A Delay-Based Admission Control Mechanism for Multimedia Support in IEEE 802.11e Wireless LANs

Bechir Hamdaoui
EECS Dept., Oregon State University
hamdaoui@eecs.oregonstate.edu

Moncef Elaoud
Motorola
moncef@motorol.com

Parameswaran Ramanathan
ECE Dept., University of Wisconsin
parmesh@ece.wisc.edu

ABSTRACT

Multimedia over IEEE 802.11 wireless local area networks (WLANs) has recently been the focus of many researchers due to its rapidly increasing popularity. Unlike their best-effort counterparts, multimedia applications have quality of service (QoS) needs typically expressed in terms of the maximum allowed delay and/or the minimum required throughput. Therefore, prior to accepting a multimedia application, the network must assure the satisfaction of its QoS requirements. In this paper, we develop a mechanism that can be used to control the admissibility of multimedia applications into WLANs. To develop the proposed mechanism, we first derive an analytical approximation of the delay experienced by packets when travelled through these networks. The analytical approximation of the delay is then used to propose an admission control mechanism for the enhanced distributed channel access (EDCA) method used by the hybrid coordination function (HCF) of IEEE 802.11e. The proposed delay-based admission control mechanism is validated via simulations of voice traffic.

Index Terms—Multimedia traffic, admission control, modelling, IEEE 802.11e WLANs, quality of service.

1 Introduction

Internet has become essential in people's lives due to its reliable functionality and availability. In addition, the demand for wireless connectivity has grown rapidly resulting in IEEE 802.11 wireless local area networks (WLANs) being successfully deployed in many public and commercial areas. IEEE 802.11 WLANs are initially designed to provide accessibility to all forms of best-effort Internet data, such as web browsing, e-mailing, and file transferring. With the increasing popularity of multimedia applications, such as voice and video, these networks are then required to support not only best-effort traffic, but also this new emerging multimedia traffic. However, unlike their best-effort counterparts, multimedia applications have quality of service (QoS) needs typically expressed in terms of the maximum allowed delay and/or the minimum required throughput. Therefore, the network must make sure that it can meet the QoS requirements of a multimedia application before accepting it.

The IEEE Task Group E has been working towards creating QoS enhancements to IEEE 802.11 [1] under the designation of IEEE 802.11e [2]. The IEEE 802.11e QoS facility provides MAC enhancements to support applications with QoS requirements. These QoS enhancements are made available to all QoS stations (QSTAs) associated with a QoS access point (QAP). The IEEE 802.11e QoS facility also includes an additional coordination function, called hybrid coordination function (HCF), that is implemented by all QSTAs. The HCF is a combination of both the DCF (Distributed Coordination Function) and the PCF (Point Coordination Function) with some QoS enhancements. The HCF uses two channel access mechanisms. The first mechanism, called enhanced distributed channel access (EDCA), is a contention based channel access method used during contention periods (CPs). The second mechanism, called HCF controlled channel access (HCCA), is a contention free based access method used during contention-free periods (CFPs).

The objective of this paper is to develop an efficient and simple mechanism that enables IEEE 802.11e EDCA WLANs with the capability of controlling the admissibility of traffic with delay requirements, such as voice and video traffic. We first derive an analytical approximation of the delay experienced by packets when delivered via these WLANs. The analytical approximation of the delay is then used to develop the admission control mechanism.

The simplicity of the proposed mechanism lies in our derived delay estimation model that it uses to decide whether to accept new flows. Because the model provides an "approximation" of the delay instead of the exact delay, which we show to be good enough for our admission control, the complexity of the analysis is reduced substantially. The efficiency of the mechanism is justifed by validating it by means of simulation studies of voice traffic. The results show that our admission control mechanism accurately estimates the maximum number of QSTAs that can be successfully admitted into the network.

In Section 2, we describe the related work. Section 3 provides an overview of the IEEE 802.11e EDCA access method. Section 4 derives an approximation of the delay experienced by packets when travelled via IEEE 802.11e EDCA networks. In Section 5, we present the proposed admission control mechanism. We validate the mechanism via simulations of voice traffic in Section 6. Finally, we conclude the paper in Section 7.

2 Related Work

The recent success of IEEE 802.11 WLANs has attracted the attention of many researchers, thereby resulting in numerous studies of various aspects, analytical, simulations, and experimental, of their performance. Bianchi's pioneering work [3], analytically modelling and evaluating the saturated throughput of these networks under the DCF mode, has both triggered and laid out the foundation for studying the legacy IEEE 802.11 protocol and its derivatives under various scenarios. Generally, the reported studies aim at characterizing, evaluating, and measuring the performances of IEEE 802.11 WLANs under various traffic conditions, such as saturated traffic [4] and non-saturated traffic [5, 6], and various traffic types, such as data traffic [7] and multimedia traffic [8, 9].

Bianchi's model [3] has also been extended/modified to study IEEE 802.11e [10–22]. In [10], the authors provide an improved version of Bianchi's model to account for IEEE 802.11e's features, and use their improved model to estimate the saturated throughput. Robinson and Randhawa [11] also study the saturated throughput of 802.11e EDCA by proposing and validating an analytical model, which, unlike [10], considers post-collision periods of the protocol. In [14], Engelstad and Osterbo analytically model the throughput of IEEE 802.11e EDCA by predicting the delay distribution via its z-transform. Using a three-dimensional Markovian approach, the work in [15] also proposes a model, capturing the saturated throughput in IEEE 802.11e EDCA networks, that can be used to compute the maximum sustainable throughput in each access category (AC) under a saturation traffic load. Most of the developed models, however, consider throughput saturation, and hence, cannot be used to control the admissibility of periodic traffic such as voice. The development of models and mechanisms that can be used to control the admissibility of multimedia traffic into IEEE 802.11e EDCA is more recent [16–23]. Generally, these reported admission control mechanisms can be classified into two categories: measurement-based [16–18] and model-based [19, 20, 22]. One measurement-based admission control scheme, called virtual MAC (VMAC), is proposed in [16]. The basic idea is to virtually run the MAC in parallel to the real MAC in order to measure the achievable throughput and delay. These virtually measured parameters are then used by QSTAs to control the admissibility of new traffic. VMAC handles "virtual packets" instead of real packets, and virtual packets are scheduled for delivery in the same way as real packets. Another local measurement-based admission control approach is proposed in [17], where each QSTA dynamically maps the measured traffic load condition into the three EDCA parameters, CWmin, CWmax, and AIFS, so that its throughput requirement is still met. Basically, these EDCA parameters are mapped by means of a linear function, called direct function mapping (DFM), whose coefficients are determined periodically (each beacon interval) through measured system parameters, such as the transmission collision ratio.

