

Service Time Approximation in IEEE 802.11 Single-Hop Ad Hoc Networks

Atef Abdrabou and Weihua Zhuang, *Senior Member, IEEE*

Abstract—This paper investigates the near-memoryless behavior of the service time for IEEE 802.11 saturated single-hop ad hoc networks. We show that the number of packets successfully transmitted by any node over a time interval follows a general distribution, which is close to a Poisson distribution with an upper bounded distribution distance. The bound on the distribution distance is almost constant and is mainly affected by some system parameters and very slightly by the number of active nodes in the network. We also show that the service time distribution can be approximated by a geometric distribution. We illustrate that the usage of discrete-time queuing analysis (M/Geo/1) near network saturation greatly simplifies the queuing analysis and leads to sufficiently accurate results for both the first order statistics and the probability distribution of the number of packets in the queuing system. Computer simulation results demonstrate that the M/Geo/1 queuing model is very accurate.

Index Terms—IEEE 802.11 MAC, ad hoc network, service time, packet delay, distribution distance, queuing model, resource allocation, call admission control

I. INTRODUCTION

IN recent years, contention-based medium access control (MAC) protocols (such as IEEE 802.11) are widely adopted in wireless local area networks (WLANs). There are many different research works that address 802.11 performance analysis in the open literature (for example, [1], [2], [3] and references therein). Nevertheless, in the bulk of the research works, the analysis of very important design parameters (such as MAC packet delay and service time) is done in terms of the first and second order statistics only. The MAC service time under consideration in this paper is defined as the delay seen by a packet from the instant of being at the head of the queue to the instant of being successfully transmitted. The service time is vital for handling any 802.11 queuing model. Our objective in this paper, in contrary to most of previous research works, is not to analyze the performance of IEEE 802.11 by including the impact of the queuing model. We mainly aim at reaching a sufficiently accurate approximation for the service time distribution that can easily be used in statistical resource allocation (call admission control and/or resource reservation) decisions. In fact, using the first and second order statistics may lead to inefficient network resource utilization or ineffective quality-of-service (QoS) provisioning.

Although not much researches in the literature study the queuing models of 802.11 [2] [4] [5] [6], we can identify

four different queuing disciplines; namely, M/G/1 [2] [7] [8], M/MMGI/1/K [4], G/G/1 [5], [9] and M/M/1/K [2], [6]. Two of these disciplines M/G/1 [2], [7] and G/G/1 [5], [9] treat the 802.11 as a server with a general service time distribution. The queuing analysis with a general service time distribution can be carried out either by (i) finding the distribution itself; (ii) using the estimated average and variance of an unknown distribution; or (iii) approximating the general distribution, if possible, to an easy-handling one such as exponential or geometric. Finding a close-form expression for the service time probability density function (PDF) is a mathematically challenging task. In fact, the distribution is complicated since, between two successful packet transmissions of any node, three different random variables (in the case of a fixed packet size) are involved; namely, the number of idle time slots, the number of collisions happened (either to other nodes or to the node under consideration), and the number of successful transmissions of the other nodes. Moreover, these random variables are not independent since the number of successful transmissions and collisions from the others nodes (between successful transmissions of a given node) depends on the backoff counter value. As the counter value increases, more successful transmissions and collisions are likely to happen. Also, the number of successful transmissions of the other nodes depends on how many collisions they suffered. On the other hand, previous analysis and simulation results indicate that a large number of packets have a very short service time and a small number of packets experience a very long one (i.e., the packet service time is not close to its average value) [2], [5], [9]–[12]. Using only the average value in resource allocation may lead eventually to a conservative estimation of the available resources, which in turn reduces the utilization of the network resources.

In this paper, we study the service time distribution for the 802.11 DCF with the request-to-send/clear-to-send (RTS/CTS) access method. We seek a simplified approximation mainly to be used for efficient statistical call admission control and resource reservation in ad hoc networks. The paper presents three related contributions that lead to the realization of this objective:

- It is shown that the service time distribution has a partial memoryless behavior. We demonstrate that the distribution of the number of packets successfully transmitted over a time interval from any of the active nodes in a saturated ad hoc network follows a general distribution that is close to the Poisson distribution with an upper bounded distribution distance. The Poisson process is a renewal counting process with a memoryless distribution for the renewal inter-arrival times [13]. We obtain this

Manuscript received July 31, 2006; revised October 4, 2006; accepted November 10, 2006. The associate editor coordinating the review of this paper and approving it for publication was R. Fantacci.

The authors are with the Centre for Wireless Communications (CWC), Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: {alotfy, wzhuang}@bber.uwaterloo.ca).
Digital Object Identifier 10.1109/TWC.2008.060530.

bound analytically using the Chen-Stein approximation method [14] and verify it by simulations. We also show that the bound is almost a constant, which depends mainly on some system parameters and very slightly on the number of active nodes in the network.

- We illustrate that the service time distribution, with its near memoryless behavior and the discrete nature shown in [2] [7] [10], can be approximated by the geometric distribution. We characterize the distribution by analytically deriving its parameter.
- We propose to use the discrete-time queuing system (M/Geo/1) as a queuing model for IEEE 802.11 single-hop ad hoc networks near saturation. We show that the average queue length and the probability distribution of the number of packets in the queuing system obtained by computer simulations match closely the analytical results obtained from the M/Geo/1 queuing system.

The significance of this research lies in the introduction of a simple approximation to the service time distribution, which can be used with sufficient accuracy in the queuing analysis and the prediction of the buffer occupancy for the sake of QoS provisioning. Distributed resource allocation mechanisms (such as call admission control) are mandatory in ad hoc networks which lack a centralized controller. This research offers a step toward a fully distributed statistical call admission control for single-hop ad hoc networks, based on the PDF of the buffer occupancy instead of just first or second order statistics. Any node with a minimal amount of information from its neighbors (i.e. the number of neighboring nodes) can determine the possibility of its call being admitted with its QoS constraints (such as delay) being satisfied without degrading the QoS provisioning of the ongoing calls.

The rest of paper is organized as follows. Section II presents some related works. We introduce the system model in Section III. In Section IV, we discuss the partial memoryless behavior of the service time in the IEEE 802.11 at saturation. Section V presents our proposed approximation to the service time and the M/Geo/1 queuing system. We verify the analysis by simulation results in Section VI. Finally, we conclude this research in Section VII.

