

# Resource Control for the Enhanced Distributed Channel Access (EDCA) Mechanism in IEEE 802.11e

Vasilios A. Siris\* and Costas Courcoubetis†

Technical Report 352 - March 2005

Institute of Computer Science (ICS)

Foundation for Research and Technology - Hellas (FORTH)

P.O. Box 1385, GR 711 10 Heraklion, Crete, Greece

vsiris@ics.forth.gr

## Abstract

We investigate the problem of efficient resource control for elastic traffic in IEEE 802.11e's Enhanced Distributed Channel Access (EDCA) mechanism. Our approach considers an economic modelling framework based on congestion pricing that captures how various factors, such as the probability of attempting to transmit a frame, the use of the basic CSMA/CA or the RTS/CTS procedure, and the physical layer transmission rate, contribute to the congestion in an IEEE 802.11e network using the EDCA mechanism. We discuss the application of the framework for achieving class-based proportional throughput differentiation, for performing explicit congestion notification (ECN) marking based on the level of congestion in the wireless channel, and for modelling the performance of TCP congestion control over EDCA. In the aforementioned application scenarios we discuss how to estimate the optimal minimum contention window and the congestion prices based on the 802.11e parameters and actual measurements, in order to efficiently utilize the wireless channel. Experiments demonstrate the application of the proposed framework, compare simulation and analytical results, and identify the gains in using the optimal minimum contention window.

**Keywords:** service differentiation, contention-based access, economic efficiency, congestion pricing

## 1 Introduction

An increasing number of users is accessing the Internet and enterprise intranets through wireless networks.

\*Also with the University of Crete, Heraklion, Greece

†Also with the Athens University of Economics and Business, Greece

This number is expected to grow dramatically with the proliferation of wireless hotspots and enterprise wireless LANs, and with the increasing use of wireless mesh topologies for building community and broadband access networks. The technology that has dominated the market is IEEE 802.11, also known as WiFi. The IEEE 802.11b standard can support transmission rates up to 11 Mbps, using the unregulated 2.4 GHz frequency band, whereas the newer standards IEEE 802.11a/g support transmission rates up to 54 Mbps. Clearly, the available bandwidth in wireless LANs is at least one order of magnitude smaller than the capacity typically available in wired networks, and there is a limited ability to increase this capacity, since it is limited by the available wireless spectrum. Moreover, emerging multimedia and packet-based services over wireless networks will have different bandwidth and delay requirements. For all the above reasons, resource control and service differentiation in wireless LANs based on IEEE 802.11 will become increasingly important.

Resource control procedures are required to be fair and adaptive to different network loads, and to achieve efficient utilization of the shared wireless channel. One approach for developing such procedures is based on economic modelling, through the use of utility functions and congestion pricing. Such an approach has been successfully applied to both wired networks and wireless, e.g. CDMA networks [13, 16, 19]. A key feature of economic models is the efficient utilization of network resources through a decentralized control approach. Moreover, since user requirements are encoded using utility functions, service differentiation can be supported. The work presented in this paper is, to the authors' knowledge, the first application of economic modelling that takes into account the specific characteristics and operation of IEEE 802.11e's contention-based EDCA mechanism.

The goal of this paper is to present and investigate an approach for efficient resource control for IEEE 802.11e's EDCA contention-based channel access mechanism. The approach considers an economic modelling framework using congestion prices, and an important feature is that it captures how various factors affect congestion; these factors include the probability of attempting to transmit a frame, which is related to the minimum contention window used in 802.11's collision avoidance procedure, the use of the CSMA/CA or the RTS/CTS procedure for accessing the shared channel, and the physical layer transmission rate. Indeed, wireless stations located at different distances from the access point can experience different levels of attenuation and shadowing, hence will have different physical layer transmission rates. Moreover, it is known that in such mixed transmission rate scenarios, the aggregate throughput of the wireless LAN can deteriorate significantly [9]. Hence, understanding how to efficiently control such networks has practical engineering importance. We note here that the performance analysis of 802.11 networks with stations having different transmission rates has received limited attention in the literature; work that considers different transmission rates includes [9, 6, 7]. Finally, by considering utility functions for encoding end user and application requirements, the proposed framework can be used to investigate the performance of transport protocols (in this paper we consider the TCP congestion control algorithm) over EDCA.

Our model can be applied in a class-based service differentiation framework, where the access point of a wireless LAN estimates the optimal minimum contention window for different service classes, taking into account the level of congestion; this has high practical significance, since the IEEE 802.11e standard defines how the minimum contention windows for various classes are communicated to the wireless stations, but not how they should depend on the network load and traffic characteristics in order to achieve efficient channel utilization. Another application of our model is to achieve efficient utilization of the wireless channel through a decentralized scheme, using explicit congestion notification (ECN) marks for signalling the level of congestion in the 802.11 wireless network, which is expressed with congestion prices. End stations react to the ECN feedback based on their requirements, which are expressed through utility functions; more specifically, end stations select their minimum contention window that maximizes their net benefit, defined as their utility minus the congestion cost they are creating, where the latter is given by the rate of ECN marks they receive.

The rest of the paper is organized as follows. In Section 2 we present a brief overview of the IEEE 802.11 MAC

layer procedures, focusing on the Enhanced Distributed Channel Access (EDCA) mechanism of the 802.11e standard. In Section 3 we discuss an analytical model for the throughput of a wireless station in an 802.11e network, which captures the case where different stations can have different minimum contention windows and different physical layer transmission rates. In Section 4 we present our framework for efficient resource control, based on the aforementioned throughput models. In Section 5 we discuss the application of the resource control framework, and in Section 6 we present experimental results that demonstrate and evaluate its use. Finally, in Section 7 we present a brief overview of related work, identifying where it differs from the work presented in this paper, and in Section 8 we conclude the paper identifying related and future research directions.

## 2 IEEE 802.11e and EDCA

In IEEE 802.11, access to the shared wireless channel is controlled through two MAC layer mechanisms: polling-based PCF (Point Coordination Function) and contention-based DCF (Distributed Coordination Function). DCF is based on CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance). As in Ethernet's CDMA/CD, with CSMA/CA prior to a frame transmission the wireless channel must be sensed idle for a time interval called interframe spacing (IFS), which can be different for different frame types; hence, for data frames the interval is a DCF-IFS (DIFS), and for acknowledgements it is a short IFS (SIFS).

The RTS/CTS procedure is used to avoid the hidden node problem, which can occur when two wireless stations that are not within the range of each other wish to transmit at the same time to the access point. According to the RTS/CTS procedure, before transmitting a data frame, a station transmits a special frame called request to send (RTS), after sensing that the wireless channel is idle for time DIFS. The RTS frame contains the addresses of both the sender and the destination. The destination, upon receiving an RTS frame responds with a clear to send (CTS) frame. After receiving the CTS frame, the sending station can send the data frame. Using the above procedure the hidden node problem is avoided, since all stations within the range of the destination will receive the CTS frame, independently of whether they are in the range of the other senders with which they can potentially collide, hence will defer their transmission. The gains in using the RTS/CTS procedure are due to collisions occurring (mainly) with RTS frames. Because the length of RTS frames is smaller than the length of data frames, the cost of collisions involving

RTS frames is smaller.

In wireless networks, unlike in Ethernet LANs, collision detection is not possible or is too costly. For this reason the DCF mechanism uses collision avoidance: Before initiating the transmission of a frame, the station selects a random backoff period from the interval  $[0, CW - 1]$ , where  $CW$  is referred to as contention window. The station waits for the channel to be idle for a total time equal to this backoff period, after which it senses the channel to see if it is idle for a DIFS interval, in which case it can transmit a data frame, when the basic CSMA/CA procedure is used, or an RTS frame, when the RTS/CTS procedure is used. The contention window has an initial value  $CW_{min}$ , and is doubled when a collision occurs, which can be detected when an acknowledgement is not received, up to the maximum value  $CW_{max}$ . Finally, when a frame is successfully transmitted, the contention window is set to its initial value  $CW_{min}$ .

IEEE 802.11e is an upcoming version of the IEEE 802.11 standard that addresses the issue of QoS support in wireless LANs. The MAC protocol of IEEE 802.11e is the Hybrid Coordination Function (HCF), which supports both contention-based and controlled channel access [8], and coexists with the basic DCF and PCF mechanisms for backward compatibility. The contention-based access of HCF is based on the Enhanced Distributed Channel Access (EDCA) mechanism, which is an extension of the DCF mechanism that enables distributed differentiated access to the wireless channel with the support of multiple access categories (AC). A higher priority access category has a smaller minimum contention window  $CW_{min}$ , thus giving it a higher probability to access the channel. Additionally, different access categories can have different values for the maximum contention window  $CW_{max}$  and the interframe spacing interval (IFS), which is now called Arbitration IFS (AIFS). It is important to note that although the IEEE 802.11e standard defines a number of parameters that can be used to achieve service differentiation, it does not define how these parameters should depend on the network load and traffic characteristics in order to efficiently utilize the shared wireless channel.