Several other proposed admission control schemes are model-based [19, 20, 22]. In [19], Banchs *et al.* propose a model-based scheme that relies on throughput requirements to control the admissibility of multimedia traffic into the network. Using a proposed analytical model, throughput requirements are first used to compute CWs of all the already accepted flows as well as the newly arrived flow. These CWs are then used to calculate the achievable throughput for each accepted station. If the calculated throughput of any of the already admitted flows and/or the new flow to be admitted is less than the measured throughput, then the new flow is rejected. In [22], a dynamic admission control (DAC) mechanism is proposed. The key idea is based on the fact that the IEEE 802.11 protocol allows multiple transmission rates to be used; a QSTA can adapt to channel quality variation by selecting the appropriate transmission rate. However, rate adaptation may increase the time needed to transmit a packet, and hence, the admitted "medium time" (time allocated by the QAP to the QSTA in question) may not be enough to fully deliver a packet. An analytical model, which dynamically computes the "medium time" for each QSTA, is first developed, and then used to admit traffic with throughput requirements. In essence, the reported admission control schemes first estimate or measure the available throughput, and then use it to decide whether to accept or reject new flows. Unlike previous schemes, our proposed model-based admission control mechanism is delay-based instead of throughput-based.

3 Overview of IEEE 802.11e EDCA

The EDCA mechanism of IEEE 802.11e [2] provides differentiated and distributed access to the wireless channel using 8 different user priorities (UPs). To provide support for the delivery of traffic with UPs at the QSTAs, the EDCA mechanism defines 4 access categories (ACs). QSTAs in each AC use an enhanced DCF (EDCF) to contend for transmission opportunities (TXOPs) all using an identical set of EDCA parameters specified by the QAP.

3.1 Enhanced DCF (EDCF)

Prior to transmitting a packet, a QSTA must first sense the medium to be idle for a minimum duration called arbitration inter-frame space (AIFS) period. Then, to reduce collision, the QSTA must wait for an additional random backoff period calculated as $b \times \tau$, where b is a number, called *backoff counter*, selected from a uniform distribution in the interval $[0, W_0 - 1]$, and τ is the length of the time slot period. W_0 is a fixed number referred to as the initial *contention window* size. While waiting, the QSTA decrements its counter by 1 every idle time slot. Every time the medium becomes busy, the QSTA must freeze its backoff counter. Once the counter is frozen, the QSTA resumes decrementing the counter by 1 every idle time slot after sensing the medium again idle for a AIFS period. When the counter reaches 0, the QSTA proceeds with transmission. In case of unsuccessful transmission, the QSTA keeps retransmitting the packet until it either succeeds or reaches a threshold number of attempts. At the i^{th} retransmission attempt, the contention window size W must equal $W_i = max\{f^i \times W_0, W_m\}$, where f is a persistent factor (typically, $f = 2$), and W_m is the maximum allowed size of the contention window. Upon a successful transmission, the contention window is reset to its initial size. When the receiving QSTA receives a non-erroneous packet, it only needs to wait for a short inter-frame space (SIFS) period—shorter than the AIFS period—before acknowledging the sending QSTA.

3.2 EDCA Parameters

The following is the EDCA set of parameters that all QSTAs belonging to the same AC must use to contend for TXOPS:

- AIFS: Each AC is associated with an arbitration inter-frame space (AIFS) period. AC whose QSTAs have higher UPs is associated with shorter AIFS period. Before transmitting, QSTAs must wait for the AIFS period associated with the AC to which they belong. Hence, QSTAs belonging to ACs with shorter AIFS periods will have higher priority of accessing the channel than those with larger AIFS ones.

- **W_0, W_m:** Each AC is associated with two contention window sizes: initial (W_0) and maximum (W_m). That is, QSTAs belonging to ACs with smaller initial and/or maximum contention window sizes will have higher priority of accessing the channel than those belonging to ACs associated with larger initial and/or maximum contention window sizes. In this work, instead of using W_m as a differentiating parameter, we will use the parameter m such that $W_m = f^m W_0$ (f is defined below). We refer to m as the number of backoff stages.

- **f:** QSTAs can further be differentiated among each other via using different values of the persistent factor f; i.e., ACs with higher priority can use smaller values of f than those with lower priority.

4 Delay Analysis

In this section, we present an analytical approximation of the delay experienced by packets when travelled via IEEE 802.11e WLANs. This approximated delay is then used to develop an admission control mechanism for IEEE 802.11e WLANs to support applications with delays requirements.

4.1 Assumptions and Notation

We consider an IEEE 802.11e [2] EDCA WLAN that consists of one QoS access point (QAP) and N wireless QoS stations (QSTAs). Our developed model is based on the following assumptions.

- Since we focus on providing admission control to flows belonging to the same AC, we consider that all the N generated traffic flows by the N QSTAs are identical, and have equal chances of accessing the medium.
- Given that the QAP has N downstream flows while each QSTA has only one upstream flow all competing with each other, the QAP ought to be given a greater share of the medium than each QSTA. In the proposed model, we allow the QAP to have EDCA parameter set different from that of a QSTA. All QSTAs, however, have identical EDCA parameter sets since all of them are treated equally. In the remainder of this paper, a subscript a associated with any parameter indicates that the parameter corresponds to the QAP[1]. For example, while W_0 denotes the initial contention window size of a QSTA, $W_{0,a}$ denotes that of the QAP.

[1]Note that in all the following sections, we only provide derivation for the case of QSTAs. Expressions corresponding to the QAP will be inferred by analogy from those derived for QSTAs.

- We assume that packet arrival processes at QSTAs' queues are all i.i.d. Let A be a random variable representing the interarrival time between packets measured at their arrivals to QSTAs' queues, and $a(t)$ be the corresponding distribution. Also, let q be the probability of finding a non-empty queue upon arrival of a packet to a QSTA's queue corresponding to the studied AC. Note that since the ultimate goal of this work is to develop a mechanism that regulates the admissibility of multimedia applications into IEEE 802.11e WLANs, we will assume that the upstream and the downstream traffic travelling through the WLAN is periodic. For convenience, we, however, start our analysis with the general distribution of the interarrival time; i.e., $a(t)$. In Subsection 4.7, we then consider and study the multimedia case by taking the periodic assumption into account. More details regarding traffic assumptions and notation are given in Subsection 4.7.

4.2 Virtual Time

Let i be a given QSTA. At any time, QSTA i (resp. QAP) may sense the wireless medium to be in one and only one of the following states:

1. With probability $p_f = p_{f,a}$, the medium is not being used for transmission by any QSTA, nor by the QAP; i.e., the medium is idle.
2. With probability p_s (resp. $p_{s,a}$), QSTA i (resp. QAP) is successfully transmitting a packet via the medium.
3. With probability q_s (resp. $q_{s,a}$), a QSTA other than i or the QAP (resp. a QSTA not the QAP) is successfully transmitting a packet via the medium.
4. With probability p_c (resp. $p_{c,a}$), a collision is occurring due to a delivery of an i's (resp. QAP's) packet.
5. With probability q_c (resp. $q_{c,a}$), a collision is occurring not due to an i's (resp. QAP's) packet transmission.

