A Novel CAC Scheme for Homogeneous 802.11 Networks

Qinglin Zhao, Danny H. K. Tsang, Senior Member, IEEE, and Taka Sakurai, Member, IEEE

Abstract—This paper proposes a new call admission control (CAC) scheme for one-hop homogeneous 802.11 DCF networks. Using the proposed scheme, we can perform admission control quickly and easily without the need for network performance measurements and complex calculations. The CAC rule is derived under asymptotic conditions, but our extensive numerical examples show that it works well for practical-sized networks with a finite retransmission limit and realistic nonsaturated traffic.

Index Terms—IEEE 802.11, homogeneous, traffic load, admission control.

I. INTRODUCTION

T

HIS paper is concerned with designing a new CAC scheme for one-hop 802.11 wireless local area networks (WLANs) employing the Distributed Coordination Function (DCF) protocol. We focus on homogeneous DCF networks, where all nodes have the same DCF protocol parameters and traffic parameters (i.e., packet length and the packet arrival rate), but the derived results pave a way to design powerful CAC schemes for more realistic heterogeneous DCF networks, as we show in the heterogeneous extension of this paper [1].

To support real-time applications in popular IEEE 802.11 or IEEE 802.11e WLANs, many CAC schemes have been proposed, such as a stochastic-delay-guarantee-based scheme [2], a measurement-assisted scheme [3]–[6], a cross-layer-based scheme [5], and a queue-stability-based scheme [6]. All these schemes have high complexity, which may limit their applicability as on-line algorithms. For instance, they all require the solution of a nonsaturated fixed-point formulation, which is not only computationally intensive but sometimes may fail because none of the algorithms guarantees the uniqueness and convergence of the fixed-point solution.

In contrast to the above CAC schemes, our CAC scheme is simple and practical. The objective of our scheme is to maintain the total traffic load below the critical traffic load, where the critical traffic load represents the threshold for buffer stability for the infinite buffer case or the threshold for excessive buffer rejection for the finite buffer case. For the example of a network with infinite node buffers, the proposed critical-load-based CAC condition is \( n \lambda \leq \mu \), where \( n \) denotes the number of nodes, \( \lambda \) denotes the packet arrival rate of each node, and \( \mu \) denotes the service rate of the network. We point out that this CAC condition is derived without placing any assumptions on the input traffic distributions. Also, \( \mu \) is determined only by the DCF protocol parameters, the packet size, and the number of nodes. It can be evaluated simply and quickly, without requiring network performance measurements and fixed-point calculations as needed in other CAC schemes [2]–[6]. Moreover, there exists signalling support for the information needed by our CAC scheme in the IEEE 802.11 standard for EDCA [7] which is an extension of DCF to multiple access categories. The traffic specification in this DCF extension contains fields for the mean flow data rate and the nominal flow packet size, which are equivalent to the flow-specific requirements that are needed by the access point (AP) in our CAC scheme.

The developed CAC scheme is based on a homogeneous and nonsaturated 802.11 DCF model that we proposed in [8]. In [8], we model nonsaturated DCF networks and show that our model compares favourably with prior models such as [9] in terms of simplicity and accuracy. Our extensive numerical examples in [8] reveal that there is a critical traffic load beyond which the system performance (such as the collision probability, the delay, and the packet loss) deteriorates dramatically. In the current paper, we perform an asymptotic analysis on the nonsaturated model to characterize the critical traffic load in an asymptotic regime where the number of stations tends to infinity and an infinite number of retransmissions is possible. Our analysis makes use of asymptotic results for saturated operation derived in [10] and [11]. Based on the asymptotic results for the critical traffic load, we devise a simple and practical CAC scheme for the homogeneous environment.

Our CAC scheme is very effective, even though the CAC rule is extrapolated from asymptotic results. Our numerical examples show that the scheme works well for practical-sized networks with a finite retransmission limit and realistic nonsaturated traffic. Moreover, we find that our scheme is effective for real-time services, despite the fact that our CAC rule is based on a queue stability criterion, which is a considerably coarser performance indicator than direct measures of quality-of-service such as delay. The queue stability criterion is adequate because DCF system performance exhibits an abrupt transition in the neighborhood of the critical traffic load, which is especially pronounced in the infinite buffer case [8]. For example, for the infinite buffer case, when the traffic load is increased beyond the critical point, the total delay will jump to several seconds from several milliseconds and the.
packet loss rate will jump to about 20% from zero. Therefore, ensuring that the proposed CAC condition is maintained can guarantee a total delay in the order of tens of milliseconds and a negligible packet loss rate for real-time applications.

