A Queuing Theoretic Model of Ad Hoc Wireless LANs

Mustafa Özdemir & A. Bruce McDonald
RWIN-Lab
Dept. of Electrical & Computer Engineering
Northeastern University
Boston, MA 02115

Abstract

An M/MMGI/1/K queuing model is developed for the analysis of IEEE 802.11 DCF using RTS/CTS. Results are based on arbitrary contention conditions, namely, collision probabilities, transmission probabilities and contention window sizes vary arbitrarily among nodes contending for channel access. This is fundamentally different from earlier work. Results are presented for the fully-connected case and validated via simulation with statistical analysis. The main contributions are the analysis of DCF and the foundation for the analysis of multi-hop scenarios. Error analysis is also carried out. A key element of the model is that complexity normally encountered is reduced by effectively restoring the independence between service times and packet inter-arrivals.

I. Introduction

The performance of the IEEE 802.11 MAC standard for wireless LANs has been the subject of numerous analyses. However, until recently there had been little work focused on the DCF using RTS/CTS. The objective of this paper is to present a generalized queuing model for Markov arrivals; this is the first such model to be proposed and has been well validated via simulation under the range of traffic conditions. The broader objective is to develop a methodology that is extensible to multi-hop ad hoc networks scenarios.

The analysis is based upon a bounded network of M/G/1/K queues, where K represents the max queue length at each node and N represents the number of potentially interfering nodes. For multiple-hop systems N represents the number of nodes within two-hops of either a sender or receiver. It represents the effective number of interfering nodes accounting for only that portion of traffic generated by the two-hop away neighbors destined for a one-hop neighbor. An important property of the model is underscored by the earlier reference to ”potentially interfering” nodes. Each of the N nodes may be in any state of a given time, namely, any node may be busy or idle, and a busy node may be in any back-off stage. In contrast to related work [3] [11] [18] [17] [9] [7], no limiting assumption are made that force uniform collision probability, uniform back-off stage distribution or saturation conditions. For the comparison reasons a new model based on [17] is proposed. The analysis assumes only Markov arrivals at each node that are pairwise and collectively independent.

An alternative characterization for the queuing model is the M/MMGI/1/K: The service times are modeled as a Markov Modulated General Independent process. In principle this system is representative of a Phase-Type (PH) service. Hence, comprehensive analysis is well adapted to matrix-geometric techniques. The difficulty of this approach, however, is in finding an accurate parametric description of the PH service. For steady-state analysis a general service distribution in which the service-times depend on *local* collision probabilities of the RTS/CTS frames and the distribution of the time to resolve them, which, itself is dependent on the number of busy nodes in contention for the channel.

For single-hop analysis knowledge of node distribution is necessary and sufficient to determine the aforementioned values: a two-dimensional Discrete-Time Markov-Chain (DTMC) that characterizes the back-off stages and collision probabilities associated with each node effectively modulate the general process, thus facilitating the estimation of the needed parameters. The nature of this model enables a node-oriented versus channel-oriented analysis. This approach is more natural and facilitates analysis of multi-hop and network scenarios. The model introduced by Fang and McDonald in [8] provides the final element underlying general multi-hop queuing analysis. A key insight given in [8] is that a node attempting to transmit acts equivalently to contention regardless of whether it originates from an adjacent node or from multiple-hops away. This observation leads to representation of a multi-hop wireless network that uses geometric arguments and numerical integration to transform the multi-hop problem into an equivalent single-hop problem.

This paper is not “yet another analysis” of IEEE

802.11 throughput and delay. The importance of the M/MMGI/1/K model presented in this paper is that it provides a foundation for the analysis of multiple-hop scenarios and can be extended to model arbitrary network configurations. Thus, it is the basis for a fundamentally different strategy for understanding and improving the effectiveness of practical networks. A key point and novel aspect of the model is that the complexity level normally encountered is reduced by effectively restoring the independence between service times and packet inter-arrivals through the DTMC formulation. Thus, direct stead-state analysis is possible using iterative techniques to estimate the model parameters for different traffic intensity levels.

Performance measurements are shown using the analytical model over IEEE 802.11b wireless networks. The goals of these measurements are to assess whether there are significant issues with 802.11b wireless networks regarding real-time applications, as it is stated in several papers ([3], [6], [20] and [15]). With this information it is intended to obtain a better understanding of wireless networks' suitability for real-time applications. Practical future applications include routing, admission control and scheduling in ad hoc networks.

It is well-known that the wireless channel is error-prone due to noise and interference in the channel. However, the impact of bit errors in the packet is not considered in the previous analytical models ([3], [9] and [17]). It is expected that when bit error rate is high, the good throughput of DCF access mechanism will degrade significantly. But it is not well-known how the bit error rate affects other performance measures such as MAC delay, jitter, blocking and dropping probabilities. Thus, the tradeoff of DCF performance metrics and packet reliability is investigated by extending the proposed model to error-prone channels.

The remainder of this paper is organized as follows: Section II describes the system model which encompasses models for the back-off algorithm, service time distribution and the M/G/1/K formulation. Performance measures are given in section III. Section IV shows how to carry out the error analysis on top of the system model. Section V applies simulation and statistical analysis to validate the analytical results and section VI depicts some sensitivity analysis results. Finally, Section VIII presents conclusions and discusses future work.

II. System Model

This section describes the basic methodology and components of the M/MMGI/1/K queuing analysis. The main parameters are identified and applied to each of the system entities—the back-off algorithm, the service

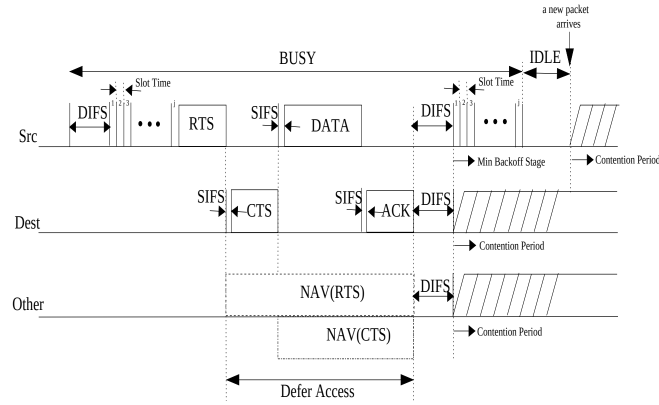

Fig. 1. IEEE 802.11 DCF RTS/CTS procedure

time distribution and, finally, the queuing model. In the present analysis the states of each node are considered independently and then coupled through an iterative process in order to evaluate the system state. Specifically, the MAC algorithm executes at each node leaving each in an arbitrary back-off stage at any time. Due to the independent and identically distributed (IID) arrival processes and assuming uniform traffic distribution and positions it is possible to evaluate the expected value of the global service time by determining the busy probability, transition probabilities and mean waiting times from the DTMC.