Studies [1, 14] have shown that different values for  $CW_{min}$  can provide differentiation in terms of throughput. On the other hand, different values for  $CW_{max}$  lead to different service only in cases of increased collisions. Moreover, different values for AIFS are effective for supporting different priorities in accessing the wireless channel. Motivated by these results, in this paper we investigate models that achieve service (throughput) differentiation through the assignment of different values of the minimum contention window  $CW_{min}$ .

### 3 Throughput model for EDCA

In this section we discuss an analytical model that gives the throughput of a station in an IEEE 802.11 wireless LAN using the EDCA mechanism. The model captures important characteristics of the corresponding protocol, namely the congestion avoidance and RTS/CTS procedures, and the case where different wireless stations have different transmission rates.

Several analytical studies that focus on the throughput of IEEE 802.11 have approximated the congestion avoidance procedure of IEEE 802.11 with a p-persistent model [5, 18]. In a p-persistent model, the probability that a wireless station tries to transmit in a time slot is  $p$ , and is independent of the success or failure of previous transmission attempts. It has been shown that the p-persistent model closely approximates the throughput of the actual congestion avoidance procedure of IEEE 802.11, when the average backoff interval is the same [5].

In particular, if  $E[CW]$  is the average contention window, then the approximate p-persistent model has a transmission probability  $p$  given by [5]

$$p = \frac{2}{E[CW] + 1}.$$

If we assume that the probability of a frame being involved in more than one collision is very small, then we have the approximation  $E[CW] \approx CW_{min}$  [18]. In IEEE 802.11e, different wireless stations can have a different minimum contention window, hence using the same arguments as above [18], the corresponding transmission probability of station  $i$  in the p-persistent model is

$$p_i = \frac{2}{CW_{min,i} + 1}. \quad (1)$$

Hence, we can estimate the throughput for a given set of stations with different values of the minimum contention window, by approximating it with the corresponding p-persistent model, where the transmission probability of a station is given by the above relation.

The MAC operation of IEEE 802.11 can be viewed in time as involving three different types of time intervals: a successful transmission interval, a collision interval, and an idle time interval. We denote the length of each interval type as  $T^{suc}$ ,  $T^{col}$ , and  $T^{idl}$ , respectively. The duration of each time interval depends on the physical layer encoding and the MAC layer operations. In particular, we have the following:

$T^{suc}$ : For the basic CSMA/CA operation in 802.11b, we

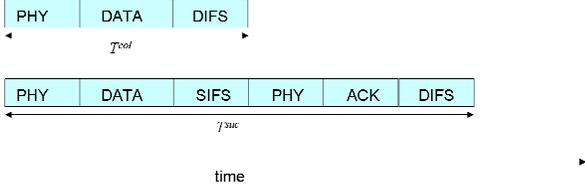


Figure 1: Time intervals for basic CSMA/CA.

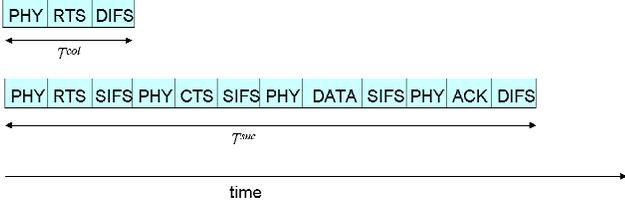


Figure 2: Time intervals for RTS/CTS.

have the following<sup>1</sup>, Figure 1:

$$T^{suc} = 2T_{PHY} + T_{SIFS} + \frac{8(34 + L)}{R} + T_{ACK} + T_{DIFS}, \quad (2)$$

where  $L$  is the frame length, 34 bytes is the MAC overhead,  $R$  is the transmission rate, and  $T_{PHY}$ ,  $T_{SIFS}$ ,  $T_{DIFS}$ ,  $T_{ACK}$  are the durations (in time units) of the physical layer overhead, the SIFS interval, the DIFS interval, and the ACK transmission time, respectively. In 802.11b, ACK, RTS, and CTS frames are transmitted at 1 Mbps, hence their transmission times are independent of the rate  $R$ . For RTS/CTS we have, Figure 2

$$T^{suc, RTS/CTS} = 4T_{PHY} + 3T_{SIFS} + \frac{8(34 + L)}{R} + T_{RTS} + T_{CTS} + T_{ACK} + T_{DIFS}, \quad (3)$$

where  $T_{RTS}$ ,  $T_{CTS}$  is the duration of an RTS, CTS frame transmission, respectively.

$T^{col}$ : For basic CSMA/CA we have, Figure 1

$$T^{col} = T_{PHY} + \frac{8(34 + L)}{R} + T_{DIFS}. \quad (4)$$

For RTS/CTS the collision interval is, Figure 2

$$T^{col, RTS/CTS} = T_{PHY} + T_{RTS} + T_{DIFS}. \quad (5)$$

$T^{idl}$ : This is equal to one time slot, whose default value for IEEE 802.11b and 802.11a is 20  $\mu$ s and 9  $\mu$ s, respectively.

Time can be viewed as a sequence of intervals of the above types. The average throughput for station  $i$ , considering a renewal assumption, can be expressed as the

<sup>1</sup>We assume that the propagation delay is very small, hence do not consider it.

ratio of the average amount of data transmitted by that station in one time interval over the average duration of a time interval [4, 5, 18].

$$x_i = \frac{E[X_i]}{E[T]}. \quad (6)$$

The average amount of data transmitted by station  $i$  in one time interval, considering a p-persistent model and assuming that the station always has a frame ready to transmit, is

$$E[X_i] = p_i \prod_{j \neq i} (1 - p_j) L,$$

where  $L$  is the frame size, which for simplicity we assume is the same for all stations. The average time interval is a weighted sum of the three types of intervals. If we assume that the intervals  $T^{suc}$  and  $T^{col}$  are normalized to the size of the idle slot time, then the average time interval becomes

$$E[T] = \sum_k p_k \prod_{j \neq k} (1 - p_j) T^{suc} + \left[ 1 - \prod_j (1 - p_j) - \sum_k p_k \prod_{j \neq k} (1 - p_j) \right] T^{col} + \prod_j (1 - p_j). \quad (7)$$

If the individual transmission probabilities  $p_i$  and the aggregate transmission probability are very small, then we have the approximation

$$\prod_{j \neq i} (1 - p_j) \approx 1 - \sum_{j \neq i} p_j = 1 - P_{-i},$$

where we have defined  $P_{-i} = \sum_{j \neq i} p_j$ . Indeed, numerical experiments indicate that the optimal aggregate transmission probability is typically less than 0.2 when the CSMA/CA procedure is used, and less than 0.35 when the RTS/CTS procedure is used. Based on the above, when all wireless stations have the same transmission rate, the average throughput  $x_i$  for station  $i$  is

$$x_i = \frac{p_i(1 - P_{-i})L}{\sum_k p_k(1 - P_{-k})T^{suc} + [P - \sum_k p_k(1 - P_{-k})]T^{col} + 1 - P}. \quad (8)$$

Note that the above expression is valid for all versions of 802.11, provided all stations have the same transmission rate. The specific version of 802.11, and whether the CSMA/CA or RTS/CTS procedure is used, will determine the values of  $T^{suc}$  and  $T^{col}$ , which we have taken to be normalized to the duration of the idle interval.

As discussed previously, in 802.11b with the RTS/CTS procedure, the transmission rate does not affect the

collision interval, since the latter involves RTS frames which are always transmitted at the basic rate (1 Mbps). Hence, for 802.11b with the RTS/CTS procedure, when different stations have a different transmission rate, the average throughput  $x_i$  is

$$x_i = \frac{p_i(1 - P_{-i})L}{\sum_k p_k(1 - P_{-k})T_k^{suc} + [P - \sum_k p_k(1 - P_{-k})]T^{col} + 1 - P}. \quad (9)$$

The last expression for the throughput of a wireless station captures a well-known property of 802.11 networks [9]: a station with a small transmission rate leads to decreased throughput not only for itself, but for all other wireless stations in the same network, independently of their transmission rate; this is because a small transmission rate for some station increases the denominator in (9) for all stations  $i$ .

Next we consider the case where the collision interval also depends on the transmission rate. If we assume that the probability for three or more frames to collide is negligible, then the average throughput  $x_i$  is

$$x_i = \frac{p_i(1 - P_{-i})L}{\sum_k p_k(1 - P_{-k})T_k^{suc} + \sum_k \sum_{j \neq k} p_k p_j \max\{T_k^{col}, T_j^{col}\} + 1 - P} \quad (10)$$

where  $T_k^{col}$  is the duration of the collision interval for station  $k$ .

## 4 Resource control model

In this section, based on the throughput model for the EDCA mechanism presented in the previous section, we present a congestion pricing framework for achieving efficient utilization of the shared wireless channel.