The scheme in [6] is the most relevant of the prior studies to our work because it is also based on a queue-stability condition. Since the scheme in [6] is developed for heterogeneous environments, we postpone a comparison with our scheme to our work because it is also based on a queue-stability condition. In [1], where all results derived in this paper (i.e., Theorem 1, Lemma 1, and Theorem 2) have been extended to the heterogeneous case. In [1], we show that our scheme is superior to that of [6] in terms of simplicity, practicality, and generality.

The rest of this paper is organized as follows. Section II develops the critical traffic load analysis for a homogeneous environment and proposes a simple and practical CAC scheme based on the analytical results. Section III verifies the theoretical results via ns2 simulation. Section IV concludes the paper.

II. ANALYSIS OF CRITICAL TRAFFIC LOAD

In this section, we first review the homogeneous nonsaturated DCF model that we developed in [8]. Then, using the nonsaturated model, we perform an asymptotic analysis of the critical traffic load assuming an arbitrary buffer size.

We now introduce the terminology and assumptions that we will use in our analysis: an 802.11 system is said to be saturated if each node always has a packet to transmit or is nonsaturated otherwise; a packet transmission is said to be finished when the packet is either successfully received at the destination node or dropped due to reaching a retransmission limit; the service time of a packet is defined as the interval between when a packet becomes the first packet in the buffer and when an acknowledgement of the packet is successfully received at the sending node; time is measured in slots unless explicitly indicated. Similar to [12] and [11], we assume that 1) all nodes reside in a single-cell network (i.e., all stations are in sensing range of each other); 2) the collision processes of the nodes can be decoupled, such that the collision probability experienced by each node is constant and identical; and 3) channel conditions are ideal so that transmission errors are a result of packet collision only. In addition, we assume that the packet inter-arrival time (in slots) is independent and identically-distributed and the buffer size is measured in units of packets. Our analysis is focussed on the basic access mode of DCF, but it could be readily extended to the RTS-CTS mode.

A. The general fixed point equation

The nonsaturated model in [8] characterizes the general attempt rate of nodes running the DCF protocol. The general attempt rate is governed by a general fixed-point equation described below.

Let \( \beta^c \) be the conditional average attempt rate per slot for each node (i.e., the ratio of the number of attempts to the time spent in backoff measured in slots) on the condition that the buffer is not empty. Let \( p_0 \) be the probability of an empty buffer. Let \( \beta (0 \leq \beta \leq 1) \) be the general (or unconditional) average attempt rate per slot for each node. Noting that the attempt rate is equal to zero on the condition that the buffer is empty, we assume that \( \beta \) is equal to the product between the system utilization \((1-p_0)\) and the conditional attempt rate \( \beta^c \):

\[
\beta = (1-p_0)\beta^c. \tag{1}
\]

On the other hand, the collision probability experienced by the tagged node, \( \gamma \), can be expressed as follows in terms of \( \beta \):

\[
\gamma = \Gamma(\beta) \triangleq 1 - (1-\beta)^{n-1}, \tag{2}
\]

where \( n (n \geq 2) \) is the number of contending nodes.

To complete the general fixed point equation governed by (1) and (2), we need to express \( \beta^c \) and \( p_0 \) in terms of \( \gamma \). We first find an expression for \( \beta^c \). Let \( R \triangleq R(\gamma) \) and \( X \triangleq X(\gamma) \) be the mean number of attempts and the mean backoff time (in slots) when a packet transmission is finished, respectively. Applying the result in [11], we have

\[
\beta^c \triangleq \beta^c(\gamma) = \frac{R(\gamma)}{X(\gamma)}, \tag{3}
\]

where

\[
R(\gamma) = 1 + \gamma + \gamma^2 + \cdots + \gamma^M, \tag{4}
\]

\[
X(\gamma) = b_0 + \gamma b_1 + \gamma^2 b_2 + \cdots + \gamma^M b_M. \tag{5}
\]

In (4), \( b_i = b^i b_0, \) for \( 1 \leq i \leq m-1 \) and \( b_i = b^m b_0 \) for \( m \leq i \leq M. \) Here, \( b_0 = \frac{C_{W_0}}{2}; \) \( C_{W_0} \) is the minimum window size (in slots); \( b (> 1) \) is the multiplier of the exponential backoff; \( m \) determines the maximum backoff window size \( C_{W_{\text{max}}} \) (i.e., \( C_{W_{\text{max}}} = 2^m C_{W_0} \)); and \( M \) is the retransmission limit.