Figure-1 illustrates how each node alternates between busy and idle periods. During a busy period the node executes the RTS/CTS protocol, transmits its data and receives an acknowledgment. If it has multiple frames to send, it may contend and transmit more than one during the same busy period. After each transmission a node goes into back-off following DIFS; other nodes continue counting down the back-off time according to IEEE 802.11 standards [4]. Given random packet arrival times and transmission success probability, the busy nodes do not generally share a common Contention Window (CW) at a given time. Moreover, the number of busy nodes that have at least one frame to send may vary from one contention period to another. Each busy node becomes idle when there are no more frames to send. A packet is reserved in the queue on arrival if at the instant of arrival, the node is non-empty. One of the basic parameters that must be evaluated is b_0, the steady-state probability that a node is busy or approximately non-empty. The state probabilities depend on the traffic rate λ. Packet inter-arrivals are assumed to include packets originating at a given node and those routed through the node. This reflects the broader objective of applying the model to generalized wireless ad hoc networks. Based on the Markov assumption the inter-arrival times are IID

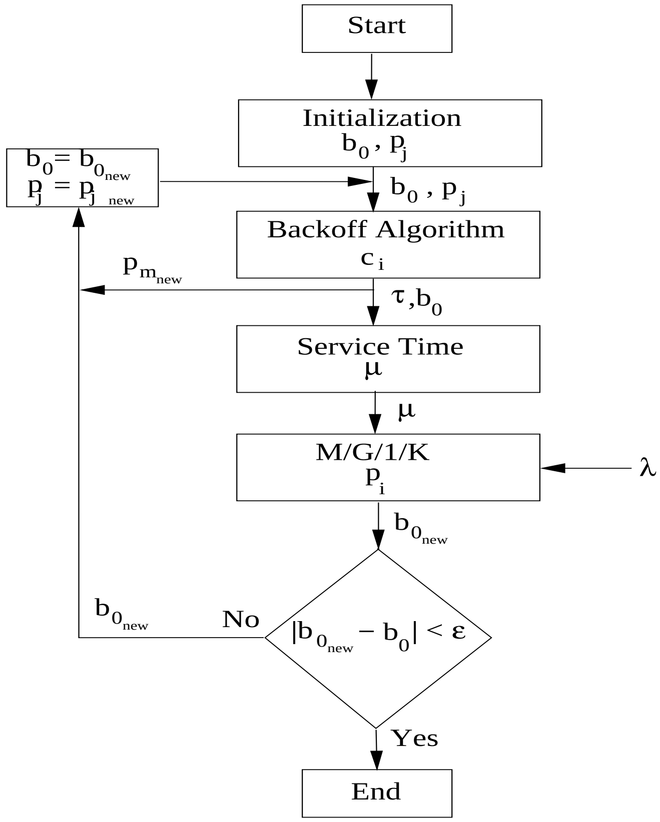

Fig. 2. IEEE 802.11 Organization

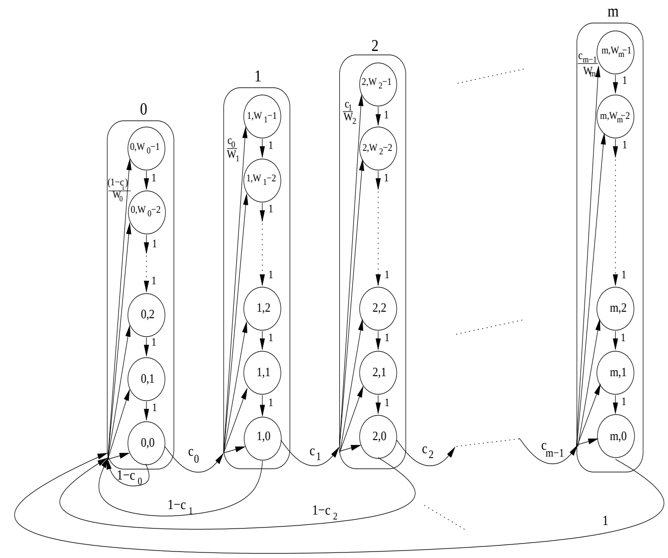

Fig. 3. DTMC model for back-off window size

with exponential distribution.

An iterative process that is dependent on the busy probabilities is used to evaluate the system behavior. Values are estimated through using an initial guess and iterative correction. The process is depicted in Figure-2 and outlined as follows:

- **Step 1:** Initialize b_0 and the probabilities that a node is in any of $m+1$ back-off stages. These must be known to determine the collision probabilities for the back-off algorithm.
- **Step 2:** Calculate the collision probabilities, c_i, for each back-off stage and the corresponding transmission attempt probability, τ.
- **Step 3:** Given c_i and τ, evaluate the packet service rate, μ, where $\frac{1}{\mu}$ is the mean packet service time.
- **Step 4:** Given μ and λ, find all the state probabilities in $M/G/1/K$ queuing system. This process will result in a new value of b_0 - the steady-state busy probability.
- **Step 5:** Repeat Steps 2, 3 and 4 until the difference between a new and previous value for b_0 is small.

The remainder of this section provides detailed explanation of the loop steps which reflect three system entities.

A. The First Modeling of Back-off Algorithm

A core contribution of this modeling is that evaluation of collision probabilities *is not limited by the assumptions* that packets collide with constant equal probability and collisions are pairwise and collectively independent over all transmission attempts [3],[11]. A more accurate characterization is invoked reflecting random variation of back-off stage among nodes and correlation between back-off time and back-off stage which depends on the time the number of collisions already experienced during a given packet transmission attempt. The generalized analysis of collision probability is achieved through a sequential process of conditioning and applying total probability.

The first step is to evaluate the conditional probabilities for two busy nodes in each back-off stage. $\alpha_{ij,k}$ is the conditional collision probability given node 1 is in stage i and node 2 in stage j, and node 1 selects slot k. For notational convenience $k \in (0, (W_i - 1))$, where W_i is the stage i contention window size. $\alpha_{00,k}$ is first considered. The following must hold: $\alpha_{00,1} = 1/W_j$ since if node 1 picks slot 1, node 2 must also pick slot 1 for a collision to occur. If node 1 picks slot 2, then there are two scenarios that result in a collision: (1) node 2 also selects slot 2, or (2) node 2 first picks slot 1 after which it becomes busy again and chooses slot 1—corresponding to the same time as node 1's initial occupancy of slot 2. Here it is assumed that any transmission attempt does not affect the busy probability, b_0. The conditional collision probability is given by the following:

$$\alpha_{00,2} = \frac{1}{W_0} + \frac{1}{W_0} b_0 \frac{1}{W_0}$$

$$= \frac{1}{W_0} + b_0 \frac{1}{W_0^2}$$

Consider the following example for finding the probability of a collision given node 1 is in back-off stage 0, node 2 is in back-off stage 0 conditioned on node 1 selecting slot $k = 3$: $\alpha_{00,3}$. There are four scenarios that can result in a collision: (1) node 2 also select the same slot—slot 3, (2) node 2 selects slot 1 becomes busy again and selects slot 2, or the reverse situation—slot 2 is selected first followed by slot 1, and, finally, (4) if node 2 selects slot 1 three times consecutively. The resulting conditional collision probability given by:

$$\begin{aligned}\alpha_{00,3} &= \frac{1}{W_0} + \frac{1}{W_0} b_0 \frac{1}{W_0} + \frac{1}{W_0} b_0 \frac{1}{W_0} + \frac{1}{W_0} b_0 \frac{1}{W_0} b_0 \frac{1}{W_0} \\ &= \frac{1}{W_0} + 2b_0 \frac{1}{W_0^2} + b_0^2 \frac{1}{W_0^3}\end{aligned}$$