II. RELATED WORKS

Studying the service time distribution of the IEEE 802.11 DCF has drawn the attention of many researchers since it is essential for performance evaluation and queuing analysis. In [5], [9], [7], [10] and [8], exact close-form expressions of the probability generating function (PGF) of the service time are derived. The PGF expressions can be converted to the PDF only numerically, which makes them not practical to use in making dynamic statistical resource allocation decisions. In [15], an approximation to the service time PGF has been given and shown to be accurate. However, the approximated PGF is a general distribution, which is not easy to handle with queuing analysis although it is simpler than the exact close-form expressions. In [16], an approximation to the asymptotic distribution of the total delay (M/G/1 queuing delay plus the service time) has been shown to follow a power law. This approximation is given under certain assumptions such as large

total delay and non-integer binary logarithm of the collision probability.

The assumption of Markovian service time in the 802.11 queuing discipline (such as M/M/1/K [2], [6]) has not been analytically verified, to the best of our knowledge. Zhai et al. in [2] compare the service time distribution obtained by simulations graphically with standard distributions and conclude that an exponential distribution may give a good approximation to the inter-arrival times of successfully received packets. Foh and Zukerman [17] and Tantra et al. [18] study the IEEE 802.11 DCF performance by modeling a WLAN with K nodes using an M/PH/1/K queuing analysis (where each node waits in the queue to be served). In [17] and [18], a phase type distribution (such as Erlang with parameter 8 in [17]) is used to approximate the service time based on simulation results and graphical comparison to the actual service time distribution (obtained by simulations). It is also assumed in [17] and [18] that every node can only keep one data frame in its queue (low utilization factor), which impacts the service time distribution for a low or medium number of nodes. Pham et al. in [6] use the M/M/1/K discipline, assuming that the service time is exponential without verifying the assumption. In [19], it is shown that the service interval distribution can converge to an exponential distribution when the number of nodes in the network is sufficiently large. However, the definition of the service interval or the service time for a node in [19] is the same as the slot time in [1]. The author in [19] uses the expression for the average slot time given in [1] to analytically describe the average service interval for a node. The operation of the IEEE 802.11 protocol is based on a slotted time. The slot time in [1] is defined as either the unit slot time (when the channel is idle), or the packet transmission time (when the channel is busy sending a packet), or the time for a collision to be detected on the channel. The service time definition under consideration in this paper is substantially different; it is also used by many other researchers [2], [3], [10]- [12].

III. SYSTEM MODEL

We consider a single-channel IEEE 802.11 single-hop ad hoc network that contains a cluster of terminals (nodes). The nodes use the DCF mechanism to access the channel. The random access employs the four-way RTS-CTS-DATA-ACK handshaking procedure. All the nodes have the same transmission range, and are randomly distributed in an area with dimensions limited to the node's transmission range. As a result, all the nodes can hear each other, and there are no hidden or exposed terminals. Only half of the nodes are active traffic sources, the other half are only receivers. The network represents a single-hop ad hoc network; every sender (active) node sends data packets to one unique receiver node. For simplicity in studying the 802.11 protocol operation, we assume that the transmitted packets may be lost only due to collision. We consider that the network is in saturation condition (i.e. all the active nodes have backlogged packets all the time) unless otherwise specified.

Following the carrier-sense multiple access with collision avoidance (CSMA/CA) protocol as described in the IEEE 802.11 standard [20], if a node has packet to transmit and senses the channel to be idle for a period of Distributed

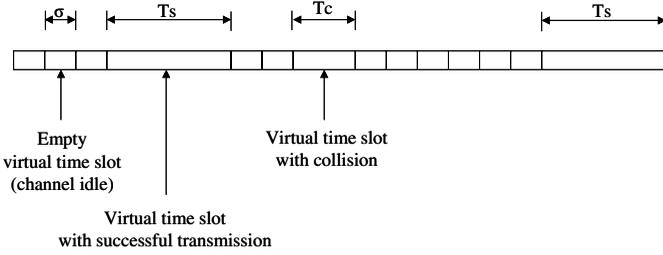


Fig. 1. Virtual time slots.

InterFrame Spacing (DIFS), the node proceeds by transmitting an RTS packet. If the channel is busy, the node defers its transmission until an idle DIFS is detected and waits for a random backoff time in order to avoid collisions. The backoff time counter is chosen uniformly in the range $[0, W_i - 1]$, where $i \in [0, m_b]$, m_b is the number of backoff stages, and W_i is the current contention window (CW) size in time slots. A time slot is the unit time in IEEE 802.11. The contention window at the first transmission of a packet is set equal to CW_{min} . After an unsuccessful transmission, the CW is doubled up to a maximum value

$$CW_{max} = 2^{m_b} \times CW_{min}. \quad (1)$$

The backoff counter decreases at every slot time when the channel is sensed idle. The counter is stopped when the channel is busy and resumed when the channel is sensed idle again for more than DIFS. A station transmits the RTS packet when its backoff timer reaches zero. If the destination station successfully receives the RTS packet, it responds with a CTS packet after a short inter-frame space (SIFS) time interval. Upon the reception of the CTS packet, the sender sends the data packet. The receiver then waits for an SIFS time interval and transmits an acknowledgment (ACK) packet. If the ACK packet is not received within a specified ACK timeout interval, the data packet is assumed lost and a retransmission will be scheduled.

We assume a fixed packet size. The packet transmission time T_s is given by [1]

$$T_s = T_{RTS} + T_{CTS} + 3 SIFS + T_{ACK} + T + DIFS \quad (2)$$

and the packet collision time is given by [1]

$$T_c = T_{RTS} + DIFS. \quad (3)$$

The symbols T_{RTS} , T_{CTS} and T_{ACK} represent the transmission times for the RTS, CTS and ACK packets, respectively; T is the data packet transmission time, which is constant for a fixed packet size.

Here we differentiate two types of time slots: physical time slot and virtual time slot. The physical time slot (the unit time) has a fixed length denoted by σ . A virtual time slot is the time during which the channel does not change its state (busy or idle) as indicated in Figure 1. A virtual time slot may contain one or more physical time slots. If the node is backing off and the channel is idle, the virtual time slot is equal to one physical slot (σ). If the channel is detected busy, the node stops

 TABLE I
 IEEE 802.11 SYSTEM PARAMETERS [20].