Next, we find the expression of \( p_0 \). According to queuing theory, \( p_0 \) is a function of the traffic intensity \( \rho \). We define \( \rho \) as follows:

\[
\rho = \lambda Y^c, \tag{6}
\]

where \( \lambda \) is the packet arrival rate per slot for each node and \( Y^c \triangleq Y^c(\gamma) \) is the mean service time (in slots) of a packet of a tagged node. Since a node will, on average, wait for a backoff time of \( X \) before it finishes a packet transmission, \( Y^c \) can be given by

\[
Y^c = X \Omega, \tag{7}
\]

where \( \Omega \triangleq \Omega(\gamma) \) is the mean time (in slots) that elapses for one decrement of the backoff counter. Note the backoff counter decreases by one for each idle time slot and is suspended when the channel is busy. \( \Omega \) is given by

\[
\Omega = \sigma (1-P_b) + (T_s + \sigma) P_s + (T_\varphi + \sigma) P_\varphi, \tag{8}
\]

where

\[
P_b = 1 - (1-\beta)^n - 1 - (1-\gamma)^{n-1}, \tag{9}
\]

\[
P_s = n \beta (1-\beta)^{n-1} = n(1 - (1-\gamma)^{\frac{n-1}{\gamma}})(1-\gamma), \tag{10}
\]

\[
P_\varphi = P_b - P_s, \tag{11}
\]

\[denote the probability of a busy slot, the probability of a successful transmission from any of the \( n \) contending nodes, and the probability of an unsuccessful transmission from any of the \( n \) contending nodes, respectively. Also, \( \sigma = 1 \) slot, and \( T_s \) and \( T_\varphi \) are the mean time (in slots) for a successful
transmission and an unsuccessful transmission, respectively. $T_\text{s}$ and $T_\text{p}$ depend on packet payload length, SIFS, DIFS, and other protocol parameters. Note we have explicitly expressed $\Omega$ in terms of $\gamma$.

Now, we have already expressed $\rho$ by (5) and (6) in terms of $\gamma$. The remaining task is to express $p_0$ in terms of $\rho$. In [8], we model only two extreme cases of the buffer size. That is, we let $p_0=1$-$\min(1, \rho)$ for an infinite buffer size and $p_0=\rho e^{-\rho}$ for a small buffer size. It is possible to generalize the model to an arbitrary buffer size $K$ by approximating the packet queuing dynamics of each node with a well-known queuing model such as the M/G/1/K or M/M/1/K queue, and finding the relevant expression for $p_0$. For example, under the M/M/1/K approximation, $p_0$ can be calculated as follows: if we assume that packet arrivals are Poisson with parameter $\lambda$ for each node and the service time of a packet of the tagged node follows an exponential distribution with mean $\gamma\mu$, $p_0$ can be calculated by [13]

$$p_0 = \frac{1}{1 + \rho + \rho^2 + \ldots + \rho^K}.$$  \hfill (9)

Our extensive simulations show that the above M/M/1/K model can adequately predict the simulation results for the collision probability and other performance metrics. Due to space restrictions, we focus in this paper on numerical examples showing the accuracy of the overall results for our CAC scheme (see Section III).

B. Asymptotic analysis of the critical traffic load point

For $K = \infty$, queuing theory tells us that $\rho \geq 1$ will result in infinite backlog and hence cause infinite delay. For $K \leq \infty$, [14] shows that a lower bound on the buffer rejection probability is $1 - 1/\rho$ for $\rho > 1$. Hence, for $K \leq \infty$, $\rho \geq 1$ will increase the rejection probability significantly.

In short, whatever the buffer size is, we should restrict the total packet arrival rate below a critical value so as to ensure $\rho < 1$. To obtain a simple expression for the critical arrival rate (or critical traffic load), we make the following assumptions.

**Asymptotic assumption:** To perform asymptotic analysis, like [11], we assume

$$n \to \infty \text{ and } m = M = \infty.$$  \hfill (10)

**Traffic assumption:** When $n \to \infty$, to ensure stable system performance, like [15], we assume

$$\lim_{n \to \infty} \frac{\lambda}{n} = 0 \text{ and } \lim_{n \to \infty} n\lambda = \Lambda,$$  \hfill (11)

where $\lambda$ is a function of $n$ and $\Lambda$ is a constant.

**Utilization assumption:** There exists a threshold $\Lambda^*$ such that when $\lim_{n \to \infty} n\lambda = \Lambda \geq \Lambda^*$, we have

$$\lim_{n \to \infty} p_0(\gamma) \triangleq p_0^+ < 1.$$  \hfill (12)

Let $\rho^+ \triangleq \lim_{n \to \infty} \rho(\gamma)$. For the G/G/1/$\infty$ queue, $\rho = 1 - p_0$ so (12) implies that $\rho^+ > 0$. Likewise, for the M/M/1/K queue, (12) implies that $\rho^+ > 0$ as can be seen from (9). In other words, we exclude the case that the traffic is so light that the system is almost always idle (i.e., $\rho^+ = 0$). Mathematically, we call a system saturated if $p_0^+ = 0$, which corresponds to the case that each node always has packets to transmit. Since $0 \leq p_0^+ < 1$ always holds under utilization assumption (12), (12) includes the saturated case (i.e., $p_0^+ = 0$) and the nonsaturated case (i.e., $0 < p_0^+ < 1$). In general, we use the notation of $x^+$ to denote $\lim_{n \to \infty} x$ in the rest of this paper.