A clear pattern emerges by continuing in this way, thus resulting in the following general expression for $\alpha_{00,k}$:

$$\alpha_{00,k} = \frac{1}{W_0} \sum_{m=0}^{k-1} \begin{pmatrix} k-1 \\ m \end{pmatrix} \left(\frac{b_0}{W_0}\right)^m \quad 1 \leq k \leq W_0 - 1$$

Given an arbitrary back-off stage a node selects a transmission slot randomly from a uniform distribution covering the window size of the stage. Let p_{s_k} be the uniformly distributed random variable that determines the probability that node 1 chooses slot k. Applying the theorem of total probability the first condition is eliminated: $c_{00,2}$ is the probability of a collision between two nodes in stage 0:

$$\begin{aligned}c_{00,2} &= \sum_{k=1}^{W_0} \alpha_{00,k} p_{s_k} \\ &= \sum_{k=1}^{W_0} \frac{1}{W_0} \sum_{l=0}^{k-1} \begin{pmatrix} k-1 \\ l \end{pmatrix} \left(\frac{b_0}{W_0}\right)^l \frac{1}{W_0} \\ &= \frac{1}{W_0^2} \sum_{k=1}^{W_0} \sum_{l=0}^{k-1} \begin{pmatrix} k-1 \\ l \end{pmatrix} \left(\frac{b_0}{W_0}\right)^l \qquad (1)\end{aligned}$$

Consider the case in which $j = 0$. Let node 1 be in stage 1 with back-off window size 64 and let node 2 be in stage 0 with window size 32. A collision occurs only if both nodes select a slot in the first half of stage 1 which is equivalent to stage 0. Hence, if node 1 selects a slot from the first half of stage 1 the collision probability is equivalent to $c_{00,2}$. This is the probability given by Equation-1. If, however, node 1 picks a slot from the second half of stage 1, then a collision can occur only if node 2 becomes busy. The corresponding collision probability is $c_{00,2}$. The total collision probability for $j = 0$ for the example is:

$$c_{10,2} = \frac{1}{2} c_{00,2} + \frac{1}{2} b_0 c_{00,2}$$

Following with the logic the result $c_{20,2}$ is given by the following, where node 1 is in stage 2, thus, has a back-off window size of 128:

$$\begin{aligned}c_{20,2} &= \frac{1}{4} c_{00,2} + \frac{1}{4} b_0 c_{00,2} + \frac{1}{4} b_0^2 c_{00,2} + \frac{1}{4} b_0^3 c_{00,2} \\ &= \frac{c_{00,2}}{4}(1 + b_0 + b_0^2 + b_0^3)\end{aligned}$$

A pattern emerges that gives the general result for collision probability $c_{i0,2}$:

$$c_{i0,2} = \frac{c_{00,2}}{2^{(i)}} \sum_{k=0}^{(2^i-1)} b_0^k \qquad (2)$$

To complete the resolution of the first condition consider the case in which $i = 0$. Let node 1 be in stage 0 and let node 2 be in stage j. Node 2 picks a slot and if there is no collision, node 2 goes to stage 0 and chooses a slot from stage 0 until collision occurs. In other words, it is assumed that a node that is at stage j starts from the stage 0 after any number of its subsequent successful attempts. Except for picking the first slot at stage j, the rest is same to $c_{00,2}$. With the help of Equation-(1), the total collision probability for $i = 0$ is:

$$\begin{aligned}c_{0j,2} &= \frac{1}{W_0 W_j} \sum_{k=1}^{W_0} \sum_{l=0}^{k-1} \begin{pmatrix} k-1 \\ l \end{pmatrix} \left(\frac{b_0}{W_0}\right)^l \\ &= \frac{W_0}{W_j} c_{00,2} \qquad (3)\end{aligned}$$

Following with the same logic and using Equation-(2) and (3), the general result is given by:

$$c_{ij,2} = \frac{W_0}{W_j} \frac{c_{00,2}}{2^{(i)}} \sum_{k=0}^{2^i-1} b_0^k \qquad (4)$$

Given the steady-state occupancy probabilities p_j, $j \in (0, m)$ for the $m+1$ back-off stages total probability is used to remove the condition on node 2 resulting in the collision probability $c_{i,2}$ conditioned on node 1 selecting stage i and a second busy node (node 2):

$$c_{i,2} = \sum_{j=0}^{m} c_{ij,2} p_j \qquad (5)$$

Assuming that the probabilities of collisions between different node pairs are independent, then the collision probability $c_{i,n}$ conditioned on node 1 selecting stage i and n busy nodes is given by the following, where N is the total number of competing nodes:

$$c_{i,n} = 1-(1-c_{i,2})^{(n-1)} \quad 2 \le n \le N \tag{6}$$

The next step is to find the transmission attempt probability, τ_n, for an arbitrary node given n nodes are busy. Let $b(t)$ and $s(t)$ be continuous time stochastic processes representing the back-off time counter and back-off stage associated with an arbitrary node at time t. Taken together $(b(t), s(t))$ is a two dimensional process. Figure-3 relates the values of c_i to the individual node back-off window sizes that are determined by this two dimensional process. Based on the independence between nodes and the exponentially increasing back-off times a DTMC formulation similar to the one first proposed in [3] and later modified in [11] is used to model the process. Referring to Figure-3 if $b_{i,k,n} = \lim_{t\to\infty} P\{s(t)=i, b(t)=k, \texttt{n busy nodes}\}$, where $i \in [0,m]$ and $k \in [0, W_i - 1]$ is the stationary distribution of the Markov chain, then all probabilities $b_{i,k,n}$ can be found using standard Markov analysis.

$$\begin{array}{ll} b_{i,0,n} = c_{i-1,n} \cdot b_{i-1,0,n} & 0 < i \le m \\ b_{i,0,n} = \prod_{k=0}^{i-1} c_{k,n} \cdot b_{0,0,n} & 0 < i \le m \end{array} \tag{7}$$

Equation-7 reflects the condition that a node proceeds to the next back-off stage only if there is a collision in the current stage and the collision probabilities are independent from one stage to the next. The suggested node short retry count $m = 7$ according to the standard [4]. Here m is equivalent to the maximum back-off stage. The MC is irreducible and ergodic, for each $k \in [0, W_i - 1]$ express $b_{i,k,n}$ as follows:

$$b_{i,k,n} = \tfrac{W_i - k}{W_i} \cdot b_{i,0,n} \quad 0 < i \le m \tag{8}$$

Equations-(7) and (8) express all $b_{i,k,n}$ values as functions of $b_{0,0,n}$ and of collision probabilities $c_{i,n}$. The normalization condition leads to the following:

$$\begin{aligned} 1 &= \sum_{k=0}^{W_i-1} \sum_{i=0}^{m} b_{i,k,n} \\ &= \sum_{i=0}^{m} b_{i,0,n} \sum_{k=0}^{W_i-1} \frac{W_i - k}{W_i} = \sum_{i=0}^{m} b_{i,0,n} \frac{W_i + 1}{2} \end{aligned} \tag{9}$$

Combining Equation-7 with the normalization condition and applying algebraic manipulation results in the following expression for $b_{0,0,n}$:

$$b_{0,0,n} = \frac{1}{\sum_{i=0}^{m} (\prod_{k=0}^{i-1} c_{k,n}) \frac{W_i+1}{2}} \tag{10}$$