System Parameter	Value
Packet payload	256 Bytes
PHY header	128 bits
ACK	112 + PHY header
RTS	160 + PHY header
CTS	112 + PHY header
Slot Time	50 μ s
SIFS	28 μ s
DIFS	128 μ s
Basic Rate	1 Mbps
Data Rate	2 Mbps
CW_{min}	16
Backoff stages (m_b)	5

decrementing its backoff counter and waits for two virtual time slots (one when the channel is busy and one when it is idle) before it starts decrementing its backoff counter again. The virtual time slot is equal to T_s if the channel is busy sending a packet successfully, or equal to T_c if a collision happened. If the node is transmitting, it takes one virtual time slot (with duration T_s if no collision happens or T_c otherwise) to finish its transmission. Thus, a virtual time slot duration is a random variable, denoted by s , that has three possible values with different probabilities as follows:

$$s = \begin{cases} \sigma, & P(s = \sigma) = 1 - P_{tr} \\ T_s, & P(s = T_s) = P_{tr}P_s \\ T_c, & P(s = T_c) = P_{tr}(1 - P_s) \end{cases} \quad (4)$$

where [1]

$$P_{tr} = 1 - (1 - \tau)^N \quad (5)$$

is the probability that the channel has at least one transmission in the considered slot time, τ is the probability that a node transmits in a randomly chosen time slot, given by

$$\tau = \frac{2(1 - 2p)}{(1 - 2p)(CW_{min} + 1) + p CW_{min}(1 - (2p)^{m_b})} \quad (6)$$

and

$$P_s = \frac{N\tau(1 - \tau)^{N-1}}{P_{tr}} \quad (7)$$

is the probability that the channel has a successful transmission. The average virtual slot time is then given by

$$E[s] = (1 - P_{tr})\sigma + P_{tr}P_sT_s + P_{tr}(1 - P_s)T_c. \quad (8)$$

IV. THE NEAR-MEMORYLESS BEHAVIOR OF IEEE 802.11

In this section, we study the behavior of the random counting process that controls the number of packets successfully transmitted by any of the contending nodes in the saturated network. We show that this counting process has a nearly memoryless behavior. We prove analytically using the Chen-Stein approximation method that the probability of the number of packets sent over a time interval by any active node follows a distribution that is close to a Poisson distribution with an upper bounded distribution distance. We also discuss the possible causes of this behavior; namely, the fairness of the IEEE 802.11 and the randomness of the CSMA/CA backoff

procedure. In the following, brief overviews of the Chen-Stein approximation and the IEEE 802.11 fairness are given for the sake of completeness.

A. Chen-Stein Approximation

The Chen-Stein approximation is a more generalized form of the “law of small numbers”. The law states that the distribution $B(n, p_b)$ can converge to the Poisson distribution P_λ , where $\lambda = np_b$, for small p_b and very large n [14] [19] as long as $B(n, p_b)$ can be represented as the sum of n independent and identically distributed Bernoulli (indicator) random variables where each indicator equals to one with probability p_b . The law of small numbers applies only to this class of variables. However, the Chen-Stein approximation method extends the law to measuring the convergence rate (the distribution distance) between P_λ and $B(n, p_b)$ as n goes large and relaxes to some extent both the identical distribution and the independence assumptions [14]. In our case, the indicator random variables are independent and identically distributed. Therefore, the distribution under consideration can be described by the random variable W as follows

$$W = \sum_{i=1}^n I_i \quad (9)$$

where I_1, I_2, \dots, I_n are independent and identically distributed random variables and

$$p_{bi} = P(I_i = 1) = E[I_i]. \quad (10)$$

According to the Chen-Stein method [14], the distribution distance between the cumulative distribution function (CDF) of the actual distribution $P(W \in A)$ and the Poisson CDF $P_\lambda(A)$, where $A \subset Z^+$, is bounded by

$$|P(W \in A) - P_\lambda(A)| \leq \frac{(1 - e^{-\lambda})}{\lambda} \sum_{i=1}^n p_{bi}^2. \quad (11)$$

B. MAC Fairness

MAC fairness refers to the ability of the link layer to allow contending nodes to equally access a channel. The CSMA/CA technique used in IEEE 802.11 is not perfectly fair, but it can achieve long-term fairness with a high fairness index [21]. This implies that the probability (the fraction of the number of times) the channel has been accessed by one node successfully can be considered to be $1/N$ on the long term, where N is the number of contending nodes. Since we aim at approximating the distribution of the number of packets successfully transmitted over a time interval, a question here is how short the time interval could be. Recently it has been shown in [22] and [23] that the IEEE 802.11 MAC intrinsically (without the hidden terminal problem) also has a short-term fairness. The short term is in the order of tens of milliseconds [22]. As our aim is to provide a tool for statistical resource allocation in order to provision QoS, the short term fairness implies the validity of our analysis for multimedia traffic sessions which usually have durations in the order of minutes [24] [25]. The short-term fairness can be proved as long as the CSMA/CA backoff procedure, as mentioned in the system model, follows the IEEE 802.11 standard [20]. We

limit our analysis only to the IEEE 802.11 standard since other implementations of the CSMA/CA backoff procedure (such as WaveLAN) is proved to be short-term unfair [21].

C. Distribution Distance

The motivation to study the distance between the distribution of the number of successfully transmitted packets and the Poisson distribution is driven by our intuition that the IEEE 802.11 has a kind of memoryless behavior. This behavior can be explained from the following two aspects:

- The number of packets successfully transmitted by any active node at saturation over a time interval is a renewal process, since the node restarts again its contention window to the minimum size after every successful transmission and it always has backlogged packets;
- The node contends for the channel with the same collision probability, irrespective of the number of retransmissions it had before, though the contention window size doubles after every retransmission [1]. In fact, the probability of a successful transmission of any node (when contending for the channel) is the same irrespective of how long it has been waiting for the transmission.

We use the Chen-Stein approximation as a tool to quantify mathematically the distribution distance. We introduce a mathematical model to the random counting process that describes the number of successfully transmitted packets over a time interval (one second for simplicity).

In this model, any active node has a number of virtual time slots with successful transmissions in one second. This number can be considered as the number of successful trials out of the total number of other virtual time slots that the channel has in one second, as illustrated in Figure 2. The figure shows the virtual time slots that contain successful transmissions for a certain active node in black, and other white slots corresponding to successful transmissions or collisions from other nodes, idle channel or collisions associated with the same node.