When sensing the medium, each one of these states may be thought of as a different time slot, each of which has different length. We call these slots *virtual time slots*. Each virtual slot may occur with a different probability. The length of each virtual time slot falls into one of the following three lengths:

T_f, T_s, and T_c. The virtual time slot corresponding to idle medium and occurring with probability p_f or $p_{f,a}$ is of length T_f; those corresponding to successful transmissions and occurring with probability p_s, q_s, $p_{s,a}$, or $q_{s,a}$ are of length T_s; and those corresponding to collided transmissions and occurring with probability p_c, q_c, $p_{c,a}$, or $q_{c,a}$ are of length T_c. Since QSTAs and the QAP all belong to the same AC and hence are likely to generate similar traffic, then we consider that T_f, T_s, and T_c are the same for all stations including the QAP[2].

If we let $E[T_{slot}]$ signify the expected length of a virtual time slot, from a QSTA's standpoint we can write

$$E[T_{slot}] = p_f T_f + (p_s + q_s)T_s + (p_c + q_c)T_c. \quad (1)$$

Since the quantities $(p_s + q_s)$ and $(p_{s,a} + q_{s,a})$ both represent the network's transmission success (the former from a QSTA's perspective and the latter from the QAP's perspective), these two quantities are then equal. Similarly, since the quantities $(p_c + q_c)$ and $(p_{c,a} + q_{c,a})$ both represent the network's transmission failure, the two quantities are also equal. Hence, $E[T_{slot,a}] = E[T_{slot}]$ (both a QSTA and the QAP see identical expected values of their experienced virtual time slots). Note that the claims we just made regarding the equality of these quantities (e.g., $p_c + q_c = p_{c,a} + q_{c,a}$) can easily be justified later in the paper when we derive the analytic expressions of the probabilities.

It is important to note that while T_f is constant ($T_f = time\_slot\_length$), both T_s and T_c depend on the value of AIFS[3] as well as other traffic and network parameters, such packet lengths and transmission rates. On the other hand, $E[T_{slot}]$ depends on T_s and T_c via Eq. (1). Hence, $E[T_{slot}]$ depends on AIFS. In our model, the parameter AIFS is then accounted for via Eq. (1).

4.3 Average Contention Window Size: $\bar{W}$, $\bar{W}_a$

Let i designate any QSTA. The window size W of a given QSTA can be modelled as a discrete-time Markovian chain with m states, each of which corresponds to a backoff stage, as shown in Fig.1. The

[2]Even when QSTAs do not generate similar traffic, our analysis is still valid; one needs then to replace T_f, T_s, and T_c by their corresponding expected values.

[3]Recall that AIFS=SIFS + $time\_slot\_length \times n = 10 + 20n\ \mu s$, where n is a tunable parameter (integer) that is typically set by the QAP to control the value of AIFS. Exact expressions of how to compute T_s and T_c as a function of AIFS, and hence as a function of n, are given in Section 6.

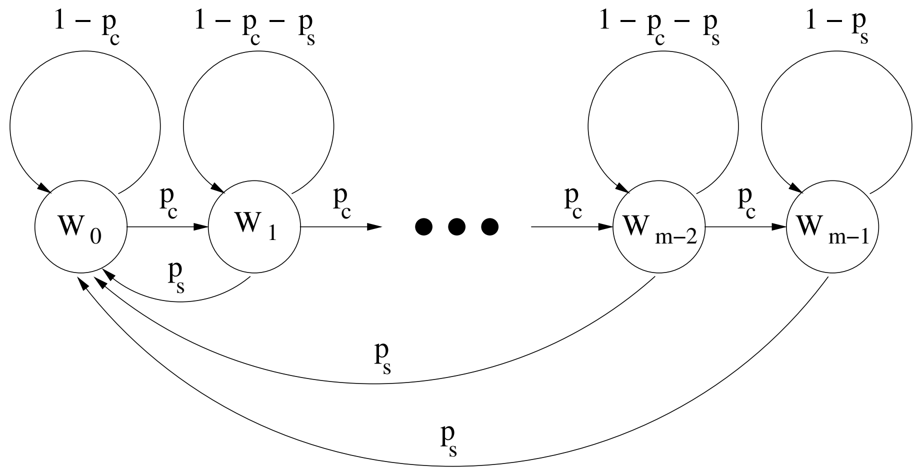


Figure 1: Discrete-time Markovian chain of the contention window size.

stationary distribution π_i, $0 \leq i \leq m-1$, of the Markovian chain can be expressed as

$$\begin{cases} \pi_i = p(1-p)^i, & 0 \leq i \leq m-2 \\ \pi_{m-1} = (1-p)^{m-1} & \end{cases}$$

where $p = \frac{p_s}{p_s+p_c}$. Therefore, the average window size $\bar{W}$ of any QSTA can be written as

$$\bar{W} = \frac{p - (f-1)f^{m-1}(1-p)^m}{1 - f(1-p)} W_0 \tag{2}$$

By analogy, one can derive the average window size $\bar{W}_a$ of the QAP as

$$\bar{W}_a = \frac{p_a - (f_a-1)f_a^{m_a-1}(1-p_a)^{m_a}}{1 - f_a(1-p_a)} W_{0,a} \tag{3}$$

where $p_a = \frac{p_{s,a}}{p_{s,a}+p_{c,a}}$.

Note that the new introduced system parameters, p and p_a, represent the fraction of the successful attempts to that of all tried ones respectively of a QSTA and the QAP.

4.4 QSTA Fractional Success Equation

Let i be any QSTA and v be any given virtual time slot. In this section, we derive expressions for the probabilities of occurrence of all the medium states that i may sense.