Each node attempts to transmit when its back-off counter reaches zero. The transmission probability τ_n can be found by considering the summation of the total back-off stage probabilities weighted by the collision probabilities. Hence, the probability that a node transmits a packet in a randomly chosen slot is:

$$\tau_n = \sum_{i=0}^{m} b_{i,0,n} = \sum_{i=0}^{m} (\prod_{k=0}^{i-1} c_{k,n}) \cdot b_{0,0,n} \tag{11}$$

Similarly, the steady-state occupancy probabilities $p_{j,n}$ for each back-off stage j are readily determined:

$$\begin{aligned} p_{j,n} &= \sum_{k=0}^{W_j-1} b_{j,k,n} = b_{j,0,n} \cdot \frac{W_j + 1}{2} \\ &= (\prod_{k=0}^{j-1} c_{k,n}) \cdot \frac{W_j + 1}{2} \cdot b_{0,0,n} \end{aligned}$$

Given the individual node busy probability, b_0, the probability that n nodes are busy, β_n, is a random variable with a binomial distribution. Thus, the steady-state occupancy probabilities p_j, conditioned on being in stage j is then determined by total probability:

$$\beta_n = \binom{N}{n} b_0^n (1-b_0)^{N-n} \tag{12}$$

$$p_j = \sum_{n=2}^{N} p_{j,n} \beta_n \tag{13}$$

B. The Second Modeling of Back-off Algorithm

The following model is proposed in order to compare it with the first model and the simulation results. Unlike the first modeling, the collision probability is limited by the assumption that packets collide with constant equal probability, c. This modeling is based on the analytical work in paper [17]. In [17], the collision probability c was derived for the saturated network case in which every node always has a packet to transmit at any given time. In this paper, it is extended to model the general case by obtaining an approximate expression for collision probabilities. In [18], the model in [17] is also extended for the general case. G/G/1 queuing model where the queue length is not involved is considered in [18] but M/G/1/K is considered in this paper to take the queue length into consideration. Evaluation of collision probabilities is based on the node model in this modeling whereas in [18] it is based on the system model which consists of N node model. Unlike that in [18], an iterative process in Figure-2 is used in order to get more accurate estimates of the collision probabilities in a similar way the first modeling of back-off algorithm uses. In section-V, both models are compared with the simulation results.

With probability c the transmission is collided and with probability $1-c$, it is successful. Hence, the number of transmissions per packet is modeled as geometrically distributed with probability of success $1-c$. The average back-off window in the saturated case is given by [17]:

$$\begin{aligned} \overline{W} &= (1-c)\frac{W_0}{2} + c(1-c)\frac{2W_0}{2} + \cdots + \\ & c^m(1-c)\frac{2^m W_0}{2} + c^{m+1}\frac{2^m W_0}{2} \\ &= \frac{1-c-c(2c)^m}{1-2c}\frac{W_0}{2} \end{aligned} \quad (14)$$

Now consider a network with n busy nodes. A packet is reserved in the queue on arrival if the node is non-empty. It is said before that the probability that the node is busy or non-empty (node utilization factor) when an arbitrary arrival occurs is b_0. Hence, for any arbitrary packet, with probability $1-b_0$, the back-off window is 0 and with probability b_0, it is reserved in the queue. Therefore, the average back-off window size for general arrival rates is given by:

$$\overline{W} = b_0\frac{1-c-c(2c)^m}{1-2c}\frac{W_0}{2} \quad (15)$$

Following the arguments of [17] and considering the fact that only busy nodes can actually collide with packets from other busy nodes and conditioned on n busy nodes, the conditional collision probability is given by:

$$c_n = 1 - \left(1 - b_0\frac{1-c_n-c_n(2c_n)^m}{1-2c_n}\frac{W_0}{2}\right)^{n-1} \quad (16)$$

Note that all other parameters can be easily found by replacing $c_{i,n}$ by c_n since all c_i are assumed to be same in the second modeling.

C. The Service Time Distribution

In the next part of the analysis it is necessary to find the the distribution of the back-off window size. Here m is the maximum back-off stage and can have a value larger or smaller than M. The back-off window size random variable W_i is uniformly distributed at each stage:

$$W_i \sim \begin{cases} U(0,(2^i W-1)) & 0 \le i \le M \\ U(0,(2^M W-1)) & i > M \end{cases} \quad (17)$$

where W is the initial back-off window size and $U(a,b)$ shows the uniform distribution between a and b. The aggregate back-off window size random variable W_n depends on the number of busy nodes n and the collision probability $c_{i,n}$:

$$\mathcal{W}_n = \sum_{i=0}^{m}\left(\prod_{k=0}^{i-1} c_{k,n}\right) W_i \quad (18)$$

Consider the following probabilities and random variables as given in [3]: P_{tr} is defined as the probability that at least one transmission occurs in a given slot time. Since n stations contend to access the medium and each station transmits with probability τ_n, $P_{tr,n}$ is given by:

$$P_{tr,n} = 1-(1-\tau_n)^n$$

The probability P_s that an occurring transmission is successful is given by the probability that a station is transmitting and the remaining $n-1$ stations remain silent, conditioned on the fact that at least one station transmits:

$$P_{s,n} = \frac{n\tau_n(1-\tau_n)^{n-1}}{1-(1-\tau_n)^n}$$

T_s is the average time that the medium is sensed busy due a successful transmission and T_c is the average time that the medium is sensed busy by each station when a collision occurs and σ is the duration of an empty slot. The values of T_s and T_c depend on the channel access mechanism. Assuming that all stations use the same channel access mechanism, T_s and T_c are defined as follows, assuming the RTS/CTS access mechanism is employed:

$$\begin{aligned} T_s &= DIFS + RTS + SIFS + CTS + SIFS + \\ & H + E[P] + SIFS + ACK + \sigma \\ T_c &= DIFS + SIFS + RTS + CTS \end{aligned}$$

Where $H = MAC_{hdr} + PHY_{hdr}$, and $E[P]$ is the average length of the frame. Let $T_{slot,n} = (1-P_{tr,n})\sigma + P_{tr,n}P_{s,n}T_s + P_{tr,n}(1-P_{s,n})T_c$. Note that collision only occurs between RTS frames and T_c is different from that in paper [3] because the CTS timeout effect is considered.

The conditional service time for an arbitrary frame given n busy nodes is determined based on the parametric model of 802.11:

$$B_n = T_s + (\mathcal{W}_n T_{slot,n})$$

The unconditional service time is found by combining this result with the binomial distribution for the probability of the number of busy nodes given in Equation-12–the mean value of service time b is used to solve the queuing problem in the next section:

$$B = \sum_{n=1}^{N} B_n \cdot \beta_n \quad (19)$$

Note that while a packet arrives to an idle node at a moment where other nodes have non-empty queues but are in back-off, it will be transmitted right away if the channel is idle for DIFS time. In the analysis,

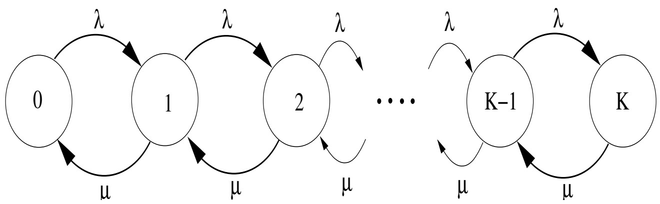

Fig. 4. M/G/1/K state transition diagram: single node case

the occurrence of such cases are disregarded since this possibility is very small. As verified in section-V-D, this consideration results in reasonably close results.