Under the assumption that the packet of any node sees a collision with constant and independent packet collision probability p , the relation between the probability τ that a node sends a packet at a random time slot and p is given by [1]

$$p = 1 - (1 - \tau)^{N-1}. \quad (12)$$

We model the number of successfully transmitted packets over a time interval of one second from a certain node as the summation of indicator random variables I_i , where

$$I_i = \begin{cases} 1, & \text{A certain node transmitted a packet} \\ & \text{successfully (no collision) in slot } i \\ 0, & \text{Otherwise} \end{cases} \quad (13)$$

with

$$P(I_i = 1) = \tau (1 - p) = \tau (1 - \tau)^{N-1} \triangleq q. \quad (14)$$

In (14), the probability is observed from the perspective of a certain node, where i is the slot index as the transmission in 802.11 MAC is time slotted. Therefore, the number of

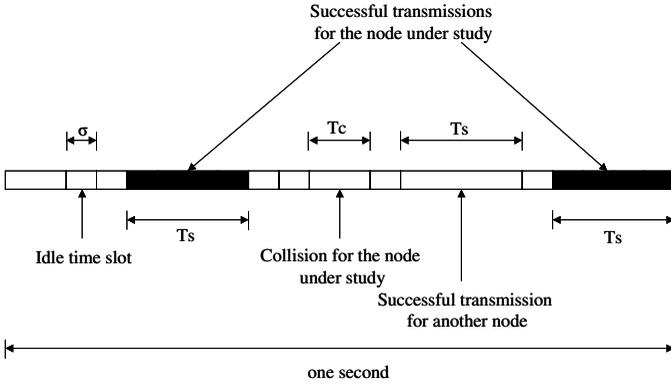


Fig. 2. Successful transmission virtual time slots for a node.

successfully transmitted packets in one second from a certain node is given by

$$W = \sum_{i=1}^M I_i \quad (15)$$

where M is a random variable represents the total number of virtual time slots in one second. Given M , the expected number of packets sent per second by a certain node is represented by

$$\begin{aligned} \lambda \triangleq E(W|M = m) &= E\left[\sum_{i=1}^m I_i | M = m\right] \quad (16) \\ &= \sum_{i=1}^m E(I_i) = m\tau(1-p). \end{aligned}$$

We assume that I_i and M are independent. The independence is reasonable since under the saturation condition M takes very large values since the backoff procedure should be executed after each collision or successful transmission. This implies the number of idle virtual time slots is much higher than the number of virtual collision slots or the number of virtual successful transmission slots. The duration of an idle time slot is in the order of tens of micro seconds as in Table I. Because of the fairness of the MAC and the saturation condition, all the active nodes are treated similarly. Hence, the distribution distance in (11) for any node is bounded by

$$\begin{aligned} &|P(W \in A|M = m) - P_\lambda(A|M = m)| \\ &\leq \frac{N(1 - e^{-\lambda})}{\lambda} \sum_{i=1}^m \tau^2 (1 - \tau)^{2(N-1)} \end{aligned}$$

By evaluating the summation and substituting the value of λ , we have

$$\begin{aligned} &|P(W \in A|M = m) - P_\lambda(A|M = m)| \\ &\leq N\tau(1 - \tau)^{N-1}(1 - e^{-\lambda}) \quad (17) \end{aligned}$$

which leads to

$$\begin{aligned} &|P(W \in A) - P_\lambda(A)| \\ &\leq N\tau(1 - \tau)^{N-1} \sum_m (1 - e^{-\lambda}) P(M = m) \end{aligned}$$

The random variable M represents the number of virtual slots within a certain time period (one second). The duration of a virtual time slot with successful transmission, T_s , is longer

than the other two types of virtual slots; namely, an idle slot and a slot with collision. Therefore, if almost all¹ the virtual slots contain successful transmission, the random variable M would take its smallest value and hence λ , which directly depends on M , would take its smallest value. In this case, M roughly takes the value $1/T_s$, which is in the order of hundreds virtual slots per second (according to the parameters given in Table I) making λ in the order of tens of packets per second. Therefore, the exponential term $e^{-\lambda}$ approaches zero and the distribution distance can be approximately bounded by

$$|P(W \in A) - P_\lambda(A)| \leq N\tau(1 - \tau)^{N-1}. \quad (18)$$

It has been shown in [1] that the saturation throughput, for a certain number of active nodes in the network, has a maximum value that can be achieved by fine tuning of the probability τ . The tuning can be done by changing the minimum size of the contention window CW_{min} and/or the number of backoff stages m_b . It can be noticed from [1] that the maximum saturation throughput for the RTS/CTS access scheme approaches the saturation throughput calculated at the standardized values for both CW_{min} and m_b [20] (i.e. 16, 32 or 64 for CW_{min} and 5 for m_b) and a sufficiently large number of nodes ($N \geq 5$). As indicated in Table I, we use the standard values for both CW_{min} and m_b (i.e. $CW_{min} = 16$ and $m_b = 5$). For these standard values, the transmission probability for a maximum throughput is approximately given by [1]

$$\tau \approx \frac{1}{NK} \quad (19)$$

where

$$K = \sqrt{\frac{T_c}{2\sigma}} \quad (20)$$

which leads to

$$N\tau(1 - \tau)^{N-1} \approx \frac{1 - e^{-1/K}}{K(e^{1/K} - 1)}. \quad (21)$$

Thus, the distribution distance becomes

$$|P(W \in A) - P_\lambda(A)| \leq \frac{1 - e^{-1/K}}{K(e^{1/K} - 1)}. \quad (22)$$

Equation (22) shows that there is an almost constant upper bound on the distribution distance. The bound depends mainly on K , which in turn depends on system parameters T_c (i.e. T_{RTS} and $DIFS$) and σ . An approximated upper bound value of 0.3 is obtained from (22) when using the IEEE system parameters given in Table I. It implies that IEEE 802.11 has a kind of near-memoryless behavior, which is aligned with our intuition, but not completely memoryless since the upper bound is not small. This is due to that IEEE 802.11 is not completely fair, as the protocol may favor the node that had a successful transmission before to transmit successfully again and again. Also, the discrete nature of the slotted operation limits the packet service time to discrete values. Therefore, the renewal counting process for successfully transmitted packets does not exactly have independent increments, which explains the deviation from the Poisson process.

¹The channel should be idle for at least one time slot between two successful transmissions.