Theorem 1 below proves a key result which is used later to analyze the asymptotic behavior of the traffic load.

**Theorem 1:** Under assumptions (10)-(12), $\beta$ and $\gamma$ have the following asymptotic properties:

\begin{align*}
(a) & \lim_{n \to \infty} \beta(\gamma) = 0. \\
(b) & \lim_{n \to \infty} \gamma = \frac{\rho^+}{\rho}. \\
(c) & \lim_{n \to \infty} n\beta(\gamma) = \ln \frac{b}{\rho-1}.
\end{align*}  \hfill (13)

**Proof:** Please see the Appendix. ■

**Remarks:** Under the saturated assumption (i.e. $p_0(\rho) \equiv 0$ or $\beta(\gamma) = \beta(\gamma)$) and the asymptotic assumption (10), [11] and [10] independently prove that the saturated collision probability $\gamma^s$ has the above asymptotic properties, where $\gamma^s = \Gamma(\beta(\gamma))$. Here, Theorem 1 proves that the general collision probability $\gamma$, where $\gamma \leq \gamma^s$ holds from (1) and (2), also has the same asymptotic properties. Significantly, Theorem 1 tells us that when $n \to \infty$, the general collision probability approaches the same limit as in the saturated case, even though the system is not saturated (i.e., $0 < p_0^+ < 1$). Note that Theorem 1 can be proved by following the proof of either Theorem 1 in [10] (as we do in the Appendix) or Theorem 7.2 in [11].

Before we set up the main result (i.e., Theorem 2), we need the intermediate results stated in Lemma 1 below. The main function of Lemma 1 is to define the service rate of the system. Let $\mu$ denote the service rate of the system in the saturated case and $\mu^*$ denote the service rate of each node in the saturated case. Since the DCF protocol is fair in the long term (i.e. each of the $n$ nodes has the same opportunity to transmit packets in a given time) intuitively, we have $\mu = n\mu^*$ or $\frac{\mu}{\mu^*} = \frac{1}{n}$. In other words, the service time of a packet in the system is equal to $1/n$ of the service time of a packet of a tagged node.

**Lemma 1:** Under assumptions (10)-(12), $R$, $\Omega$, $\mu$, and $\rho^+$ have the following asymptotic properties:

\begin{align*}
(a) & \lim_{n \to \infty} R(\gamma) = \frac{b}{\mu^*-1}. \\
(b) & \lim_{n \to \infty} \Omega(\gamma) = \Omega^* \triangleq \sigma + \frac{\mu^*}{\mu} + \frac{(b-1)\ln\frac{b}{\mu^*}}{\mu^*}(T_\text{s} - T_\text{p}). \\
(c) & \lim_{n \to \infty} \frac{\gamma^s(n)}{n} = \frac{1-p_0^+}{\mu^*}, \text{ where } \mu = \frac{\rho^+}{\mu^*} \frac{b}{b-1} \frac{\ln\frac{b}{\mu^*}}{\mu^*} \text{ and } \frac{\mu}{\mu^*} = \frac{2\ln 2}{\mu^*} + \frac{1}{\mu^*} \text{ if } b = 2 \text{ and } T_\text{s} = T_\text{p}. \\
(d) & \rho^+ \triangleq \lim_{n \to \infty} \rho(\gamma) = \frac{1-p_0^+}{\mu^*} \lambda^*, \text{ where } \lambda^* = \frac{\sigma}{\mu^*} + \frac{k^*}{\mu^*}. \\
(e) & \Phi(x, K) \text{ is decreasing with respect to } x, \text{ where } \Phi(x, K) \triangleq \frac{1+x^+e^{\sigma K^+}+x^+e^{\sigma K^+}}{1+x^+e^{\sigma K^+}+x^+e^{\sigma K^+}}, \quad x \geq 0, \text{ and } K = 1, 2, \ldots \text{.}
\end{align*}

**Proof:** Please see the Appendix. ■

**Remarks:** (i) Lemma 1 (c) gives an expression for the asymptotic service time of the system $\lim_{n \to \infty} \gamma^s(n)$. In the special case of saturation, $p_0^+ = 0$ and $\gamma = \gamma^s$ so we have $\lim_{n \to \infty} \gamma^s(n) = \frac{1}{\mu^*}$, which is consistent with our definition of $\mu$. In the general case, the asymptotic service time of the system is less than that in the saturated case by the factor $1-p_0^+$.
which is the system utilization when $n \to \infty$. (ii) Lemma 1 (c) also shows that when $n \to \infty$, the service rate $\mu = \frac{\rho^+ \ln 2}{K}$, where $\rho^+$ is the system utilization when $n \to \infty$. It is worth noting that $\mu$ is determined only by the DCF protocol parameters (such as $b$, DIFS, ACK, and so on) and the packet size (which is used to calculate $T_s$ and $T_\pi$).