We consider the case of elastic traffic, for which users value the average throughput of data transfer. The utility for such a user  $i$  has the form

$$U_i(x_i),$$

where  $x_i$  is the average throughput for user  $i$  (for simplicity, we assume that one user corresponds to one wireless station, hence we use the terms user and wireless station interchangeably), which depends on his transmission probability  $p_i$  through (8), (9) or (10).

Consider the global problem of maximizing the aggregate utility (social welfare), in a wireless network with a set of users  $N$

$$\begin{aligned} & \text{maximize} && \sum_i U_i(x_i) \\ & \text{over} && \{p_i \geq 0, i \in N\}. \end{aligned} \quad (11)$$

If  $U_i(\cdot)$  is differentiable and strictly concave, then the necessary conditions for the maximization in (11) are

$$\frac{\partial \sum_i U_i(x_i)}{\partial p_i} = \frac{\partial U_i(x_i)}{\partial p_i} + \sum_{j \neq i} \frac{\partial U_j(x_j)}{\partial p_i} = 0, \quad (12)$$

for  $i \in N$ . The above condition holds only when the optimal is achieved for transmission probabilities in the interior of  $[0, 1]$ . Experiments show that this is indeed the case for utility functions we have considered and parameter values that correspond to IEEE 802.11.

In the following sections we first consider the case where all wireless stations have the same transmission rate, and then the case where different stations can have different transmission rates; for the latter, we have two subcases: in the first, only the successful transmission interval  $T^{suc}$  depends on the transmission rate, whereas in the second both the successful transmission interval  $T^{suc}$  and the collision interval  $T^{col}$  depends on the transmission rate.

### 4.1 Equal transmission rates

In this case the throughput for a wireless station is given by (8). Substituting this equation in (12), and after some mathematical manipulations<sup>2</sup> where we compute the partial derivative  $\frac{\partial U_j(x_j)}{\partial p_i}$ , we find that the necessary conditions for the global optimum are approximately

$$\frac{\partial U_i(x_i)}{\partial p_i} = L \frac{(1 - P)^2 T^{suc} + P(2 - P)T^{col}}{E[T]^2} \sum_j U_j' p_j, \quad (13)$$

for  $i \in N$ , where  $P = \sum_j p_j$  and  $E[T]$  is given by (7); if  $p_i \ll P$ , which will hold when there is a large number of stations,  $E[T]$  can be approximated by

$$E[T] = P(1 - P)T^{suc} + P^2 T^{col} + 1 - P.$$

In order to solve the global optimization problem in a distributed manner, we need to first define the following user problem for station  $i$

$$\begin{aligned} & \text{maximize} && U_i(x_i) - \lambda p_i \\ & \text{over} && p_i \geq 0, \end{aligned} \quad (14)$$

where  $\lambda$  is the shadow (or congestion) price. The necessary condition for the user maximization problem is

$$\frac{\partial U_i(x_i)}{\partial p_i} = \lambda.$$

From the last equation and (13), the user and global problems coincide if we define  $\lambda$  as

$$\lambda = L \frac{(1 - P)^2 T^{suc} + P(2 - P)T^{col}}{E[T]^2} \sum_j U_j' p_j. \quad (15)$$

<sup>2</sup>The full proofs and mathematical derivations will appear in an extended version of this paper.

In (14) the congestion price is in terms of the transmission probability  $p_i$ . Application of the proposed model may require that the congestion price is defined in terms of the achieved throughput; this is the case if the congestion cost is signalled using a mechanism such as ECN marking, which we discuss in Section 5.2. Substituting (8) in (14) and combining the result with (15), we have the following expression for the user problem

$$\begin{aligned} & \text{maximize} && U_i(x_i) - \mu x_i \\ & \text{over} && p_i \geq 0, \end{aligned} \quad (16)$$

where, for the user and global problems to coincide, the shadow price  $\mu$  is defined as

$$\mu = \frac{(1-P)^2 T^{suc} + P(2-P)T^{col}}{(1-P)E[T]} \sum_j U'_j p_j. \quad (17)$$

Computing  $\mu$  directly using the above equation requires knowing the utility of all users. In Section 5.2 we discuss how the congestion price can be determined by an access point without this knowledge.

It is interesting to note that the above model differs from other models for wireless networks, such as CDMA networks [13, 16, 19], in that there is no resource capacity constraint. The shadow price is due solely to the congestion, in terms of the cost due to frame collisions.

When all users have the same utility function, then from (16) we have  $U' = \mu$ , which together with (17) gives the following necessary condition for optimality

$$P = \frac{(1-P)E[T]}{(1-P)^2 T^{suc} + P(2-P)T^{col}}, \quad (18)$$

hence the optimal aggregate transmission probability  $P$  is independent of the utility function. The above equation can be solved analytically, in which case we find

$$P = \frac{\sqrt{T^{col}} - 1}{T^{col} - 1}. \quad (19)$$

The last equation suggests the following interesting result: For a large number of stations, all having the same utility, when the aggregate transmission probability is much smaller than one, then the optimal aggregate transmission probability that maximizes the economic efficiency of the wireless channel is independent of the specific utility function and the successful transmission interval, and depends only on the collision interval normalized to the idle time slot.

#### 4.2 Different transmission rates: $T^{suc}$ depends on transmission rate and $T^{col}$ is independent of transmission rate

Next we consider the case where different wireless stations have different transmission rates, which affect the

duration of the successful transmission interval  $T^{suc}$ , but not the duration of the collision interval  $T^{col}$ . Recall that this is the case of 802.11b, when the RTS/CTS procedure is used. The throughput for station  $i$  is now given by (9). Substituting (9) in (12), and after some manipulations, we find that the necessary conditions for the global optimization problem are approximately

$$\frac{\partial U_i(x_i)}{\partial p_i} = L \frac{(1-P)^2 T_i^{suc} + P(2-P)T^{col}}{E[T]^2} \sum_j U'_j p_j, \quad (20)$$

for  $i \in N$ , where  $P = \sum_j p_j$  and  $E[T]$  is given by (7) for different successful transmission intervals  $T_i^{suc}$ . When  $p_i \ll P$ ,  $E[T]$  can be approximated by

$$E[T] = (1-P) \sum_j p_j T_j^{suc} + P^2 T^{col} + 1 - P.$$

As in the previous subsection, in order to solve the global optimization problem in a distributed manner, we define the following user problem

$$\begin{aligned} & \text{maximize} && U_i(x_i) - (\lambda_1 T_i^{suc} + \lambda_2) p_i \\ & \text{over} && p_i \geq 0, \end{aligned} \quad (21)$$

where we see that now the congestion price can be different for different users, and depends on factors  $\lambda_1$ ,  $\lambda_2$ , and the successful transmission interval  $T_i^{suc}$ . The necessary condition for the user maximization problem is

$$\frac{\partial U_i(x_i)}{\partial p_i} = \lambda_1 T_i^{suc} + \lambda_2.$$

From the last equation and (20), the user and global problems coincide if we define  $\lambda_1, \lambda_2$  as

$$\begin{aligned} \lambda_1 &= \frac{L(1-P)^2}{E[T]^2} \sum_j U'_j p_j \\ \lambda_2 &= \frac{LP(2-P)T^{col}}{E[T]^2} \sum_j U'_j p_j. \end{aligned} \quad (22)$$

In (21) the congestion price is in terms of the transmission probability  $p_i$ . Substituting (9) in (21), and combining the result with (22) we have the following expression for the user problem

$$\begin{aligned} & \text{maximize} && U_i(x_i) - (\mu_1 T_i^{suc} + \mu_2) x_i \\ & \text{over} && p_i \geq 0, \end{aligned} \quad (23)$$

where, for the user and global problems to coincide,  $\mu_1, \mu_2$  are defined as

$$\begin{aligned} \mu_1 &= \frac{(1-P)}{E[T]} \sum_j U'_j p_j \\ \mu_2 &= \frac{P(2-P)T^{col}}{(1-P)E[T]} \sum_j U'_j p_j. \end{aligned} \quad (24)$$

The congestion price in (23) contains two components: The first component  $\mu_1 T_i^{suc}$  contains the factor  $\mu_1$  which depends on the level of congestion in the wireless channel, and the duration of a successful transmission  $T_i^{suc}$ . The second component  $\mu_2$  is related to the level of congestion in the wireless channel. The interpretation of the above is that the congestion cost for a wireless station depends, in addition to its throughput, also on the duration of the successful transmission interval, which in turn depends on the station's transmission rate. For stations with the same throughput, the station with the lower transmission rate, which will have a longer successful transmission interval, will induce a higher congestion cost; this can be explained by noting that a higher successful transmission interval results in the wireless channel being occupied for a longer time. The relative importance of the successful transmission interval is determined by the ratio  $\mu_1/\mu_2$ .