D. M/MMGI/1/K Queuing model

Let λ be the packet arrival rate and denote by $B(x)$ and b the distribution and expected value of packet service time respectively (see Equation-19). The maximum number of packets that can be accommodated at any node in the system (including the one in service) at any time is given by $K < \infty$. Those packets that arrive when K packets are already present are dropped. The initial throughput γ for an arbitrary node can be expressed in terms of the probability of packet loss, which is equivalent to the probability of an arrival find K in the system. Since the arrival process is Poisson this is equivalent to the time average of finding K in the system: P_k.

$$\gamma = \lambda(1 - P_K) \tag{20}$$

The traffic intensity or *offered load* is defined as ρ:

$$\rho = \lambda b = \frac{\lambda}{\mu} \tag{21}$$

The offered load is distinguish from the carried load, which is defined as ρ' and accounts for those packets lost due to buffer overrun. Thus, it represents the fraction of time the server is busy. It can also be viewed as is the probability that a server is busy at an arbitrary time; this probability is equivalent to the node busy probability, b_0, from the previous section:

$$\rho' = \gamma b = \rho(1 - P_K) \tag{22}$$

Let $E[L]$ be the mean number of messages in the system at an arbitrary time. From Little's theorem applied to those message that are accepted in the system, we have the relation:

$$E[L] = \gamma E[T] \tag{23}$$

system, including the one in services (called the queue size). We first apply the method of the imbedded Markov chain to obtain the queue size distribution immediately after service completion. We choose a set of imbedded Markov point at those point in time when packets leave the system after service completion. Let L_n be the number of packets left behind in the system immediately the nth Markov point, where $n = 1, 2, ...$, and let the steady-state probability distribution for $L_n; n = 1, 2, ...$ be

$$\pi_k = \lim_{n\to\infty} Prob[L_n = k] \quad 0 \le k \le K-1 \tag{24}$$

Note that L_n cannot be K, because when a packet leaves the system, it cannot leave behind a completely full system; at least one waiting position must be empty. Let $p_{jk}; 0 \le j, k \le K-1$ be the probabilities of the state transitions in the Markov chain $L_n; n = 1, 2, 3, ...$:

$$p_{jk} = Prob[L_{n+1} = k | L_n = j] \tag{25}$$

independent of n (a homogeneous Markov chain). They are given by

$$\begin{aligned} p_{0k} &= a_k & 0 \le k \le K-2 \\ &= \textstyle\sum_{l=K-1}^{\infty} a_l & k = K-1 \end{aligned} \tag{26}$$

and, for $1 \le j \le K-1$, we have

$$\begin{aligned} p_{jk} &= a_{k-j+1} & 0 \le k \le K-2 \\ &= \textstyle\sum_{l=K-j}^{\infty} a_l & k = K-1 \end{aligned} \tag{27}$$

where

$$a_k = \textstyle\int_0^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x} \partial B(x) \quad k = 0, 1, 2, \ldots \tag{28}$$

is the probability that k packets arrive during a service time.

The steady-state equations for state transitions are given by

$$\pi_k = \textstyle\sum_{j=0}^{K-1} \pi_j p_{jk} \quad 0 \le k \le K-1 \tag{29}$$

$$\sum_{k=0}^{K-1} \pi_k = 1 \tag{30}$$

Substituting (26) and (27) into (29), we obtain:

$$\pi_k = \pi_0 a_k + \textstyle\sum_{j=1}^{k+1} \pi_j a_{k-j+1} \quad 0 \le k \le K-2 \tag{31}$$

$$\pi_{K-1} = \pi_0 \left(\sum_{k=K-1}^{\infty} a_k \right) + \sum_{j=1}^{K-1} \pi_j \left(\sum_{k=K-j}^{\infty} \right) \tag{32}$$

Note that (32) is redundant. Equations (30) and (31) provide K independent equations for K unknowns $\pi_k; 0 \le k \le K-1$.

An efficient algorithm for computing $\pi_k; 0 \le k \le K-1$ can be given in terms of

$$\pi'_k = \tfrac{\pi_k}{\pi_0} \quad 0 \le k \le K-1 \tag{33}$$

[5]. It is easy to see from (31) that $\pi'_k; 0 \le k \le K-1$ can be recursively calculated as follows:

$$\pi'_0 = 1 \tag{34}$$

$$\pi'_{k+1} = \tfrac{1}{a_0} \left(\pi'_k - \textstyle\sum_{j=1}^{k} \pi'_j a_{k-j+1} - a_k \right) \quad 0 \le k \le K-2 \tag{35}$$

and π_0 is found from (30) as

$$\pi_0 = \left(\sum_{k=0}^{K-1} \pi'_k \right) \tag{36}$$

Thus we get $\pi_k; 0 \leq k \leq K-1$ from (33). The computational complexity of this algorithm is of the order $O(K^2)$.

Let P_k be the probability that there are k messages in the system at an arbitrary time, where $k = 0, 1, 2, \ldots, K$. From [5] and [16] P_k can be evaluated in terms of the steady-state probabilities. Moreover, this leads to an expression for the value of P_K:

$$P_k = \frac{\pi_k}{\pi_0 + \rho} \quad 0 \leq k \leq K-1 \tag{37}$$

$$P_K = 1 - \frac{1}{\pi_0 + \rho} \tag{38}$$

Substituting Equation-38 into Equation-22 gives the following expression for the busy probability or node utilization factor:

$$\rho' = \frac{\rho}{\pi_0 + \rho} = b_0 \tag{39}$$

III. Performance

While 802.11b networks have proved their appropriateness for best effort traffic, i.e. e-mail, browsing, chat or file transfer, their lack of QoS support makes it questionable whether the use of real-time multimedia applications, such as voice communication, is possible in these wireless networks. For QoS applications, there are many system measures to be decided according to the application. The ones to be measured are MAC delay, jitter, delay, variance of waiting time, throughput, dropping probability, and blocking probability. For instance, in order to establish a successful voice session, three critical network parameters must be kept under certain levels. These parameters are loss, jitter and delay.

Hence, since all these measures are of such importance to obtain a good perceived quality in QoS applications, it is interesting to know what causes the deterioration of these parameters in a wireless network, and how these parameters are affected by the 802.11b standard design. Whilst in wired networks the deterioration is generally caused by congestion, in wireless networks it can be caused both by degradation of the signal (due to fading or interference), the bad choice of system parameters and congestion.