We are now ready to present the main result. Theorem 2 gives sufficient conditions on the total packet arrival rate to ensure that $\rho^+ \leq 1$ holds. It also gives a result showing the effect on $\rho^+$ of increasing the buffer size $K$.

**Theorem 2:** Under assumptions (10)-(12), either $\rho^+ \leq 1$ or $\partial \rho^+ / \partial K > 0$ holds if the following conditions on the total packet arrival rate are satisfied:

(a) When $K = \infty$, $\rho^+ \leq 1$ holds if $n\lambda \leq \mu$ for all $n$.
(b) When $K < \infty$, $\rho^+ \leq 1$ holds if $n\lambda \leq \mu(K+1)/K$ for all $n$ under the M/M/1/K queue approximation.
(c) When $K < \infty$, $\partial \rho^+ / \partial K > 0$ holds if $\rho^+ \neq 1$ under the M/M/1/K queue approximation.

**Proof:**

(a) If $n\lambda \leq \mu$ for all $n$, we have $\Lambda = \lim_{n \to \infty} n\lambda \leq \mu$. Then $\rho^+ = (1 - p_0^+)\Lambda / \mu \leq \Lambda / \mu \leq 1$ holds for an arbitrary buffer size including $K = \infty$.

(b) Under the M/M/1/K queue approximation, $p_0$ is given by (9). Taking $n \to \infty$ for both sides in (9), we have

$$p_0^+ = \frac{1}{1 + \rho^+ + \ldots + (\rho^+)^K}.$$ 

and hence

$$\frac{(1 - p_0^+)}{\rho^+} = \frac{1 + \rho^+ + \ldots + (\rho^+)^{K-1}}{1 + \rho^+ + \ldots + (\rho^+)^K} = \Phi(\rho^+, K).$$

On the other hand, from Lemma 1 (d), we have

$$\frac{(1 - p_0^+)}{\rho^+} = \frac{\mu}{\Lambda}.$$ 

Then, we have

$$\Phi(\rho^+, K) = \frac{\mu}{\Lambda},$$

(14)

If $n\lambda \leq \mu(K+1)/K$ for all $n$, we have $\Lambda = \lim_{n \to \infty} n\lambda \leq \mu(K+1)/K$. Thus,

$$\Phi(\rho^+, K) = \frac{\mu}{\Lambda} \geq K/(K + 1) = \Phi(1, K).$$

Using Lemma 1 (e), we have $\rho^+ \leq 1$ from $\Phi(\rho^+, K) \geq \Phi(1, K)$. In other words, $\rho^+ \leq 1$ holds if $n\lambda \leq \mu(K+1)/K$.

(c) Under the M/M/1/K queue approximation, from (14), $\Phi(\rho^+, K) = \mu / \Lambda = constant$. According to the implicit function theorem, there exists an implicit function $\rho^+ = \rho^+(K)$ in equation $\Phi(\rho^+, K) = \mu / \Lambda$. Taking the derivative of $\rho^+$ with respect to $K$, we have

$$\frac{d\rho^+}{dK} = -\frac{\Phi'_K(\rho^+, K)}{\Phi'_{\rho^+}(\rho^+, K)}.$$ 

(15)

On one hand, since $\Phi(x, K)$ is decreasing with respect to $x$, we have

$$\Phi'_x(\rho^+, K) < 0.$$ 

(16)

On the other hand, rewriting $\Phi(x, K) = \frac{\mu}{\rho^+ - x}$, we have

$$\Phi'_K(\rho^+, K) = -\frac{(\rho^+)^K \ln \rho^+}{1 - \rho^+} \frac{1}{(1 - (\rho^+)^{K+1})^2}.$$ 

(17)

Then, from (15) to (17), $\frac{d\rho^+}{dK}$ has the same sign as $\frac{(\rho^+)^K \ln \rho^+}{1 - \rho^+}$. Therefore, we must have $\frac{d\rho^+}{dK} > 0$ for all values of $\rho^+$ except $\rho^+ = 1$.