- The probability p_f is defined to be the probability of finding the medium idle at v. It can be expressed as

$$p_f \;=\; (1 - \frac{q_a}{\bar{W}_a})(1 - \frac{q}{\bar{W}})^N$$

- The probability p_s of i being successfully transmitting via the medium at v can similarly be computed as

$$p_s \;=\; \frac{q}{\bar{W}}(1 - \frac{q_a}{\bar{W}_a})(1 - \frac{q}{\bar{W}})^{N-1} \tag{4}$$

- The probability q_s of either a QSTA other than i or the QAP being successfully transmitting via the medium at v can similarly be computed as

$$\begin{aligned} q_s \;=\; & (N-1)\frac{q}{\bar{W}}(1 - \frac{q_a}{\bar{W}_a})(1 - \frac{q}{\bar{W}})^{N-1} \\ & + \frac{q_a}{\bar{W}_a}(1 - \frac{q}{\bar{W}})^N \end{aligned}$$

- The probability p_c that a collision in the medium is occurring at v due to i's packet transmission can similarly be computed as

$$p_c \;=\; \frac{q}{\bar{W}}(1 - (1 - \frac{q_a}{\bar{W}_a})(1 - \frac{q}{\bar{W}})^{N-1})$$

- The probability q_c that a collision in the medium is occurring at v provided that i is not transmitting can similarly be computed as

$$\begin{aligned} q_c \;=\; & 1 - \frac{q}{\bar{W}} - (N-1)\frac{q}{\bar{W}}(1 - \frac{q_a}{\bar{W}_a})(1 - \frac{q}{\bar{W}})^{N-1} \\ & -(1 - \frac{q}{\bar{W}})^N \end{aligned}$$

It is worth noting that all these probabilities depend on q, q_a, p and p_a only. Using the fact that $p = \frac{p_s}{p_s + p_c}$, we derive the following equation (we call it *QSTA Fractional Success Equation*).

$$p = (1 - \frac{q_a}{\bar{W}_a})(1 - \frac{q}{\bar{W}})^{N-1} \tag{5}$$

4.5 QAP Fractional Success Equation

Similarly, one can now derive expressions for the probabilities ($p_{f,a}$, $p_{s,a}$, $q_{s,a}$, $p_{c,a}$, and $q_{c,a}$) of occurrence of all the medium states that the QAP may sense. For example, the probability $p_{s,a}$ of the QAP being successfully transmitting via the medium at a given virtual time slot can be written as

$$p_{s,a} = \frac{q_a}{\bar{W}_a}(1-\frac{q}{\bar{W}})^N \tag{6}$$

Now using the fact that $p_a = \frac{p_{s,a}}{p_{s,a}+p_{c,a}}$, we derive a second equation (we call it *QAP Fractional Success Equation*)

$$p_a = (1-\frac{q}{\bar{W}})^N \tag{7}$$

of only p, p_a, and q as unknowns.

One point that requires attention is that both QSTA Fractional Success Equation and QAP Fractional Success Equation contain only the unknowns p, p_a, q, q_a. Hence we just derived two equations with 4 unknowns all of which (and only them) are needed to compute all the probabilities. It will become clear later that these probabilities are essential to compute in order for us to approximate the delay experienced by packets. Hence we still need two more equations to solve for the unknowns p, p_a, q, and q_a.

4.6 Service Time

The delay that packets experience when delivered through a WLAN has two components: queuing delay, and service delay (hereafter, the service delay will be referred to as service time). The queueing delay is the amount of time that packets spend on the queue; i.e., the time between the moment packets arrive at the queue and the moment packets arrive at the head of the queue. The service time is the amount of time counted from the moment packets reach the head of queue until they arrive successfully at their destinations.

Let S be a random variable representing the service time at a QSTA in terms of the number of virtual time slots. Similarly, S_a will denote that of the QAP. Let i be a given QSTA. Provided that at this instant a packet just reached the head of the queue of i, S for this packet is then the number of virtual time slots needed for i to successfully deliver the packet via the wireless medium. The probability mass function (pmf) of S and S_a can be written as $P[S = k] = p_s(1-p_s)^{k-1}$ and $P[S_a = k] = p_{s,a}(1-p_{s,a})^{k-1}$ for $k = 1, 2, \ldots$ where p_s and $p_{s,a}$ are given respectively in Eqs. (4) and (6). One can approximate the probability that the service time S (resp. S_a) of a packet delivered by a QSTA (resp. the QAP) equals T seconds as $p_s(1-p_s)^{k-1}$ (resp. $p_{s,a}(1-p_{s,a})^{k-1}$) where $k = \lfloor \frac{T}{E[T_{slot}]} \rfloor$.

Note that the service time is computable only if we know the probabilities p_s and $p_{s,a}$ which also depend only on the four unknowns p, p_a, q, and q_a. To solve for these unknowns, we need two more equations.

4.7 Lindley's Equations

The service times S and S_a derived in Section 4.6 are both Geometric with parameters p_s and $p_{s,a}$, expected values $\frac{1}{p_s}$ and $\frac{1}{p_{s,a}}$, and cumulative mass functions (cmf) $P[S \leq k] = 1 - (1-p_s)^k$ and $P[S_a \leq k] = 1 - (1-p_{s,a})^k$ for $k = 1, 2, \ldots$, respectively. We approximate the discrete geometric random variables S and S_a by the exponential random variables whose parameters are p_s and $p_{s,a}$. Now by assuming an independence between A and S, from a QSTA's standpoint, the network can be thought of as a G/M/1 system with an average interarrival rate, λ, and an average service rate, p_s. Using the G/M/1 based Lindley's Equation [24], we obtain the following two equations

$$q = A^*(p_s - p_s q) \tag{8}$$

and

$$q_a = A_a^*(p_{s,a} - p_{s,a} q_a), \tag{9}$$

where A^* and A_a^* are the Laplace Transforms of the interarrival distributions of packets respectively at a QSTA ($a(t)$) and the QAP ($a_a(t)$).

Recall that the ultimate goal of this work is to use the developed delay-based analysis to derive an admission control mechanism to regulate the admissibility of application with QoS requirements, such as multimedia applications, into IEEE 802.11e WLANs. Therefore, as mentioned in Subsection 4.1, we consider the periodic traffic model in this work since it captures most of the traffic generated by multimedia applications. In the remainder of this paper, we then assume that both the upstream and the downstream traffic travelling through the WLAN is periodic with an interarrival rate λ. We further consider that the average interarrival rate at the QAP is $\lambda_a = N \times \lambda$. Therefore, Eqs. (8) and (9) can be written as

$$\frac{1-q}{log_e(q)}\frac{q}{\bar{W}}(1-\frac{q_a}{\bar{W}_a})(1-\frac{q}{\bar{W}})^{N-1} = -\lambda \tag{10}$$

and

$$\frac{1-q_a}{log_e(q_a)}\frac{q_a}{\bar{W}_a}(1-\frac{q}{\bar{W}})^{N} = -\lambda N \tag{11}$$

Now we derived four nonlinear equations (Eqs. (5), (7), (10), and (11)) with four unknowns (p, p_a, q, and q_a). One can use traditional numerical methods to solve this system of equations. In this work, we use MATLAB to do so. Once the system of equations is solved, the solution can be used to determine the probabilities p_s and $p_{s,a}$ which are needed to determine the service times. However, a solution obtained by solving the system of non-linear equations is meaningless unless the network is stable.

