Modeling 802.11e for data traffic parameter design

Peter Clifford, Ken Duffy, John Foy, Douglas J. Leith and David Malone

Hamilton Institute, NUI Maynooth, Ireland
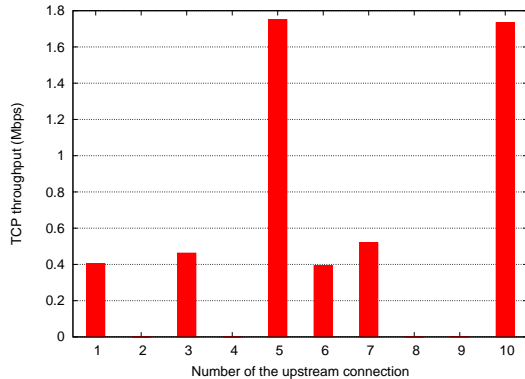
Fig. 1. Competing TCP uploads, 10 stations (NS2 simulation, 802.11 MAC, 300s duration, parameters as in Table I).

Abstract— **This paper introduces a finite load multi-class 802.11e EDCF model that is simple enough to be explicitly solvable. The model is nevertheless flexible enough to model the impact of 802.11e parameters on the prioritization of realistic traffic. We emphasize that a modeling framework which allows nonsaturated sources is essential in the study of realistic traffic. We apply the model to a situation of practical interest: competing TCP flows in an infrastructure network. The model allows us to make a principled selection of 802.11e parameters to resolve problems highlighted in this scenario. Model predictions and parameter selections are validated against simulation and experiment. The model is shown to be accurate and the parameters effective.**

I. Introduction

802.11a/b/g has been extremely successful, but is not without shortcomings, which has motivated the definition of the 802.11e extensions to the basic 802.11 MAC. For instance, it is known that cross-layer interactions between the 802.11 MAC and the flow/congestion control mechanisms employed by TCP can lead to gross unfairness between competing flows, and indeed sustained flow lockout (e.g. [1][2]). For example, consider an 802.11b 11Mbps network consisting of laptops trying to upload large data files using TCP. Figure 1 plots simulated throughput achieved by each station. The existence of gross unfairness is clearly evident.

It is widely recognized that the 802.11a/b/g MAC requires greater flexibility to alleviate difficulties such as those in the example above, and consequently the new 802.11e standard allows tuning of MAC parameters that have previously been constant. Although the 802.11e standard provides adjustable parameters within the MAC layer, the challenge is to understand how best to use this flexibility to achieve enhanced network performance.

The 802.11e MAC has been the subject of empirical studies (e.g [3][4][5]) and multi-class 802.11e models do already exist (e.g. [6][7][8]). However, these models are strictly confined to saturated conditions; that is, where every station always has a packet to send. To understand the operation of 802.11e in the context of realistic traffic we argue that saturated models are inadequate and that it is *essential* to model traffic sources with finite (nonsaturated) demands. For example, saturated models are not able to capture the behavior exhibited in the example above. Data traffic such as web and email, which constitutes the vast majority of traffic on current networks, is typically bursty in nature. Even long-lived data traffic such as large file transfers are problematic for saturated modeling as delayed acking is ubiquitous in TCP receivers and means that TCP ACK (acknowledgement) transmission is nonsaturated even if the TCP sender is itself saturated. In the context of traffic prioritization, we note that when high priority traffic is low-rate or on-off this leads to different prioritization schemes from situations where the high priority traffic is greedy or saturated. To see this, observe that when high-priority traffic is saturated, strict prioritization schemes cause high-priority traffic to swamp the network. Strict priorization (plus admission control) is, in contrast, a standard approach when high priority traffic is low rate, such as voice. Note that the saturation of a wireless station is logically distinct from whether the network is heavily or lightly loaded. It is possible for a network of saturated stations to be lightly loaded if there are only a small number of stations and, conversely, a network of nonsaturation stations may be heavily loaded if there are many stations.

We also note that interesting features of 802.11 MAC behavior only emerge in nonsaturated conditions. For example, a saturated model cannot predict maximum network throughput as it is well known that for CSMA/CA random access schemes of the type used in 802.11 the throughput is generally not a monotonic function of offered load ([9]). That is, there exists a pre-saturation throughput peak. This occurs in 802.11a/b/g [10] and will be shown in this article to persist in 802.11e.

The main contribution of this paper is a multi-class 802.11e EDCF finite-load model that is simple enough to be explicitly solvable, but complex enough to accurately predict the throughputs of unsaturated traffic, and the impact of the three most significant 802.11e MAC parameters on traffic prioritization: TXOP, AIFS and CW_{min}. In particular, modeling the effect of AIFS introduces difficulties, but its inclusion is fundamental in understanding the full power of 802.11e prioritization. We demonstrate the value of the model by using it to determine settings of the 802.11e parameters that restore

fairness to the TCP flows in Figure 1.

II. RELATED WORK

A number of finite-load models of the 802.11a/b/g DCF exist, including [10][11][12][13] [14][15][16][17][18]. None of these models support multiple traffic classes differentiated by all the variable 802.11e MAC parameters: AIFS, $\mathrm{CW_{min}}$ and TXOP. As noted previously, [6][7][8] develop models of the 802.11e EDCF, but these are confined to saturated traffic conditions, and thus are unsuited for the design of prioritization schemes under realistic traffic conditions.

With regard to TCP unfairness, early work [19] studies the impact of path asymmetries in both wired and wireless networks. More recently [20][21] specifically consider TCP unfairness issues in 802.11 WLANs. All of these authors seek to work within the constraints of the basic 802.11 MAC, not utilizing the flexibility of 802.11e. In [1][2][22], the authors use 802.11e functionality to restore TCP fairness. As that work was conducted without a finite load 802.11e model, the proposed parameter settings were derived empirically.

III. IEEE 802.11 AND 802.11E CSMA/CA

The 802.11 MAC layer CSMA/CA mechanism employs a binary exponential back-off algorithm to regulate access to the shared wireless channel. On detecting the wireless medium to be idle for a period $DIFS$, each station initializes a counter to a random number selected uniformly in the interval $[0, CW - 1]$. Time is slotted and this counter is decremented once during each slot that the medium is observed idle. A significant feature is that the countdown halts when the medium becomes busy and resumes after the medium is idle again for a period $DIFS$. Once the counter reaches zero the station attempts transmission and can transmit for a duration up to a maximum time TXOP (defined to be one packet without 802.11e). If two or more stations attempt to transmit simultaneously, a collision occurs. Colliding stations double their CW (up to a maximum value), select a new back-off counter uniformly and the process repeats. After successful transmission, CW is reset to its minimal value $\mathrm{CW_{min}}$ and a new countdown starts regardless of the presence of a packet at the MAC. If a packet arrives at the MAC after the countdown is completed, the station senses the medium. If the medium is idle, the station attempts transmission immediately; if it is busy, another back-off counter is chosen from the minimum interval. This bandwidth saving feature is called post-back-off. The new 802.11e MAC enables the values of $DIFS$ (called AIFS in 802.11e), $\mathrm{CW_{min}}$ and TXOP to be set on a per-class basis for each station. That is, traffic is directed to up to four different queues at each station, with each queue assigned different MAC parameter values.

IV. 802.11E EDCF FINITE-LOAD MODEL

As it will suffice for the applications presented in this paper, we assume there are two AIFS values, $\mathrm{AIFS_1}$ and $\mathrm{AIFS_2}$. We divide our stations into two classes by AIFS value. Within each class, stations can have distinct arrival rates, $\mathrm{CW_{min}}$

values, and so forth. Without loss of generality, those in the class 1 are assumed to have an AIFS smaller than or equal to those in class 2. Stations in each class are modeled by Markov chains of distinct structure whose transition probabilities are functions of their system parameters. The Markov chains are coupled by the operation of the network.