**Remarks:** (i) We formally define the critical traffic load to be the maximal value of $n\lambda$ which, according to Theorem 2, is still sufficiently small to ensure $\rho^+ \leq 1$. Theorem 2 (a) states that for an infinite buffer, the critical traffic load is equal to $\mu$. Theorem 2 (b) states that for a finite buffer of size $K$, the critical traffic load is equal to $\mu(K+1)/K$. The results are mutually consistent, since $\mu(K+1)/K$ becomes $\mu$ when $K \to \infty$. Note that Theorem 2 (a) is derived without any restrictions on the input traffic distributions, while Theorem 2 (b) is derived under Markovian assumptions. (ii) Theorem 2 (c) manifests that under the same traffic loads, increasing the buffer size of each node has the effect of increasing the traffic intensity. Indeed, increasing the buffer size means that more packets can be backlogged in the buffer. These backlogged packets contend for the channel, increasing the average service time observed by each node, and thereby increasing the traffic intensity. (iii) Theorem 2 (a) and Theorem 2 (b) show that the larger the buffer, the smaller the critical traffic load. The reason for this is evident from Theorem 2 (c): increasing the buffer size increases the mean service time and therefore, to ensure $\rho^+ < 1$, we need to compensate by decreasing the total packet arrival rate.

Applying Theorem 2 (a), we immediately have

**The critical-load-based CAC scheme:** For a one-hop homogeneous DCF network with infinite node buffers, the CAC rule is

$$n\lambda \leq \mu,$$

where $\mu$ is defined in Lemma 1 (c).

In an implementation of our scheme, the admission control function would reside in a central entity, such as the AP in an infrastructure WLAN. We propose that the AP would acquire knowledge of the three parameters needed for our CAC scheme (namely $\lambda$, $L$, and $n$) by making use of the admission control signaling mechanism already specified in EDCA [7]. The EDCA signaling mechanism defines a request/response handshake whereby a station requests the AP for admission and learns of the decision via a response message. The request message has fields for the traffic specification (TSPEC) of the intended flow, including the packet arrival rate $\lambda$ and the packet size $L$. The use of the signaling mechanism makes it trivial for the AP to keep track of the total number $n$ of active nodes plus the node requesting admission. In a one-hop ad-hoc WLAN there is no AP, but since every node can hear all other nodes in range, the information transmitted in the network, one of the nodes could be elected to act as the central entity that performs the admission control. Note that while EDCA specifies an admission control signaling mechanism, admission control is optional and the standard leaves the admission control algorithm for vendor definition.

**III. Model verification**

In this section, we demonstrate the efficiency of the proposed CAC scheme for homogeneous DCF networks. We use the 802.11 simulator in ns2 version 2.28 [16] as a validation tool, and set the protocol parameters to the default values for
802.11b, as listed in Table I; note that in the table, a slot is equal to 20 $\mu$s and $\delta$ denotes the propagation delay. In our simulation, we used the NOAH routing protocol [11] and we removed some bugs in the 802.11 simulator, as reported in [17]. Each simulation value is an average over 5 simulation runs, where each run was for 100 seconds. Unless otherwise specified, each node and the AP have an buffer size of 1000 packets, which is used to mimic the infinite buffer. We present the theoretical results under the assumption of $T_s = T_{\text{th}}$, but this assumption can be removed, as explained in [8].

The proposed CAC scheme is applicable for a wide range of traffic types. However, due to space limitations, we focus on real-time traffic to demonstrate that our scheme works well. We ran two experiments for homogeneous environments. The first experiment involves real-time data streams with CBR characteristics and we consider the one-hop star network with an AP, where the AP only acts as the receiver of data packets from all nodes. The second experiment involves voice streams with on-off characteristics and we consider the one-hop ad hoc network. According to the relevant ITU standard [18], the preferred one-way transmission delay for a real-time stream is 150 ms. On the other hand, according to [19], the Internet delay and coding delay are about 100 ms and 25 ms, respectively. Therefore, we set the maximum admissible mean total delay per packet in the one-hop WLAN to 25 ms.

We compare our theoretical results for the critical-load-based CAC scheme (denoted by ana_CLB) with the simulation results (denoted by sim) in terms of the maximum number of connections, $N_{\text{max}}$, which can be admitted without violating the respective constraints. For the theoretical maximum number of connections, $N_{\text{max}}$ is constrained by $N_{\text{max}} \lambda < \mu$, where $\mu$ is defined in Lemma 1 (c); for the simulated maximum number of connections, $N_{\text{max}}$ is constrained by the above-mentioned delay requirement for both voice and data.

That is, if accepting a new connection will violate the admitted total delay of 25 ms, $N_{\text{max}}$ is set to the current number of connections excluding the new connection.

The first experiment investigates the performance for 5 different types of real-time data streams: AP-50, AP-48, AP-46, AP-44, and AP-42, as defined in Table II (a). Fig. 1 (a) compares the theoretical and simulation results for $N_{\text{max}}$. The theoretical results underestimate the simulated maximum number of streams defined by the delay constraint, but only by a small margin, indicating that our proposed CAC rule is very efficient. To highlight that underestimating is preferable to overestimating, we plot simulation results for the mean total packet delay versus the number of data streams in Fig. 1 (b). For each type of stream, we see that the total delay jumps to thousands of milliseconds from several milliseconds when the number of streams exceeds $N_{\text{max}}$.