### 4.3 Different transmission rates: Both $T^{suc}, T^{col}$ depend on transmission rate

Next we consider the case where different wireless stations have different transmission rates, which affect both the duration of the successful transmission interval  $T^{suc}$  and the duration of the collision interval  $T^{col}$ . The throughput for station  $i$  is now given by (10). Substituting (10) in (12), and after some manipulations, we find that the necessary conditions for the global optimization problem are approximately

$$\frac{\partial U_i(x_i)}{\partial p_i} = L \frac{(1-P)^2 T_i^{suc} + 2P(1-P)E_i[T^{col}] + P^2 E[T^{col}]}{E[T]^2} \sum_j U'_j p_j, \quad (25)$$

for  $i \in N$ , where  $E[T^{col}]$  is the average collision interval and  $E_i[T^{col}]$  is the average collision interval when a frame from station  $i$  is involved in the collision. If the percentage of stations transmitting at different rates, e.g. 11, 5.5, and 2 Mbps are known and equal to  $\phi_{11}, \phi_{5.5}$ , and  $\phi_2$  respectively, then  $E_i[T^{col}]$  can be approximated by ( $R_i$  is the transmission rate for station  $i$ )

$$E_i[T^{col}] = \begin{cases} \phi_{11} T_{11}^{col} + \phi_{5.5} T_{5.5}^{col} + \phi_2 T_2^{col} & \text{if } R_i = 11 \text{ Mbps} \\ (\phi_{11} + \phi_{5.5}) T_{5.5}^{col} + \phi_2 T_2^{col} & \text{if } R_i = 5.5 \text{ Mbps} \\ T_2^{col} & \text{if } R_i = 2 \text{ Mbps} \end{cases}$$

and  $E[T^{col}]$  can be approximated by

$$E[T^{col}] = \phi_{11} E_{11}[T^{col}] + \phi_{5.5} E_{5.5}[T^{col}] + \phi_2 E_2[T^{col}],$$

where  $T_x^{col}$  is the duration of the collision interval for a station with transmission rate  $x$ , and  $E_x[T^{col}]$  is the average collision interval when a frame from a station with transmission rate  $x$  is involved in the collision.

Following a similar approach with the previous two subsections, we can define the user problem

$$\begin{aligned} & \text{maximize} && U_i(x_i) - (\lambda_1 T_i^{suc} + \lambda_2 E_i[T^{col}] + \lambda_3) p_i \\ & \text{over} && p_i \geq 0. \end{aligned} \quad (26)$$

The necessary condition for the user maximization problem is

$$\frac{\partial U_i(x_i)}{\partial p_i} = \lambda_1 T_i^{suc} + \lambda_2 E_i[T^{col}] + \lambda_3.$$

From the last equation and (25), the user and global problems coincide if we define  $\lambda_1, \lambda_2, \lambda_3$  as

$$\begin{aligned} \lambda_1 &= \frac{L(1-P)^2}{E[T]^2} \sum_j U'_j p_j \\ \lambda_2 &= \frac{L2P(1-P)}{E[T]^2} \sum_j U'_j p_j \\ \lambda_3 &= \frac{LP^2 E[T^{col}]}{E[T]^2} \sum_j U'_j p_j. \end{aligned} \quad (27)$$

In (21) the congestion price is in terms of the transmission probability  $p_i$ . Substituting (9) in (26), and combining the result with (27) we have the following expression for the user problem

$$\begin{aligned} & \text{maximize} && U_i(x_i) - (\mu_1 T_i^{suc} + \mu_2 E_i[T^{col}] + \mu_3) x_i \\ & \text{over} && p_i \geq 0, \end{aligned} \quad (28)$$

where, for the user and global problems to coincide,  $\mu_1, \mu_2, \mu_3$  are defined as

$$\begin{aligned} \mu_1 &= \frac{(1-P)}{E[T]} \sum_j U'_j p_j \\ \mu_2 &= \frac{2P}{E[T]} \sum_j U'_j p_j \\ \mu_3 &= \frac{P^2 E[T^{col}]}{(1-P)E[T]} \sum_j U'_j p_j. \end{aligned} \quad (29)$$

From the above analysis, similar to the case where only the successful transmission probability was affected by the transmission rate, the congestion price for a station depends on the level of congestion in the wireless channel and on the station's transmission rate.

## 5 Application

In this section we investigate the application of the models presented in the previous section to class-based throughput differentiation, to ECN marking, and for investigating the performance of TCP over EDCA. Due to space limitations, we focus mainly on the model presented in Section 4.2.

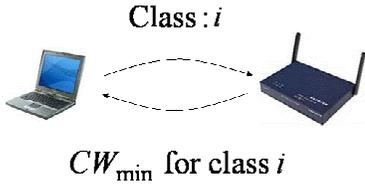


Figure 3: Class-based service differentiation. Stations select a class during their association with an access point. The access point computes the minimum contention window for each class, and communicates them to the wireless stations.

### 5.1 Class-based differentiation with proportional sharing

In this section we apply the model presented in the previous section to the case of proportional resource sharing [11], where resources are shared in proportion to weights or willingness-to-pay factors; different service classes can be assigned different such weights. A wireless station can select a class, e.g. during its association with an access point. Different classes would correspond to different values of the minimum contention window. The values of the minimum contention window can be computed at the access point and, as indicated in the 802.11e standard, periodically communicated to the wireless stations using beacon frames, Figure 3.

In the case of proportional sharing, the utility for user  $i$  is given by [11]

$$U_i(x_i) = w_i \log x_i, \quad (30)$$

where  $w_i$  is the weight or willingness-to-pay factor. Based on (9), the throughput  $x_j$  for station  $j$  can be written as

$$x_j = \frac{p_j(1-P)L}{E[T]}. \quad (31)$$

With the above, the sum  $\sum_j U'_j p_j$  in (24) is computed as follows:

$$\sum_j U'_j p_j = \sum_j \frac{w_j}{x_j} p_j = \frac{E[T]}{(1-P)L} \sum_j w_j.$$

Substituting the last equation in (24), and after some manipulations, we have the following expressions for  $\mu_1, \mu_2$ :

$$\begin{aligned} \mu_1 &= \frac{\sum_j w_j}{L} \\ \mu_2 &= \frac{P(2-P)T^{col}}{(1-P)^2} \frac{\sum_j w_j}{L}. \end{aligned} \quad (32)$$

From (23) and (30) we have

$$\frac{\partial U_i(x_i)}{\partial x_i} = \mu_1 T_i^{suc} + \mu_2 \Rightarrow x_i = \frac{w_i}{\mu_1 T_i^{suc} + \mu_2}.$$

Substituting (9) and (32) in the last equation we obtain

$$p_i = \frac{w_i}{\sum_j w_j} \frac{(1-P)E[T]}{(1-P)^2 T_i^{suc} + P(2-P)T^{col}}. \quad (33)$$

The last expression can be used to compute the optimal transmission probabilities for the wireless stations, based on the willingness-to-pay factors, the successful transmission duration for each station  $T_i^{suc}$ , and the collision duration  $T^{col}$ . The optimal minimum contention window  $CW_{min,i}$  can then be computed from (1).

If all stations have the same transmission rate, then (33) reduces to (18), hence the optimal aggregate transmission probability  $P$  should satisfy (19). Note that this is true even when different stations have different weights, provided the utility is logarithmic. In this case the transmission probability for a station  $i$  is given by

$$p_i = \frac{w_i}{\sum_j w_j} P. \quad (34)$$

As before, the optimal minimum contention window  $CW_{min,i}$  can be computed from (1). The above model for the case of identical transmission rates is similar to the one presented in [2]; the two approaches will be compared in Section 6.2.

When the number of stations is large the above system becomes incentive compatible, i.e. each station obtains a higher net benefit by declaring its true value of  $w_i$ . Suppose that  $w_i$  is this true value with  $x_i$  being the resulting throughput, and  $\tilde{w}_i, \tilde{x}_i$ , are some corresponding non-truthful values. Then for incentive compatibility we need

$$w_i \log x_i - w_i \geq \tilde{w}_i \log \tilde{x}_i - \tilde{w}_i.$$

But if  $\mu, \tilde{\mu}$  are the congestion prices when all stations are truthful and when all stations but station  $i$  are truthful respectively, we have that  $x_i = w_i/\mu$  and  $\tilde{x}_i = \tilde{w}_i/\tilde{\mu}$ . Now our incentive compatibility condition becomes

$$\log(\tilde{w}_i/w_i) + \log(\tilde{\mu}/\mu) \leq \tilde{w}_i/w_i - 1,$$

which holds when the number of stations is large, since in this case  $\tilde{\mu}/\mu$  tends to 1.

### 5.2 ECN marking

In this section we assume that stations receive congestion feedback according to (24), and respond to it by selecting the optimal transmission probability, hence minimum contention window, according to (23). However,

the application of (24) requires that the access point knows the utility for all users. An alternative is to use a tatonnement process as follows: the components  $\mu_1$  and  $\mu_2$  are expressed as a function of a factor  $M$ ,

$$\begin{aligned}\mu_1 &= (1 - P)M \\ \mu_2 &= \frac{P(2 - P)T^{col}}{(1 - P)}M.\end{aligned}\quad (35)$$

The factor  $M$  is increased or decreased in the direction where the aggregate throughput increases. From  $\mu_1, \mu_2$ , the access point can compute the following congestion price  $\mu_i$  for each user  $i$ :

$$\mu_i = \mu_1 T_i^{suc} + \mu_2. \quad (36)$$

If all stations have the same transmission rate then, according to the model presented in Section 4.1, the feedback  $\mu$  is the same for all stations; hence,  $\mu = M$ , where as before the factor  $M$  is increased or decreased in the direction where the aggregate throughput increases.