V. SERVICE TIME APPROXIMATION

Our service time approximation stems from the mathematical model introduced in the previous section for the counting process of the number of successfully transmitted packets. Here, we approximate the random length of the virtual time slot by its average value $E[s]$. By this approximation, the number of virtual time slots over a time interval becomes a deterministic value. Thus, the counting process now describes the number of virtual time slots with successful transmissions (successful trials) out of the total number of virtual time slots (total number of trials) over a time interval t , which is the typical binomial process

$$B(t) = \sum_{i=1}^{\lfloor t/E[s] \rfloor} I_i \quad (23)$$

where I_i is defined in (13). It can be shown that, for a binomial process, the time between successive events (the service time in our case) follows a geometric distribution [13]. The geometric distribution is a probability distribution for discrete random variables, and suits well the discrete slotted nature of IEEE 802.11. In addition, it has a memoryless nature [26]. This can be intuitively explained: the fact that we have done n trials and got failures does not change how many more times we still have to try to get the next success.

Therefore, the probability that the service time equals n virtual time slots is given by

$$P\{t_s = n\} = q(1-q)^{n-1} \quad (24)$$

where the distribution parameter q is the successful transmission probability given by (14).

As a result, the average service time μ can be obtained from

$$\frac{1}{\mu} = \sum_{i=1}^{\lfloor 1/E[s] \rfloor} E[I_i] = \frac{\tau \cdot (1-p)}{E[s]} \quad (25)$$

which is consistent with (14). The service time distribution given by (24) is discrete with an exponential-like decay that really resembles the actual distributions shown in [2], [7] and [10]. Moreover, the average value given in (25) is consistent with the expressions obtained by the previous researchers [12] [27] for the average packet delay when substituting the value of τ by (6).

A. M/Geo/1 Queuing Model

Here, we propose using the discrete-time queuing discipline (M/Geo/1) as a queuing model for nearly saturated IEEE 802.11 single-channel ad hoc networks (with Poisson traffic sources). This model describes a queuing system with a Poisson arrival process, and an output server (channel) that is subjected to interruptions controlled by a geometric distribution [28]. The output channel is capable of sending one packet successfully per unit service interval (virtual time slot). The probability that the output channel is available (i.e. available to send the packet successfully) at saturation is given by q , which is defined by (14). The probability the channel is blocked (cannot send the packet successfully) for exactly $(n-1)$ consecutive service intervals at saturation is given by (24). We do the queuing analysis at near saturation (with

utilization factor ρ very close to but less than 1) to guarantee the stability of the queue and also to take the advantage of high network throughput [1]. At saturation, the queue may not be stable and hence it is impractical for resource allocation and QoS provisioning. We note that, in IEEE 802.11, the service rate of the queuing system is not constant but depends on the arrival rate. In most queuing systems, we can simply choose the arrival rate for a required ρ value. However, in IEEE 802.11, when the arrival rate increases toward saturation, the service rate decreases until it reaches a saturation value. In fact, the saturation service rate is the minimum achievable value before the queue becomes totally unstable. The M/Geo/1 queuing model is sufficiently accurate for $0.98 \leq \rho < 1$, which is very close to saturation ($\rho = 1$). As ρ decreases, the approximation error increases. The service rate (the number of successfully transmitted packets per virtual slot) in a non-saturated case, denoted by β , can be calculated with sufficient accuracy using the method described in [29]. Actually, Cai et al. give basic equations in [29] that can be solved to get the average service rate and the collision probability p for a certain utilization factor ρ and a certain arrival rate in non saturated conditions. We use the service rate obtained from the solution of those equations to get the average queue length and the probability distribution of the buffer occupancy based on the M/Geo/1 analysis given by (26)-(27). We assume an infinite buffer model for simplicity. The assumption is reasonable due to the huge available capacity of the latest memory chips, e.g. those used in small devices such as personal digital assistants (PDAs). Based on [28], the average queue length for the Poisson arrivals and geometric service time can be exactly calculated by

$$L_q = \frac{\rho(2-\alpha)}{2(1-\rho)} \quad (26)$$

where α is the number of packet arrivals per virtual slot and ρ is the utilization factor given by

$$\rho = \frac{\alpha}{\beta}.$$

The probability distribution of m packets in the queuing system can be approximated (as the average virtual slot time is small) by [30]

$$p_m \approx \begin{cases} \frac{(1-\gamma)\gamma^m}{1-\beta(1-\gamma)}, & m > 0 \\ \frac{(1-\beta)(1-\gamma)}{1-\beta(1-\gamma)}, & m = 0 \end{cases} \quad (27)$$

where

$$\gamma = \frac{\alpha(1-\beta)}{\beta(1-\alpha)}.$$

In the next section, we verify by computer simulations that both the average queue length and the probability distribution of the number of packets given in (26)-(27) are very accurate.

VI. SIMULATION RESULTS

We verify our analysis using the ns-2 simulator [31]. The simulation model simulates nodes moving in an unobstructed plane following the *random waypoint* model [32] with a maximum speed of 1 m/s. In the model, a node chooses its speed and its destination randomly and then moves to the

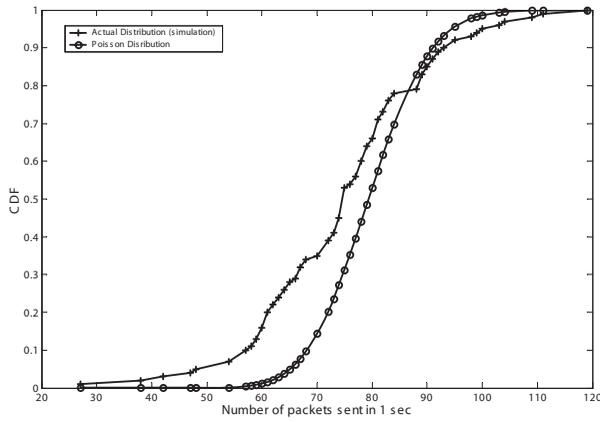


Fig. 3. The actual CDF and the Poisson CDF for the number of successfully transmitted packets in one second (5 nodes).

