effect [12].

The packet length is $L_e = 12000$ bits for elastic flows and $L_r = 1000$ bits for rigid flows. Each MAC queue can hold up to 50 packets and the BA two-way handshake is used unless otherwise stated. The MAC parameters are taken from the IEEE 802.11b specification, and the physical data rate under consideration is 2Mbps. The constant backoff after successes is $V = 16$.

In order to fully validate a MAC protocol, it is required to show that it delivers acceptable performance for both elastic and rigid flows. In [4], a comprehensive study of the coexistence of elastic and rigid flows in IEEE 802.11 networks is presented. The simulator used in that paper has been enhanced to support also CSMA/ECA and has been used to obtain the results which are presented in the following section. It is based on the Component Oriented Simulation Toolkit (COST) [8].

3.1 Rigid and Elastic Flows

In a simplification of the myriad of traffic patterns that can be found in a wireless network, we consider only two kinds of flows: elastic and rigid.

Elastic flows are characterized by the fact that they have a clear tendency to consume all the bandwidth that is available in the network. They are typically associated to the use of the Transport Control Protocol (TCP) at the transport layer. At the MAC layer, they manifest as saturated stations. Web traffic, email, and peer-to-peer file interchange are good examples of elastic flows.

Rigid flows consume a fixed amount of bandwidth and are often encapsulated by the User Datagram Protocol (UDP) at the transport layer. During normal (uncongested) network operation, rigid flows do not saturate the station. On the contrary, the MAC queue remains empty for most of the time. A single packet is periodically received from the upper layer and, after the packet is serviced, the queue remains empty until a new packet arrives. Nevertheless, if the network is highly loaded and cannot transmit all the packets arriving from the upper layers, the MAC queues quickly build up and packet loss occurs due to queue overflow. If that is the case, we say that the network is congested. Voice over IP (VoIP) is an example of a service that uses rigid flows.

Elastic and rigid flows have different requirements regarding the MAC layer. When elastic flows are considered, the focus is placed on maximizing the throughput. In contrast, the goal in a network that forwards rigid flows is to prevent congestion and the associated packet loss.

4 Performance Results

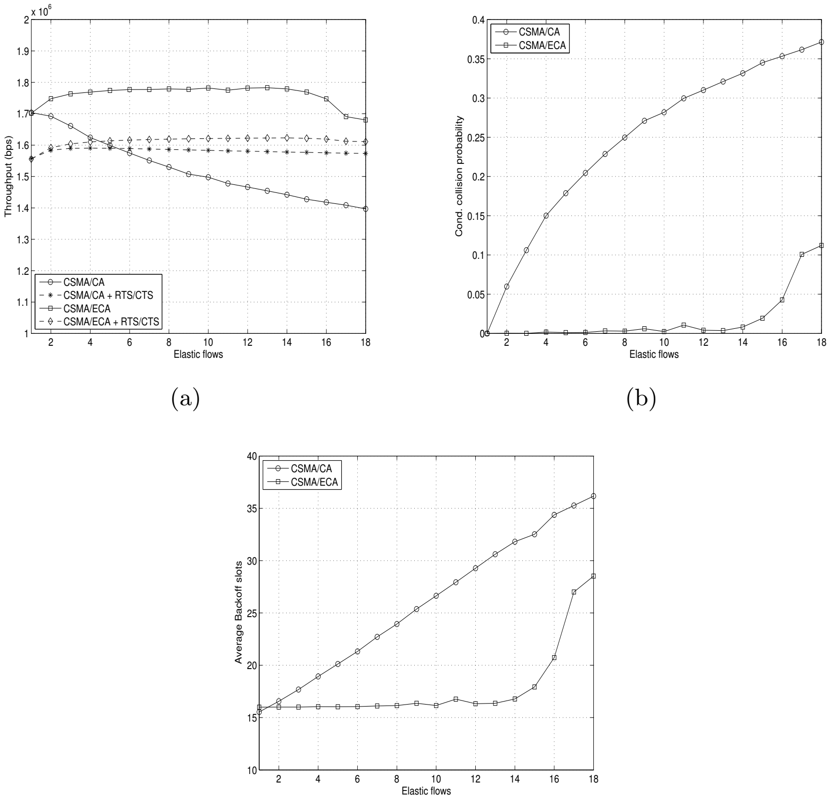
This section presents a simulation assessment of the performance of CSMA/ECA in scenarios with elastic flows, rigid flows and a combination of both.

4.1 Results for Elastic Flows

In Fig. 2, the throughput, conditional collision probability and expected backoff are plotted for an increasing number of elastic flows. The figure compares the performance of the proposed CSMA/ECA mechanism and the legacy CSMA/CA. It can be observed that CSMA/ECA maintains a constant (maximum) throughput as the number of contending stations increases. In contrast, the aggregated throughput of CSMA/CA is penalized when the number of contenders increases.

Note that the throughput when there is one single flow is the same in CSMA/CA and CSMA/ECA. The advantage of CSMA/ECA is that collisions cannot occur between two stations that successfully transmitted. This advantage cannot manifest when there is only one saturated station.

The higher throughput achieved by CSMA/ECA is a consequence of the lower number of collisions and the lower average backoff value. Since collisions are effectively suppressed, the backoff value is always $V = 16$ (See Fig. 2c). In the plots, it can be observed that there is a turning point when the number of active stations is equal to the deterministic backoff value V. At this point collisions can no longer



(a)

(b)

(c)

Fig. 2. Performance results for elastic flows.

be avoided and the performance of CSMA/ECA is degraded. If the number of flows continued to increase, the curves for CSMA/ECA would tend asymptotically to the ones obtained for CSMA/CA.

4.2 Results for Rigid Flows

In Fig. 3 the aggregated throughput and conditional collision probability for rigid flows are plotted for both CSMA/ECA and CSMA/CA. The simulations are performed for 50Kbps and 100Kbps flows.

Throughput plots for rigid flows (Figs. 3a and 3c) are read as follows: while the throughput grows linearly with the number of flows, it means that the network can absorb the traffic offered by the stations. As soon as the throughput deviates from the linear growth, it is a symptom that congestion has appeared and packets are lost in the buffer queues.

While the number of flows is small and the contention is low, the MAC queues remain empty for most of the time. Thus, after a successful transmission, there is not a second packet to transmit, and the CSMA/ECA rule that states that a deterministic backoff is used after successes never applies.

For this reason, it can be observed that, when the network is lightly loaded, the performance metrics delivered by CSMA/ECA are exactly the same as the ones that can be obtained from CSMA/CA. Nevertheless this situation changes when the load increases and the network approaches congestion.

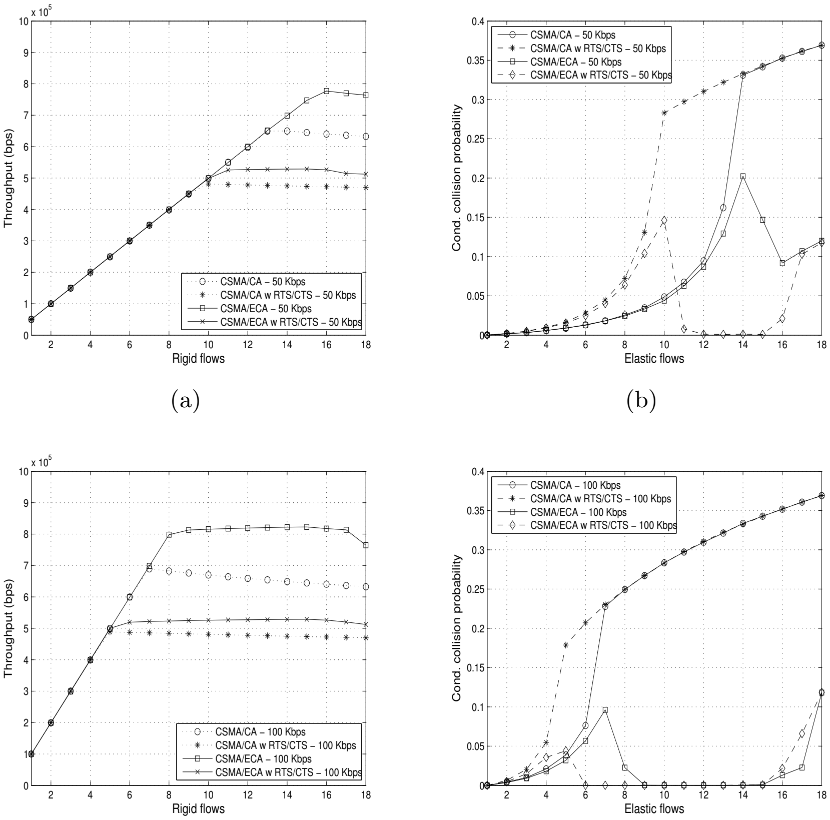
At this point, the MAC queues build up and, as the probability to find more than one packet in the queue increases, the CSMA/ECA rule for deterministic backoff after successes applies and hence the collisions are reduced or even suppressed.

The performance boost obtained by CSMA/ECA thanks to the suppression of collisions allows the network to satisfactorily support more rigid flows than CSMA/CA.

4.3 Results for the Coexistence of Elastic and Rigid Flows

In this scenario, a single elastic flow coexists with an increasing number of rigid flows. Fig. 4a depicts the aggregated throughput obtained by the rigid flows while Fig. 4b is the throughput of the elastic flow. These plots show that the advantages of CSMA/ECA for both elastic and rigid flows are also apparent in mixed scenarios. The two kinds of traffic benefit from the fact that CSMA/ECA is used.