Equation (35) requires estimation of the aggregate transmission probability  $P$ . Recall from (8) that the throughput is proportional to the transmission probability, hence the aggregate transmission probability  $P$  can be estimated from

$$P = \frac{X \cdot p_{AP}}{X_{AP}},$$

where  $X$  is the aggregate throughput, i.e. the sum of the throughput in the uplink and the downlink,  $X_{AP}$  is the transmission throughput from the access point to the wireless stations, i.e. the downlink throughput, and  $p_{AP}$  is the transmission probability at the access point. Both  $X_{AP}$  and  $X$  can be measured at the access point.

The congestion price given by (36) can be communicated to the wireless stations either explicitly, by piggy-backing it on beacon frames that are periodically sent by the access point, or using ECN marking, Figure 4. Packet marking using the ECN bit in the IP header is a communication mechanism for signalling to the end systems the level of congestion in the network. ECN can be used for conveying congestion information from both the wireless network and the wired network, hence achieving seamless congestion control over different network technologies.

When transmission rates differ, from (23) and (35) the marking probability  $\mathcal{P}_i$  for station  $i$  should satisfy

$$\mathcal{P}_i = \frac{\mu_1 T_i^{suc} + \mu_2}{p_{mrk}}.$$

The value of  $p_{mrk}$  (price-per-mark) should be such that for the range of user demands expected, the marking probability  $\mathcal{P}_i$  is in the range  $[0, 1]$ .

From the above, the marking probability for station  $i$  should depend, in addition to the level of congestion in

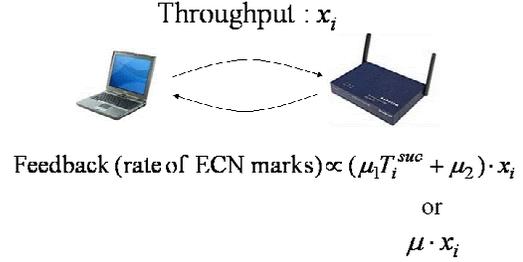


Figure 4: ECN marking feedback. The rate of marks is proportional to the congestion price, which is different for stations with different transmission rates, and the station's packet rate.

the wireless channel, also on station  $i$ 's transmission rate, which determines  $T_i^{suc}$ . The reason for this dependence, as discussed in Section 4.2, is that a lower transmission rate results in a longer frame transmission time, hence the wireless channel is occupied for longer time.

### 5.3 TCP over EDCA

In this section we assume that the ECN marking scheme described in the previous subsection is applied, and investigate the performance of the resource control model presented in Section 4 when users respond in a manner similar to TCP's congestion control algorithm. Indeed, TCP congestion control can be viewed as having the following implicit utility [12]:

$$U_{TCP}(x) = -\frac{2}{RTT^2 x}, \quad (37)$$

where  $RTT$  is the round trip time. If we assume that all stations encounter the same  $RTT$  and consider that the throughput  $x_i$  is given by (31), then the sum  $\sum_j U_j' p_j$  is computed as follows:

$$\begin{aligned}\sum_j U_j' p_j &= \sum_j \frac{2}{RTT^2 x_j^2} p_j = \frac{2E[T]}{RTT^2 L(1 - P)} \sum_j \frac{1}{x_j} = \\ &= \frac{2E[T]^2}{RTT^2 L^2(1 - P)^2} \sum_j \frac{1}{p_j}.\end{aligned}$$

Substituting the last equation in (24), and after some manipulations, we have the following:

$$\begin{aligned}\mu_1 &= \frac{2E[T]}{RTT^2 L^2(1 - P)} \sum_j \frac{1}{p_j} \\ \mu_2 &= \frac{2P(2 - P)E[T]T^{col}}{RTT^2 L^2(1 - P)^3} \sum_j \frac{1}{p_j}.\end{aligned}\quad (38)$$

From (23) and (37) we have

$$\frac{\partial U_i(x_i)}{\partial x_i} = \mu_1 T_i^{suc} + \mu_2 \Rightarrow \frac{2}{RTT^2 x_i^2} = \mu_1 T_i^{suc} + \mu_2.$$

Substituting (31) and (38) in the last equation we obtain

$$p_i = \sqrt{\frac{(1-P)E[T]}{(1-P)^2T_i^{suc} + P(2-P)T^{col}} \frac{1}{\sum \frac{1}{p_j}}}. \quad (39)$$

The last expression can be used to compute the optimal transmission probabilities for the stations, based on the successful transmission duration for each station  $T_i^{suc}$  and the collision duration  $T^{col}$ . The optimal minimum contention window  $CW_{min,i}$  can be computed from (1).

If all stations have the same transmission rate, then the last equation reduces to (18), which is identical to the corresponding equation for proportional sharing. Indeed, as discussed in Section 4.1, when all stations have the same utility and the same transmission rate, then the optimal aggregate transmission probability is independent of the utility, and is given by (19).

Above we assumed that all TCP flows had the same round trip time. If this is not the case, and  $RTT_i$  is the round trip time for flow  $i$  and assuming that each station has one flow, then (39) becomes

$$p_i = \frac{1}{RTT_i} \sqrt{\frac{(1-P)E[T]}{(1-P)^2T_i^{suc} + P(2-P)T^{col}} \frac{1}{\sum \frac{1}{RTT_j^2 p_j}}}$$

If the transmission rate for all stations is the same, then from the last equation we have that the optimal transmission probability is inversely proportional to the round trip time. If a station has multiple TCP flows, then the last equation could be used to estimate the transmission probability for each flow; the station's total transmission probability, which determines its minimum contention window, will be the sum of these probabilities.

## 6 Experiments

In this section we present experimental results that demonstrate and evaluate the models presented in the previous two sections, and in particular

- we investigate how various factors such as the CSMA/CA and RTS/CTS procedures, the utility function, and the network load affects the level of congestion,
- we compare simulation and analytical results for estimating the aggregate throughput and for estimating the optimal minimum contention window,
- we investigate the gains in using the optimal minimum contention window, and

Table 1: Parameters of IEEE 802.11b

Parameter	Value (in $\mu s$ )
Slot Time	20
$T_{DIFS}, T_{SIFS}$	50, 10
$T_{PHY}$	192
$T_{ACK} = T_{CTS}, T_{RTS}$	112, 160

- we investigate a simple model capturing the closed-loop interaction between the access point and the wireless stations, when the latter select their transmission probability (equivalently, the minimum contention window).

The simulation results were obtained using the ns-2 simulator, with the module<sup>3</sup> documented in [21] for implementing the EDCA mechanism. The simulation experiments considered UDP traffic with packet size 1044 bytes, which includes the UDP/IP headers, and the results are the average of 3 runs, each for 300 seconds.

The parameters of IEEE 802.11b required to compute the intervals  $T^{suc}$ ,  $T^{col}$ , and  $T^{idl}$  are shown in Table 1. The parameters for IEEE 802.11a are shown in Table 2. To compute the time intervals for IEEE 802.11a we need to modify (2)-(5). In particular, the duration of the transmission of a data frame in IEEE 802.11a is [10]

$$T_{DATA} = 4 \left\lceil \frac{22 + 8(34 + L)}{N_{DBPS}} \right\rceil,$$

where  $N_{DBPS}$  is 24, 48, 96, 216 for OFDM-6, OFDM-12, OFDM-24, and OFDM-54 respectively. Moreover,  $T_{ACK}$ ,  $T_{RTS}$ , and  $T_{CTS}$  in 802.11a can be computed from

$$T_{ACK} = T_{CTS} = 4 \left\lceil \frac{22 + 8 \cdot 14}{N_{DBPS}} \right\rceil, \quad T_{RTS} = 4 \left\lceil \frac{22 + 8 \cdot 20}{N_{DBPS}} \right\rceil.$$

Note that ns-2 does not implement the specific procedure of 802.11a for handling the physical layer overhead, hence in Section 6.2 where we compare simulation with analytical results we consider a slightly different definition of the above time intervals.