States in the Markov chain model for class 1 stations are labeled by a pair of integers (i, k) or $(0, k)_e$. The variable i represents the back-off stage, which is incremented (to a possible maximum m) when attempted transmission results in collision and set to 0 when transmission is successful. After attempted transmission the variable k is chosen randomly with a uniform distribution on the integers in the range $[0, W_i - 1]$, where $W_i = 2^i W_0$ and W_0 is the minimum contention window. While the medium is idle, k is decremented. If a packet is present, transmission is attempted when $k = 0$. The empty states $(0, k)_e$ represent the station when it does not have a packet to send. After successful transmission if a higher layer does not provide a packet, the MAC layer continues to decrement k to 0. If a packet arrives during the countdown, the station switches to the appropriate $(0, k)$ state. Otherwise, if countdown has ended with no packet, the station is in the state $(0, 0)_e$. When a higher layer provides a packet, the station senses the medium. If the medium is sensed idle, transmission is attempted immediately. If the medium is sensed busy, a stage 0 back-off is initiated, now with a packet.

The chain for class 2 stations has to be augmented because their larger AIFS value results in class 1 stations counting down before class 2 stations treat the medium as idle. Let D be the integer number of slots difference in the AIFS of class 2 and AIFS of class 1. We model the behavior of a class 2 stations with a three dimensional Markov chain indexed (i, k, d) and $(0, k, d)_e$ if the MAC layer is empty, i.e. there is no packet in the MAC. The variable $d \in \{0, \dots, D\}$ represents hold states for class 2. That is, $d > 0$ represents states in which the class 2 stations cannot decrement k while class 1 flows do, as they are not treating the medium as idle. When in a hold state class 2 stations must count up to D before returning to a non-hold state with $d = 0$.

Our main assumptions are the same as in [23][6][10]. We assume there are no hidden stations and errors are only caused by collisions. Conditioned on attempted transmission, each station has a fixed probability of collision irrespective of the network's history. In addition, as in [10][11][12], for each station there is a fixed probability of a packet arriving to the MAC during transitions in the Markov chains. In Section IV-C we relate the model arrival probability to the real offered load. In the following two subsections we define the transition probabilities for each station's Markov chain. Within a given class, these chains have the same structure. Calculations based on their stationary distributions lead to the equations in Section IV-C that determine the model's solution.

Let n_1 be the number of stations in class 1 and n_2 the number in class 2. We denote by $p_i^{(1)}$, $i \in \{1, \dots, n_1\}$ the probability that station i in class 1 will experience a collision given it is attempting transmission and by $q_i^{(1)}$ be

the probability the MAC receives a packet during a state-transmission in the chain. We define $p_i^{(2)}$ and $q_i^{(2)}$, $i \in \{1, \ldots, n_2\}$, similarly for class 2 station i. We denote the probability that station i in class 1 attempts transmission by $\tau_i^{(1)}$ and by $\tau_i^{(2)}$ the probability that station i in class 2 attempts transmission, conditioned on it not being in hold state. For notational convenience we suppress subscripts when describing each individual station's Markov chain.

A. Class 1 stations' Markov chain

For a station in class 1, let p be the probability of collision given attempted transmission, τ be the probability of transmission and q be the probability a higher layer presents a packet to the MAC. The transition probabilities of a class 1 station's Markov chain are listed in full below. They are determined by straight-forward logic. For $0 < k < W_i$, $0 < i \le m$ we have $P((i, k-1)|(i, k)) = 1, P((0, k-1)_e|(0, k)_e) = 1 - q$ and $P((0, k-1)|(0, k)_e) = q$. For $0 \le i \le m$ and $k \ge 0$ we have $P((0, k)_e|(i, 0)) = ((1-p)(1-q))/W_0, P((0, k)|(i, 0)) = ((1-p)q)/W_0$, and $P((\min(i+1, m), k)|(i, 0)) = p/W_{\min(i+1,m)}$. The most complex transitions occur from the $(0, 0)_e$ state where

$$
\begin{aligned}
P((0,0)_e|(0,0)_e) &= 1 - q + \frac{q(1-p)(1-p)}{W_0}, \\
0 < k < W_0, \quad P((0,k)_e|(0,0)_e) &= \frac{q(1-p)(1-p)}{W_0}, \\
0 \le k < W_1, \quad P((1,k)|(0,0)_e) &= \frac{q(1-p)p}{W_1}, \\
0 \le k < W_0, \quad P((0,k)|(0,0)_e) &= \frac{qp}{W_0}.
\end{aligned}
$$

B. Class 2 stations' Markov chain

We begin by identifying the probability that this class 2 station observes the medium is silent with the probability that it would not have a collision if it attempted transmission, as $1 - p = \prod_{i=1}^{n_1}(1 - \tau_i^{(1)}) \prod_j (1 - \tau_j^{(2)})$, where the second product is over all class 2 stations other than the one under consideration. Define P_{S_1} to be the probability that all class 1 stations are silent

$$
P_{S_1} = \prod_{i=1}^{n_1} (1 - \tau_i^{(1)}). \tag{1}
$$

The transition probabilities of a class 2 station's Markov chain are listed in full below. We start with transitions from non-hold states. For $0 < k \le W_i - 1$ and $i > 0$ we have

$$
\begin{aligned}
P((i, k-1, 0) \mid (i, k, 0)) &= 1 - p, \\
P((i, k, 1) \mid (i, k, 0)) &= p, \\
P((0, k-1, 0)_e \mid (0, k, 0)_e) &= (1-p)(1-q), \\
P((0, k-1, 0) \mid (0, k, 0)_e) &= (1-p)q, \\
P((0, k, 1)_e \mid (0, k, 0)_e) &= p(1-q), \\
P((0, k, 1) \mid (0, k, 0)_e) &= pq.
\end{aligned}
$$

For $k \ge 0$ and $i \ge 0$,

$$
\begin{aligned}
P((0, k, 1)_e \mid (i, 0, 0)) &= \frac{(1-p)(1-q)}{W_0}, \\
P((0, k, 1) \mid (i, 0, 0)) &= \frac{(1-p)q}{W_0}, \\
P((\min(i+1, m), k, 1) \mid (i, 0, 0)) &= \frac{p}{W_{\min(i+1,m)}}.
\end{aligned}
$$

The final set of non-hold states we need to consider if the window counter reaches 0 and there is still no packet to send. We deal with them in a way that enables us to give the explicit expression in Equation (4), below, for the probability of not being in a hold state. We refine $(0, 0, k)_e$ further into the states $(0, 0, k)_{e,sense}$ and $(0, 0, k)_{e,trans}$. In $(0, 0, 0)_{e,sense}$ the station has no packet and is sensing if the medium is busy. If it is busy it goes to a hold state. If it is idle and no packet arrives, it remains in $(0, 0, 0)_{e,sense}$, but if a packet arrives it goes to the second new state $(0, 0, 0)_{e,trans}$. In $(0, 0, 0)_{e,trans}$ the source transmits. Regardless of what happens (collision, successful transmission), the state that follows is a hold state. The hold states $(0, 0, k)_{e,sense}$ and $(0, 0, k)_{e,trans}$, $k > 0$, are kept separate because if an arrival occurs while in $(0, 0, k)_{e,sense}$, any k, a new back-off is initiated on departing from the hold states. This necessitates the introduction of a new arrival probability q_h, the probability a packet arrives at the MAC at some stage during transitions from $(0, 0, 1)_{e,sense}$ to successful departure from $(0, 0, D)_{e,sense}$. It is not necessary to give an expression for q_h in terms of q, as it cancels out before our final equations, but simplifies the derivation. Thus

$$
\begin{aligned}
P((0, 0, 1)_{e,sense} \mid (0, 0, 0)_{e,sense}) &= p, \\
P((0, 0, 0)_{e,sense} \mid (0, 0, 0)_{e,sense}) &= (1-p)(1-q), \\
P((0, 0, 0)_{e,trans} \mid (0, 0, 0)_{e,sense}) &= (1-p)q, \\
k > 0, \; P((0, k, 1)_e \mid (0, 0, 0)_{e,trans}) &= \frac{(1-p)(1-q)}{W_0}, \\
P((0, k, 0)_{e,sense} \mid (0, 0, 0)_{e,trans}) &= \frac{(1-p)(1-q)}{W_0}, \\
P((0, k, 1) \mid (0, 0, 0)_{e,trans}) &= \frac{(1-p)q}{W_0}, \\
P((1, k, 1) \mid (0, 0, 0)_{e,trans}) &= \frac{p}{W_1}.
\end{aligned}
$$