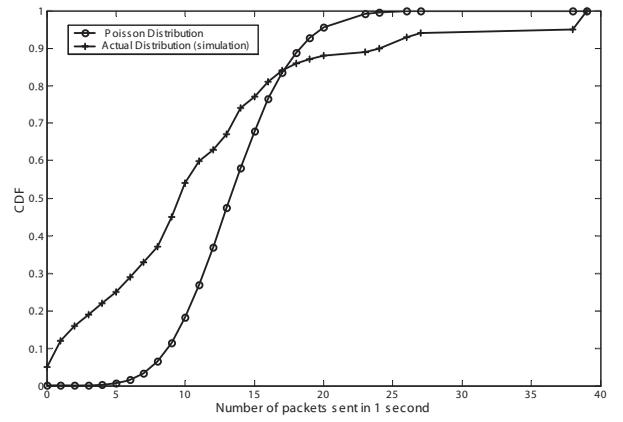


Fig. 5. The actual CDF and the Poisson CDF for the number of successfully transmitted packets in one second (30 nodes).

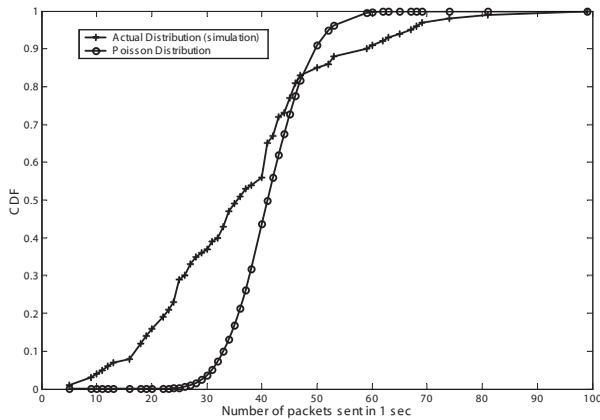


Fig. 4. The actual CDF and the Poisson CDF for the number of successfully transmitted packets in one second (10 nodes).

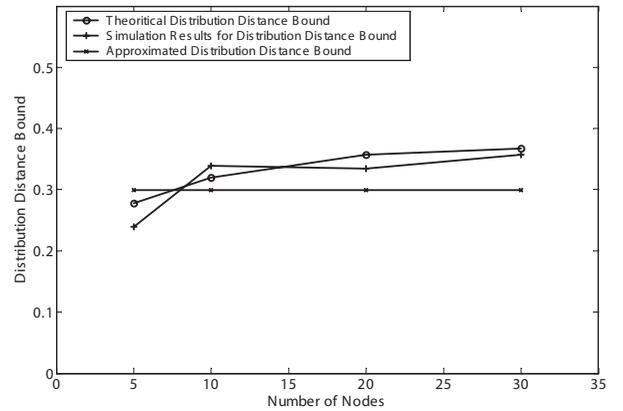


Fig. 6. Distribution distance upper bound.

destination. The simulation is done for a network having a variable number of mobile nodes over an area of 250×250 m². The node radios have a transmission range of 250 m and a carrier-sense range of 550 m. Only half of the nodes are active traffic sources, the other half are only receivers. The network represents a single-hop ad hoc network; every sender sends data packets to one unique receiver. To verify the distribution distance bound (22), we use constant bit rate traffic sources with a high data rate to force the active nodes to be in a saturation state (always have backlogged packets). For the queuing analysis verification, we use Poisson traffic sources.

Basically, the ns-2 simulator uses the WaveLAN implementation for medium access control. This MAC implementation has two main differences from the IEEE 802.11 standard as follows: (i) The backoff counter does not stop, when a transmission of another node is in progress; (ii) The CSMA/CA implementation is short-term unfair [21] since the node doubles its backoff window if it sensed the channel busy after its backoff counter is already decremented to zero. This gives a higher chance for the node currently transmitting a packet to continue transmission. According to the IEEE 802.11 standard [20], the node doubles its contention window only when collision is detected. Therefore, we modified the ns-2 implementation to comply with the standard. In the following, we verify the distribution distance analysis and the queuing

analysis. Table I gives the system parameter values used in the analysis and simulations.

A. Distribution distance verification

We measure the probability distribution of the number of packets successfully transmitted by any node over a duration of one second for different number of active source nodes, namely, 5, 10 and 30 nodes. Figures 3-5 show the cumulative distribution function (CDF) of the number of successfully transmitted packets for the different numbers of active nodes, respectively. For comparison, the corresponding Poisson distribution is also included. Figure 6 shows a comparison between the calculated upper bounds for different numbers of active source nodes (5, 10, 20 and 30 nodes) using (18) and (22) respectively, and the results from the computer simulations. The figure shows a close match between the analysis and simulation results. The upper bound, as can be seen from the figure, is almost constant and slightly affected by the number of active nodes. The figure also shows that the upper bound for different number of active nodes is very close to the approximated theoretical value obtained from (22). From Figures 3-5, it can be seen that the upper bound has been reached mainly at a small number of packets, which reflects a higher probability of a long service time than the exponential distribution. When the number of packets increases, getting closer to the average and larger, the distribution distance becomes smaller than the upper bound. The difference between the distributions results

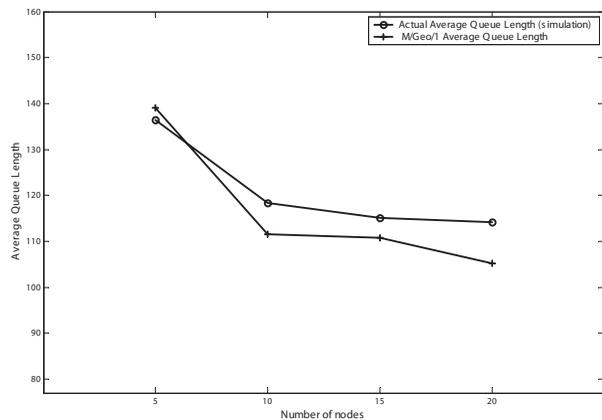


Fig. 7. Average queue length.

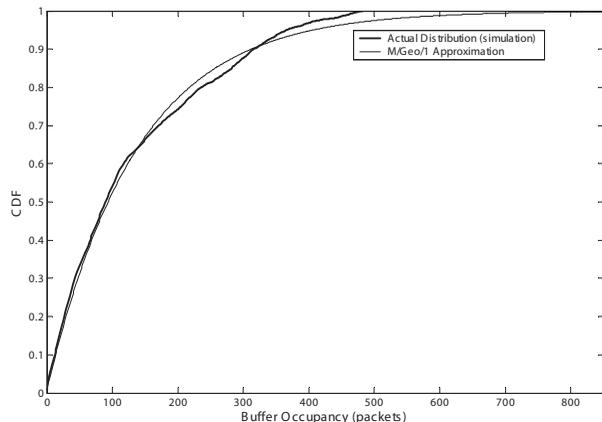


Fig. 8. The CDF of the number of packets in the actual queuing system and the M/Geo/1 queue (5 nodes).

from the discrete nature of the service time distribution, in addition to the fact that the service time is not completely memoryless.