### 6.1 Impact of various factors on the level of congestion

We first consider the model presented in Section 4.2, focusing solely on how various factors affect the level of congestion and not considering the closed-loop interaction between the network and the users (wireless stations); we investigate the latter in Section 6.4. Figure 5(a) shows the two congestion price components

<sup>3</sup>“An IEEE 802.11e EDCF and CFB simulation model for ns-2”, [http://www.tkn.tu-berlin.de/research/802.11e\\_ns2/](http://www.tkn.tu-berlin.de/research/802.11e_ns2/)

Table 2: Parameters of IEEE 802.11a

Parameter	Value (in $\mu\text{s}$ )
Slot Time	9
$T_{DIFS}, T_{SIFS}$	34, 16
$T_{PHY}$	20

$\mu_1 T^{suc}$  and  $\mu_2$  in (23), as a function of the minimum contention window  $CW_{min}$ , when users have a logarithmic utility, in which case  $\mu_1, \mu_2$  are given by (32). Observe that the first component  $\mu_1 T^{suc}$  does not depend on  $CW_{min}$ ; this is due to the logarithmic form of the utility, and is not the case for a TCP-like utility as we will see later. On the other hand,  $\mu_2$  increases as  $CW_{min}$  decreases, i.e. as the level of contention for the wireless channel increases. Also observe in Figure 5(a) that  $\mu_1 T^{suc}$  is larger when the RTS/CTS procedure is used, whereas  $\mu_2$  is smaller; both of these behaviors are due to the fact that  $T^{suc}$  and  $T^{col}$  increases and decreases, respectively, when the RTS/CTS procedure is used.

Next we consider the model in Section 5.3, where stations react in a manner similar to the TCP congestion control algorithm. Figure 5(b) shows the two congestion price components  $\mu_1 T^{suc}$  and  $\mu_2$ , given by (38). Observe that now  $\mu_1 T^{suc}$  increases for smaller values of  $CW_{min}$ .

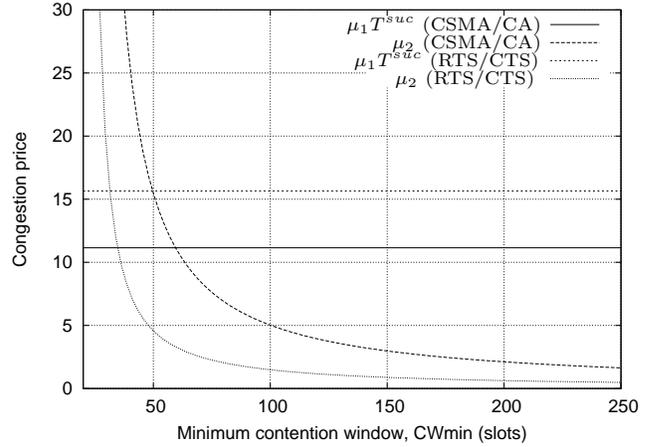
Finally, Figure 6 shows that, when  $CW_{min}$  is constant, both  $\mu_1 T^{suc}$  and  $\mu_2$  increase with the number of wireless stations. Moreover, observe that the increase of  $\mu_2$  is higher compared to that of  $\mu_1 T^{suc}$ , hence the ratio  $\mu_1 T^{suc} / \mu_2$  decreases as the number of stations increases.

## 6.2 Comparison of simulation and analytical results

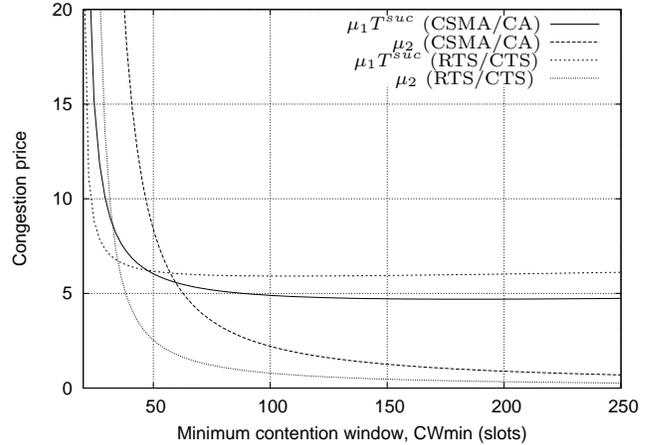
Figure 7 compares for IEEE 802.11b<sup>4</sup> (11 Mbps) with CSMA/CA, and in the case all stations have the same transmission rate, the aggregate throughput estimated from (8) with the value found using simulation. Figure 8 shows the corresponding results for 802.11a<sup>5</sup> at 24 Mbps. Observe that the analytical results follow the simulation results very well. Observe that in both Figures 7 and 8 the analytical model underestimates the throughput for small values of  $CW_{min}$ . This is due to the p-persistent model that we considered for obtaining the analytical expression of the throughput; this model does not capture the exponential backoff procedure of

<sup>4</sup>For the analytical results in Figure 7 we have considered 28 bytes MAC overhead, since this is the size used in ns-2.

<sup>5</sup>Due to different handling of the physical layer headers in 802.11a, compared to 802.11b which is the model implemented in ns-2, for Figure 8 we consider the intervals  $T_{PHY} = 16 + 40/6 \mu\text{s}$ ,  $T_{DATA} = 4 \left\lceil \frac{8(28+L)}{N_{DBPS}} \right\rceil$ ,  $T_{ACK} = 4 \left\lceil \frac{8 \cdot 14}{N_{DBPS}} \right\rceil$ ; see also [21].



(a) Logarithmic utility



(b) TCP-like utility

Figure 5: Effect of CSMA/CA and RTS/CTS on congestion for different  $CW_{min}$ .  $C = 11$  Mbps,  $N = 10$ 

802.11, which can improve the throughput in cases of high contention; such high contention occurs for small values of  $CW_{min}$ . Nevertheless, what is important for the work in this paper is that there is good agreement between the optimal  $CW_{min}$  estimated using analysis and using simulation. Hence, Table 3 compares the simulation with the analytical approach based on (8) for estimating the optimal  $CW_{min}$ , when the latter obtains discrete values that are powers of 2. Observe that both the analytical and simulation-based selection of the optimal  $CW_{min}$  agree, except for the case  $C = 24$  Mbps and  $N = 10$  users. Nevertheless, note that using in this case the value 128 indicated by the analysis, rather than the value 64 indicated by the simulation, yields a 3% difference in the aggregate throughput.

It is important to note here that using a minimum contention window smaller than the optimal does not only

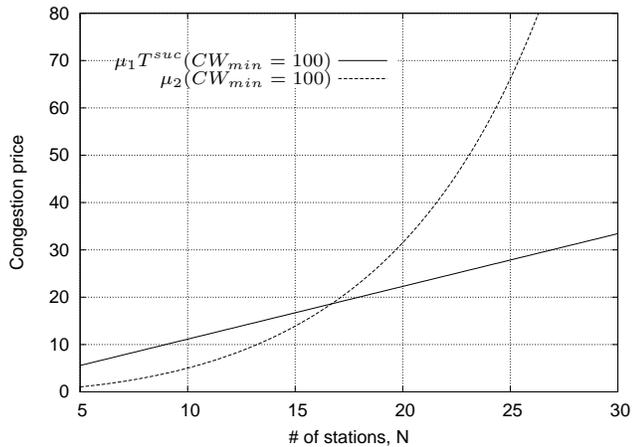


Figure 6: Effect of number of stations on congestion for logarithmic utility.  $C = 11$  Mbps,  $CW_{min} = 100$

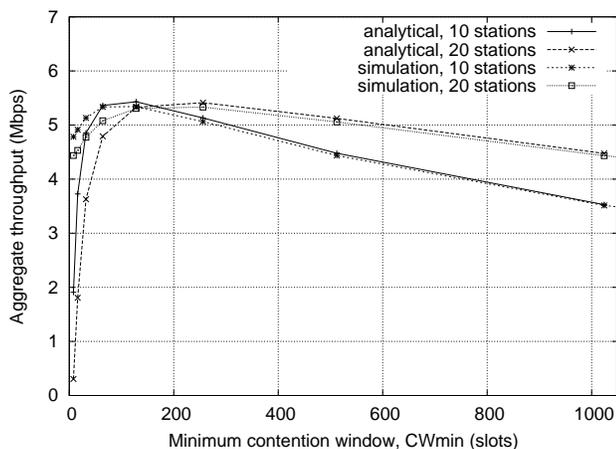


Figure 7: Simulation and analytical results for 11 Mbps.

have an impact on the aggregate throughput, which Figures 7 and 8 show to be a reduction of up to about 15%. Indeed, a smaller  $CW_{min}$ , especially when there is a large number of stations, results in higher channel contention, which can lead to short term unfairness [3]; this occurs because after a collision, the station that has successfully transmitted a frame has a higher probability of transmitting subsequent frames, since it would have decreased its contention window to the minimum value (this is referred to as the *channel capture effect*).

Table 4 compares, for the case where all wireless stations have the same transmission probability, the analytical expression for estimating the transmission probabilities proposed in this paper using (19) and (34), with the approach proposed in [2]. The results in Table 4 show that the optimal values of  $CW_{min}$  estimated using the two approaches are very close.