Turning our attention to transitions from hold states. For $k \ge 0$

$$
\begin{aligned}
P((0, k, 0) \mid (0, 0, D)_{e,sense}) &= P_{S_1} \frac{1}{W_0} q_h, \\
P((0, 0, 0)_{e,sense} \mid (0, 0, D)_{e,sense}) &= P_{S_1}(1 - q_h), \\
P((1, k, 0) \mid (0, 0, D)_{e,trans}) &= P_{S_1} \frac{p}{W_1}, \\
P((0, k, 0) \mid (0, 0, D)_{e,trans}) &= P_{S_1} \frac{1-p}{W_0} q, \\
k > 0, \; P((0, k, 0)_e \mid (0, 0, D)_{e,trans}) &= P_{S_1} \frac{1-p}{W_0}(1-q), \\
P((0, 0, 0)_{e,sense} \mid (0, 0, D)_{e,trans}) &= P_{S_1} \frac{1-p}{W_0}(1-q),
\end{aligned}
$$

For $1 \le j < D$,

$$
\begin{aligned}
P((0, 0, j+1)_{e,sense} \mid (0, 0, j)_{e,sense}) &= P_{S_1}, \\
P((0, 0, j+1)_{e,trans} \mid (0, 0, j)_{e,trans}) &= P_{S_1}.
\end{aligned}
$$

For $1 \le j \le D$,

$$
\begin{aligned}
P((0, 0, 1)_{e,sense} \mid (0, 0, j)_{e,sense}) &= (1 - P_{S_1}), \\
P((0, 0, 1)_{e,trans} \mid (0, 0, j)_{e,trans}) &= (1 - P_{S_1}).
\end{aligned}
$$

For $1 \leq j < D$,

$$P((0,k,j+1)_e \mid (0,k,j)_e) = P_{S_1}(1-q),$$
$$P((0,k,j+1) \mid (0,k,j)_e)) = P_{S_1}q.$$

For $k > 0$,

$$P((0,k-1,0)_e \mid (0,k,D)_e) = P_{S_1}(1-q),$$
$$P((0,k-1,0) \mid (0,k,D)_e) = P_{S_1}q.$$

For $k > 0$, $1 \leq j \leq D$,

$$P((0,k,1)_e \mid (0,k,j)_e) = (1-P_{S_1})(1-q),$$
$$P((0,k,1) \mid (0,k,j)_e) = (1-P_{S_1})q.$$

For $1 \leq j < D$, $k \geq 0$,

$$P((i,k,j+1) \mid (i,k,j)) = P_{S_1},$$
$$k > 0, \ P((i,k-1,0) \mid (i,k,D)) = P_{S_1},$$
$$P((\min(i+1,m),k,0) \mid (1,0,D)) = \frac{P_{S_1}p}{W_{\min(i+1,m)}},$$
$$P((0,k,0) \mid (i,0,D)) = \frac{P_{S_1}(1-p)q}{W_0},$$
$$P((0,k,0)_e \mid (i,0,D)) = \frac{P_{S_1}(1-p)(1-q)}{W_0},$$
$$P((i,k,1) \mid (i,k,j)) = (1-P_{S_1}).$$

C. Relating station and network models

Solving the Markov Chains for their stationary distributions, as outlined in the Appendix, leads to a relation between p and τ. Irrespective of class and station, each triple $(p,q,\tau) = (p_j^{(i)}, q_j^{(i)}, \tau_j^{(i)})$, $i \in \{1,2\}$ and $j \in \{1,\ldots,n_i\}$, is related by

$$\tau = \frac{1}{\eta}\left(\frac{q^2 W_0}{(1-q)(1-p)(1-(1-q)^{W_0})} - \frac{q^2(1-p)}{1-q}\right), \tag{2}$$

where the normalization constant η is

$$\begin{aligned}\eta &= \frac{qW_0}{1-(1-q)^{W_0}} + \frac{qW_0(qW_0+3q-2)}{2(1-q)(1-(1-q)^{W_0})} + (1-q)\\ &+ \frac{q(W_0+1)(p(1-q)-q(1-p)^2)}{2(1-q)}\\ &+ \frac{pq^2}{(1-q)2(1-p)}\left(\frac{W_0}{1-(1-q)^{W_0}} - (1-p)^2\right)\\ &\left(\frac{2W_0(1-p-p(2p)^{M-1})}{(1-2p)} + 1\right)\end{aligned} \tag{3}$$

and W_0 is the minimum contention window.

The hold probability P_{h} is the same for all class 2 stations, because if one them is in a hold state, they all are. A combination of the stationary distribution of class 2 stations' Markov chains and the fixed probability of collision assumption gives an expression for P_{h}. If D is zero, then the hold probability P_{h} is zero, otherwise it is

$$P_{\text{h}} = \frac{(1 - \prod_{j=1}^{n_1}(1-\tau_j^{(1)})\prod_{j=1}^{n_2}(1-\tau_j^{(2)}))\sum_{i=1}^{D}P_{S_1}^{-i}}{1 + (1 - \prod_{j=1}^{n_1}(1-\tau_j^{(1)})\prod_{j=1}^{n_2}(1-\tau_j^{(2)}))\sum_{i=1}^{D}P_{S_1}^{-i}}, \tag{4}$$

where P_{S_1} is defined in Equation (1). From the network model it is possible to deduce the following non-linear equations, (5) and (6), that couple all stations in the network. Their solution completely determines $p_j^{(i)}$ and $\tau_j^{(i)}$, from which throughputs and other performance metrics can be determined: for $i \in \{1,\ldots,n_1\}$

$$p_i^{(1)} = 1 - \prod_{j \neq i}(1-\tau_j^{(1)})(P_{\text{h}} + (1-P_{\text{h}})\prod_{j=1}^{n_2}(1-\tau_j^{(2)})) \tag{5}$$

and for $i \in \{1,\ldots,n_2\}$

$$p_i^{(2)} = 1 - \prod_{j=1}^{n_1}(1-\tau_j^{(1)})\prod_{j \neq i}(1-\tau_j^{(2)}). \tag{6}$$

Having solved for all $p_j^{(i)}$, $\tau_j^{(i)}$ and P_{h} we can determine station throughputs.

D. Throughput

The length of each state in the Markov chain is not a fixed period of real time. Each state may be occupied by a successful transmission, a collision or the medium being idle. To convert between states and real time, we must calculate the expected time spent per state, which is given by

$$\begin{aligned}E_s &= (1-P_{tr})\sigma + \sum_{i=1}^{n_1}P_{s:i}^{(1)}T_{s:i}^{(1)} + \sum_{i=1}^{n_2}P_{s:i}^{(2)}T_{s:i}^{(2)}\\ &+ \sum_{r=2}^{n_1}\sum_{1 \leq k_1^{(1)} < \cdots < k_r^{(1)} \leq n_1} P_{c:k_1^{(1)}\ldots k_r^{(1)}}^{(1)} T_{c:k_1^{(1)}\ldots k_r^{(1)}}\\ &+ \sum_{s=2}^{n_2}\sum_{1 \leq k_1^{(2)} < \cdots < k_s^{(2)} \leq n_2} P_{c:k_1^{(2)}\ldots k_s^{(2)}}^{(2)} T_{c:k_s^{(2)}\ldots k_s^{(2)}}\\ &+ \sum_{r=1}^{n_1}\sum_{s=1}^{n_2}\sum_{\substack{1 \leq k_1^{(1)} < \cdots < k_r^{(1)} \leq n_1\\1 \leq k_1^{(2)} < \cdots < k_s^{(2)} \leq n_2}} P_{c:\substack{k_1^{(1)}\ldots k_r^{(1)}\\k_1^{(2)}\ldots k_s^{(2)}}}^{(1)\,(2)} T_{c:\substack{k_1^{(1)}\ldots k_r^{(1)},\\k_1^{(2)}\ldots k_s^{(2)}}}\end{aligned} \tag{7}$$

where:

$$P_{tr} = 1 - \prod_{i=1}^{n_1}(1-\tau_i^{(1)})\prod_{i=1}^{n_2}(1-\tau_i^{(2)})$$

is the probability at least one station attempts transmission; σ is the slot-time;

$$P_{s:i}^{(1)} = \tau_i^{(1)}\prod_{j \neq i}(1-\tau_j^{(1)})(P_{\text{h}} + (1-P_{\text{h}})\prod_{j=1}^{n_2}(1-\tau_j^{(2)}))$$