Fig. 4c shows the conditional collision probability as perceived by the rigid flows. CSMA/ECA significantly reduces the chances of collision for both 50Kbps and 100Kbps flows. As the number of simultaneous flows increases, the packet service time also increases. As a consequence, the probability that a station holds multiple packets in its queue is higher. When there is more than one packet in the queue, CSMA/ECA actuates to lower the collision probability.



(a) (b)

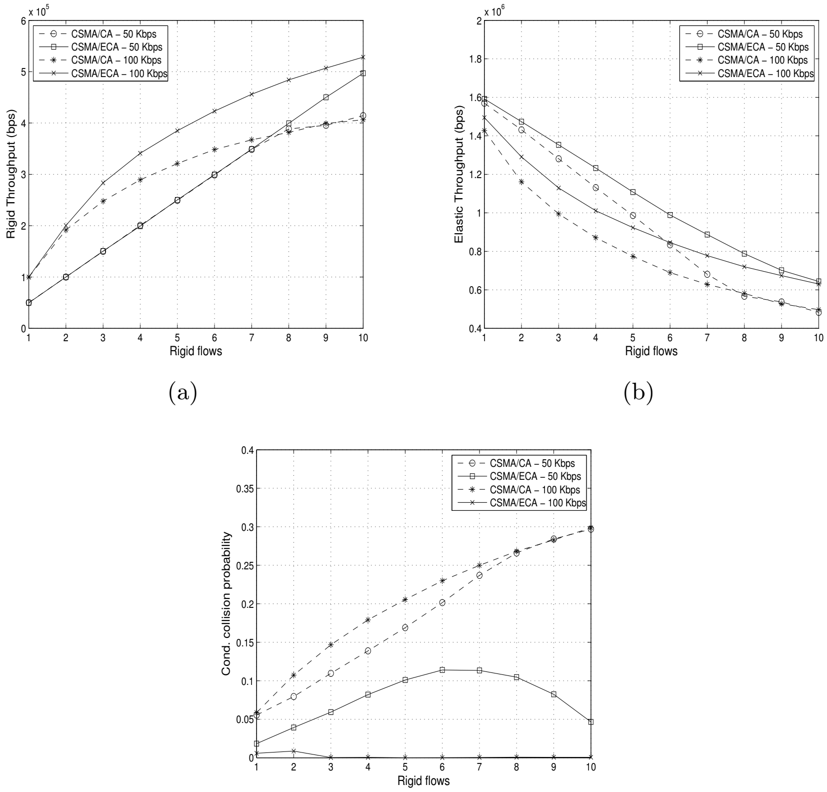
(c) (d)

Fig. 3. Performance results for rigid flows

Note that, since CSMA/ECA as presented in this paper does not support traffic differentiation, the presence of elastic flows is detrimental for the performance of the rigid ones. Although it is possible to combine CSMA/ECA with prioritization mechanisms, it is out of the scope of the present paper.

4.4 The impact of RTS/CTS on the performance

RTS/CTS minimizes the time wasted due to collisions, but increases the channel access overhead because of the additional control packets. In Fig. 2a, the impact of the RTS/CTS mechanism on the performance of elastic flows can be observed. CSMA/CA + RTS/CTS out-



(a) (b)

(c)

Fig. 4. Performance results for one elastic flow coexisting with multiple rigid flows

performs CSMA/CA + BA when the number of contenders is large. However, when CSMA/ECA is used, the four-way-handshake mechanism offers little advantage. Since the collisions are already prevented by the enhanced collision avoidance mechanism, RTS/CTS penalizes the throughput because of the associated overhead.

From the results, it is clear that the best performance is obtained by CSMA/ECA combined with BA. Nevertheless, if RTS/CTS is to be used for reasons out of the scope of this paper (*e.g.* to prevent the hidden terminal effect), CSMA/ECA still presents a performance advantage when compared with CSMA/CA.

The effect of RTS/CTS on rigid flows is depicted in Fig. 3. Because of the four-way-handshake, the time required to transmit each packet is significantly increased. As a consequence, the number of packets that can traverse the network in a given time interval is reduced. Thus, a network using the RTS/CTS mechanism can support a lower number of rigid flows than a network using BA. A final observation is that CSMA/ECA also outperforms CSMA/CA when RTS/CTS is used.

5 Conclusion

This article assesses the performance of CSMA/ECA in single-hop ad-hoc networks. CSMA/ECA is a modification of CSMA/CA that uses deterministic backoff values after successful transmissions, which reduces the chances of collision. In order to validate the goodness of CSMA/ECA, it is necessary to show that it delivers higher performance for the most common kinds of traffic: elastic flows and rigid flows. Throughout the article, the performance metrics of CSMA/ECA have been compared with those delivered by CSMA/CA

Simulation has been used to evidence that CSMA/ECA delivers higher throughput when elastic flows are considered. Regarding rigid flows, CSMA/ECA allows for a larger number of simultaneous flows before reaching the congestion condition. In a mixed scenario that includes both rigid and elastic flows, CSMA/ECA still attains higher throughput for the elastic flows and increased protection for the rigid flows. In summary, CSMA/ECA outperforms CSMA/CA in all considered scenarios.

Acknowledgment

The authors are thankful to the anonymous reviewers for their helpful comments and to the organizers for their kind invitation.

References

1. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, 1999 Edition (Revised 2007).
2. J. Barcelo, B. Bellalta, C. Cano, and M. Oliver. Learning-BEB: Avoiding Collisions in WLAN. In *Eunice Summer School*, 2008.

3. J. Barcelo, B. Bellalta, C. Cano, A. Sfairopoulou, M. Oliver, and J. Zuidweg. Traffic Prioritization for Carrience Sense Multiple Access with Enhanced Collision Avoidance. In *IEEE ICC (MACOM)*, 2009.
4. B. Bellalta, M. Oliver, M. Meo, and M. Guerrero. A Simple Model of the IEEE 802.11 MAC Protocol with Heterogeneous Traffic Flows. In *EUROCON*, volume 2, 2005.
5. G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE J. Sel. Areas Commun.*, 18(3):535–547, 2000.
6. G. Bianchi and I. Tinnirello. Kalman Filter Estimation of the Number of Competing Terminals in an IEEE 802.11 Network. In *IEEE Infocom*, volume 2, pages 844–852, 2003.
7. F. Cali, M. Conti, E. Gregori, and P. Aleph. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Trans. Netw.*, 8(6):785–799, 2000.
8. G. Chen and B.K. Szymanski. COST: Component-oriented simulation toolkit. In *Winter Simulation Conference*, 2002.
9. J. Choi, J. Yoo, S. Choi, and C. Kim. EBA: An Enhancement of the IEEE 802. 11 DCF via Distributed Reservation. *IEEE Trans. Mobile Comput.*, 4(4):378–390, 2005.
10. A. Lopez-Toledo, T. Vercauteren, and X. Wang. Adaptive Optimization of IEEE 802.11 DCF Based on Bayesian Estimation of the Number of Competing Terminals. *IEEE Trans. Mobile Comput.*, 5(9):1283, 2006.
11. C. Wang, B. Li, and L. Li. A New Collision Resolution Mechanism to Enhance the Performance of IEEE 802.11 DCF. *IEEE Trans. Veh. Commun.*, 53(4):1235, 2004.
12. A. Zahedi and K. Pahlavan. Natural hidden terminal and the performance of the wireless LANs. volume 2, 1997.

Carrier Sense Multiple Access with Enhanced Collision Avoidance: a Performance Analysis

J. Barcelo, B. Bellalta, C. Cano, A. Sfairopoulou, M. Oliver

Universitat Pompeu Fabra

Abstract. Carrier Sense Multiple Access with Enhanced Collision Avoidance (CSMA/ECA) is a recently proposed modification to the well-known CSMA/CA protocol. By using a deterministic backoff after successful transmissions, the number of collisions decreases. This article presents a model that captures the behaviour of CSMA/ECA in both saturated and non-saturated scenarios. The results, which are validated by simulations, show that CSMA/ECA effectively prevents collisions and, therefore, it can deliver a higher throughput than CSMA/CA.

1 Introduction

Carrier Sense Multiple Access with Enhanced Collision Avoidance (CSMA/ECA) [4] is a novel Medium Access Control (MAC) mechanism for Wireless Local Area Networks (WLANs). The differentiating property of CSMA/ECA when compared to CSMA/CA is the fact that, in the former, a deterministic backoff value is used after successful transmissions. The backoff is always selected randomly in CSMA/CA.

The backoff value is used to defer transmission attempts. The idea of using a deterministic backoff value after successes was already introduced in [2]. That paper presents expressions for the performance of CSMA/ECA in saturation conditions. Saturation implies that all the stations have always a packet ready to transmit, and greatly simplifies the analysis.

The behaviour of CSMA/ECA substantially varies with the load conditions. In lightly loaded scenarios, the MAC queues are empty and, therefore, random backoffs are used just as in CSMA/CA. Conversely, in highly loaded scenarios, the queues are full. The packets arrive to the queue's Head-Of-Line (HOL) after the successful transmission of the previous packet and thus a deterministic backoff is used for a packet's first transmission attempt. The highly loaded scenario can be analyzed under the saturation assumption.

Regarding to the protocol evaluation, there are different ways to assess the performance of MAC protocols. In this article we focus on the throughput (S) and the channel status probabilities, which are defined as the fraction of slots that are empty (P_e), contain a successful transmission (P_s) or contain a collision (P_c).

After this introductory section, the remainder of the article is organized as follows. Sec. 2 reviews previous approaches in solving the problem of collisions in WLANs. Then, Sec. 3 briefly describes CSMA/ECA, comments on previous results and defines the scope of the present article. The main contribution of the article, which is a comprehensive model for CSMA/ECA is presented in Sec. 4. The results derived from the model, which are validated by simulations, provide insights on the behaviour of the protocol and are discussed in Sec. 5. Finally, some concluding remarks are provided in Sec. 6.