B. M/Geo/1 queuing system verification

We calculate the average queue length and the probability distribution of the number of packets in the queuing system using (26) and (27). Figure 7 compares the average queue length for different numbers of active nodes based on both the analysis and simulations. It is observed that, the difference between the analytical and simulation results is small (around 5% to 7%). Figures 8-10 show a comparison between the simulation and analysis results for the CDF of the number of packets in the queuing system for 5, 10 and 20 active source nodes, respectively. The two distributions in each of the figures are in a close match. The result indicates that the geometric distribution is effective in approximating the actual service time. We notice that, although the counting process of the successful transmitted packets deviates from the true memoryless behavior, the deviation does not significantly affect the queuing analysis when the discrete memoryless distribution is considered for the service time. As a result, we suggest to use the M/Geo/1 analysis as a tool for statistical QoS provisioning (such as statistical call admission control). The accurate match between the analytical and simulation results of the probability distribution of the buffer occupancy

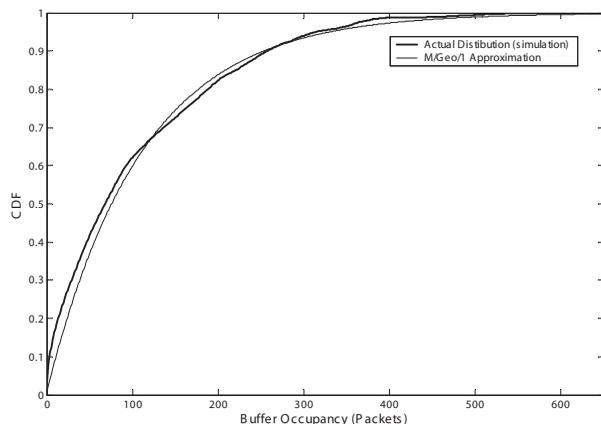


Fig. 9. The CDF of the number of packets in the actual queuing system and the M/Geo/1 queue (10 nodes).

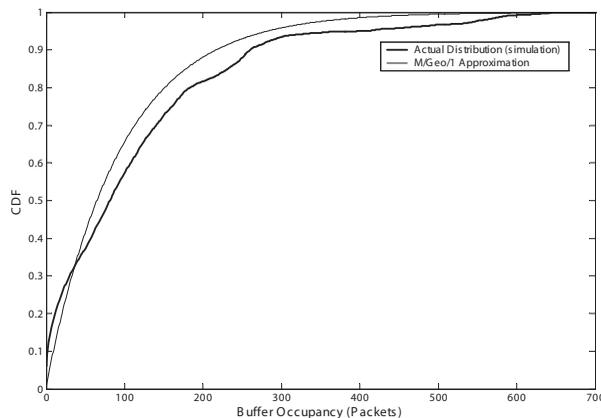


Fig. 10. The CDF of the number of packets in the actual queuing system and the M/Geo/1 queue (20 nodes).

implies that the M/Geo/1 analysis can be used to provide stochastic QoS guarantees for any type of traffic flows (as long as their arrival process can be modeled approximately as a Poisson process).

VII. CONCLUSION

This paper aims at reaching a simplified and sufficiently accurate approximation for the service time distribution in IEEE 802.11 nearly saturated single-hop ad hoc networks. The approximated distribution can be used in statistical resource allocation for efficient resource utilization and QoS provisioning. Through investigating the memoryless behavior of the service time, we have shown that the number of successful packet transmissions by any node in the network over a time interval has a probability distribution that is close to Poisson by an upper bounded distribution distance. By using the Chen-Stein approximation, we calculate the bound and illustrate that it depends mainly on some system parameters and slightly on the number of active nodes. Further we propose to use the geometric distribution with the appropriate parameter as an approximation of the probability distribution of the actual discrete service time. We illustrate that a discrete-time queuing discipline (M/Geo/1) can be used as a queuing model for IEEE 802.11 ad hoc networks (fed by Poisson traffic sources). The analytical results and computer simulation results show a very close match not only in the

average queue length but also in the probability distribution of the number of packets in the queuing system.

ACKNOWLEDGEMENTS

This work was supported by the Natural Science and Engineering Research Council (NSERC) of Canada.