is the probability station i in class 1 successfully transmits; $T_{s:i}^{(1)}$ is the time taken for a successful transmission from station i in class 1;

$$P_{s:i}^{(2)} = (1-P_{\text{h}})\tau_i^{(2)}\prod_{j=i}^{n_1}(1-\tau_j^{(1)})\prod_{j \neq i}^{n_2}(1-\tau_j^{(2)})$$

is the probability station i in class 2 successfully transmits; $T_{s:i}^{(2)}$ is the time taken for a successful transmission from

station i in class 2;

$$P^{(1)}_{c:k_1^{(1)}...k_r^{(1)}} = \prod_{i=1}^{r}\tau^{(1)}_{k_i^{(1)}}\prod_{i\neq k_1^{(1)}...k_r^{(1)}}(1-\tau_i^{(1)})$$
$$(P_{\mathrm{h}}+(1-P_{\mathrm{h}})\prod_{j=1}^{n_2}(1-\tau_j^{(2)}))$$

is the probability that only the class 1 stations labeled $k_1^{(1)}$ to $k_r^{(1)}$ experience a collision by attempting transmission, while class 2 stations are in a hold state or are not attempting transmission;

$$P^{(2)}_{c:k_1^{(2)}...k_s^{(2)}} = (1-P_{\mathrm{h}})\prod_{i=1}^{s}\tau^{(2)}_{k_i^{(2)}}\prod_{i\neq k_1^{(2)}...k_s^{(2)}}(1-\tau_i^{(2)})$$
$$\prod_{j=1}^{n_1}(1-\tau_j^{(1)})$$

is the probability that only the class 2 stations labeled $k_1^{(2)}$ to $k_s^{(2)}$ experience a collision by attempting transmission, while class 1 are not attempting transmission;

$$P^{(1)\,(2)}_{c:\substack{k_1^{(1)}...k_r^{(1)}\\k_1^{(2)}...k_s^{(2)}}} = (1-P_{\mathrm{h}})\prod_{i=1}^{r}\tau^{(1)}_{k_i^{(1)}}\prod_{j=1}^{s}\tau^{(2)}_{k_j^{(2)}}$$
$$\prod_{i\neq k_1^{(1)}...k_r^{(1)}}(1-\tau_i^{(1)})\prod_{j\neq k_1^{(2)}...k_s^{(2)}}(1-\tau_j^{(2)})$$

is the probability that only the class 1 stations labeled $k_1^{(1)}$ to $k_r^{(1)}$ and class 2 stations labeled $k_1^{(2)}$ to $k_s^{(2)}$ experience a collision by attempting transmission; and

$$T_{c:\substack{k_1^{(1)}...k_r^{(1)}\\k_1^{(2)}...k_s^{(2)}}}$$

is the time taken for their collision.

For example, using the basic 802.11b MAC values found in Table I with station i having payload E_i Bytes,

$$\text{Header} = \text{PLCP+MAC+CRC+IP},$$

$$T_{s:i} = E_i+\text{Header}+\delta+\text{SIFS+ACK+PLCP}+\delta+\text{AIFS}_1,$$

$$T_{c:\substack{k_1^{(1)}...k_r^{(1)}\\k_1^{(2)}...k_s^{(2)}}} = \max_{\substack{i=k_1^{(1)},...,k_r^{(1)}\\k_1^{(2)},...,k_s^{(2)}}} T_{c:i},$$

and

$$T_{c:i} = E_i+\delta+\text{Header}+\text{SIFS+ACKTimeout},$$

where for 802.11b ACKTimeout is the time taken for an ACK plus PLCP plus δ plus DIFS, making $T_{s:i} = T_{c:i}$.

E. Relating offered load to model parameters

If a station is saturated and always has a packet to send, then q is 1. If the station is not saturated, then to a first level of approximation q is the probability that at least one packet arrives at the MAC during E_s is one minus the probability that the first inter-packet time is greater than E_s. For example, when inter-packet arrival times to the MAC are exponentially

	Duration(μs)
Slot time, σ	20
Propagation delay, δ	1
DIFS (AIFS=0)	50
SIFS (Short Inter Frame Space)	10
PLCP Header @1Mbps	192
MAC Header 24 Bytes @1Mps	17.5
CRC Header 4 Bytes @1Mps	2.9
IP Header 20 Bytes @11Mbps	14.5
MAC ACK 14 Bytes @1Mbps	112
E_i payload 80 Bytes @11Mbps	58.2
E_i payload 560 Bytes @11Mbps	407.3
E_i payload 1500 Bytes @11Mbps	1090.9

TABLE I

802.11B MAC VALUES, BASIC RATE 1MBPS AND DATA RATE 11MBPS.

distributed with rate λ, then q and λ are related by $q = 1 - \exp(-\lambda E_s)$. With packet-length E Bytes, the station's offered load is

$$8E\lambda = \frac{-\log(1-q)8E}{E_s} \text{ Mbps.} \qquad (8)$$

Similar calculations are possible with other inter-arrival time distributions.

V. MODEL VALIDATION

We start with a representative selection of figures that demonstrate the model's throughput prediction accuracy through comparison with TU-Berlin's [24] 802.11e NS2 packet-level simulator. Consider a peer-to-peer 11Mbps network whose parameter values are given in Table I. The network consists of 10 class 1 station and 20 class 2 stations transmitting 560 Byte packets. Each station in class 2 station offers 4 times the load of each class 1 station. In simulation interface buffers are short and packets arrive with exponentially distributed inter-arrival times. The model is solved independently with q_1 and q_2 determined by Equation (8).

The graphs in Figure 2 show throughput for a station in each class versus offered load. The AIFS for class 1 is the 802.11b $DIFS$ value; the AIFS for class 2 is $DIFS$ plus D slots of length δ. Stations in each class have $\text{CW}_{\min} = 32$. Not only is the accuracy of the model apparent, but also the strong prioritization effect of AIFS. The related graphs in Figure 3 show throughput for a station in each class versus offered load for a range of minimum contention window pairs. For these graphs the AIFS value of each class are the same. Prioritization is still apparent and again the model makes remarkably accurate predictions. In all presented graphs throughput is not a monotonic function of offered load, that there is pre-saturation throughput peak, and the model predicts this effect. We have performed a large range of validation experiments, matching simulation to model predictions, and similar accuracy was obtained in all of them.

VI. 802.11E PARAMETER DESIGN

The 802.11e extensions to the 802.11 DCF allow a number of CSMA/CA parameters to be adjusted on a per class basis. Our motivation in developing an 802.11e finite-load model is
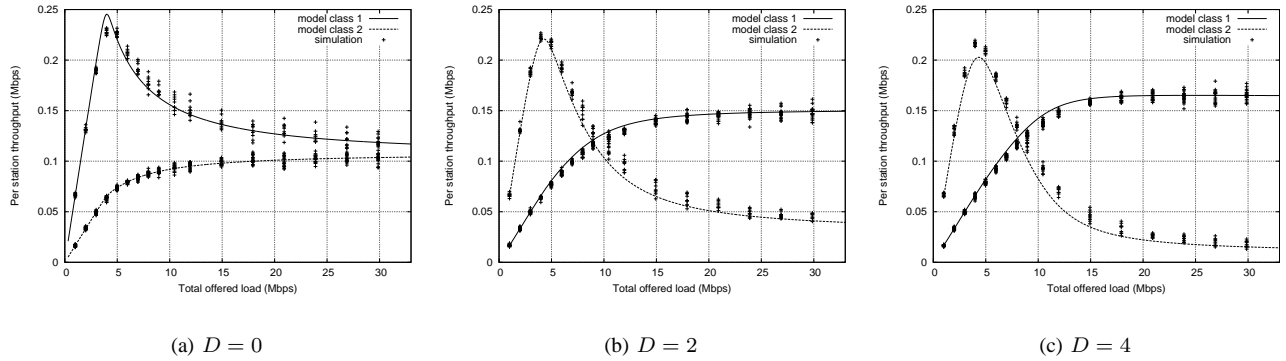
(a) $D = 0$ (b) $D = 2$ (c) $D = 4$

Fig. 2. Throughput for a station in each class vs. offered load. 10 class 1 stations offering one quarter the load of 20 class 2 stations. Range of D values, the difference in AIFS between class 2 and class 1 (NS2 simulation and model predictions, 802.11e MAC, 11Mbps PHY, 100s duration, MAC parameters as in Table I).