4.8 Network Stability Conditions

As for all G/M/1 queueing systems, in order for the network to be stable, the average interarrival rates of packets at a QSTA and at the QAP must be smaller than their corresponding average service rates. That is, $\lambda \leq p_s$ and $\lambda \leq \frac{p_{s,a}}{N}$. Hence, the following two equations (we call them *Network Stability Equations*)

$$\frac{q}{\bar{W}}(1-\frac{q_a}{\bar{W}_a})(1-\frac{q}{\bar{W}})^{N-1} \geq \lambda \tag{12}$$

and

$$\frac{q_a}{\bar{W}_a}(1-\frac{q}{\bar{W}})^{N} \geq \lambda N \tag{13}$$

must be satisfied—when solving the system of nonlinear equations, the solution must then satisfy Eqs. (12) and (13).

4.9 Queueing Delays

So far we have modelled packet delays due to service times. In this section, we model the second component of the delay: queueing delay. For a G/G/1 system, the average queueing delay (AQD) is shown [25] to satisfy

$$\text{AQD} \le \frac{\sigma_a^2 + \sigma_s^2}{2(\frac{1}{\lambda} - \frac{1}{\mu})} \tag{14}$$

where λ (resp. σ_a^2) and μ (resp. σ_s^2) are the averages (resp. variances) of respectively the interarrival rates and service rates of the G/G/1 system. Note that the above upper-bound holds provided that the network is stable—i.e., the solution to the nonlinear system satisfies Eqs. (12) and (13).

Let Q and Q_a denote the average queueing delays experienced by packets respectively at a QSTA and the QAP. Now by applying Eq. (14) to the studied periodic traffic, the delays Q and Q_a can be upper-bounded by

$$\bar{Q} \equiv \frac{1 - p_s}{2p_s^2(\frac{1}{\lambda E[T_{slot}]} - \frac{1}{p_s})} E[T_{slot}] \tag{15}$$

for the case of a QSTA, and by

$$\bar{Q}_a \equiv \frac{1 - p_{s,a}}{2p_{s,a}^2(\frac{1}{\lambda N E[T_{slot}]} - \frac{1}{p_{s,a}})} E[T_{slot}] \tag{16}$$

for the QAP case. Note that, given the number of stations N and given the average interarrival rate of traffic, the two upper bounds on the average queueing delays depend on the probabilities p_s, $p_{s,a}$, and $E[T_{slot}]$ only; all of which are determined once a solution to the nonlinear system is found.

5 The Proposed Admission Control Mechanism

Based on the analytic delay approximation derived in the previous section, we now describe an admission control mechanism that can be used to admit and/or reject flows with QoS requirements in IEEE 802.11e wireless LANs. We assume that packets of multimedia applications are considered to be successfully delivered over the wireless LAN and hence of useful and acceptable QoS if they arrive at their destinations within a total budget delay δ_{total}. This budget is a parameter that is typically specified by the applications. For now we can think of δ_{total} as the maximum allowed delay that packets can experience via the wireless LAN due to both the queueing and the service delays. Later in the validation section we show how to compute such budget delay when considering other delays that packets encounter during their end-to-end trips. We also assume that multimedia applications can tolerate a certain ratio of unsuccessfully delivered packets while the quality of application is still considered acceptable. Let ζ denote such ratio.

Given the periodic traffic model, the QAP derives the interarrival packet rate λ and the time slot lengths T_f, T_s, and T_c. For a given AC, the QAP also knows the EDCA Parameters W_0, $W_{0,a}$, m, m_a, f, f_a and the two QoS parameters δ_{total} and ζ, all of which are specified a priori. Suppose that the QAP is currently serving N stations. When a new QSTA desires to establish a connection with the QAP, the following admission control procedure takes place:

1. The new QSTA i requests channel access by sending its request to the QAP. The request must contain the QoS parameters, $\delta_{total,i}$ and ζ_i, regarding the new application that needs to be admitted.

2. The QAP will admit the requesting station if doing so does not violate the QoS requirements of neither an already admitted station, nor the requesting station. The QAP does the following:

   (a) If $\delta_{total,i} < \delta_{total}$ or $\zeta_i < \zeta$, go to Step 3.

   (b) Increments N by one and solves the system of nonlinear equations formed by Eqs. (5), (7), (10), and (11) to determine p, p_a, q, and q_a.

   (c) Verifies whether the solution obtained in Step 2b satisfies the Network Stability Conditions given by Eqs. (12) and (13). If any of the two conditions are not satisfied, go to Step 3.

   (d) Computes the quatities p_s, $p_{s,a}$, and $E[T_{slot}]$ using Eqs. (4), (6), and (1), respectively.

   (e) Computes the upper bounds $\bar{Q}$ and $\bar{Q}_a$ on the queueing delays by using Eqs. (15) and (16). If $\bar{Q} \geq \delta_{total}$ or $\bar{Q}_a \geq \delta_{total}$, go to Step 3.

   (f) Calculates the service budget delay $\delta_{service} = \delta_{total} - \bar{Q}$ and $\delta_{service,a} = \delta_{total} - \bar{Q}_a$ respectively of a QSTA and the QAP. If any of the two inequalities,

$$p_s(1-p_s)^{\lfloor \frac{\delta_{service}}{E[T_{slot}]} \rfloor - 1} \leq \zeta$$

   or

$$p_{s,a}(1-p_{s,a})^{\lfloor \frac{\delta_{service,a}}{E[T_{slot}]} \rfloor - 1} \leq \zeta,$$

   are not satisfied, go to Step 3.

Table 1: Packet parameters in accordance with IEEE 802.11b [1]

	PHY	MAC	RAT/UDP	LOAD	SIFS	ACK	AIFS
Size (Bytes)	24	28	40	L	–	14	–
Rate (Mpbs)	1	11	11	11	–	1 or 2	–
Time (μs)	192	20.36	29.09	$\frac{L\times 8}{11}$	10	$192+56$	$10+20n$

(g) The QAP admits the new request; it then sends an acceptance message to the new requesting station. Skip step 3.

3. The QAP rejects the new request by sending a rejection message to the requesting station.

6 Validation of the Admission Control Mechanism

To validate the proposed admission control mechanism, we simulate an IEEE 802.11e EDCA wireless LAN that consists of one QAP and multiple QSTAs. The simulated IEEE 802.11e EDCA wireless LAN is an extension of IEEE 802.11b DCF wireless LAN with QoS enhancements. Packet parameters used in this simulation are taken from the IEEE 802.11b standard [1] and summarized in Table 1. The AIFS parameter is equal to (SIFS + $time\_slot\_length \times n = 10 + 20n$) μs where n is an integer that will be varied here in this section (recall that the parameter n is typically set by the QAP). We simulated voice traffic that was generated in accordance with the ITU-T G711 codec [8] which samples voice at 64 kilo bits per second. Sampled bits are then grouped into packets. This grouping process is called packetization. In this validation, we consider two packetization types, 10-millisecond and 20-millisecond, which are supported by ITU-T G711 [8] and respectively correspond to voice packet payloads of $L = 80$ and $L = 160$. Using Table 1, T_f, T_s, and T_c can be computed respectively as 20, $(509.5+20n+\frac{L\times 8}{11})$, and $(529.5+20n+\frac{L\times 8}{11})$ all in μs. Table 2 summarizes the virtual time slot length corresponding to each of the two packetization types: 10-millisecond and 20-millisecond.