REFERENCES

- [1] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Select. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.
- [2] H. Zhai, Y. Kwon, and Y. Fang, "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs," *Wireless Commun. Mob. Comput.*, vol. 4, pp. 917–931, 2004.
- [3] P. Raptis, V. Vitsas, K. Paparrizos, P. Chatzimisios, A. C. Boucouvalas, and P. Adamidis, "Packet delay modeling of IEEE 802.11 Wireless LANs," in *Proc. Intl. Conf. on Cyber. Info. Tech. Sys. Apps. CITSA 2005*, July 2005.
- [4] M. Özdemir and A. McDonald, "An M/MMG/1/K queuing model for IEEE 802.11 ad hoc networks," in *Proc. 1st ACM Intl. Workshop on Perf. Eval. of Wirel. Ad Hoc Sensor Ubiquit. Net. PE-WASUN '04*, Apr. 2004, pp. 107–111.
- [5] O. Tickoo and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks," in *Proc. INFOCOM 2004*, vol. 2, Mar. 2004, pp. 1404–1413.
- [6] P. Pham, S. Perreau, and A. Jayasuriya, "New cross-layer design approach to ad hoc networks under Rayleigh fading," *IEEE J. Select. Areas Commun.*, vol. 23, no. 1, pp. 28–39, Jan. 2005.
- [7] Y. Zheng, K. Lu, D. Wu, and Y. Fang, "Performance analysis of IEEE 802.11 DCF in binary symmetric channels," in *Proc. IEEE GLOBECOM 2005*, vol. 5, Dec. 2005, pp. 3144–3148.
- [8] P. Engelstad and O. Sterb, "Analysis of the total delay of IEEE 802.11e EDCA and 802.11 DCF," in *Proc. IEEE ICC 2006*, June 2006.
- [9] O. Tickoo and B. Sikdar, "A queueing model for finite load IEEE 802.11 random access MAC," in *Proc. ICC 2004*, vol. 1, June 2004, pp. 175–179.
- [10] A. Zanella and F. De Pellegrini, "Statistical characterization of the service time in saturated IEEE 802.11 networks," *IEEE Commun. Lett.*, vol. 9, pp. 225–227, Mar. 2005.
- [11] P. Raptis, K. Paparrizos, P. Chatzimisios, and A. C. Boucouvalas, "Packet delay distribution of the IEEE 802.11 distributed coordination function," in *Proc. 6th IEEE Intl. Symp. World of Wirel. Mob. Multimed. Net. WoWMoM'05*, June 2005, pp. 299–304.
- [12] P. Chatzimisios, A. C. Boucouvalas, and V. Vitsas, "IEEE 802.11 packet delay – a finite retry limit analysis," in *Proc. GLOBECOM 2003*, vol. 2, 2003, pp. 950–954.
- [13] S. Ross, *Introduction to Probability Models*, 7th ed. Harcourt Academic Press, 2000.
- [14] C. Goldschmidt, "The Chen-Stein method for convergence of distributions", masters-level essay, University of Cambridge, UK, 2000; [online] <http://www.statslab.cam.ac.uk/~cag27/chen-stein.ps.gz>.
- [15] A. Banchs, "Analysis of the distribution of the backoff delay in 802.11 DCF: a step towards end-to-end delay guarantees in WLANs," in *Proc. QoFIS 2004*, LNCS 3266, Sept. 2004, pp. 64–73.
- [16] P. Jacquet, A. Naimi, and G. Rodolakis, "Routing on asymptotic delays in IEEE 802.11 wireless ad hoc networks," in *Proc. RAWNET 2005*, Apr. 2005.
- [17] C. Foh and M. Zukerman, "Performance analysis of the IEEE 802.11 MAC protocol," in *Proc. European Wireless 2002*, Feb. 2002.
- [18] J. Tantra, C. Foh, I. Tinnirello, and G. Bianchi, "Analysis of the IEEE 802.11e EDCA under statistical traffic," in *Proc. IEEE ICC 2006*, June 2006.
- [19] S. Sitharaman, "Modeling queues using poisson approximation in IEEE 802.11 ad hoc networks," in *Proc. Loc. Metro. Area Net. LANMAN 2005*, Sept. 2005, pp. 1–6.
- [20] IEEE standard for wireless LAN medium access control (MAC) and physical layer (PHY) specifications, ISO/IE 8802-11: 1999(E), Aug. 1999.
- [21] C. E. Koksal, H. Kassab, and H. Balakrishnan, "An analysis of short-term fairness in wireless media access protocols," in *Proc. ACM SIGMETRICS*, June 2000.
- [22] G. Berger-Sabbatel, A. Duda, M. Heusse, and F. Rousseau, "Short-term fairness of 802.11 networks with several hosts," in *Proc. 6th IFIP/IEEE Intl. Conf. Mob. Wirel. Communi. Net.*, Oct. 2004, pp. 263–274.
- [23] G. Berger-Sabbatel, A. Duda, O. Gaudoin, M. Heusse, and F. Rousseau, "Fairness and its impact on delay in 802.11 networks," in *Proc. GLOBECOM '04*, vol. 5, Dec. 2004, pp. 2967–2973.
- [24] C. Trabelsi, "Access protocol for broadband multimedia centralized wireless local area networks," in *Proc. Second IEEE Symp. on Comp. and Communi.*, July 1997, pp. 540–544.
- [25] W. Song, H. Jiang, W. Zhuang, and X. Shen, "Resource management for QoS support in cellular/WLAN interworking," *IEEE Network*, vol. 19, pp. 12–18, Sept. 2005.
- [26] J. Walrand (Nov. 2003), EECS 126 - probability and random processes; [online] available: <http://robotics.eecs.berkeley.edu/wlr/126/w11.htm>
- [27] P. Chatzimisios, A. C. Boucouvalas, and V. Vitsas, "Packet delay analysis of IEEE 802.11 MAC protocol," *Elect. Lett.*, vol. 39, Sept. 2003, pp. 1358–1359.
- [28] J. Hsu, "Buffer behavior with Poisson arrivals and geometric output processes," *IEEE Trans. Commun.*, vol. 22, pp. 1940–1941, Dec. 1974.
- [29] L. X. Cai, X. Shen, J. Mark, L. Cai, and Y. Xiao, "Voice capacity analysis of WLAN with unbalanced traffic," *IEEE Trans. Veh. Technol.*, vol. 55, pp. 752–761, May 2006.
- [30] M. Woodward, *Communication and Computer Networks: Modeling with Discrete-Time Queues*. Los Alamitos, CA: IEEE Computer Society Press, 1994.
- [31] The VINT Project, The UCB/LBNL/VINT Network Simulator-ns (version 2). <http://mash.cs.berkeley.edu/ns>.
- [32] J. Broch, D. Maltz, D. Jonthson, Y. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad-hoc network routing protocols," in *Proc. ACM/IEEE Mobicom'98*, pp. 85–97.



Atef Abdrabou received the B.Sc. degree in 1997 and the M.Sc. degree in 2003, both in electrical engineering, from Ain Shams University, Cairo, Egypt. He is currently working toward his Ph.D. degree at the Department of Electrical and Computer Engineering, University of Waterloo, Ontario, Canada. His current research interests include quality-of-service provisioning and resource allocation for multi-hop ad hoc wireless networks.



Weihua Zhuang (M'93-SM'01) received the B.Sc. and M.Sc. degrees from Dalian Maritime University, China, and the Ph.D. degree from the University of New Brunswick, Canada, all in electrical engineering. Since October 1993, she has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, where she is a Professor. She is a co-author of the textbook *Wireless Communications and Networking* (Prentice Hall, 2003). Her current research interests include wireless communications and networks, and radio

positioning.

Dr. Zhuang is a licensed Professional Engineer in the Province of Ontario, Canada. She is the Chair of the Vehicular Technology Chapter, IEEE Kitchener-Waterloo Section. She received the Outstanding Performance Award in 2005 and 2006 from the University of Waterloo for outstanding achievements in teaching, research, and service, and the Premier's Research Excellence Award (PREA) in 2001 from the Ontario Government for demonstrated excellence of scientific and academic contributions. She is the Editor-in-Chief of *IEEE Transactions on Vehicular Technology*, and an Editor of *IEEE Transactions on Wireless Communications*, *EURASIP Journal on Wireless Communications and Networking*, and *International Journal of Sensor Networks*.