The second experiment considers the voice over WLAN scenario. In this experiment, we investigate 5 different types of voice streams, which are the same as used in [6]: G.711-100, G.711-50, iLBC, G.729, and G.723a, as defined in Table II (b).
In our simulation, each voice connection is between a pair of nodes. The voice traffic consists of on and off periods, where the on period of one node corresponds to the off period of the other, and the on and off periods are exponentially distributed with mean 1.5 seconds. Periods of less than 240 ms are increased to 240 ms in length to reproduce the minimum talkspurt period. In the analytical model, we use Poisson traffic with the same arrival rate to approximate the on-off traffic.

The reason that we can adopt this approximation is because [8] demonstrated that 802.11 nonsaturated performance is determined by the mean arrival rate of the input traffic and is largely insensitive to the input traffic arrival distribution.

Fig. 2 (a) tabulates the theoretical and simulation results for $N_{\text{max}}$, and Fig. 2 (b) shows the expanded simulation results. As observed in the first experiment, the proposed CAC rule for ad hoc networks is very efficient, and the error leads to a conservative policy.

### IV. Conclusion

This paper proposes a critical-offered-load-based CAC scheme for homogeneous DCF networks. The proposed CAC scheme is not only simple, practical, robust, but also compatible with IEEE 802.11 EDCA standard. The work in this paper serves as a stepping stone for the design of powerful CAC schemes for more realistic heterogeneous DCF networks.

### V. Appendix

In this appendix, we present the proof of Theorem 1 and Lemma 1.

**Proof of Theorem 1:**

[11] has shown that under asymptotic assumption (10), (3) becomes

\[
\beta^c(\gamma) = \begin{cases} 
\frac{1-\gamma}{b_0(1-\gamma)}, & 0 \leq \gamma < \frac{1}{b} \\
0, & \gamma \geq \frac{1}{b} 
\end{cases}, \tag{18}
\]

Thereafter, focusing on $\gamma \in [0, \frac{1}{b}]$, we prove (13) by following the proof of Theorem 1 in [10]. From (1), (2), and (18), we have

\[
\beta(\gamma) = 1 - (1 - \gamma)^{\frac{1}{\gamma b}} = (1 - p_0) \frac{(1 - \gamma b)}{b_0(1 - \gamma)}. \tag{19}
\]

Taking an infinite limit of $n$, using utilization assumption (12), and noting

\[
\lim_{n \to \infty} \beta(\gamma) = 0 = (1 - p_0)^{\frac{(1 - \gamma b)}{b_0(1 - \gamma)}}, \tag{19}
\]

we have $\lim_{n \to \infty} (1 - \gamma b) = 0$, which implies

\[
\lim_{n \to \infty} \gamma = 1/b. \tag{20}
\]

Now, since $\lim_{n \to \infty} \beta(\gamma) = 0$ and $\lim_{n \to \infty} \gamma = 1/b$, taking (2) to the limit as $n \to \infty$,

\[
\frac{1}{b} = \lim_{n \to \infty} \gamma = 1 - \lim_{n \to \infty} (1 - \beta(\gamma))^n = 1 - \lim_{n \to \infty} e^{n \ln(1-\beta(\gamma))} = 1 - \lim_{n \to \infty} e^{n \beta(\gamma) \ln(1-\beta(\gamma))} = 1 - e^{-\lim_{n \to \infty} n \beta(\gamma)}.
\]

Thus, we have

\[
\lim_{n \to \infty} n \beta(\gamma) \to \ln \frac{b}{b - 1}. \tag{21}
\]

From (19), (20), and (21), we conclude the proof. ■

**Proof of Lemma 1:**

(a) The conclusion holds immediately from (4) and Theorem 1 (b).

\[
\lim_{n \to \infty} R(\gamma) = \lim_{n \to \infty} \frac{1}{1 - \gamma} = \frac{b}{b - 1}. \tag{22}
\]

(b) The conclusion holds immediately from (6), (8), and Theorem 1.

\[
\lim_{n \to \infty} P_b(\gamma) = \lim_{n \to \infty} 1 - (1 - \gamma)^{\frac{1}{\gamma b}} = \frac{1}{b}, \tag{23}
\]

\[
\lim_{n \to \infty} P_s(\gamma) = \lim_{n \to \infty} n \beta(\gamma) \cdot (1 - \gamma) = \frac{b - 1}{b} \ln \frac{b}{b - 1}, \tag{24}
\]

\[
\lim_{n \to \infty} \Omega(\gamma) \to \lim_{n \to \infty} [\sigma + P_s T_s + P_b T_b] = \Omega^*. \tag{25}
\]