The simulator used in this work is similar to the ns2 tool, except that ours is specifically custom-made for simulating voice traffic over IEEE 802.11 wireless LANs. The simulator, written in C, was developed in Telcordia few years ago, and has been used intensively by Telcordia to simulate and evaluate the performance of voice applications over IEEE 802.11 wireless LANs [9]. The simulator models a single IEEE 802.11 wireless LAN consisting of one access point (QAP) and many stations

Table 2: Virtual time slot lengths

$\delta_p(ms)$	rate(kbps)	L(Bytes)	$T_f(\mu s)$	$T_s(\mu s)$	$T_c(\mu s)$
10	64	80	20	$567.6 + 20n$	$587.6 + 20n$
20	64	160	20	$625.8 + 20n$	$645.8 + 20n$

(QSTAs). The simulator also includes a voice packet-generation model that we describe now. Voice packets are generated between two parties, A and B, according to a four-state (mutual silence, single talk A, single talk B, and double talk) Markov model. The durations of each state are exponential random variables with means 456 ms (mutual silence), 854 ms, (single talk), and 226 ms (double talk), and transition probabilities of 0.5 from either mutual silence or double talk to either of the single-talk states, 0.4 from single talk to mutual silence, and 0.6 from single talk to double talk. The simulator takes several parameters that can be varied from one run to another. These include: 1) voice packetization rates; 2) number of QSTAs; 3) duration of time to simulate; 4) MAC queue limits; and 5) maximum number of retransmissions per packet.

6.1 End-to-End Budget Delay

Recall that multimedia packets are delay-sensitive; i.e., a voice packet is considered to be successfully delivered if it reaches its destination within a budget delay. During its end-to-end trip, a packet experiences three types of delays before reaching its destination:

1. Packetization delay, δ_p: this delay corresponds to the amount of time needed to gather all samples of the packet before transmission; i.e., δ_p equals 10 and 20ms respectively for the two studied types of packetizations.

2. Coding/decoding delay, δ_c: this delay is the time needed for coding the packet before its transmission or decoding it after its reception. In the simulations, we set δ_c to $5ms$.

3. Wireless propagation delay, δ_{LAN}: this delay is the time needed for the packet to traverse a wireless LAN once the station (either a QSTA or the QAP) is given access to the medium to transmit the packet. Note that this delay does not account for queuing delays, nor collision delay; these two delays will be accounted for later. In the remainder of this section, we assume that the propagation delay, δ_{LAN}, is negligible ($\delta_{LAN} = 0$).

Table 3: Delay parameters

$\bar{D}(ms)$	$\delta_p(ms)$	$\delta_c(ms)$	$\delta_{LAN}(us)$	$\delta_{WAN}(ms)$	$\delta_{total}(ms)$
200	10	5	0	50	130
200	20	5	0	50	120

4. Wired propagation delay, δ_{WAN}: this delay is the amount of time that takes the packet to travel through the wired backbone of the Internet; the time amount needed to travel from one wired end of the voice stream to the other wired end (a wired end could be the QAP). In the simulations, we set δ_{WAN} to $50ms$ [8].

We consider that voice streams are initiated between one wired station and one wireless station (QSTA). That is, voice packets travel through only one single wireless LAN[4] during their end-to-end trips. (One end of the stream is a wireless station (QSTA) and the other end of the stream is a wired station.) We also consider that packets are received within an acceptable quality of service if their end-to-end delay does not exceed $\bar{D} = 200ms$ [26]. Hence, the total budget delay δ_{total}, that packets could experience when traversing the wireless LAN, is

$$\delta_{total} = \bar{D} - \delta_p - 2\delta_c - \delta_{LAN} - \delta_{WAN}.$$

Recall that δ_{total} is the maximum allowed time (resulting from both the queueing and the service delays) that packets could take when travelled via the wireless LAN if the quality of the voice communication is not to be violated. Table 3 summarizes the values of δ_{total} corresponding to each of the two packetization types. We also assume $\zeta = 0.02$ where again ζ is the ratio of unsuccessfully delivered packets to the total number of delivered packets that a station could tolerate while the quality of application is still considered acceptable.

6.2 Validation Scenarios and Method

We consider two types of traffic streams: unidirectional and bidirectional. In the unidirectional type, QSTAs can only receive traffic from the Internet via the QAP; i.e., there is no uplink traffic from QSTAs to the QAP. This, for example, corresponds to the case wherein QSTAs are running video or music

[4]Note that the validation of the mechanism does not depend on whether only one end of the stream is a wireless LAN or both ends are wireless LANs. A validation of the mechanism in one of these scenarios suffices.

streaming applications (traffic flows in the downlink direction only). In the bidirectional type, traffic flows in both directions—uplink and downlink; i.e., QSTAs can send packets to and receive packets from the QAP. The bidirectional type reflects multimedia applications that generate traffic in both directions such as interactive voice applications. For each one of the two types of traffic streams, we consider the two types of packetizations: 10 and 20ms. For each validation scenario (stream type/packetization type), we also consider different sets of EDCA parameters of the network.