2 Related Work on Collision Prevention

In wired networks, a collision can be detected when the voltage on the wire surpasses a certain threshold. This is not the case in wireless networks. Since wireless equipment uses the same antenna to transmit and receive, it cannot sense the channel while transmitting. The immediate consequence is that collisions cannot be detected until the transmission has finished. The lack of positive acknowledgement after a transmission is an indication that a collision might have occurred.

Since collisions are long, they waste a substantial amount of channel time and represent a serious limitation to WLANs performance. In the following, we present different approaches to address this problem.

A four-way-handshake floor-acquisition mechanism can be used to prevent collisions among long data packets. Specifically, Request-To-Send and Clear-To-Send (RTS/CTS) control packets are sent before the data packet transmission. Although collisions are still possible, these collisions occur among short control packets, thus minimizing the impact of collisions on the network performance. The RTS/CTS mechanism offers the additional value of preventing the hidden terminal problem. Nevertheless, the use of RTS/CTS has also a negative side: the two additional control packets required for a data packet transmission imply an increased overhead that reduces the efficiency of the network. Simulation results in [4] show that the four-way-handshake provides better performance than a two-way-handshake when the number of contenders is high.

Another approach to limit the collision probability is to reduce the transmission probability (τ). A negative side effect is an increased number of empty slots. Actually, the optimal (τ) can be computed if the number of contenders [13] and the packet size are known. In [2, 8] the fundamental limits of the performance of CSMA with an optimized transmission probability are explored.

To attain a performance higher than the above mentioned limit, the selection of the transmission slot can no longer be completely random. Several studies present solutions based on that direction. A mechanism called Blackburst [14] introduces a variable length preamble before the actual data transmission. If different stations are willing to transmit, the station with the longer preamble wins access to the channel. This is not very efficient, as the preambles represent a non-negligible overhead.

The announcement of the backoff values is proposed in [10], thereby allowing the different stations to choose different backoff values and prevent collisions. The only caveat is that the MAC headers have to be modified and the stations have to keep track of the other stations' backoff values.

Reservation is another alternative to prevent collisions. In Reservation-Aloha [11], PRMA [12] and DBRA [9], different MAC protocols that use reservation are suggested. One of the shortcomings of using reservation is that the coexistence with CSMA/CA in the same network is unfeasible, because CSMA/CA is unaware of any reservation.

3 Motivation, Scope and Previous Work on CSMA/ECA

The familiarity of the readership with the IEEE 802.11 [1] protocol is assumed. A description of terms such as Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA), Binary Exponential

Backoff (BEB), minimal and maximum congestion window (CW_{min} and CW_{max}), maximum backoff stage (m) and retransmission limit (R) can be found in the literature. All those concepts are defined in [7], which presents comprehensive description of the MAC protocol of IEEE 802.11 and a model valid for saturated scenarios.

The only new parameter that is introduced in CSMA/ECA with respect to CSMA/CA is the deterministic backoff value after successful transmissions (V). As described in [4], to guarantee the smooth coexistence in the same network of CSMA/ECA and CSMA/CA protocols, this parameter is chosen to be equal to the expected backoff value for the first transmission attempt in CSMA/CA:

$$V = \lceil E\left[\mathcal{U}[0, CW_{min} - 1]\right]\rceil = \lceil (CW_{min} - 1)/2\rceil, \tag{1}$$

where $\lceil\cdot\rceil$ is the ceiling operator and $E[\cdot]$ is the expectation operator.

Although several MAC protocols for WLAN have been proposed, the challenge of finding a protocol that could be a successor of CSMA/CA remains open. A requirement for the successor is backward compatibility with the legacy standard, since a large number of networks and devices have been deployed and simultaneously replacing or upgrading all of them is unfeasible.

CSMA/ECA is backward compatible with CSMA/CA since $V \in [0, CW_{min}]$ and therefore it is an acceptable backoff value for the first transmission attempt in CSMA/CA. The remainder of behaviour of CSMA/ECA is identical to CSMA/CA. Hence, the presence of CSMA/ECA stations will not perturb the normal operation of a CSMA/CA network.

In [2] it is shown that there is a fundamental limit on the efficiency of completely random access protocols, in which the transmission slot is chosen without using any prior information. Then, it is explained that CSMA/ECA can overcome that limit by using a random behaviour after failures (to trigger a change) and a deterministic behaviour after successes (to stabilize the system).

In [4], simulations are used to assess the performance of CSMA/ECA in saturated, non-saturated and hybrid (a combination of saturated and non-saturated) scenarios. CSMA/ECA is shown to perform equal or better than CSMA/CA in all the considered scenarios. Specifically, the two protocols deliver the same throughput in the scenarios in which the network is able to absorb all the offered traffic. However, when the traffic load overwhelms the network, CSMA/ECA performs better than CSMA/CA. Traffic prioritization in CSMA/ECA is addressed in [3].

Other issues, such as simulations of the coexistence of CSMA/CA and CSMA/ECA stations, the behaviour of the protocol in lossy channels or a model to study the convergence time before reaching the collision-free operation are out of the scope of the present paper. They are the subject of a separate article.

The present article is focused on presenting a comprehensive model that captures the stationary behaviour of CSMA/ECA for both lightly and heavily loaded scenarios. The model is also accurate in predicting the transition from non-saturation to saturation.

In a first approach, the performance of CSMA/ECA can be described as being equal to the performance of CSMA/CA for lightly loaded scenarios. The effectiveness of CSMA/CA in non-saturated scenarios has already been studied in the literature [5], and those models also apply to CSMA/ECA when the MAC queues are empty. For highly loaded and stationary scenarios (saturation), the analysis in [2] is valid. We briefly revisit that analysis for the sake of completeness. It is valid when the number of active contenders[1] (ς) is not greater than the deterministic backoff value used after successes (V).

Thanks to the fact that a deterministic backoff is used, two stations that successfully transmitted in their last transmission attempt cannot collide among them. This leads to a progressive reduction of the number of collisions and, after all stations have consecutively successfully transmitted, the operation of the system is collision-free and deterministic (See Fig. 1). Notice that the stationary behaviour of the system is cyclic and each cycle comprehends V slots.

The fraction of successful slots during the collision-free operation is:

$$P_s^{(CF)} = \frac{\varsigma}{V}, \tag{2}$$

---

[1] Note that the number of active contenders is different from the number of terminals registered to the network. As an example, in a network of 40 nodes with an activity rate of 10%, the expected number of active contenders is 4.

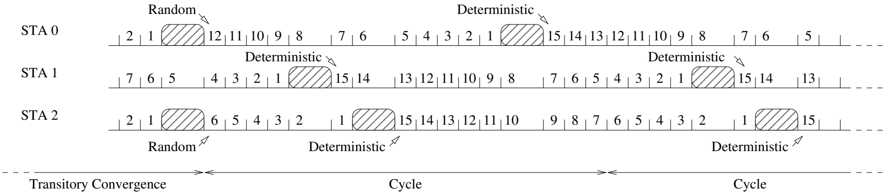


Fig. 1. Three wireless stations contend for the channel using CSMA/ECA. The shaded boxes represent transmissions and the numbers are the backoff counters of the stations. They use a random backoff after collisions and a deterministic one after successes.

where the superscript *CF* indicates *collision-free* operation and the subscript s stands for *successful*. The rest of slots are empty:

$$P_e^{(CF)} = \frac{V - \varsigma}{V}. \tag{3}$$

Since the system is collision-free, the fraction of slots containing collisions is zero $P_c^{(CF)} = 0$. The values of $P_s^{(CF)}$ and $P_e^{(CF)}$ are linear with the number of contending stations. This would be clearly observed in the results presented in Sec. 5.

4 A Model for CSMA/ECA

The model presented in this section is based on the following assumptions. Only one-hop communications are considered, and all the stations can hear each others' transmissions. This implies that there are no hidden or exposed terminals. Moreover, the channel does not introduce any errors. The model is only valid for a number of contenders up to V, and when all the stations present the same traffic patterns (load, packet size) and use the same data rate.

Since the behaviour of CSMA/ECA differs depending on the queue occupation (ρ), it is necessary to develop a model that includes the queue occupation in the analysis. The model presented here relies on the previous work in [5], which is a model for CSMA/CA that takes into account the queues.

The idea behind the model that will be presented in the following is taking the analysis of [5] for the unsaturated case and the analysis in [2] for the saturated case. Then, the equations of the former model are weighted by $(1-\rho)$ and the equations of the latter are weighted by ρ.

The traffic that is offered by the network layer of one station to the MAC layer is

$$\nu = \lambda X. \tag{4}$$

where λ is the packet arrival rate and X is the packet service time. The MAC queue is modelled as a M/M/1/K queue, which is a single-server queue with Poisson arrival process and exponentially distributed service time, with finite queue length K. Although the actual packet service time in WLANs is not exponentially distributed, the assumption is made for the sake of tractability.

The probability of a packet being discarded after finding the queue full is:

$$p_B = \frac{(1-\nu)\nu^K}{1-\nu^{K+1}}. \tag{5}$$

The fraction of the packets that are not discarded due to queue overflow are served:

$$\rho = \nu\left(1 - p_B\right). \tag{6}$$

Note that ρ can also been interpreted as the probability that a tagged station has one or more packets to transmit.

If a packet transmission fails, the packet is retransmitted until it is correctly decoded and acknowledged by the receiver. Under the ideal channel assumption, a transmission fails only if a collisions happens. Let p_{cc} be the probability of collision conditioned to the fact that a tagged station is attempting transmission. Then, the average number of transmission attempts required to successfully transmit one packet can be approximated [6] by:

$$A \approx \frac{1}{1-p_{cc}}. \tag{7}$$

And the average service time can be computed as:

$$X = A \cdot B \cdot \omega + (A-1)\,T_c + T_s, \tag{8}$$

where B is the average number of backoff slots between transmission attempts and ω is the average duration of a waiting slot.