MAC Delay and Jitter: The average MAC delay for a successfully transmitted packet is defined to be the time interval from the time the packet is at the head of its MAC queue ready to be transmitted, until an acknowledgment for this packet is received. If a packet is dropped because it has reached the specified retry limit, the MAC delay time for this packet will not be included into the calculation of the average MAC delay. The average packet MAC delay, provided that this packet is not discarded, is obtained easily by:

$$\begin{aligned} P_{r,n} &= \frac{\prod_{k=0}^{i-1} c_{k,n} - \prod_{k=0}^{m} c_{k,n}}{1 - \prod_{k=0}^{m} c_{k,n}} \\ E[D|n] &= T_s + P_{r,n} \frac{W_i}{2} T_{slot,n} \\ E[D] &= \sum_{n=1}^{N} E[D|n] \beta_n \end{aligned} \tag{40}$$

where $(1 - \prod_{k=0}^{m} c_{k,n})$ is the probability that the packet is not dropped and $P_{r,n}$ is the probability that a packet that is not dropped reaches the i stage given that n busy nodes.

Jitter is defined as the variance of interpacket arrival times compared to the inter-packet times of the original transmission. In other words, it is the variance of MAC delay. A buffer is commonly used to absorb this variation, at the cost of some delay. Jitter's effect in some QoS applications can be harmful if it is not kept under certain level, because it can lead either to additional packet losses or to additional delay. The jitter is found using (40) and assuming all W_i's are independent:

$$\begin{aligned} Var(D|n) &= \left[P_{r,n} T_{slot,n} \right]^2 Var(W_i) \\ Var(D) &= \sum_{n=1}^{N} Var(D|n) \beta_n^2 \end{aligned} \tag{41}$$

where W_i is found from (17):

$$Var(W_i) = \begin{cases} \frac{(2^i W - 1)^2}{12} & 0 \leq i \leq M \\ \frac{(2^M W - 1)^2}{12} & i > M \end{cases} \tag{42}$$

Delay: Delay is the time that data packet takes to travel from the sender to the receiver. This parameter is important in QoS applications, because over a certain value it lessens the interactivity between QoS entities. Finding the mean number in the system, $E[L]$, is straight forward:

$$E[L] = \sum_{k=1}^{K} k P_k = \frac{\sum_{k=1}^{K-1} k \pi_k}{\pi_0 + \rho} + K \left(1 - \frac{1}{\pi_0 + \rho} \right) \tag{43}$$

Based on Little's Law and [16] the mean delay is given by:

$$E[T] = \frac{1}{\lambda} \left[\sum_{k=1}^{K-1} k \pi_k + K(\pi_0 + \rho - 1) \right] \tag{44}$$

Variance of Waiting Time: Delay is the sum of waiting time and the service time. From (44) the mean waiting time is given by:

$$E[W] = E[T] - b \tag{45}$$

From [16], the second moment of the waiting time is given by:

$$E[W^2] = (K-1)\left[(K-2)b^2 + E[B^2] - \frac{Kb\pi_0}{\lambda} - \frac{2b}{\lambda}\sum_{k=1}^{K-1} k\pi_{K-k}\right] + \frac{1}{\lambda^2}\sum_{k=1}^{K-1} k(k+1)\pi_{K-k} \tag{46}$$

Where $E[B^2]$ is the second moment of the service time and can be easily found in section II-C. Thus, the variance is given by:

$$Var[W] = E[W^2] - E[W]^2 \tag{47}$$

Throughput: A packet loss can be defined, in the scope of QoS applications, as a packet that never reaches its destination, or a packet that arrives too late and thus cannot be used to play out the multimedia content. It is directly related to throughput.The initial throughput was given by (20). From the back-off algorithm if a packet collides m times it is dropped. Hence, the *adjusted* throughput Γ is given by:

$$\begin{aligned} \Gamma_n &= \frac{\lambda}{\pi_0 + \rho} \cdot \left(1 - \prod_{i=0}^{m} c_{i,n}\right) \\ \Gamma &= \sum_{n=1}^{N} \Gamma_n \beta_n \end{aligned} \tag{48}$$

There are two more measures to be considered for QoS applications in terms of packet reliability. One is the blocking probability which is the probability of blocking of a packet that arrived to the MAC layer from upper layer when the MAC layer buffer is full. The other one is the dropping probability which shows that some packets, given that they are not blocked, will be dropped since it has reached the specified retry limit. They are given by:

$$P_{block} = P_K = 1 - \frac{1}{\pi_0 + \rho} \tag{49}$$

$$P_{drop} = (1 - P_K)\prod_{i=0}^{m} c_i \tag{50}$$

IV. Error Analysis

The ideal channel model is assumed in the analysis up to this point. However, during the previous research work, the impact of bit error rate on the performance of 802.11 DCF is not taken into consideration for the whole spectrum of traffic load. In [12] and [14], the saturation conditions are assumed to see the effects of channel errors. In [13] the specific wireless channel models [10] are considered and in [2] channel models are presented. The more general error analysis is developed here and it can be applied to any kinds of wireless channel models.

The proposed analytical model predicts 802.11 DCF protocol performance very accurately so the effect of errors on the performance of the access scheme, in terms of all measures mentioned in the previous section, is explored by the extension of error analysis to this model.

Error analysis is divided into two distinct parts in the analytical model. First, the impact of bit errors on the collision probabilities c_i in the back-off algorithm is analyzed based on bit error rate P_b and length of data packet P. Second, using new values of c_i, the rest of the analytical model is pursued.

For basic access mechanism, a data packet needs retransmission if any one bit in the packet is corrupted. For RTS/CTS mechanism, RTS and CTS packets may also be corrupted. For convenience, a variable P_e is defined as a probability that a back-off occurs in a station due to bit errors in packets. It is further assumed that bit errors randomly appear in the packets. In the case of RTS/CTS access mechanism, it is easy to get:

$$P_e = 1 - (1 - P_b)^{RTS+CTS+P+ACK} \tag{51}$$

Since error analysis considers transmission errors, c_{e_i} is now the collision-error probability that a transmitted packet encounters a collision or is received in error provided that no collision occurs and is given by:

$$c_{e_{i,n}} = c_{i,n} + (1 - c_{i,n})P_e \tag{52}$$

This new probability c_{e_i} is used instead of c_i in the analytical model between (13) and (51). For example, τ and p_j is found using c_{e_i} rather than c_i in the back-off algorithm. Also, it is used in the service time distribution, M/G/1/K queueing model and performance metrics.

V. Simulation Model and Validation

The main objective of the simulation model is to provide an unbiased and systematic performance analysis of the queueing analytical model. This section presents the simulation environment, the system parameters, design of the simulation, discussion of the statistical analysis of the simulation results, and model validation.

A. Environment Model

The environment model provides a framework to simulate the network topology and the behavior of network nodes. The primary design requirement is to provide an environment capable of modeling WLAN and to support the features of IEEE 802.11 MAC protocol. The simulation environment consists of three models: the network model, the node model and the process model. Several simulation environments are available that can be used for modeling IEEE 802.11 MAC Wireless LANs. The discrete-event simulator OPNET is chosen to implement the simulation model. OPNET Modeler is the industry's leading environment for network modeling and simulation, allowing you to design and study WLAN networks with unmatched flexibility and scalability. Modelers object-oriented modeling approach and graphical editors mirror the structure of WLAN networks.

OPNET Modeler is based on a series of hierarchical editors that directly parallel the structure of real networks protocols, here WLAN. One of them is the Project Editor which graphically represents the topology of a WLAN network that is the network model. Network model consists of node and link objects, configurable via dialog boxes. Node and link models is dragged and dropped from the editor's object palettes to build the network model. The Project Editor provides geographical context, with physical characteristics reflected appropriately in simulation of WLAN networks.