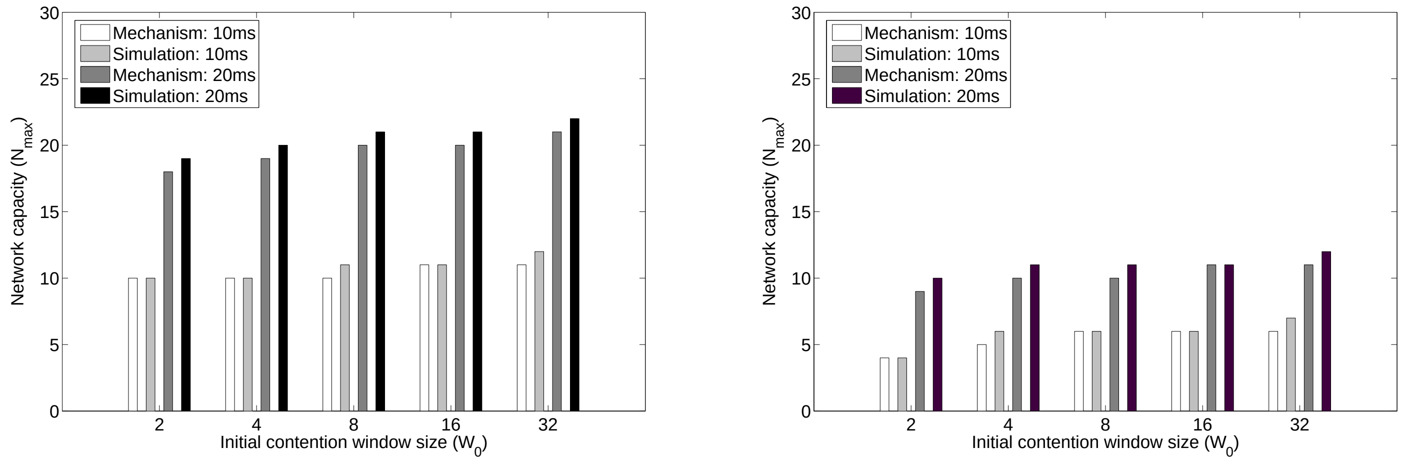
We define the capacity of an IEEE 802.11e wireless LAN, N_{max}, to be the largest number of QSTAs that can be supported while satisfying all of their QoS requirements. That is, once admitted into the wireless LAN, a QSTA must successfully deliver (transmit and/or receive) at least 98% ($= 1 - \zeta$) of the total delivered packets. Again, a packet is considered to be successfully delivered if it reaches the destination without violating its end-to-end delay requirement. To validate our proposed admission control mechanism, we proceed as follows: For each validation scenario,

- Fix the EDCA parameter set ($AIFS$, W_0, W_m, and f),
- Estimate the capacity of the IEEE 802.11e wireless LAN via our admission control mechanism (proposed in Section 5),
- Estimate the capacity of the IEEE 802.11e wireless LAN via simulation. The simulation consists of estimating the capacity by gradually increasing the number of QSTAs in the wireless LAN until at least one of existing QSTAs does not meet its required packet success ratio ($1 - \zeta = 0.98$).

6.3 Validation Results

We now present the validation results of our proposed admission control mechanism by comparing the accuracy of the network capacity obtained via the mechanism to that obtained via simulations. This is done for different EDCA parameter sets.

Figs. 2, 3, and 4 show the network capacity obtained under both the admission control mechanism and simulations for different sets of the EDCA parameters. In Fig. 2, we vary the initial contention window size (W_0) while fixing n to 2 and f to 2; in Fig.3, we vary the arbitration inter-frame space (AIFS) while fixing W_0 to 8 and f to 2; in Fig.4, we vary the persistent factor (f) while fixing W_0 to 8 and n to 2. In all these validation scenarios, the maximum contention window size W_m is set to



(a) Unidirectional traffic

(b) Bidirectional traffic

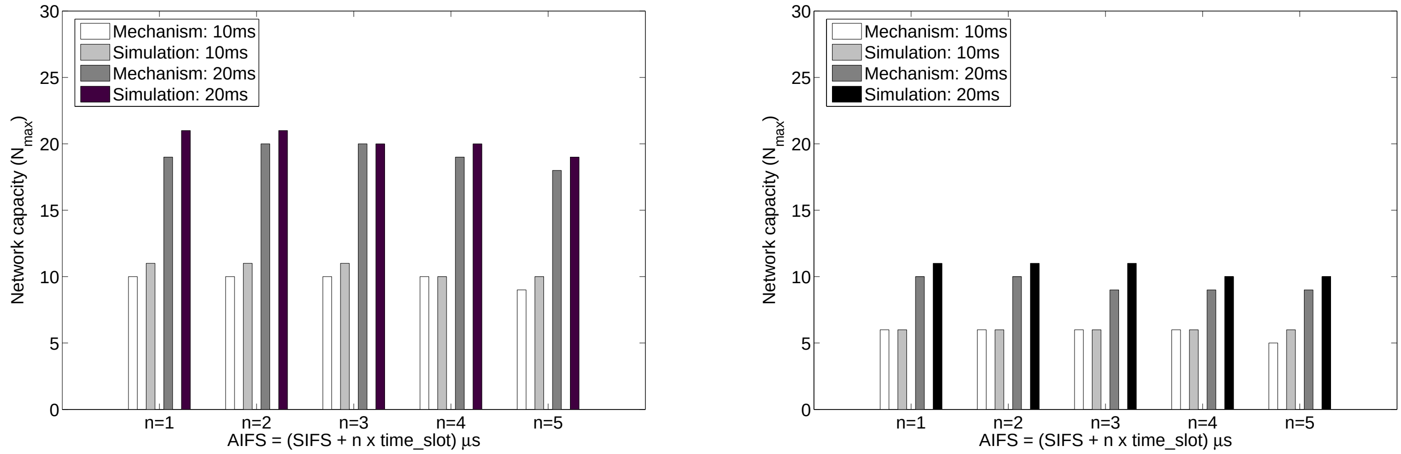
Figure 2: Validation under different values of W_0: $W_m = 64$, $f = 2$, $n = 2$

64. Figures labelled (a) show results of the unidirectional traffic type whereas those labelled (b) show results of the bidirectional traffic type.

The figures show that our model accurately estimates the maximum number of QSTAs that can be admitted into the wireless LAN. The figures also show that the network capacities (N_{max}) estimated by our mechanism are within 1 of those estimated via simulations. Furthermore, our admission control mechanism either accurately calculates the capacity or underestimates it by no more than two users. Overestimating the capacity is generally more detrimental to the network operation. Hence our developed model and admission control mechanism performs well under a variety of simulation scenarios.