T_s and T_c are the average duration of successful and collision slots, respectively, and are a function of the packet length [7].

In CSMA/CA, the conditional collision probability can be computed as a function of the transmission probability τ. Assuming that there are ς stations contending for the channel, the probability that $\varsigma - 1$ stations remain silent while a tagged station is transmitting is $(1-\tau)^{\varsigma-1}$. Thus, the conditional collision probability is $1-(1-\tau)^{\varsigma-1}$.

Nevertheless, in CSMA/ECA the backoff is selected randomly only when the packet finds the MAC queue empty or after a failed transmission attempt. Taking into account that two stations that use a deterministic backoff cannot collide among them, the expression for the conditional collision probability in CSMA/ECA is approximated by:

$$p_{cc} \approx \rho\left(1-(1-\tau)^{(1-\rho)\varsigma}\right) + (1-\rho)\left(1-(1-\tau)^{\varsigma-1}\right). \tag{9}$$

Notice that $p_{cc} = 0$ for $\rho = 1$.

The separation between two transmission attempts is the expected number of the backoff (B) plus all those slots that the queue remains empty. To compute the number of slots that the queue remains empty, the time that elapses from the end of the transmission until a new packet arrives to the queue ($\frac{(1-\rho)X}{\rho}$) is divided by the average duration of a waiting slot (ω).

The transmission probability is computed as the inverse of the separation between transmission attempts:

$$\tau = \frac{1}{B + \frac{(1-\rho)X}{\rho\omega}} \tag{10}$$

In CSMA/CA, the expected number of backoff slots B_{CA} is $\frac{CW_{min}-1}{2}$ if the transmission succeeds at the first attempt, which occurs with probability $1 - p_{cc}$. The contention window doubles for successive transmission attempts, up to a maximum value of $2^m CW_{min}$, where m is the maximum backoff stage. Thus the expected number of backoff slots is [15]:

$$\begin{aligned} B_{CA} = {} & (1-p_{cc})\frac{CW_{min}-1}{2} + \\ & + p_{cc}(1-p_{cc})\frac{2\cdot CW_{min}-1}{2} + \dots \\ & + p_{cc}^{m}(1-p_{cc})\frac{2^m\cdot CW_{min}-1}{2} + \\ & + p_{cc}^{m+1}\frac{2^{m+1}CW_{min}-1}{2} = \\ = {} & \frac{1-p_{cc}-p_{cc}(2p_{cc})^m}{1-2p_{cc}}\frac{CW_{min}}{2} - \frac{1}{2} \end{aligned} \tag{11}$$

In CSMA/ECA, the expected number of backoff slots B is computed as:

$$B = (1-\rho)B_{CA} + \rho V \tag{12}$$

Finally, it is necessary to compute the average duration of the waiting slots:

$$\omega = p_e T_e + p_s T_s + p_c T_c. \tag{13}$$

T_e is the duration of an empty slot and it is defined in the standard. p_e, p_s and p_c are the probabilities that a station observes an empty, successful and collision slot while it is decrementing its backoff.

The probability of observing a successful slot is the probability that there is one transmission attempt that does not result in collision. The station that is observing the channel does not transmit. Therefore, a success occurs if one of the $\varsigma - 1$ remaining stations transmits while the other $\varsigma - 2$ remain silent. In the following expression, ρ is used as a weighting factor to take into account that those stations that are saturated cannot collide among them:

$$\begin{aligned} p_s = {} & \rho(\varsigma-1)\tau(1-\tau)^{(1-\rho)(\varsigma-2)} + \\ & (1-\rho)(\varsigma-1)\tau(1-\tau)^{\varsigma-2}. \end{aligned} \tag{14}$$

The collision probability is approximated to the probability that two stations simultaneously transmit. As collisions cannot occur in the deterministic collision-free mode of operation, a factor $(1-\rho)$ is included:

$$p_c \approx (1-\rho)\frac{(\varsigma-1)(\varsigma-2)}{2}\tau^2(1-\tau)^{\varsigma-2}. \tag{15}$$

And the probability of observing an empty slot can be obtained as:

$$p_e = 1 - p_s - p_c. \tag{16}$$

These last equations complete the model, which is solved using fixed-point iteration. However, the metrics of interest in the present article are the probabilities of empty, successful and collision slot, and also the throughput. Notice that the probabilities for the channel status are similar to the probability of observing a channel in a given status while backing off. The notation P_e, P_s and P_c is used for the former, while p_e, p_s and p_c are used for the latter.

$$\begin{aligned} P_s = {} & \rho\varsigma\tau(1-\tau)^{(1-\rho)(\varsigma-1)} + \\ & (1-\rho)\varsigma\tau(1-\tau)^{\varsigma-1}. \end{aligned} \tag{17}$$

$$P_c \approx (1-\rho)\frac{\varsigma(\varsigma-1)}{2}\tau^2(1-\tau)^{\varsigma-2}. \tag{18}$$

$$P_e = 1 - P_s - P_c. \tag{19}$$

Finally, considering that each packet contains L bits of payload, the system throughput of the network is computed as

$$S = \frac{\varsigma\rho L}{X} \tag{20}$$

5 Results and Discussion

This section presents two sets of results. First, the model from the previous section is validated by means of simulations. Conclusions regarding the behaviour of CSMA/ECA are derived from the results. After that, the simulation results of CSMA/ECA are compared to those obtained from CSMA/CA, in order to highlight the advantages of the former.

The simulator[2] only implements the MAC protocol and it is oblivious to upper and lower layer functionality. The channel does not

---

[2] The interested reader is encouraged to contact the authors to obtain the source code and further explanations.

introduce errors and there is no capture effect or hidden/exposed terminal (all the nodes are in carrier sensing range of one another). The MAC parameter values are those of IEEE 802.11b , using a 2 Mbps data rate. Each packet carries L=1000 bits of useful data, and all the stations present the same traffic pattern, with an exponentially distributed packet interarrival time. A different simulation is performed by each number of simultaneous contenders, ranging from 2 to 16. Each simulation lasts for 500s, and only the last 100s are used for gathering data, to ensure stationary operation of the system.

Some of the tests are performed for both 80 Kbps flows and 130 Kbps flows. Since the packet size is fixed, a larger flow translate to a higher number of packets per second. The flows are always rigid (i.e. constant-bit-rate or CBR) and do not incorporate any self-regulating mechanisms to reduce their sending rate in the case of congestion. Therefore, when the network cannot absorb all the generated traffic, packet loss occurs due to buffer overflow.

5.1 Model Validation

The model presented in the previous section is used to compare the probabilities of successful, empty and collision slot (P_s, P_e and P_c) and the results are compared to those obtained in simulations. Fig. 2 and Fig. 3 show the results for 80 Kbps and 130 Kbps, respectively. In each plot, two modes of operation can be differentiated. When the number of flows is low, there is no saturation. All the generated data successfully transmitted through the network and the MAC queues are empty. For a low number of flows, almost all the slots are empty. As the number of flows increase, the number of empty slots quickly diminish, since a successful slot is much larger than an empty slot.

The rightmost side of the figures is characterized by the nodes being saturated. As the network can no longer absorb all the generated traffic , the MAC queues build up and eventually overflow. Since $\rho = 1$, deterministic backoffs are used, and the curves are linear. The expressions presented in Sec. 3 are valid for saturation conditions.

As expected, the network nodes saturates for a lower number of flows when the flows are larger. The transition occurs for a number of simultaneous flows around 11 for the 80 Kbps case and around 7 for the 130 Kbps case.

The main result is the observation that CSMA/ECA prevents collisions from occurring when the network is saturated.

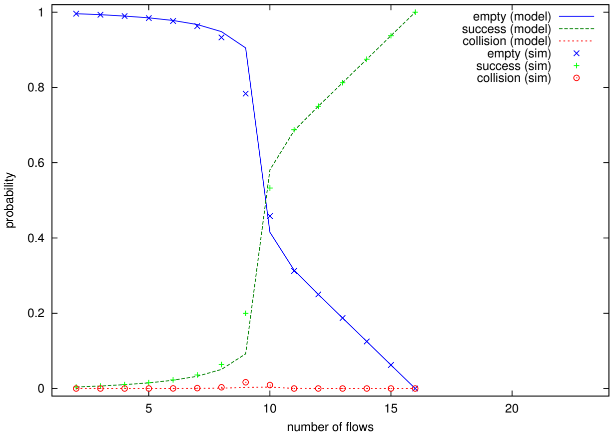


Fig. 2. The fraction of empty, successful and collision slots when each station transmits a flow of 80 Kbps using CSMA/ECA.

In Figs. 4 and 5, the throughput for the 80 Kbps and 130 Kbps case is plotted. The system throughput increases linearly with the number of flows in the unsaturated region. On the contrary, in the saturated region, the system throughput is flat and almost independent of the number of competing terminals. The throughput remains almost flat even though, for a larger number of competitors, the fraction of successful slots is greater and the fraction of empty slots is lower. The explanation is that the empty slots are orders of magnitude shorter than the successful ones and, therefore, they have a limited impact on the throughput.