(c) From (1), (3), and (6), we have

\[
\frac{Y^c(\gamma)}{n} = \frac{X(\gamma) \Omega(\gamma)}{n} = \frac{1}{n \beta(\gamma)} R(\gamma) \Omega(\gamma) \tag{24}\]

\[
= \left[1 - p_0 R(\gamma) \Omega(\gamma) \right]. \tag{24}
\]
Taking the limit as \( n \to \infty \) and applying Theorem 1 and Lemma 1 (a) (b), we have
\[
\lim_{n \to \infty} \frac{Y_n(\gamma)}{n} = (1 - p_0^+)/\mu.
\]
(d) The conclusion holds immediately from Lemma 1 (c).
\[
\rho(\gamma) = \frac{n \rho(\gamma)}{n} = \frac{n \lambda}{n} \cdot \frac{Y_n(\gamma)}{n}.
\]  
(25)
Taking the limit as \( n \to \infty \) and applying Lemma 1 (c), we have
\[
\rho^+ = \lim_{n \to \infty} \rho(\gamma) = (1 - p_0^+)/\mu.
\]
(e) The conclusion holds immediately from (26).
\[
\Phi(x, K) = 1 - \frac{x^K}{1 + x + \cdots + x^K} = 1 - \frac{1}{1 + \frac{1}{x} + \cdots + \frac{1}{x^n}}.
\]  
(26)

\section*{REFERENCES}


Qinglin Zhao received the PhD degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005. From October 2006 to August 2009, he worked as a postdoctoral researcher in the Department of Electronic and Computer Engineering, the Hong Kong University of Science and Technology (HKUST), HK SAR. Since September 2009, he has been with the Faculty of Information Technology, Macau University of Science and Technology, Avenida Wei Long, Taipa, Macau, China. His research interests include wireless and mobile networks, mobility management, and performance analysis.

Danny H. K. Tsang (M’82-SM’00) received the B.Sc. degree in mathematics and physics from the University of Winnipeg, Canada, in 1979, and the B.Eng. and M.A.Sc. degrees both in electrical engineering from the Technical University of Nova Scotia, Canada, in 1982 and 1984, respectively. He also received the Ph.D. degree in electrical engineering from the Moore School of Electrical Engineering at the University of Pennsylvania, U.S.A., in 1989.

Upon the completion of his PhD degree, he joined the Department of Mathematics, Statistics and Computing Science of Dalhousie University in Canada where he was an Assistant Professor in the Computing Science Division. While he was in Canada, he also registered as a Professional Engineer in the province of Nova Scotia. He has joined the Department of Electronic & Computer Engineering at the Hong Kong University of Science and Technology since summer of 1992. He is now a professor in the department. His current research interests include Internet quality of service, P2P video streaming over the Internet, and wireless multimedia networks.

He was the General Chair of the IFIP Broadband Communications’99 held in Hong Kong and received the Outstanding Paper from Academia Award at the IEEE ATM Workshop’99. He also served as Technical Program Committee (TPC) member of ITC18-20, ACM Multimedia 2002, ICON2002, Opticom’2002, ICC 2006 and INFOCOM 1994-96 and 2006-07. He was a Guest Editor for the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS’ special issue on Advances in P2P Streaming Systems and an Associate Editor for the JOURNAL OF OPTICAL NETWORKING published by the Optical Society of America. He is currently an Associate Technical Editor for the IEEE COMMUNICATIONS MAGAZINE. During his leave from HKUST in 2000-2001, Dr. Tsang assumed the role of Principal Architect at Sycamore Networks in the United States. He was responsible for the network architecture design of Ethernet MAN/WAN over SONET/DWDM networks. He invented the 64b/65b encoding scheme (US Patent number 6,952,405) for Transparent GFP in the T1X.1.5 standard which was advanced to become the ITU G.GFP standard. The coding scheme has now been adopted by International Telecommunication Union (ITU)’s Generic Framing Procedure recommendation GFP-T (ITU-T G.7041/Y.1503).

Taka Sakurai (M’02) received the B.Sc. degree in mathematical sciences and B.E. degree in electrical engineering from the University of Adelaide in 1988 and 1989, respectively, and the Ph.D. degree in electrical engineering from the University of Melbourne in 2003. From 1991 to 1997, he was a Research Engineer at Telstra Research Laboratories. Subsequently, he held research and development roles at NEC and Lucent Technologies. From 2003 to 2005, he was with the Department of Electrical and Electronic Engineering, University of Melbourne. Currently, he is with the Chief Technology Office of Telstra Corporation, and an Honorary Fellow of the University of Melbourne. His research interests are in the areas of performance analysis of next-generation wireless networks, design of MAC protocols for wireless LANs and sensor networks, and computational probability.