(a) $W_0^{(1)} = 32, W_0^{(2)} = 16$ (b) $W_0^{(1)} = 32, W_0^{(2)} = 64$ (c) $W_0^{(1)} = 32, W_0^{(2)} = 256$
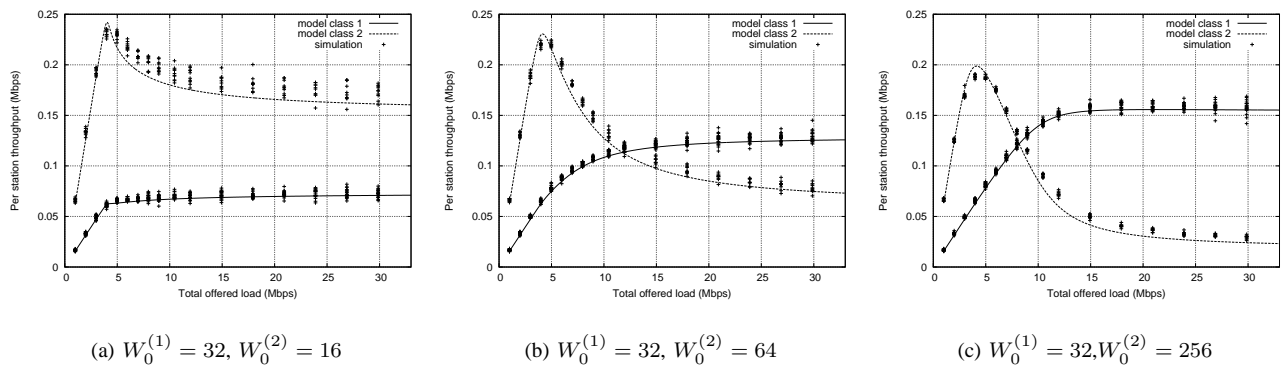
Fig. 3. Throughput for a station in each class vs. offered load. There are 10 class 1 stations each offering one quarter the load of 20 class 2 stations. Range of $\mathrm{CW_{min}}$ values (NS2 simulation and model predictions, 802.11e MAC 11Mbps PHY, 100s duration, MAC parameters as in Table I).

to establish an analytic basis for the design of strategies for adjusting the 802.11e parameters. Before considering parameter design for a situation of practical interest we first briefly discuss the impact of the impact of the following key 802.11e parameters: TXOP, AIFS and $\mathrm{CW_{min}}$.

The effect of TXOP is easy to understand. TXOP controls the number of packets (more precisely, the time allowed for packet transmission) that can be sent at a transmission opportunity. Increasing TXOP proportionately increases the relative throughput of stations, providing that they have data to send. The effect on performance of the AIFS parameter is more complex. To understand the influence of the AIFS parameter recall from Section III that the MAC countdown halts when the wireless medium becomes busy and resumes after the medium is idle again for a period AIFS. In addition to the initial delay of AIFS before countdown starts, a station accumulates an additional AIFS delay for every packet sent on the medium by other stations, leading to a reduction in the number of transmission opportunities that can be gained by a station as its AIFS is increased. This effect is, load dependent. When the network is lightly loaded, we AIFS differences have little impact on throughput. However, as the network load increases, stations with longer AIFS are rapidly penalized. This load dependent behavior can be seen by comparing

Figure 2 (a) and (b), when the network load is low, the graphs are quite similar. When the load is high, there is a dramatic change in the throughput achieved by the lower-rate class. The effect is even more apparent in Figure 2 (c). The impact on throughput of the $\mathrm{CW_{min}}$ parameter is simple. When the network is saturated, we expect the throughput of a station to be roughly inversely proportional to its $\mathrm{CW_{min}}$ value, as $\mathrm{CW_{min}}$ is proportional to how long a station must wait between transmissions. In Figure 3 (a), (b) and (c) the $\mathrm{CW_{min}}$ parameters are in the ratio 0.5, 2 and 8 respectively. Looking at the throughput ratios for heave loads, we find that they are approximately 0.4, 1.7 and 7 respectively, confirming this intuition. We note that the tuning of the $\mathrm{CW_{min}}$ parameter in the 802.11e standard is coarse, as the parameter is constrained to be a power of two, which limits its utility.

One might expect changing $\mathrm{CW_{min}}$ for 20 stations would have an impact, even in lightly loaded situations, due to a combination of impact on collision rates and extra time spent in backoff. However, we see from the figures that there is only a small impact on throughput until the network load becomes significant. It is interesting to note that unlike AIFS, 802.11e permits the values of $\mathrm{CW_{min}}$ to be reduced below its default value in 802.11. This can be used to increase the peak throughput of class 2 in Figure 3 (a) beyond that in Figure 2

(a). We turn now to the design of 802.11e parameter selection strategies for the example from the Introduction, preventing lockout of competing flows in data networks.

A. Mitigating TCP/MAC cross-layer interactions

Existing work on 802.11e is largely driven by quality of service requirements of applications such as VoIP. However, network traffic is currently dominated by data traffic (web, email, media downloads, etc.) carried by the TCP reliable transport protocol. Although lacking the time critical aspect of voice traffic, data traffic server-client applications do place significant quality of service demands on the wireless channel. Within the context of infrastructure WLANs in office and commercial environments, there is a requirement for efficient and reasonably fair sharing of wireless capacity between competing data flows.

As noted in the Introduction, cross-layer interactions between the 802.11 MAC and the flow/congestion control mechanisms employed by TCP leads to gross unfairness between competing flows, and even sustained lockout, see Figure 1. At the transport layer, to achieve reliable data transfers TCP receivers return ACK packets to the data sender confirming safe arrival of data packets. During TCP uploads, wireless stations queue data packets to be sent over the channel to their destination and the returning TCP ACK packets are queued at the wireless AP to be sent to the source station. TCP's operation implicitly assumes that the forward (data) and reverse (ACK) paths between a source and destination have similar packet transmission rates. The basic 802.11 MAC layer, however, enforces station-level fair access to the wireless channel. That is, n stations competing for access to the wireless channel are each able to secure approximately a $1/n$ share of the total available transmission opportunities. With n wireless stations and one AP, each station (including the AP) is able to gain only a $1/(n+1)$ share of transmission opportunities. By allocating an equal share of packet transmissions to each wireless station, with TCP uploads the 802.11 MAC allows $n/(n+1)$ of transmissions to be TCP data packets yet only $1/(n+1)$ to be TCP ACK packets. For larger numbers of stations this MAC layer action leads to substantial forward/reverse path asymmetry at the transport layer. Asymmetry in the forward and reverse path packet transmission leads to significant queueing and dropping of TCP ACKs disrupts the TCP ACK clocking mechanism, hinders congestion window growth and induces repeated timeouts. Repeated timeouts can lead to a persistent situation where flows are completely starved for long periods.

The requirement is clearly to prioritize the access point (AP) so that sufficient bandwidth is available for returning TCP ACKs to be transmitted back to the data stations. Following [1][2][22] we consider prioritizing the AP such that TCP ACKs have essentially unrestricted access to the wireless medium. The rationale for this approach to differentiating the AP makes use of the transport layer behavior. Namely, allowing TCP ACKs unrestricted access to the wireless channel does not lead to the channel being flooded. Instead, it ensures that the volume of TCP ACKs is regulated by the transport layer rather than the MAC layer. In this way the volume of TCP ACKs will be matched to the volume of TCP data packets, thereby restoring forward/reverse path symmetry at the transport layer. When the wireless hop is the bottleneck, data packets will be queued at wireless stations for transmission and packet drops will occur there, while TCP ACKs will pass freely with minimal queueing i.e. standard TCP semantics are recovered.

Our design requirement is that TCP ACK loss rate be no more than around 1–2%. We assume that a bandwidth-delay product of buffering is used at the wireless stations for the data traffic (recall that traffic classes are queued separately in 802.11e and it is, for example, possible for voice traffic to have small buffers and data traffic larger buffers). When operating normally the AIMD action of TCP congestion control will maintains a non-empty buffer at each wireless station, which can therefore be modeled as saturated traffic sources. In order to model TCP ACKs arriving at the access point, we adjust the arrival probability at the access point to be half the total successful transmission rate for data packets from the stations. This models TCP delayed ACKing, where a TCP receiver sends an ACK packet for approximately every second data packet.