5.2 The Advantage of CSMA/ECA

This subsection compares the performance of CSMA/CA and CSMA/ECA when 80 Kbps flows are considered. CSMA/ECA presents a clear advantage that can be observed in Fig. 6, which represents the probabilities P_e, P_s and P_c. The round markers are for CSMA/CA and the

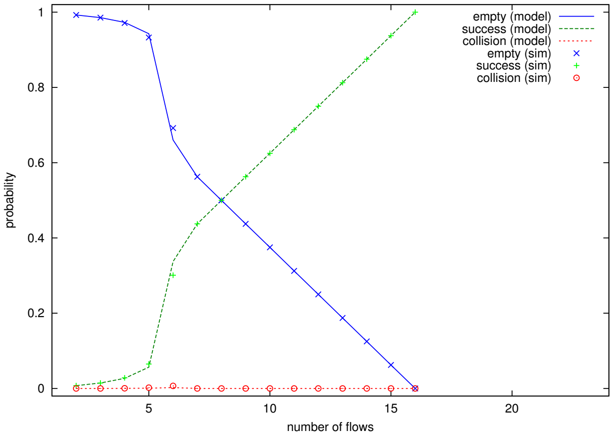


Fig. 3. The fraction of empty, successful and collision slots when each station transmits a flow of 130 Kbps using CSMA/ECA.

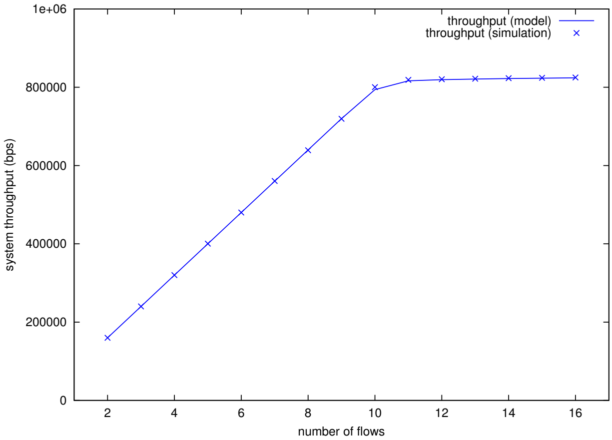


Fig. 4. The attained system throughput for an increasing number of simultaneous 80 Kbps flows using CSMA/ECA.

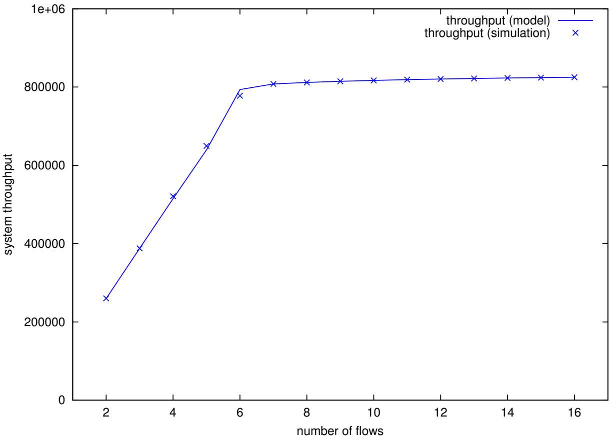


Fig. 5. The attained system throughput for an increasing number of simultaneous 130 Kbps flows using CSMA/ECA.

triangles for CSMA/ECA. The solid line are the fraction of empty slots, the dashed line the fraction of successful slots and the dotted line the fraction of collision slots. The CSMA/CA results are obtained using the model in [5] and the model presented in Sec. 4 is used for CSMA/ECA.

In the non-saturated region, the behaviour of CSMA/CA and CSMA/ECA is similar. However, as the saturation region is reached, the two protocols react distinctly. CSMA/CA lowers the transmission probability, thus preventing the number of successful slots from increasing. Even worse, in CSMA/CA collisions appear and waste a substantial fraction of the channel time.

As opposed to CSMA/CA, CSMA/ECA reacts positively to saturation. The fraction of successful slots linearly increase with the number of flows and collisions are prevented. The deterministic backoff after successful transmissions allows collision-free operation in saturation conditions.

Fig. 7 compares the throughput attained by the two protocols under comparison. The behaviour of the CSMA/CA and CSMA/ECA is similar when the network is not saturated and can deliver all the traffic that is offered. Nevertheless, under saturation conditions, CSMA/ECA delivers a higher throughput thanks to the fact that collisions are prevented.

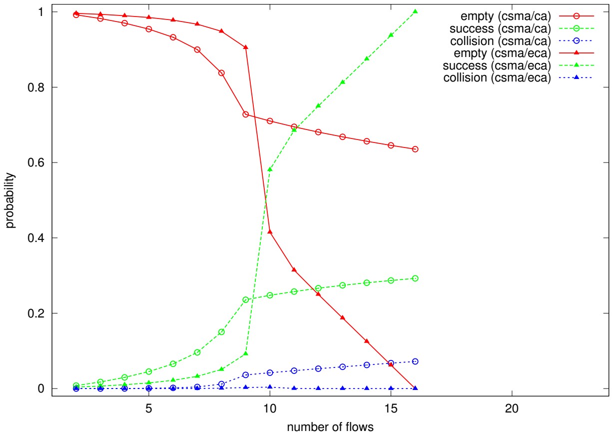


Fig. 6. Model results that compare the channel status probabilities for CSMA/ECA and CSMA/CA

It can be observed that, even for CSMA/ECA, the maximum throughput is far from the nominal data rate that is used (2 Mbps). This is typical of IEEE 802.11 networks and it is caused by the protocol overhead, which includes headers, preambles, link layer acknowledgements and inter-frame spaces.

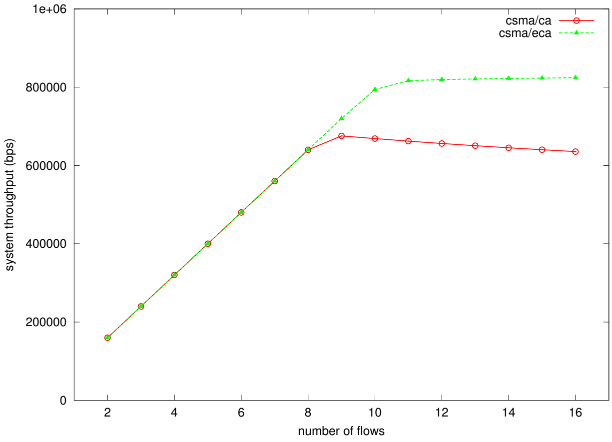


Fig. 7. Model results that compare the throughput of CSMA/ECA and CSMA/CA

6 Conclusion

The present article shortly reviews collision prevention approaches in WLANs and focuses on CSMA/ECA. In CSMA/ECA a deterministic backoff is selected afters successful transmissions, thereby preventing that those stations that have successfully transmitted in their previous attempt collide among them in their next transmission attempt.

The main contribution of the paper is the presentation of an analytical model that captures the behaviour of the protocol in saturated and non-saturated scenarios. The results, that are validated by simulations, show that CSMA/ECA prevents collisions in the stationary operation of the network. The collision-free operation translates to higher throughput.

References

1. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, 1999 Edition (Revised 2007).
2. J. Barcelo, B. Bellalta, C. Cano, and M. Oliver. Learning-BEB: Avoiding Collisions in WLAN. In *Eunice Summer School*, 2008. Available at www.dtic.upf.edu/∼jbarcelo/papers/barcelo2008lba.pdf .
3. J. Barcelo, B. Bellalta, C. Cano, A. Sfairopoulou, M. Oliver, and J. Zuidweg. Traffic Prioritization fo Carrience Sense Multiple Access with Enhanced Collision Avoidance. In *MACOM*, 2009. Available at www.dtic.upf.edu/∼jbarcelo/papers/barcelo2009tpc.pdf .
4. J. Barcelo, B. Bellalta, A. Sfairopoulou, C. Cano, and M. Oliver. CSMA with Enhanced Collision Avoidance: a Performance Assessment. In *IEEE VTC Spring*, 2009. Available at www.dtic.upf.edu/∼jbarcelo/papers/bellalta2009vtc.pdf .
5. B. Bellalta, M. Oliver, M. Meo, and M. Guerrero. A Simple Model of the IEEE 802.11 MAC Protocol with Heterogeneous Traffic Flows. In *IEEE Eurocon*, Belgrade, Serbia and Montenegro, 2005.
6. J.C. Bellamy. *Digital Telephony.* Wiley-Interscience, 1982.
7. G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE J. Sel. Areas Commun.*, 18(3):535–547, 2000.
8. F. Cali, M. Conti, E. Gregori, and P. Aleph. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Trans. Netw.*, 8(6):785–799, 2000.
9. E. Choi and W. Lee. Distributed Backoff Reservation and Scheduling for Collision Mitigation in IEEE 802.11 WLANs. In *ICOIN*, pages 1–5, 2008.
10. J. Choi, J. Yoo, S. Choi, and C. Kim. EBA: An Enhancement of the IEEE 802. 11 DCF via Distributed Reservation. *IEEE Trans. Mobile Comput.*, 4(4):378–390, 2005.
11. W. Crowther, R. Rettberg, D. Walden, S. Ornstein, and F. Heart. A System for Broadcast Communication: Reservation-ALOHA. In *Hawaii Int. Conf. Syst. Sci*, pages 596–603, 1973.

12. DJ Goodman, RA Valenzuela, KT Gayliard, and B. Ramamurthi. Packet Reservation Multiple Access for Local Wireless Communications. *IEEE Trans. Commun.*, 37(8):885–890, 1989.
13. A. Lopez-Toledo, T. Vercauteren, and X. Wang. Adaptive Optimization of IEEE 802.11 DCF Based on Bayesian Estimation of the Number of Competing Terminals. *IEEE Trans. Mobile Comput.*, 5(9):1283, 2006.
14. JL Sobrinho and AS Krishnakumar. Quality-of-service in ad hoc carrier sense multiple access wirelessnetworks. *IEEE J. Sel. Areas Commun.*, 17(8):1353–1368, 1999.
15. YC Tay and KC Chua. A Capacity Analysis for the IEEE 802.11 MAC Protocol. *Wireless Networks*, 7(2):159–171, 2001.