The next one in the hierarchical editors is the Node Editor which captures the architecture of a network device or system by depicting the flow of data between functional elements, called "modules." Each module can generate, send, and receive packets from other modules to perform its function within the node. Modules typically represent applications, protocol layers, algorithms and physical resources, such as buffers, ports, and buses. Modules are assigned process models (developed in the Process Editor) to achieve any required behavior.

The last one is the Process Editor which uses a powerful finite state machine (FSM) approach to support specification, at any level of detail, of protocols, resources, applications, algorithms, and queuing policies. States and transitions graphically define the progression of a process in response to events. Each state of a process model contains C/C++ code, supported by an extensive library of functions designed for protocol programming. Each FSM can define private state variables and can make calls to code in user-provided libraries. FSMs are dynamic and can be spawned (by other FSMs) during simulation in response to specific events. Dynamic FSMs dramatically simplify specification of protocols that manage a scalable number of resources or sessions. The Process Editor can be used to develop entirely new process models, or the models in OPNET Technologies Model Library can be used as a starting point. IEEE 802.11 Wireless module from the model library is used in the simulation model.

B. Simulation Design

Poisson arrival traffic model is employed at each node during simulation run. Smaller traffic rates are used for light load and larger rates for the heavy load. For each set of experiments the number of nodes and packet lengths were fixed. The model assumed an ideal channel and fully connected network, which matched the assumptions of the analytical model.

802.11 Parameters	**Values**
PYH Layer Specification	DSSS
Channel Transmission Rate	1Mbits/sec
CW_{min}	32
CW_{max}	1024
m	4
σ	20μs
H	6Bytes
DIFS	50μs
SIFS	10μs
RTS	44Bytes
CTS	38Bytes
ACK	38Bytes

TABLE I

IEEE 802.11 SYSTEM PARAMETERS VALUES

System parameters are ones that remain fixed to predetermined values throughout the experiments. They reflect a notion of something that does not change, at least within the scope of the experiments. The system parameter values are given in the IEEE 802.11 MAC layer implementation in OPNET modeler, given in table-I.

C. Statistical Simulation Analysis

The data obtained from a simulation should be considered as an instance of the random variables they represent. Hence these data may potentially exhibit high variation and differ a lot from the represented random variables' characteristics [1]. As such, the data must be analyzed in terms of interval statistics. This section describes the approach adopted for computing confidence intervals and ensuring independence of observations and steady-state.

1) Confidence Interval: Let $X_1, X_2, ..., X_n$ be n observations from conducting n independent experiments (runs of the simulation model). Then it is wanted to estimate the steady-state mean $\mu = E(X)$. Let the sample mean be denoted by $\bar{X}$:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{53}$$

According to the strong law of large numbers, this sample mean is an unbiased estimator of μ [1]. Similarly, let the corresponding sample variance be:

$$S^2 = \frac{\sum_{i=1}^{n} [X_i - \bar{X}]^2}{n-1} \tag{54}$$

This is an unbiased estimator of the variance σ^2 of the random variable X. Then, if n is sufficiently large, an approximate $100(1-\alpha)$ percent confidence interval for μ is defined as the time interval given by:

$$\bar{X} \pm t_{n-1,1-\alpha/2} \cdot \sqrt{\frac{S^2}{n}} \tag{55}$$

where $t_{n-1,1-\alpha/2}$ is the upper $1-\alpha/2$th critical point for the t distribution with $n-1$df (degree of freedom) [1]. The underlying physical meaning is that given n sufficiently large, the probability that the derived confidence interval contains μ is $100(1-\alpha)$. Hence if one independently constructs a very large number of confidence intervals, each based on n observations, the proportion of these that cover μ is expected to be $1-\alpha$.

A number of $n = 10$ independent observations is employed in order to get an estimator $\bar{X}$ of the mean μ of each output in the simulation analysis. Calculation of 95% confidence interval is used to provide us with the interval within which the mean μ lies with 0.95 probability.

2) The Problem of Initial Transient: Special consideration needs to be given to the choice of initial conditions of a simulation. Since it is difficult to choose such conditions to represent the steady-state distribution of the simulated system, they bias the output distribution of the simulation. As time progresses, this bias becomes less evident, and the simulation output distribution converges to the distribution of the simulated system.

Typically, an initial small part of the simulation output data is truncated as it is considered biased by the initial conditions. The corresponding time interval is known as the transient period. The simulation reaches steady-state after this period, meaning that the output distribution approaches the probability distribution of the system state. The technique most often suggested for dealing with this problem is called *warming up the model* or *initial-data deletion*. The idea is to delete some number of observations from the beginning of a run and to use only the remaining observations to estimate μ.

The simplest and most general technique for determining the transient period is a graphical procedure due to Welch [19]. Its specific goal is to determine a time index l such that $E(X_i) \approx \mu$ for $i > l$, where l is the warmup period. In general, it is very difficult to determine l from a single replication due to the inherent variability of the process. As a result, Welch's procedure is based on making n independent replications of the simulation (here $n = 10$) and the procedure how to find the moving average is given in [1]. Figure 5 shows the end-to-end observations and moving average with window size $w = 60$ as a function of instance. The simulation is run for 60 seconds, a time much longer than the anticipated transient interval. As seen in the figure, the simulation outcome before $l = 1000$ shall not be considered. In section V-D, all initial transient parts of the simulation results are ignored according to the mentioned procedure.

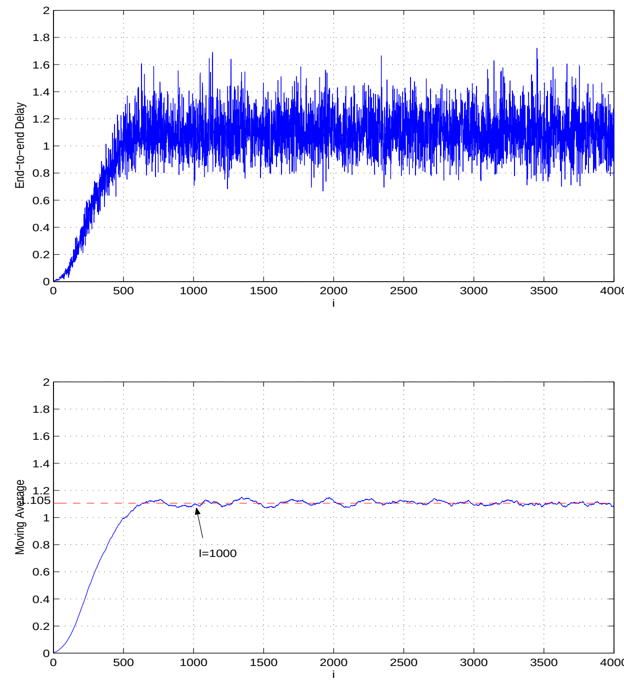

Fig. 5. Moving average with w=60 for the end-to-end delay

3) Independent Observation Acquisition: In order to estimate the mean μ of the desired output, a number n of such observations are needed. These should be statistically independent.