Figure 4 illustrates model predictions the TCP ACK loss rate at the AP and the data throughput achieved by the wireless stations. Results are shown for ten wireless stations uploading data and receiving nonsaturated TCP ACKs back. Again, observe that adjustment of $\mathrm{CW_{min}}$ alone is unable bring the TCP ACK loss rate below the required target. Use of AIFS alone requires the use of extremely large values (AIFS > 20 slots) with associated impact on data throughput (falling by > 20%). We therefore consider the combined use of AIFS and $\mathrm{CW_{min}}$. Values of $(\mathrm{CW_{min}}, \mathrm{AIFS})$ which give a loss rate of close to 1% are (1,2), (2,4), (4,7), (8,12), ... Note that data throughput decreases as the value of AIFS increases, see Figure 4(a). We therefore seek to use a small AIFS value. Throughput is, however, not a monotonic function of AIFS for small values of AIFS. The settings $\mathrm{CW_{min}} = 2$ for the AP and AIFS = 4 for the data stations maximize throughput. In the next section, the performance of this scheme is confirmed by experiment.

VII. EXPERIMENTAL SETUP

Recently, hardware supporting a useful subset of the 802.11e functionality has become available. Thus we investigate the behavior of our parameterization in a realistic network rather than through simulations. Our WLAN consists of a desktop PC acting as an AP, and 12 embedded Linux boxes, based on the Soekris net4801, acting as client stations. Each station is equipped with an Atheros 802.11a/b/g PCI card with an external antenna. The system hardware configuration is summarized in Table II. Each station uses a Linux 2.6.8.1 kernel and a MADWiFi wireless driver modified to allow adjustment of the 802.11e CW_{min}, $AIFS$ and $TXOP$ parameters. Vendor specific features are disabled. Tests are
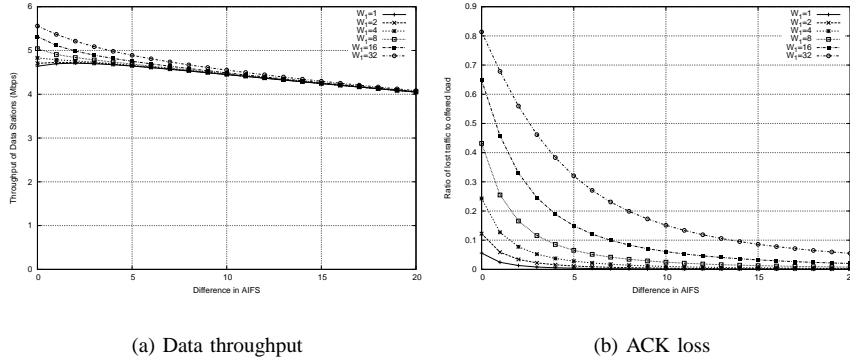
(a) Data throughput (b) ACK loss

Fig. 4. 10 stations (1500 byte packets) and AP transmitting (60 byte packets) at half achieved data rate (Model predictions, 802.11e MAC, 300s duration, parameters as in Table I).

Hardware		
$1\times$ AP	Dell GX 280	2.8Ghz P4
$12\times$ station	Soekris net4801	266Mhz 586
WLAN	D-Link DWL-G520	Atheros AR5212
Buffers	default	used
TCP	64KB	1MB
interface tx	199 packets	10 packets
driver tx	200 packets	10 packets

TABLE II

EXPERIMENTAL SETUP

performed using the 802.11b physical maximal transmission rate of 11Mbps with RTS/CTS disabled. The configuration of network buffers is detailed in Table II. In particular, we have increased the size of the TCP buffers to ensure that we see true AIMD behavior (with small TCP buffers, TCP congestion control is effectively disabled as the TCP congestion window is determined by the buffer size rather than the network capacity). We have carried out tests investigating the impact of the size of interface and driver queues and obtain similar results for a range of settings. Further details of this setup are described in [22].

Experimental results are shown in Figure 5, with and without prioritization. It can be seen that fairness between TCP uploads is restored. For other network scenarios, the model's predictions can be used in the same way to determine optimal MAC settings; we have verified that the suggested values for AIFS and $\mathrm{CW_{min}}$ are a good choice for a broad range of situations.

VIII. CONCLUSIONS

We have introduced an 802.11e CSMA/CA model that is simple enough to be explicitly solvable, but complex enough to accurately predict data throughput. We have shown that the model provides insight into the importance of different 802.11e parameters. Modeling nonsaturated traffic sources allows us to take an analytic approach to the design of prioritization schemes for practical situations and realistic traffic. To demonstrate how the model can be used to make principled selection of 802.11e parameters, we use the model to resolve serious cross-layer interactions between transport

layer congestion control and the standard 802.11 MAC in an upload scenario. In this case the selected parameter sets are demonstrated to be effective in practice. We note that solving the analytic model's equations is substantially faster than packet-level simulation and enables efficient investigation of the 802.11e parameter space. In this way, the model can be used as a design tool to help overcome the standard 802.11 MAC's known drawbacks [1].

APPENDIX

First we make observations that aid in the determination of the stationary distribution, to enable us relate p and τ for a class 1 station. With $b(i,j)$ and $b(0,j)_e$ denoting the stationary probability of being in states (i,j) and $(0,j)_e$, as b is a probability distribution we have

$$\sum_{i=0}^{m}\sum_{j=0}^{W_i-1} b(i,j) + \sum_{j=0}^{W_0-1} b(0,j)_e = 1. \tag{9}$$

We will write all probabilities in term of $b(0,0)_e$ and use the normalization in Equation (9) to determine $b(0,0)_e$. We have the following relations. To be in the sub-chain $(1,j)$, the following must have occurred: a collision from state $(0,0)$ or an arrival to state $(0,0)_e$ followed by detection of an idle medium and then a collision, so that $b(1,0) = b(0,0)p + b(0,0)_eq(1-p)p$. For $i > 1$ we have $b(i,0) = p^{i-1}b(1,0)$ and so

$$\sum_{i\geq 1} b(i,0) = \frac{b(1,0)}{1-p} = \frac{b(0,0)p + b(0,0)_eq(1-p)p}{1-p}. \tag{10}$$

The keystone in the calculation is then the determination of $b(0,W_0-1)_e$. Transitions into $(0,W_0-1)_e$ from $(0,0)_e$ occur if there is an arrival, the medium is sensed idle and no collision occurs. Transitions into $(0,W_0-1)_e$ also occur from $(i,0)$ if no collision and no arrival occurs

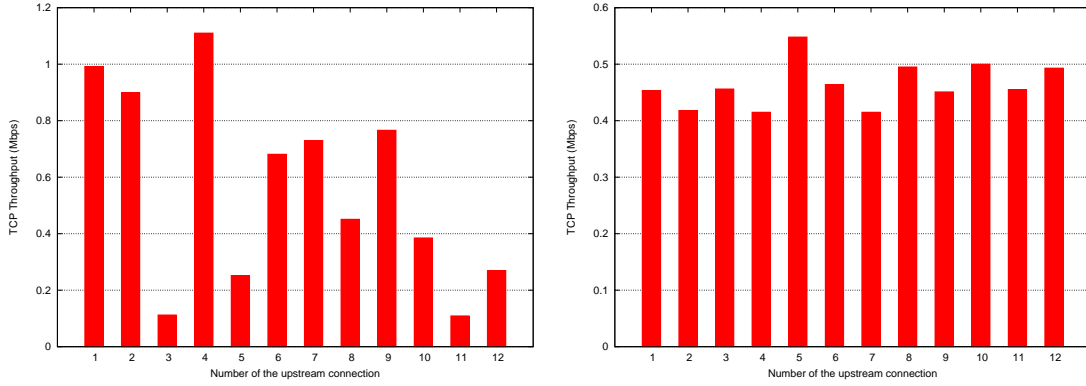$$b(0,W_0-1)_e = b(0,0)_e\frac{q(1-p)^2}{W_0} + \frac{(1-p)(1-q)}{W_0}\sum_{i\geq 0} b(i,0). \tag{11}$$

Fig. 5. Competing TCP uploads, 12 stations experiment without and with prioritization (802.11e MAC, 300s duration).