Traffic Prioritization for Carrier Sense Multiple Access with Enhanced Collision Avoidance

J. Barcelo, B. Bellalta, C. Cano, A. Sfairopoulou, M. Oliver, J. Zuidweg

Universitat Pompeu Fabra

Abstract. Carrier Sense Multiple Access with Enhanced Collision Avoidance (CSMA/ECA) is a simple MAC protocol for wireless local area networks that significantly outperforms CSMA/CA. This paper addresses traffic prioritization in CSMA/ECA using different minimum contention windows for different traffic classes. The conditions under which the collision-free operation is possible are presented and expressions for the steady-state channel efficiency are provided. Simulation results show that CSMA/ECA improves the channel utilization achieved by all traffic classes whenever the conditions for collision-free operation are satisfied.

1 Introduction

In wireless local area networks (WLANs), the radio channel is a broadcast medium that needs to be shared among the participating nodes. In each wireless station, there is a sublayer of the data link layer called Medium Access Control (MAC) which is in charge of the channel access arbitration.

The MAC layer implements a multiple access protocol to make it possible for several wireless stations to share the radio channel. MAC protocols can be classified in scheduled access and random access mechanisms. In scheduled access, each participating node is assigned (either statically or dynamically) a certain amount of radio resources to transmit.

The negative side of static scheduled access is that it is not suitable to accommodate bursty traffic. This can be solved by dynamically scheduled access, that assigns the resources to those stations that have data ready to transmit. However, dynamically scheduled access requires additional signalling and a centralized entity in charge of radio resource management.

As opposed to scheduled access, random access protocols are appropriate for bursty traffic and can be executed in a distributed manner. In random access protocols the radio resources are not explicitly assigned and thus it can occur that two or more stations try to use the same resources simultaneously. If this is the case, a collision occurs and the information conveyed in the messages involved in the collision may be lost. The goal of random access MAC protocols is to maximize the channel efficiency while reducing the chances of collision.

Additionally, in certain scenarios, it is necessary that the MAC layer offers traffic differentiation. As an example, in infrastructure deployments, prioritizing the access point might alleviate the uplink-downlink unfairness.

This article introduces Carrier Sense Multiple Access with Enhanced Collision Avoidance (CSMA/ECA) which is a random access MAC protocol that improves channel efficiency by reducing the chances of collision. The main contribution of the article is to show that it is possible to achieve traffic prioritization in CSMA/ECA. Simulation results evidence that CSMA/ECA always outperforms CSMA/CA, even when traffic differentiation is applied.

After this introductory section, the remainder of the paper is organized as follows. The current literature on random access protocols and traffic differentiation at the MAC layer is briefly reviewed in Section 2. Then, in Section 3, channel efficiency is presented as a metric to assess the performance of MAC protocols. CSMA/ECA, which is the protocol of interest in this paper, is introduced in Section 4. In Section 5, traffic prioritization in CSMA/ECA is explained. The validity of the suggested approach is supported by simulation results in Section 6. Finally, Section 7 concludes the paper.

2 Related Work

Random access MAC protocols for wireless networks dates back to the 70's, with the deployment of the ALOHA network [3] at the University of Hawaii. In ALOHA, the possibility of packet collisions places an upper bound on the channel efficiency of $1/2e$. The collisions are extremely costly in wireless networks, since the nodes involved in the collision cannot detect this circumstance until the transmission has finished. Therefore, an amount of channel time equal to the longest transmission is wasted in each collision.

The division of the channel time into slots (s-ALOHA [15]) and the idea of listening to the medium before transmitting (CSMA [13]) reduce the chances of collision and, hence, substantially increase the efficiency of WLANs.

The success of the IEEE 802.11 standard family [1] fostered a new wave of research on MAC protocols for WLANs. This standard specifies DCF (Distributed Coordination Function) as the medium access procedure that is used by the stations to gain access to the medium.

In [5] and [16] the performance of DCF is modelled for the case in which all stations are saturated (i.e. the stations have always a packet ready to transmit) and the channel does not introduce errors. The maximum theoretical performance that can be attained by adjusting the parameters of DCF is derived in [7]. An alternative computation of the theoretical upper bound is presented in [4].

Several works have presented new mechanisms to increase the performance of DCF. A recurrent approach is to adjust the backoff windows as a function of the number of contenders [6,14]. In [8], it is proposed that the stations announce their backoff values in order to avoid collisions. This last approach can surpass the aforementioned theoretical maximum thanks to the fact that the backoffs are selected with prior knowledge about other stations' intentions to transmit. Nevertheless, it requires a modification of the protocol headers and, hence, it is not compatible with current implementations of the standard.

Another method that may perform beyond the upper theoretical limit of DCF is Reservation-ALOHA [9], in which reservation is used to decrease the number of collisions. The slots are grouped in frames and a successful transmission in one slot implies a reservation for the same slot in the following frame. The negative aspects are the fact that this approach places a limit on the maximum number of stations that can be active in a network and the incompatibility with DCF.

In [4], the authors propose the use of a deterministic backoff after successful transmissions. This simple modification reduces the number of collisions and allows the protocol to achieve a performance that exceeds the theoretical upper bound associated to DCF. Moreover, this solution is compatible with current implementations of the protocol IEEE 802.11.

None of the previous proposals addresses traffic prioritization issues, which are required to prevent the uplink-downlink unfairness [12]. An enhancement of DCF to support traffic differentiation is defined in the IEEE 802.11e [2] standard amendment for quality of service. It introduces the Enhanced Distributed Channel Access (EDCA) that classifies the traffic in different queues with different parameter configuration in order to obtain traffic differentiation. The performance analysis of EDCA has motivated many research efforts, such as the ones presented in [10, 11]. However, EDCA still suffers from the inefficiency problems associated to collisions.

3 Channel Efficiency

Throughout this paper, it is assumed that the stations are saturated (they always have a packet ready to be transmitted) and the channel time is divided into slots. Ideal channel conditions are assumed as well.

Every contending station keeps a backoff counter (See Fig. 1) which is set whenever a packet arrives to the head-of-line of the MAC queue. Then, the backoff counter is decremented in every slot and, when the backoff counter reaches zero, the station transmits the packet.

If no station transmits in a given slot, the slot remains empty. When one (and only one) node transmits, the transmission is successful. Finally, if two or more contenders transmit in the same slot, a collision occurs and it is assumed that the data contained in the packets is lost.

The channel efficiency is defined as the fraction of time devoted to successful transmissions and is computed as:

$$\phi = \frac{P_s T_s}{P_e T_e + P_s T_s + P_c T_c}, \tag{1}$$

where P_e, P_s and P_c are the empty, success and collision probabilities, respectively. And T_e, T_s and T_c are the duration of an empty,

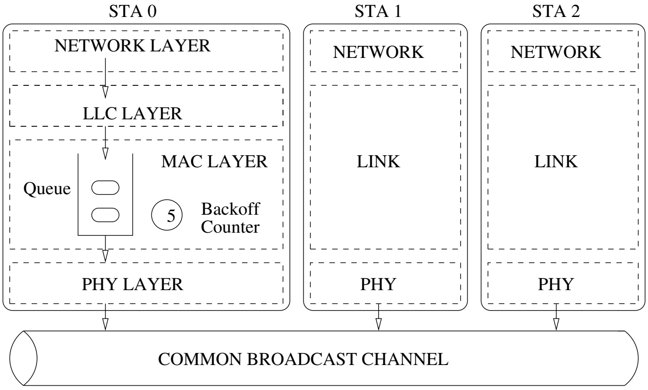


Fig. 1. The Medium Access Control uses a backoff counter to defer the transmissions.

successful and collision slot, respectively. The duration of empty slots is typically orders of magnitude below the duration of busy slots. If constant packet length is assumed, then the duration of collisions and successful slots is approximately the same ($T_s \approx T_c$).

4 CSMA/ECA: a Novel Medium Access Protocol

DCF uses CSMA with Collision Avoidance (CSMA/CA) combined with a Binary Exponential Backoff (BEB). A station with a packet ready to transmit listens to the channel for a Distributed InterFrame Space (DIFS). If the channel is sensed idle, the station transmits. Otherwise, the station waits until the channel is idle. Then, it generates a random backoff value and, finally, it waits for that number of slots before transmitting. The backoff countdown is frozen while the channel is sensed busy.

Let us emphasize that, under the saturation assumption, *all* the transmission attempts are delayed a *random* number of slots, independently of the result of the last transmission attempt (either a success or a collision). This is because the standard specifies that a station must separate two consecutive transmissions by a random backoff, even if the channel is sensed idle for a DIFS after the first transmission.

In [4], it was suggested to choose a deterministic backoff after successful transmissions. We refer to this last solution as CSMA with Enhanced Collision Avoidance (CSMA/ECA). In CSMA/ECA, when a station joins the contention or its last transmission attempt resulted in a collision, it chooses a random backoff B:

$$B \sim \mathcal{U}[0, \min(2^a \cdot CW_{min}, CW_{max}) - 1], \tag{2}$$

where $\mathcal{U}$ represents the uniform distribution. CW_{min} and CW_{max} are the minimal and maximum contention windows, respectively. The number of transmission attempts for the current packet is denoted as a. Note that a equals 0 for the first transmission attempt.

After a successful transmission attempt, the backoff is chosen deterministically as:

$$V = \lfloor E\left[\mathcal{U}[0, CW_{min} - 1]\right]\rfloor = \lfloor (CW_{min} - 1)/2 \rfloor, \tag{3}$$

where $\lfloor\cdot\rfloor$ is the floor operator and $E[\cdot]$ is the expectation operator.