7 Conclusion and Future Work

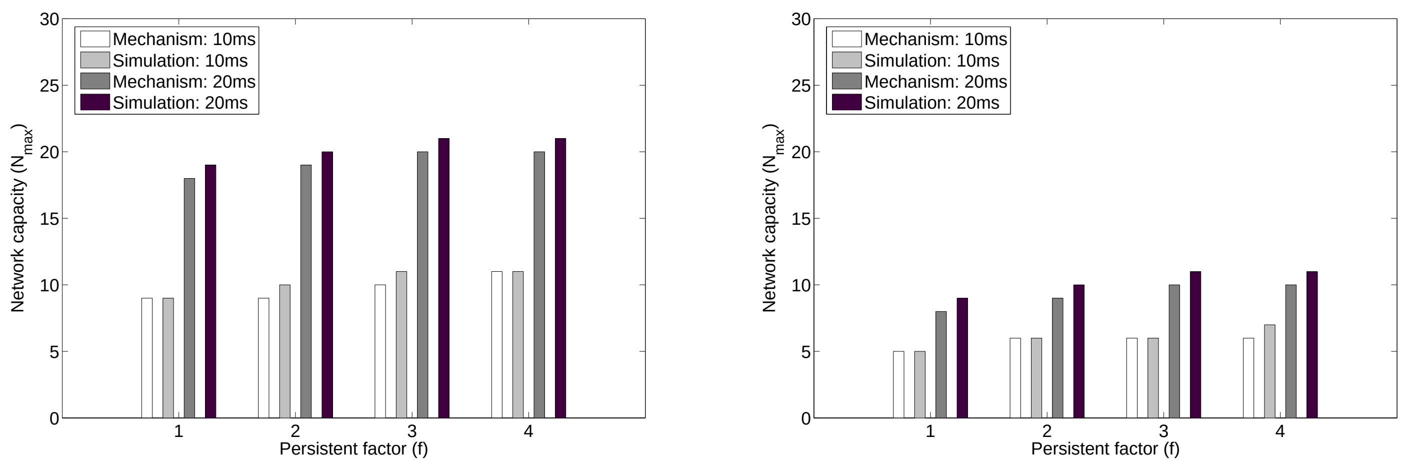
This paper presents a mechanism that IEEE 802.11e wireless LANs can use to control the admissibility of multimedia applications. The proposed mechanism is derived based on an analytical approximation of delays experienced by packets when delivered via the wireless LAN. The admission control mechanism is validated through simulations of voice traffic and shown to accurately estimate the capacity of the wireless LAN under different EDCA parameter sets. The validations are carried out for several realistic



(a) Unidirectional traffic

(b) Bidirectional traffic

Figure 3: Validation under different values of AIFS: $W_m = 64$, $f = 2$, $W_0 = 8$



(a) Unidirectional traffic

(b) Bidirectional traffic

Figure 4: Validation under different values of the persistent factor f: $W_m = 64$, $N = 2$, $W_0 = 8$

scenarios. Both the unidirectional and bidirectional traffic types are considered. While the former type reflects applications such as music/video streaming wherein traffic flows in the downlink only, the latter type represents applications in which traffic flows in both the uplink and the downlink such as interactive voice applications.

While this work focused on development of admission control for flows belonging to the same AC (intra-AC), one has to also consider developing admission control algorithms to differentiate among flows across different ACs (inter-AC). Recall that the four ACs of EDCA are meant to be: one AC for background traffic, one AC for best-effort traffic, one AC for voice traffic, and the other AC for video traffic. Therefore, by appropriately setting the AIFS parameter of each AC, one can provide differentiation across the four types of traffic. The challenge, however, is how to determine the AIFS values of all ACs so that the overall performance is optimized (i.e., the overall achievable network throughput is maximized) while still differentiating among flows with different priorities. There is still an apparent need for developing Inter-AC admission control mechanisms.

References

[1] *IEEE 802.11 Standard. Wireless LAN medium access control (MAC) and physical (PHY) layer specifications*, 1999.

[2] *Part11. Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 8: Medium access control (MAC) quality of service enhancements.*, 2005.

[3] G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535, 547 2000.

[4] F. Cali, M. Conti, and E. Gregori, “IEEE 802.11 wireless LAN: Capacity analysis and protocol enhancement,” in *Proceedings of IEEE INFOCOM*, 1998, pp. 142–149.

[5] 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, August 2005.

[6] D. Pong and T. Moors, "Call admission control for IEEE 802.11 contention access mechanism," in *Proceedings of IEEE GLOBECOM*, 2003.

[7] O. Tickoo and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks," in *Proceedings of IEEE INFOCOM*, 2004.

[8] M. Elaoud and P. Agrawal, "Voice capacity in IEEE 802.11 networks," in *Proceedings of IEEE PIMRC*, 2004, pp. 78–82.

[9] M. Eiger, M. ELaoud, and D. Famolari, "The effect of packetization on voice capacity in IEEE 802.11b networks," in *Proceedings of Consumer Communication and Networking Conference*, 2005.

[10] J. Hui and M. Devetsikiotis, "Desinging improved MAC packet schedulers for 802.11e WLAN," in *Proceedings of IEEE GLOBECOM*, 2003.

[11] J. W. Robinson and T. S. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 5, pp. 917–928, June 2004.

[12] S. Choi, J. Prado, and S. Mangold, "IEEE 802.11 e contention-based channel accesss (EDCF) performance evaluation," in *Proceedings of IEEE ICC*, May 2003, vol. 2, pp. 151–1156.

[13] S. Sehrawat, R. P. Bora, and D. Harihar, "Performance analysis of QoS supported by enhanced distributed channel access (EDCA) mechanism in IEEE 802.11e," in *Int'l Multiconference on Engineers and Computer Scientists IMECS*, 2006, pp. 943–948.

[14] P.E. Engelstad and O.N. Osterbo, "Delay and throughput analysis of IEEE 802.11e EDCA with starvation prediction," in *IEEE Conference on Local Computer Networks*, 2006.

[15] Z. Tao and S. Panwar, "Throughput and delay analysis for the IEEE 802.11e enhanced distributed channel access," *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 596–603, April 2006.

[16] M. Barry, A. T. Campbell, and A. Veres, "Distributed control algorithms for service differentiation in wireless packet networks," in *IEEE Infocom*, 2001.

[17] Y. Xiao and H. Li, "Local data control and admission control for QoS support in wireless ad-hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 5, pp. 1558–1578, September 2004.

[18] Y. Xiao and H. Li, "Voice and video transmissions with global data parameter control for the IEEE 802.11e enhanced distributed channel access," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, no. 11, pp. 1041–53, November 2004.

[19] A. Banchs, X. Perez-Costa, and D. Qiao, "Providing throughput guarantees in IEEE 802.11e wireless LANs," in *International Teletraffic Congestion*, 2003.

[20] N. Hegde, A. Proutiere, and J. Roberts, "Evaluating the voice capacity of 802.11 wlan under distributed control," in *IEEE Workshop on Local and Metropolitan Area Networks*, 2005.

[21] D. Gao, J. Cai, and K. N. Ngan, "Admission control in IEEE 802.11e wireless LANs," *IEEE Network*, vol. 19, no. 4, pp. 6–13, July-Aug. 2005.

[22] H. Yoon, J. Kim, and D-Y. Shin, "Dynamic admission control in IEEE 802.11e EDCA-based wireless home network," in *IEEE Comsumer Communications and Networking Conference*, 2006.

[23] Y. Xiao, "QoS guarantee and provisioning at the contention-based wireless MAC layer in the IEEE 802.11e wireless LANs," *IEEE Wireless Communications*, vol. 13, no. 1, pp. 14–21, February 2006.

[24] Leonard Kleinrock, *Queueing Systems. Volume 1: Theory*, John Wiley & Sons, New York, 1975.

[25] Leonard Kleinrock, *Queueing Systems. Volume 2: Computer Applications*, John Wiley & Sons, New York, 1976.

[26] *ITU-T Recommendations*, 1996.