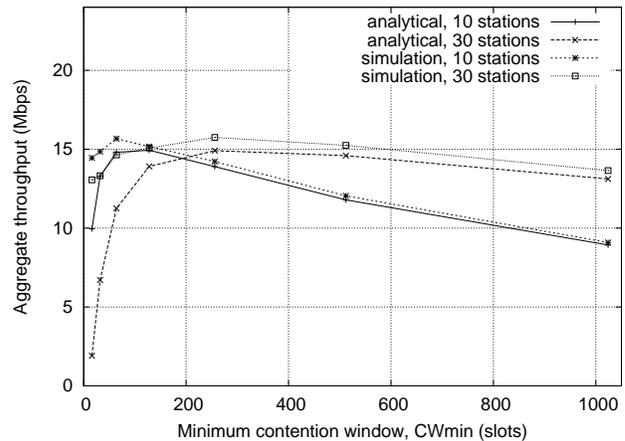


Figure 8: Simulation and analytical results for 24 Mbps.

Table 3: Optimal  $CW_{min}$  (discrete values, powers of 2) based on analysis and simulation, for IEEE 802.11b (11 Mbps) and IEEE 802.11a (24 Mbps)

C	N	analysis	simulation
11	10	128	128
11	20	256	256
24	10	128	64
24	30	256	256

### 6.3 Gains in using the optimal $CW_{min}$

Next we investigate the gains, in terms of the aggregate throughput, in using the optimal minimum contention window  $CW_{min}$  for a different number of wireless stations, for the same number of stations but with a different mix (stations with different weights), and for stations with different transmission rates. We consider the weight-based proportional sharing model in Section 5.1. The aggregate throughput results presented in this subsection were obtained using (8), when all transmission rates are equal, and (9), when transmission rates differ.

Table 4: Optimal transmission probability  $p$  and  $CW_{min}$  for IEEE 802.11b (11 Mbps) and 802.11a (24 Mbps).

C	N	Access	Eq. (19)		Approach in [2]	
			$p$	$CW_{min}$	$p$	$CW_{min}$
11	10	CSMA/CA	0.0123	162	0.0130	153
11	10	RTS/CTS	0.0182	109	0.0194	102
11	20	CSMA/CA	0.0061	325	0.0063	316
11	20	RTS/CTS	0.0091	218	0.0094	212
24	10	CSMA/CA	0.0128	155	0.0136	153
24	30	CSMA/CA	0.0043	467	0.0044	458

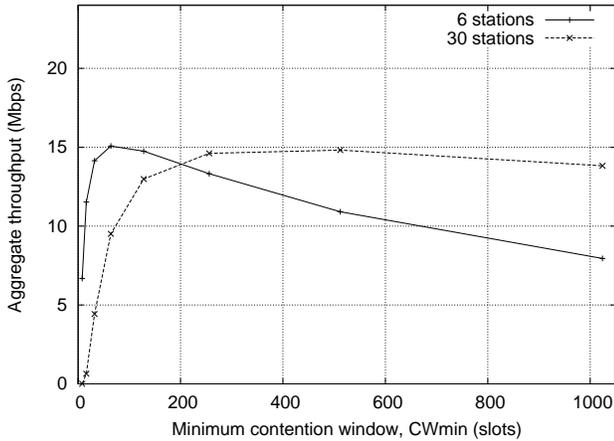


Figure 9: Throughput as a function of  $CW_{min}$  for 6 and 30 stations. Percentage of “hi” and “low” class stations is 1/3 and 2/3,  $w_{hi} = 2$ ,  $w_{low} = 1$ ,  $C = 24$  Mbps.

Figure 9 shows the aggregate throughput in the case of 6 and 30 stations. Observe that the selection of  $CW_{min}$  has a significant impact on the aggregate throughput. Hence, the optimal  $CW_{min}$  for 30 stations is 512. If this value is also used for 6 stations, then the aggregate throughput would be less by approximately 27%, compared to the case where the optimal  $CW_{min}$  for 6 users is used, which is 64.

Figure 10 shows the aggregate throughput in the case of 30 stations, and for a different mix of “hi” and “low” users, with weights 3 and 1 respectively. Observe that the optimal  $CW_{min}$  in the case of 28 stations with weight 3 and 2 stations with weight 1 is 1024. If the same value is used in the case of 2 stations with weight 3 and 28 stations with weight 1, then the throughput would be less by approximately 10%, compared to the case where the corresponding optimal  $CW_{min}$  is used, which is 256.

Table 5 shows the optimal  $CW_{min}$  when stations have different transmission rates; we assume that all stations have the same weight factor. Observe that  $CW_{min}$  is smaller for stations with a higher transmission rate. The table also shows that if we do not take into account the different transmission rates, but rather consider that all

Table 5: Optimal  $CW_{min}$  and throughput  $R$  for IEEE 802.11b with RTS/CTS when stations have different transmission rates (11 Mbps and 2 Mbps).

Eq. (33)		Eq. (19)		
$N_{11}/N_2$	$CW_{min,11}/CW_{min,2}$	$R$	$CW_{min,11}/CW_{min,2}$	$R$
2/8	51/147	2.12	104/104	1.78
10/20	186/531	2.48	332/332	1.96

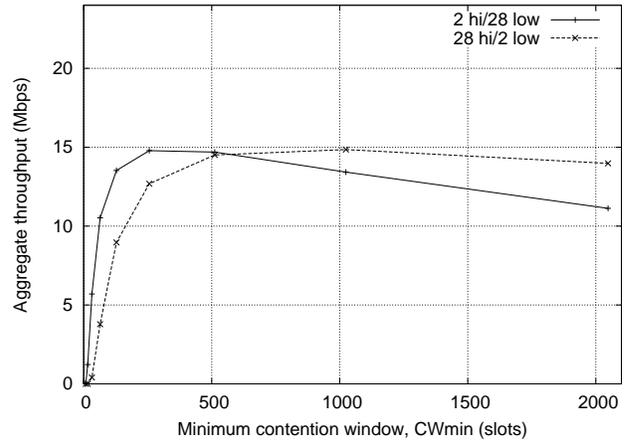


Figure 10: Throughput as a function of  $CW_{min}$  for 30 stations.  $w_{hi} = 3$ ,  $w_{low} = 1$ ,  $C = 24$  Mbps.

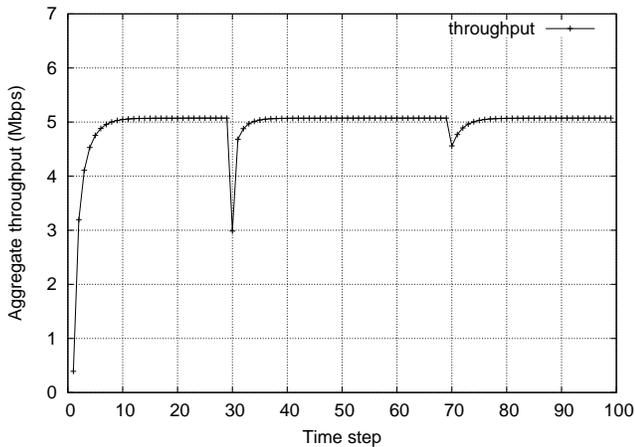
stations have the same  $CW_{min}$  with the optimal aggregate transmission probability satisfying (19), then the aggregate throughput is much lower.

The experiments of this subsection demonstrate that there are important gains in using the optimal minimum contention window, which depends on the number of stations, and their weights and transmission probabilities.

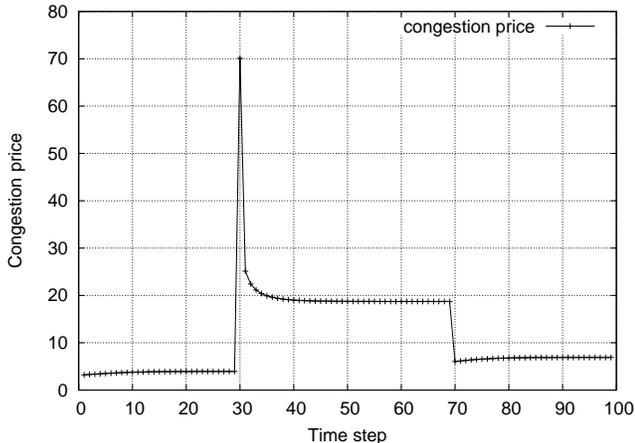
#### 6.4 Closed-loop interaction between access point and wireless stations

In this subsection we investigate the closed-loop interaction between the access point and wireless stations, when prices are explicitly signalled to the stations. We assume that one user corresponds to each station, and has a logarithmic utility (30). The access point estimates the congestion price from (15) and communicates it to the end users, which respond by selecting their transmission probabilities to maximize their net benefit (14). We assume that the access point can accurately estimate the aggregate transmission probability used in (15).

We consider transmission capacity  $C = 11$  Mbps, and channel access according to CSMA/CA. The frame size is  $L = 1044$  bytes. We consider two types of users, with weights  $w_{hi} = 3$  and  $w_{low} = 1$  respectively. Initially, there are 5 users of each type. Time is assumed to be discrete. At time  $t_1 = 30$ , we assume that 25 more users of type “hi” enter the network, hence giving a total of 30 users of type “hi”. Then, at time  $t_2 = 70$ , 20 users of type “hi” depart the system, leaving 10 users of type “hi”, and 5 users of type “low”. Figure 11(a) shows the aggregate throughput, computed using (8), as a function of time. Observe that the system reacts to changes of the number of users, quickly reaching the equilibrium, where the aggregate throughput achieves its maximum value.