In CSMA/ECA, those stations that successfully transmitted in their last transmission attempt cannot collide among them. Furthermore, after all the stations have consecutively successfully transmitted, the system adopts a deterministic behaviour and collisions disappear. This is true as long as the channel time is discretized and all the stations decrease their respective backoff counters simultaneously (i.e. in saturation and ideal channel conditions).

The behaviour of CSMA/ECA can be described as a transitory (or convergence) process followed by a steady-state collision-free operation. CSMA/ECA outperforms CSMA/CA during both the transitory and the steady-state [4]. The convergence period can be interpreted as a random search to reach the collision-free operation. This paper focuses on the steady-state performance.

The steady-state operation of CSMA/ECA is characterized by its periodical and deterministic behavior. For a number of active contenders equal to ς such that $\varsigma \leq V + 1$, each cycle contains $V + 1$ slots, ς of them are successful transmissions. Throughout a cycle, a contender decreases its backoff counter until zero, transmits once, sets the counter to V, and decreases its backoff counter again. At any given moment all the stations have different backoff values and, by the beginning of each cycle, each station has recovered the same backoff value that it had at the beginning of the previous cycle.

Therefore, when the steady-state collision free operation is reached, the channel efficiency is:

$$\phi = \frac{\varsigma \cdot T_s}{\varsigma \cdot T_s + (V + 1 - \varsigma) \cdot T_e} \; ; \; \varsigma \leq V + 1. \tag{4}$$

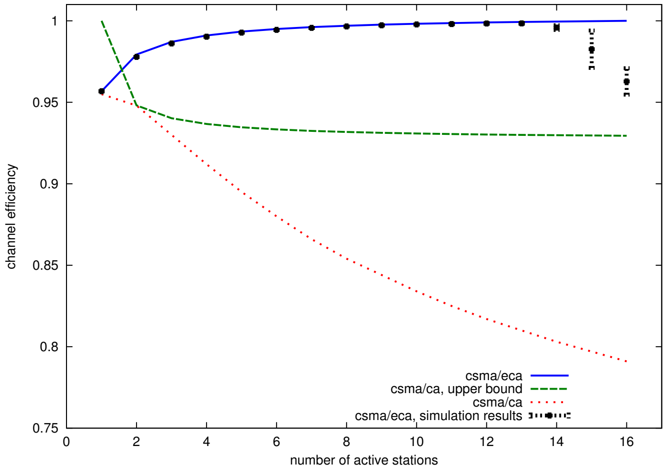


Fig. 2. Performance of CSMA/ECA compared to CSMA/CA

Fig. 2 compares the performance of CSMA/ECA with the performance of DCF and the maximum theoretical performance of DCF as presented in [4, 7]. The values of the contention windows are $CW_{min} = 32$ and $CW_{max} = 1024$. The packet length is 1500 bytes and the data rate is 2Mbps. The remainder of the parameters follow the standard IEEE 802.11b.

It is noteworthy that the channel efficiency of CSMA/ECA can exceed the upper bound associated to DCF thanks to the fact that the selection of the backoff values is not always random.

CSMA/ECA delivers extremely good results for a numbers of active contenders below the deterministic backoff value V. A station is considered to be actively contending for the channel whenever it has data to be transmitted. It has to be highlighted that the number of active contenders can be significantly lower than the number of stations registered to the network, since in data communications many traffic sources are characterized by its bursty behaviour.

When the number of contenders exceeds V, the performance of CSMA/ECA tends asymptotically to the performance of CSMA/CA. As the value of V is 15 in the simulations presented in Fig. 2, a small performance drop can be perceived for 15 and 16 active contenders.

5 CSMA/ECA with Traffic Differentiation

The standard amendment for quality of service IEEE 802.11e [2] states that different classes of traffic are directed to different queues, and each queue is assigned different contention parameters. There are three means to attain traffic differentiation: namely, the Arbitration InterFrame Space ($AIFS$), the transmission opportunity ($TXOP$) and the adjustment of contention windows (CW_{min} and CW_{max}).

When $TXOP$ is used, a station that wins the contention is allowed to transmit multiple packets. The application of $TXOP$ in CSMA/ECA is straightforward, since this parameter does not have any impact in the contention procedure. $TXOP$ simply affects the duration of a successful slot. Thus it can be concluded that $TXOP$ is a valid option for traffic differentiation in CSMA/ECA.

$AIFS$ modifies the time that the stations have to listen to an idle channel before start decrementing the backoff counter. Specifically, this time is computed as:

$$AIFS = DIFS + n \cdot T_e, \tag{5}$$

where n takes different values for the different queues. The utilization of $AIFS$ violates the assumption that all the stations decrement their backoff simultaneously and, hence, it is impractical in CSMA/ECA.

Finally, traffic differentiation can also be achieved by choosing different contention windows[1] for high and low-priority traffic (CW_{min}^{high} and CW_{min}^{low}, respectively). This approach is valid for CSMA/ECA whenever CW_{min}^{high} is an integer divisor of CW_{min}^{low}.

---

[1] The adjustment of CW_{max} has little effect on the prioritization. Therefore, only the adjustment of CW_{min} is considered.

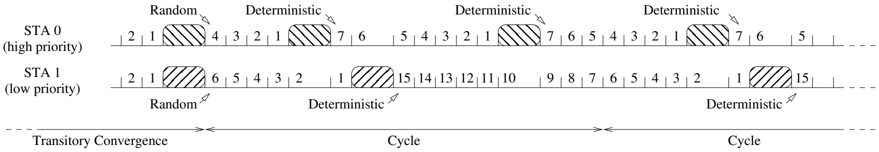


Fig. 3. Two stations with different priorities contend for the channel using CSMA/ECA. The shaded boxes represent transmissions and the numbers are the values of the backoff counters.

Since the values of the contention windows are selected as powers of 2, the aforementioned condition generally holds true. As an example, the values $CW_{min}^{high} = 16$ and $CW_{min}^{low} = 32$ can be considered for high and low-priority traffic. The deterministic backoff values after successes can be computed using (3) and are $V^{high} = 7$ and $V^{low} = 15$.

Fig. 3 depicts two stations with different priorities contending for the channel using CSMA/ECA. The channel time is divided into slots and transmissions are represented as shaded boxes. The values of the backoff counters of each station are also included in the figure. Finally, there are labels indicating whether the backoff values have been selected randomly or deterministically. The time advances from left to right and the slot length is not represented in scale. Recall from Section 3 that busy slots are actually orders of magnitude longer longer than the empty ones.

In the figure, after the initial collision, the two stations randomly choose the backoff value as expressed in (2). After successes, a deterministic backoff computed as in (3) is used. STA 0, which is the high-priority station, uses a value of 7 after successes while STA 1 uses a value of 15. After a transitory convergence, the system behaves in a periodical collision-free fashion. Note that in steady-state operation the high-priority station sends twice as much packets as the low-priority station.

The length of the cycle is also indicated in the figure. Each cycle comprehends 16 slots, including one transmission by the low priority station and two transmissions by the high priority station. Note that the backoff values of the stations in the first slot of the second cycle are exactly the same as in the first slot of the first cycle.

In general, when two traffic priorities in the stationary regime are considered, the behaviour of the system is periodic and a cycle contains $V^{low}+1$ slots. In each cycle there will be one successful transmission by each of the low priority stations and $\frac{V^{low}+1}{V^{high}+1}$ successful transmissions by each of the high priority station. As occurred with plain CSMA/ECA, the collision-free operation can be reached only if the total amount of transmissions per cycle is below $V^{low}+1$.

Let ς^{high} and ς^{low} be the number of high and low priority stations, respectively. Then, the number of successful transmissions per cycle by each of the traffic classes is τ^{high} and τ^{low}:

$$\tau^{high} = \varsigma^{high}\frac{V^{low}+1}{V^{high}+1} \ \ ; \ \ \tau^{low} = \varsigma^{low}. \tag{6}$$

And the condition for collision-free operation can be expressed as:

$$\tau^{high} + \tau^{low} \leq V^{low} + 1. \tag{7}$$

If the condition in (7) is satisfied, then the channel utilization by each of the traffic classes is:

$$\phi^{high} = \frac{\tau^{high}\cdot T_s}{(\tau^{high}+\tau^{low})T_s + (V+1-\tau^{high}-\tau^{low})T_e}, \tag{8}$$

$$\phi^{low} = \frac{\tau^{low}\cdot T_s}{(\tau^{high}+\tau^{low})T_s + (V+1-\tau^{high}-\tau^{low})T_e}. \tag{9}$$

And the overall channel efficiency is the addition of (8) and (9) :

$$\phi = \frac{\left(\tau^{high}+\tau^{low}\right)T_s}{(\tau^{high}+\tau^{low})T_s + (V+1-\tau^{high}-\tau^{low})T_e}. \tag{10}$$

6 Channel Utilization Results

The presented MAC protocol has been simulated[2] in ideal channel conditions and in the absence of hidden terminals for a range of active stations from 2 to 20. In each scenario, half of the stations are low-priority while the other half are high-priority. Each simulation lasts for 1,000,000 slots and is repeated ten times. Average

---

[2] Octave has been used for the simulations. The source code is available upon request.

results and 95% confidence intervals have been computed for each scenario. However, the confidence intervals are often too small to be seen in the figures. For each number of contenders, simulations are performed for CSMA/CA and CSMA/ECA to compare the prioritization properties and the channel efficiency that is achieved by each protocol.

Fig. 4 depicts the channel utilization obtained by the high-priority and the low-priority groups of stations. The figure also shows the aggregated channel utilization, which is equivalent to the channel efficiency. For both CSMA/CA and CSMA/ECA, the high-priority stations obtain approximately twice as much channel time as the low-priority ones. It can also be observed that CSMA/ECA clearly outperforms CSMA/CA in terms of channel efficiency.