Combining Equations (10) and (11) gives

$$b(0, W_0 - 1)_e = b(0,0)_e \frac{(1-p)q(1-pq)}{W_0} + b(0,0)\frac{1-q}{W_0}.$$

We then have for $W_0 - 1 > j > 0$, $b(0,j)_e = (1-q)b(0,j+1)_e + b(0, W_0 - 1)_e$, with $b(0,j)_e$ on the left hand side replaced by $qb(0,0)_e$ if $j = 0$. Straight forward recursion leads to expressions for $b(0,j)_e$ in terms of $b(0,0)_e$ and $b(0,0)$, and we find

$$\frac{b(0,0)_e}{b(0,0)} = \frac{1-q}{q}\left(\frac{1-(1-q)^{W_0}}{qW_0 - (1-p)(1-pq)(1-(1-q)^{W_0})}\right). \quad (12)$$

Thus the second sum in Equation (9) $\sum_{j=0}^{W_0-1} b(0,j)_e = b(0,0)_e(qW_0)/(1-(1-q)^{W_0})$. The $(0,j)$ chain can then be tackled, starting with the relation $b(0, W_0 - 1) = \sum_{i \geq 0} b(i,0)\frac{(1-p)q}{W_0} + b(0,0)_e \frac{qp}{W_0}$. Recursion leads to

$$\sum_{j=0}^{W_0-1} b(0,j) = b(0,0)_e \left[\frac{q}{1-q}\frac{W_0+1}{2}\right.$$
$$\left(\frac{q^2 W_0}{1-(1-q)_0^W} + p(1-q) - q(1-p)^2\right)$$
$$\left. + \frac{qW_0(qW_0 + q - 2)}{2(1-(1-q)_0^W)} + 1 - q\right].$$

Using Equation (12) we can determine $b(1,0)$ in terms of $b(0,0)_e$:

$$b(1,0) = b(0,0)_e \frac{pq^2}{1-q}\left(\frac{W_0}{1-(1-q)_0^W} - (1-p)^2\right).$$

Finally, the normalization (3) gives

$$\begin{aligned} 1/b_{(0,0)_e} &= (1-q) + \frac{q^2 W_0(W_0+1)}{2(1-(1-q)^{W_0})} \\ &+ \frac{q(W_0+1)}{2(1-q)}\left(\frac{q^2 W_0}{1-(1-q)^{W_0}} + \right. \\ &\quad \left. p(1-q) - q(1-p)^2\right) \\ &+ \frac{pq^2}{2(1-q)(1-p)}\left(\frac{W_0}{1-(1-q)^{W_0}} - (1-p)^2\right) \\ &\quad \left(2W_0 \frac{1-p-p(2p)^{m-1}}{1-2p} + 1\right). \end{aligned} \quad (13)$$

The main quantity of interest is τ, the probability that the station is attempting transmission. A station attempts transmission if it is in the state $(i, 0)$ (for any i) or if it is in the state $(0,0)_e$, a packet arrives and the medium is sensed idle. Thus $\tau = q(1-p)b(0,0)_e + \sum_{i \geq 0} b(i,0)$, which reduces

to Equation (2), where $1/b(0,0)_e = \eta$, given in Equation (13) so that τ is expressed solely in terms of p, q, W_0 and m.

For class 2 stations τ_2 is the probability that the station will attempt transmission in a typical slot, conditioned on it not being a hold state. With this conditioning in force, the class 2 stations' Markov chain is of identical structure to that of class 1. Thus the same relationship, Equation (2), holds between p and τ for both classes. We do, however, have to consider the stationary distribution of the class 2 chain to calculate P_h, the probability that a class 2 station is in a hold state. Let $c(i,j,k)$, $c(0,j,k)_e$, $c(0,0,k)_{e,sense}$ and $c(0,0,k)_{e,trans}$ denote the stationary distribution of the class 2 Markov chain. Our introduction of the states $(0,0,k)_{e,sense}$ and $(0,0,k)_{e,trans}$ allows for a simple deduction of P_h. It enables the division in states into those in which a source attempts transmission, which we call firing states, $\{(i,0,0); i \geq 0\} \cup (i,0,0)_{e,trans}$ and all other states, which we call non-firing states. We establish relations that write the stationary probability of hold-states in terms of non-hold states. Consider states that have a packet. Firstly, for those where we are not in a hold state: $c(1,0,0) = c(0,0,0)p + c(0,0,0)_{e,sense}q(1-p)p$; $i > 0$, $c(i+1,0,0) = c(i,0,0)p$; $i > 0$, $j \geq 0$, $c(i,j,0) = c(i,0,0)(W_i - j)/W_i$; and for $j > 0$, $c(0,j,0) = \frac{W_0-j}{W_0}(c(0,0,0) + (q(1-p)q + pq_h)c(0,0,0)_{e,sense}) + q\sum_{n \geq j}\sum_{m=0}^{m=D} c(0,n,m)_e$.

Next we consider states where we have a packet and are in a hold state. For non-firing, non-stage 0 backoff states, with $i, j > 0$ and $1 \leq k < D$, $c(i,j,k) = P_{S_1}c(i,j,k-1)$ and $c(i,j,1) = c(i,j,0)p + (1 - P_{S_1})\sum_{k=1}^{D} c(i,j,k)$. For firing, non-stage 0 backoff states, with $i, j > 0$ and $1 \leq k < D$, $c(i,0,k) = P_{S_1}c(i,0,k-1)$ and $c(i,0,1) = c(i,0,0) + (1 - P_{S_1})\sum_{k=1}^{D} c(i,0,k)$. For non-firing, stage 0 backoff states, with $j > 0$ and $1 \leq k < D$,

$$\begin{aligned} c(0,j,k+1) &= P_{S_1}c(0,j,k) + qP_{S_1}c(0,j,k)_e, \\ c(0,j,1) &= c(0,j,0)p + (1 - P_{S_1})\sum_{k=1}^{D} c(0,j,k) \\ &\quad + pqc(0,j,0)_e \\ &\quad + q(1 - P_{S_1})\sum_{k=1}^{D} c(0,j,k)_e. \end{aligned}$$

Firing, stage 0 backoff states, with $1 \leq k < D$, $c(0,0,1) = c(0,0,0) + (1 - P_{S_1})\sum_{k=1}^{D} c(0,0,k)$, $c(0,0,k+1) = P_{S_1}c(0,0,k)$ and $c(0,0,k+1)_e = P_{S_1}c(0,0,k)_e$. Next consider states that don't have a packet. With $j \geq 1$ and

$1 \leq k < D$,

$$
\begin{aligned}
c(0, W_0 - 1, 0)_e &= (1-q)\frac{c(0,0,0)_e(1-p)q + c(0,0,0)}{W_0}, \\
c(0, j, k+1)_e &= (1-q)P_{S_1}c(0, j, k)_e, \\
c(0, j, 1)_e &= (1-q)pc(0, j, 0)_e \\
&\quad + (1-q)(1-P_{S_1})\sum_{k=1}^{D}c(0, j, k)_e,
\end{aligned}
$$

and with $0 \leq j < W_0$

$$
c(0, j, 0)_e = \sum_{n=1}^{W_0 - j}(1-q)^n(1-q)\frac{c(0,0,0)_e(1-p)q + c(0,0,0)}{W_0}.
$$

Since for $1 \leq k < D$

$$
\begin{aligned}
c(0, 0, k+1)_{e,trans} &= P_{S_1}c(0, 0, k)_{e,trans}, \\
c(0, 0, 1)_{e,trans} &= c(0, 0, 0)_{e,trans} \\
&\quad + (1-P_{S_1})\sum_{k=1}^{D}c(0, 0, k)_{e,trans}, \\
c(0, 0, k+1)_{e,sense} &= P_{S_1}c(0, 0, k)_{e,sense}, \\
c(0, 0, 1)_{e,sense} &= pc(0, 0, 0)_{e,sense} \\
&\quad + (1-P_{S_1})\sum_{k=1}^{D}c(0, 0, k)_{e,trans},
\end{aligned}
$$

and for $j > 0$

$$
\begin{aligned}
c(0, j, k)_e + c(0, j, k) &= P_{S_1}(c(0, j, k-1) + c(0, j, k-1)_e), \\
c(0, j, 1)_e + c(0, j, 1) &= p(c(0, j, 0)_e + c(0, j, 0)) \\
&\quad + (1-P_{S_1}) \\
&\quad \sum_{k=1}^{D}(c(0, j, k) + c(0, j, k)_e),
\end{aligned}
$$

and

$$
\sum_{k=1}^{D}(c(0, j, k)_e + c(0, j, k)) = \sum_{k=1}^{D}\frac{p(c(0, j, 0) + c(0, j, 0)_e)}{P_{S_1}^{D+1-k}}.
$$