(a) Throughput



(b) Congestion price

Figure 11: Throughput and price as a function of time. Initially  $N_{hi} = N_{low} = 5$ , at time  $t = 30$ ,  $N_{hi} = 30$ , and at time  $t = 70$ ,  $N_{hi} = 10$ .  $w_{hi} = 3$ ,  $w_{low} = 1$ ,  $C = 11$  Mbps

Figure 11(b) shows the behavior of the congestion price with time. As expected, the congestion price is higher when there are more users in the network.

The focus of the experiment presented in this subsection is solely on demonstrating the application of the framework presented in this paper, in a decentralized setting capturing a simplified model of the closed-loop interaction between the access point and wireless stations. The investigation of more complex and realistic scenarios, which include asynchronous operation of wireless stations, propagation delays, and measurement errors, is left for future work.

## 7 Related work

In this section we discuss other related work, identifying where it differs from the work presented in this paper. The work of [5] considers a p-persistent model and the work of [4] considers a two dimensional Markov chain model for the backoff procedure of 802.11. Based on this, the authors derive the optimal contention window for the backoff mechanism such that the total throughput is maximized. The models considered in [5, 4] for the backoff procedure are more detailed than the one considered in this paper, which is accurate when the number of stations is high and when the probability of a frame being involved in more than one collision is negligible; Nevertheless, the throughput model we consider still captures important characteristics of 802.11, while remaining simple enough to yield tractable calculations when used in an economic framework context. Additionally, our work differs from [5, 4] in that we consider maximizing the aggregate efficiency of the wireless network, hence takes into account the user requirements expressed with utility functions. Moreover, the above work assumes that all stations have the same transmission rate. Work on performance analysis of 802.11 with multiple transmission rates includes [9, 6, 7]: [9] focuses on estimating the achieved throughput, whereas [6] additionally considers the problem of admission control and [7] the issue of limiting the negative effect of slow transmission rate stations by having them transmit smaller frames.

The work of [18, 2] proposes an approach for achieving weighted fairness in IEEE 802.11e. The throughput model considered is similar to the model considered in our work. Moreover, the weighted fairness model is similar to the proportional sharing model considered in Section 5.1, albeit considering different approximations. The model presented in this paper is more general in that it considers the case where different wireless stations can have different transmission rates, and the general case where user requirements are encoded through utility functions; the latter enables the framework to be applied for investigating the performance of transport protocols (in this paper we consider TCP) over EDCA.

The work of [17] investigates the problem of fairness and weighted resource sharing in wireless networks using models based on utility functions, and proposes rate control schemes applied by end systems, to achieve the specific resource sharing model. Although some general characteristics of the wireless channel are taken into account, such a location-dependent contention and inaccurate channel state information, the specific characteristics of the 802.11 MAC operation are not considered.

The problem of service differentiation in IEEE 802.11 networks is investigated in a number of papers [20, 1, 14,

15]. These works do not quantify the degree of differentiation, nor do they relate it to some specific resource sharing model. Moreover, the above work typically does not consider the aggregate throughput that is achieved by the various schemes, and focuses solely on the individual station throughput. Both of these issues are a main focus of the work presented in this paper.

## 8 Conclusions

We presented a congestion pricing framework for efficient resource control of elastic traffic in IEEE 802.11e's EDCA mechanism that captures how various factors, such as the probability of attempting to transmit a frame, which is related to the value of the minimum contention window used in 802.11's collision avoidance procedure, the use of the CSMA/CA or the RTS/CTS procedure, and the physical layer transmission rate, affect congestion. The framework can be applied for achieving class-based proportional throughput differentiation, for performing ECN marking based on the level of congestion in the wireless channel, and for modelling the performance of TCP congestion control over EDCA.

Ongoing experimental work is investigating the application of the framework presented in this paper for TCP traffic, and in the presence of measurement errors and propagation delays. Ongoing research is investigating models for resource control when both the contention-based EDCA and the polling-based HCCA (HCF Controlled Channel Access) mechanisms coexist, consideration of other control variables (this paper considered the minimum contention window), and the extension to multihop wireless networks.

## Acknowledgments

This work has been supported in part by British Telecommunications, UK.

The authors would like to thank P. Alafouzou, D. Evangelatou, and S. Karagiorgou for assisting in the experiments reported in the paper.

## References

- [1] I. Aad and C. Castelluccia. Differentiation Mechanisms for IEEE 802.11. In *Proc. of IEEE INFOCOM'01*, April 2001.
- [2] A. Banchs, X. Perez-Costa, and D. Qiao. Providing Throughput Guarantees in IEEE 802.11e Wireless LANs. In *Proc. of the 18th International Teletraffic Congress (ITC - 18)*, North Holland, 2003. Elsevier Science B. V.
- [3] G. Berger-Sabbatel, A. Duda, O. Gaudoin, M. Heusse, and F. Rousseau. Fairness and its Impact on Delay in 802.11 Networks. In *Proc. of IEEE GLOBECOM*, November 2004.
- [4] G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE J. Select. Areas Commun.*, 18(3):535–547, March 2000.
- [5] F. Cali, M. Conti, and E. Gregori. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Trans. on Networking*, 8(6):785–799, 2000.
- [6] M. Ergen and P. Varaiya. Throughput Analysis and Admission Control in IEEE 802.11a. *Mobile Networks and Applications (MONET)*, ACM-Kluwer. To appear.
- [7] M. Ergen and P. Varaiya. Throughput Formulation and WLAN Optimization in Mixed Data Rates IEEE 802.11 DCF Mode. In *Proc. of Workshop on Computer-aided Modeling, Analysis, and Design of Commun. Links and Networks, held in conjunction with IEEE GLOBECOM'04*, November 2004.
- [8] D. Gu and J. Zhang. QoS Enhancements in IEEE 802.11 Wireless Local Area Networks. *IEEE Commun. Mag.*, pages 120–124, June 2003.
- [9] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. A Performance Anomaly of IEEE 802.11. In *Proc. of IEEE INFOCOM'03*, March 2003.
- [10] J. Jun, P. Peddabachagari, and M. Sichitiu. Theoretical Maximum Throughput of IEEE 802.11 and its Applications. In *Proc. of the 2nd IEEE Intl Symposium on Network Computing and Applications*, April 2003.
- [11] F. P. Kelly. Charging and Rate Control for Elastic Traffic. *European Trans. on Telecommunications*, 8:33–37, January 1997.
- [12] F. P. Kelly. Mathematical Modelling of the Internet. In B. Engquist and W. Schmid, editors, *Mathematics Unlimited - 2001 and Beyond*, Heidelberg, 2000. Springer-Verlag. Available at: <http://www.statslab.cam.ac.uk/~frank/>.
- [13] J. W. Lee, R. R. Mazumdar, and N. B. Shroff. Downlink Power Allocation for Multi-Class CDMA Wireless Networks. In *Proc. of IEEE INFOCOM'02*, June 2002.
- [14] A. Lindgren, A. Almquist, and O. Schelen. Quality of Service Schemes for IEEE 802.11 Wireless LANs An Evaluation. *Mobile Networks and Applications*, 8(3):223–235, June 2003.
- [15] S. Mangold, S. Choi, G. R. Hiertz, O. Klein, and B. Walke. Analysis of IEEE 802.11e for QoS Support in Wireless LANs. *IEEE Wireless Communications*, pages 40–50, Dec 2003.
- [16] P. Marbach and R. Berry. Downlink Resource Allocation and Pricing for Wireless Networks. In *Proc. of IEEE INFOCOM'02*, June 2002.

- [17] T. Nandagopal, T.-E. Kim, X. Gao, and V. Bharghavan. Achieving MAC Layer Fairness in Wireless Packet Networks. In *Proc. of ACM MOBICOM 2000*, August 2000.
- [18] D. Qiao and K. G. Shin. Achieving Efficient Channel Utilization and Weighted Fairness for Data Communications in IEEE 802.11 WLAN under DCF. In *Proc. of IEEE/IFIP IWQoS'02*, May 2002.
- [19] V. A. Siris. Resource Control for Elastic Traffic in CDMA Networks. In *Proc. of ACM MOBICOM 2002*, September 2002.
- [20] A. Veres, A. T. Campbell, M. Barry, and L.-H. Sun. Supporting Service Differentiation in Wireless Packet Networks using Distributed Control. *IEEE J. Select. Areas in Commun.*, 19(10):2094–2104, October 2001.
- [21] S. Wietholter and C. Hoene. Design and Verification of an IEEE 802.11e EDCF Simulation Model in ns-2.26. Technical Report TKN-03-019, Technical University of Berlin, November 2003.