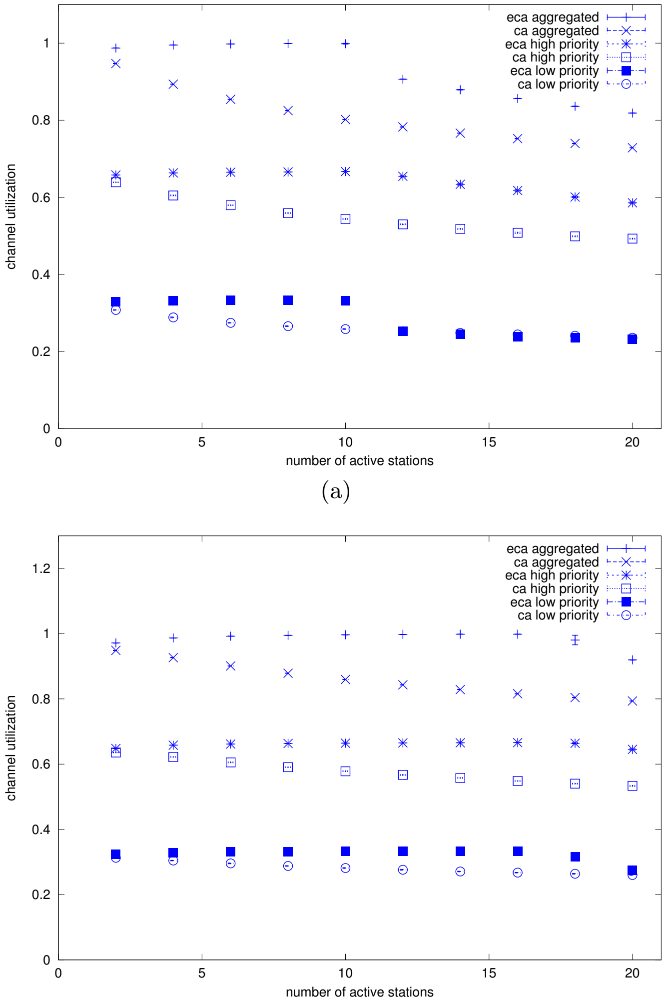
It is noteworthy the thresholding effect in the CSMA/ECA results when it is not possible for the system to completely avoid collisions and performance decays. This effect occurs for 12 contenders in Fig. 4.(a) and 20 contenders in Fig. 4.(b). For a greater number of contenders, the collisions are particularly pernicious for the low-priority traffic, while high-priority traffic roughly maintains its share of channel time.

By comparing Fig. 4.(a) and Fig. 4.(b), it can be observed that larger contention windows can accommodate more contenders. From the results, it can also be concluded that CSMA/ECA presents traffic differentiation properties similar to the ones offered by CSMA/CA.

7 Conclusions

CSMA/ECA is a novel medium access protocol that delivers higher channel efficiency than CSMA/CA. By choosing a deterministic backoff after successful transmissions, the number of collisions is reduced. Under certain conditions, it is even possible to attain collision-free operation.

In this work, CSMA/ECA is extended to support traffic prioritization by means of different contention windows. By choosing the windows sizes as sufficiently large powers of two, collisions can be prevented. Expressions for the channel utilization by each of the traffic classes have been provided and simulations show that, in the



(a)

(b)

Fig. 4. Simulation results that compare the performance of CSMA/CA and CSMA/ECA with traffic differentiation for (a) $CW_{min}^{high} = 16$ and $CW_{min}^{low} = 32$ (b) $CW_{min}^{high} = 32$ and $CW_{min}^{low} = 64$

collision-free mode of operation, all traffic classes obtain better performance in CSMA/ECA than in CSMA/CA.

In conclusion, it has been shown that CSMA/ECA can provide the same traffic differentiation properties as CSMA/CA, while offering greater overall performance.

References

1. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, 1999 Edition (Revised 2007).
2. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications; Amendment: Medium Access Control(MAC) Quality of Service Enhancements, 2005.
3. N. Abramson. The ALOHA System–Another Alternative for Computer Communications. *Cluster Computing*, 5:187–201, 1970.
4. J. Barcelo, B. Bellalta, C. Cano, and M. Oliver. Learning-BEB: Avoiding Collisions in WLAN. In *Eunice Summer School*, 2008. Available at www.dtic.upf.edu/∼jbarcelo/papers/barcelo2008lba.pdf.
5. G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE J. Sel. Areas Commun.*, 18(3):535–547, 2000.
6. G. Bianchi and I. Tinnirello. Kalman Filter Estimation of the Number of Competing Terminals in an IEEE 802.11 Network. In *IEEE Infocom*, volume 2, pages 844–852, 2003.
7. F. Cali, M. Conti, E. Gregori, and P. Aleph. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Trans. Netw.*, 8(6):785–799, 2000.
8. J. Choi, J. Yoo, S. Choi, and C. Kim. EBA: An Enhancement of the IEEE 802. 11 DCF via Distributed Reservation. *IEEE Trans. Mobile Comput.*, 4(4):378–390, 2005.
9. W. Crowther, R. Rettberg, D. Walden, S. Ornstein, and F. Heart. A System for Broadcast Communication: Reservation-ALOHA. In *Hawaii Int. Conf. Syst. Sci*, pages 596–603, 1973.
10. P.E. Engelstad and O.N. Østerbø. Non-saturation and saturation analysis of IEEE 802.11 e EDCA with starvation prediction. In *ACM MSWiM*, pages 224–233. ACM New York, NY, USA, 2005.
11. J. Hui and M. Devetsikiotis. A unified model for the performance analysis of IEEE 802.11 e EDCA. *IEEE Trans. Commun.*, 53(9):1498–1510, 2005.
12. F. Keceli, I. Inan, and E. Ayanoglu. Weighted Fair Uplink/Downlink Access Provisioning in IEEE 802.11 e WLANs. In *IEEE ICC*, pages 2473–2479, 2008.
13. L. Kleinrock and F. Tobagi. Packet Switching in Radio Channels: Part I–Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics. *Communications, IEEE Transactions on [legacy, pre-1988]*, 23(12):1400–1416, 1975.
14. A. Lopez-Toledo, T. Vercauteren, and X. Wang. Adaptive Optimization of IEEE 802.11 DCF Based on Bayesian Estimation of the Number of Competing Terminals. *IEEE Trans. Mobile Comput.*, 5(9):1283, 2006.
15. LG Roberts. Aloha Packet System with and without Slots and Capture, ASS Notes 8, Advanced Research Projects Agency. *Network Information Center, Stanford Research Institute*, 1972.

16. YC Tay and KC Chua. A Capacity Analysis for the IEEE 802.11 MAC Protocol. *Wireless Networks*, 7(2):159–171, 2001.

Final Remarks

The larger the island of knowledge, the longer the shoreline of wonder. ∼Ralph W. Sockman

This quote reflects the author's feelings in completing a Ph.D. program. Every single advancement or accomplishment implies new open questions. These final remarks briefly review the open issues and challenges that call for further research.

The most obvious is the need for a testbed. However, the lack of prototypes for research and development makes it difficult implement arbitrary MAC protocols. It is likely that the collaboration of manufacturers is required to modify the firmware of the wireless cards.

It is also necessary to prepare a CSMA/ECA module for a well-known network simulator. This module will facilitate the task of other researchers willing to test the performance of CSMA/ECA in their own works. Current efforts (mainly by Diego Saez) are focused on developing a CSMA/ECA module for Network Simulator 3 (NS-3), which is expected to be the network simulation platform of choice in the upcoming years.

Hopefully, the testbed and simulation tools will provide more insight on the interaction of CSMA/ECA with the rest of the layers of the protocol stack. The promising results obtained in this thesis gives us optimism that CSMA/ECA will deliver equal or better performance than CSMA/CA in all the considered scenarios.

One scenario of particular interest is a multi-hop mesh network, which will be supported by the upcoming standard amendment IEEE 802.11s. In multi-hop networks, it is no longer true that all the stations have the same vision of the channel, thus the backoff counters do not decrement simultaneously. Our belief is that, to attain collision-free operation in multi-hop networks, there is an additional requirement: It is necessary that the length of the slot is fixed (i.e. the same for empty, successful and collision slots). However, more evidence and study is required to fully understand the issue.

Throughout the articles, we have used a deterministic backoff value that is approximately equal to the half of the minimum contention window. It seemed a natural choice, since it maintains the expected backoff value before the first transmission attempt for a given packet. This ensures that CSMA/ECA outperforms CSMA/CA in any considered scenario. Many colleagues where concerned about the fact that this value was not optimal. There are many trade-offs in choosing the right value for V. Larger values offer a faster transition to collision-free operation and can accommodate a larger number of concurrent stations. However, if V is too large, it might slightly harm the performance when there is only one active station, due to the high number of empty slots. The optimal value of V would be a function of several time-varying factors, such as the number of stations, traffic patterns and channel conditions. Therefore, as occurs with the contention windows sizes, it is sensible to use a value that delivers an acceptable performance for the most common scenarios. As in IEEE 802.11e, the access point can distribute the value of V using broadcast beacons. In any case, it seems reasonable to choose a value that is substantially larger than the expected number of simultaneous contenders, to ensure fast recovery after channel errors or the entrance of a new contender.

The disrupting effect of a channel error or a new entrance can be prevented by introducing an additional degree of memory to the protocol. Specifically, if the stations use a deterministic value V for two consecutive backoffs after a successful transmission, the system will remain in the stationary collision-free operation with high probability, even in lossy channels and highly dynamical environments in which the nodes constantly join and leave the contention.

This concludes the dissertation. Thank you for reading so far and, if you liked the idea, please spread the word!