Therefore we have $\sum_{k=1}^{D}c(0, j, k) = \sum_{k=1}^{D}\frac{p}{(P_{S_1})^k}c(i, j, 0)$. By similarly considerations we have $\sum_{k=1}^{D}c(i, j, k) = \sum_{k=1}^{D}\frac{p}{(P_{S_1})^k}c(i, j, 0)$ for $i > 0, j > 0$ and $\sum_{k=1}^{D}c(i, 0, k) = \sum_{k=1}^{D}\frac{1}{(P_{S_1})^k}c(i, 0, 0)$ for $i > 0$. Thus, using the normalization of the stationary distribution, we have an expression that does not include hold states. Moreover, the first term consists of non-hold non-firing states and the second term consists of non-hold firing states:

$$
\begin{aligned}
1 &= \left(c(0, 0, 0)_{e,sense} + \sum_{i>0}\sum_{j\geq 0}c(i, j, 0)\right. \\
&\quad + \sum_{j>0}(c(0, j, 0) + c(0, j, 0)_e))\left(1 + p\sum_{k=1}^{D}\frac{1}{(P_{S_1})^k}\right) \\
&\quad + \left.(c(0, 0, 0)_{e,trans} + \sum_{i\geq 0}c(i, 0, 0))\left(1 + \sum_{k=1}^{D}\frac{1}{(P_{S_1})^k}\right)\right).
\end{aligned}
$$

Defining $C_{\text{non-firing}}$ to be the probability of being in a non-firing non-hold state $C_{\text{non-firing}} := c(0, 0, 0)_{e,sense} + \sum_{i>0}\sum_{j\geq 0}c(i, j, 0) + \sum_{j>0}(c(0, j, 0) + c(0, j, 0)_e)$, C_{firing} to be the probability of being in a firing state $C_{\text{firing}} := c(0, 0, 0)_{e,trans} + \sum_{i\geq 0}c(i, 0, 0)$ and the sum $S := \sum_{k=1}^{D}(P_{S_1})^{-k}$, we have

$$
1 = C_{\text{non-firing}}[1 + pS] + C_{\text{firing}}[1 + S]. \tag{14}
$$

As $\tau = C_{\text{firing}}/(C_{\text{firing}} + C_{\text{non-firing}})$ and $1 - P_{\text{h}} = C_{\text{firing}} + C_{\text{non-firing}}$. Dividing Equation (14) by $C_{\text{firing}} + C_{\text{non-firing}}$ and using the expressions for $1 - P_{\text{h}}$ and τ, we have $P_{\text{h}} = 1 - (\tau(1+S) + (1-\tau)(1+pS))^{-1}$. Recalling $P_{S_1} = (1-\tau_1)^{n_1}$ and $1 - p = (1-\tau_1)^{n}_1(1-\tau_2)^{n_2-1}$ leads to Equation (4).

REFERENCES

[1] D.J. Leith and P. Clifford, "Using the 802.11e EDCF to achieve TCP upload fairness over WLAN links," in *WiOPT, Trento, Italy*, 2005.

[2] D.J. Leith and P. Clifford, "TCP fairness in 802.11e WLANS," in *IEEE WirelessCom 2005, Maui, Hawaii, USA*, 2005.

[3] Qiang Ni, L. Romdhani, and T. Turletti, "A survey of QoS enhancements for IEEE 802.11 wireless LAN," *Wireless Communications and Mobile Computing*, vol. 5, no. 4, pp. 547–566, 2004.

[4] P. Gopalakrishnan, D. Famolari, and Toshikazu Kodama, "Improving WLAN voice capacity through dynamic priority access," in *IEEE GLOBECOM*, 2004, vol. 5, pp. 3245–3249.

[5] Y. Xiao, H. Li, and S. Choi, "Protection and guarantee for voice and video traffic in IEEE 802.11e wireless LANs," in *IEEE INFOCOM*, 2004, vol. 3, pp. 2152–2162.

[6] R. Battiti and Bo Li, "Supporting service differentiation with enhancements of the IEEE 802.11 MAC protocol: models and analysis," Tech. Rep. DIT-03-024, University of Trento, 2003.

[7] J.W. Robinson and T.S. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *IEEE JSAC*, vol. 22, no. 5, pp. 917–928, 2004.

[8] Z. Kong, D. H.K. Tsang, B. Bensaou, and D. Gao, "Performance analysis of IEEE 802.11e contention-based channel access," *IEEE JSAC*, vol. 22, no. 10, pp. 2095–2106, 2004.

[9] D. Bertsekas and R. Gallager, *Data Networks*, Prentice–Hall, 1987.

[10] K. Duffy, D. Malone, and D.J. Leith, "Modeling the 802.11 Distributed Coordination Function in non-saturated conditions," *IEEE Communications Letters*, vol. 9, no. 8, pp. 715–717, 2005.

[11] D. Malone, K. Duffy, and D.J. Leith, "Modeling the 802.11 Distributed Coordination Function with heterogenous finite load," in *RAWNET, Trento, Italy*, 2005.

[12] P. Clifford, K. Duffy, D.J. Leith, and D. Malone, "On improving voice capacity in 802.11 infrastructure networks," in *IEEE WirelessCom 2005, Maui, Hawaii, USA*, 2005.

[13] G-S. Ahn, A. T. Campbell, A. Veres, and L-H. Sun, "Supporting service differentiation for real-time and best-effort traffic in stateless wireless ad hoc networks (SWAN)," *IEEE Transactions on Mobile Computing*, vol. 1, no. 3, pp. 192–207, 2002.

[14] M. Ergen and P. Varaiya, "Throughput analysis and admission control in IEEE 802.11a," *ACM-Kluwer MONET Special Issue on WLAN Optimization at the MAC and Network Levels*, 2004.

[15] Bo Li and R. Battiti, "Analysis of the IEEE 802.11 DCF with service differentiation support in non-saturation conditions," *Lecture notes in Computer Science*, vol. 3266, pp. 64–73, 2004.

[16] G.R. Cantieni, Qiang Ni, C. Barakat, and T. Turletti, "Performance analysis under finite load and improvements for multirate 802.11," *Computer Communications*, vol. 28, no. 10, pp. 1095–1109, 2005.

[17] A.N. Zaki and M.T. El-Hadidi, "Throughput analysis of IEEE 802.11 DCF under finite load traffic," in *First International Symposium on Control, Communications and Signal Processing*, 2004, pp. 535–538.

[18] L. Bononi, M. Conti, and E. Gregori, "Runtime optimization of IEEE 802.11 wireless LANs performance," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, no. 1, pp. 66–80, 2004.

[19] H. Balakrishnan and V. Padmanabhan, "How network asymmetry affects TCP," *IEEE Communications Magazine*, pp. 60–67, 2001.

[20] A. Detti, E. Graziosi, V. Minichiello, S. Salsano, and V. Sangregorio, "TCP fairness issues in IEEE 802.11 based access networks," 2005.

[21] S. Pilosof, R. Ramjee, Y. Shavitt, and P. Sinha, "Understanding TCP fairness over wireless LAN," in *INFOCOM, San Francisco, USA*, 2003.

[22] A.C.H. Ng, D. Malone, and D.J. Leith, "Experimental evaluation of TCP performance and fairness in an 802.11e test-bed," in *ACM SIGCOMM Workshops*, 2005.

[23] G. Bianchi, "Performance analysis of IEEE 802.11 distributed coordination function," *IEEE JSAC*, vol. 18, no. 3, pp. 535–547, 2000.

[24] S. Wiethölter and C. Hoene, "Design and verification of an IEEE 802.11e EDCF simulation model in ns-2.26," Tech. Rep. TKN-03-019, Technische Universität Berlin, 2003.