A straightforward approach is to conduct n simulation run, employing distinct random number seeds for different runs. This is known as the method of replication/deletion. Each simulation runs for an amount of time $l+k$. After deleting the transient part of the simulation up to the time l, X_i is obtained as the mean of all observations of the corresponding output during the last time interval k. The duration of the intervals k should be long enough to assure that the mean $\bar{X}_j$ of the observations X_i from jth replicaton will actually be an estimate of the mean μ of the desired output. An approximately unbiased point estimator $\bar{X}$ for μ is given by the mean of $\bar{X}_j$, $E(\bar{X}_j)$ and an approximate $100(1-\alpha)$ percent confidence interval for μ is given by $\bar{X} \pm t_{n-1,1-\alpha/2}\sqrt{\frac{S^2}{n}}$ where $\bar{X}$ and S^2 are computed from Equations (53) and (55).

Load(Kbps/node)	81.92	90.11	98.30	106.49	114.68
95%CI(high)	0.0793	0.1089	0.1128	0.1140	0.1144
Simulation	0.0573	0.0869	0.0908	0.0920	0.0925
95%CI(low)	0.0354	0.0649	0.0688	0.0700	0.0705
Analysis1	0.0379	0.0945	0.0953	0.0954	0.0954
Analysis2	0.0193	0.0417	0.0942	0.0943	0.0943

TABLE II

ANALYTICAL AND SIMULATION RESULTS WHEN THE NUMBER OF NODES IS 10 AND THE PACKET SIZE IS 1024 BYTE

D. Model Validation Results

The theoretical analysis presented in this paper is validated using the simulation model described in the previous section.

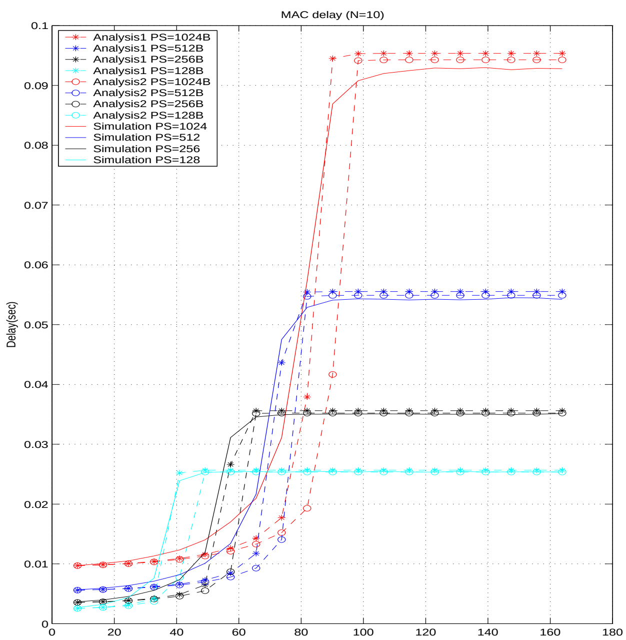

Fig. 6. MAC Delay vs Load

The comparison between the analytical and simulation result of MAC delay for 1024, 512, 256 and 128 bytes packet sizes where the number of node is 10 is shown in Figure-6, taking into consideration of the same traffic load. In the figures, the first modeling of back-off algorithm is represented by 'Analysis1' and the second one by 'Analysis2'. In Table-II the corresponding 95% confidence interval is presented for the MAC delay for 1024 bytes packet size. It is observed that 'Analysis1' fits better in the confidence interval and closer the simulation mean values in the transition part when compared to 'Analysis2'. Hence, with 95% certainty 'Analysis1' outperforms 'Analysis2' with respect to MAC delay. Another observation is that at bigger packet sizes the MAC delay becomes higher during saturation but the network gets congested with lighter traffic load at smaller packets. It is clear that at saturation the packet collision is almost same for all packet sizes, therefore, the bigger transmission times, the bigger MAC delay. It is also obvious that the packet arrival rate is higher with smaller packets for the same traffic load. Hence, the higher arrival rate, the earlier congestion starts regardless of the packet size.

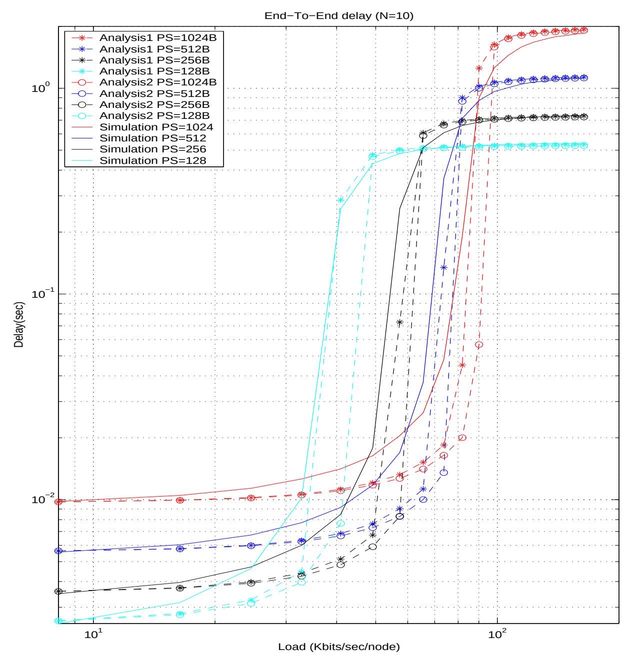

Fig. 7. Delay vs Load

Figure-7 also shows that the simulation validates the analysis in terms of delay under the full spectrum of traffic load. Each simulation point represents the average of 10 independent steady-state replications. The observations are same as in the previous figure since there is a direct relation between MAC delay and end-to-end delay.

In Figure-8 the comparison of the analytical and simulation throughput results versus the traffic load is depicted. It is evident that throughput deteriorates with smaller traffic load for smaller packets. This is also because of the early congestion of the network at smaller packets.

Figure-9 shows the comparison of the analytical and simulation results for the collision probability versus the traffic load for the case of the first stage in the back-

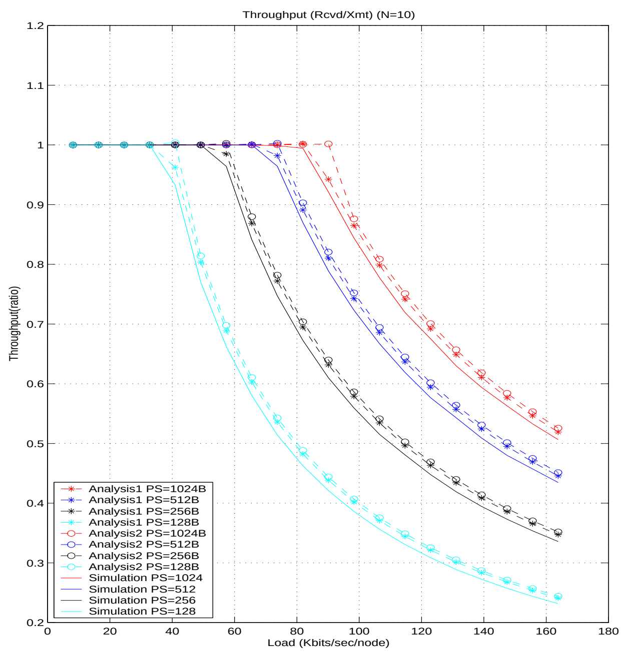

Fig. 8. Throughput vs Load

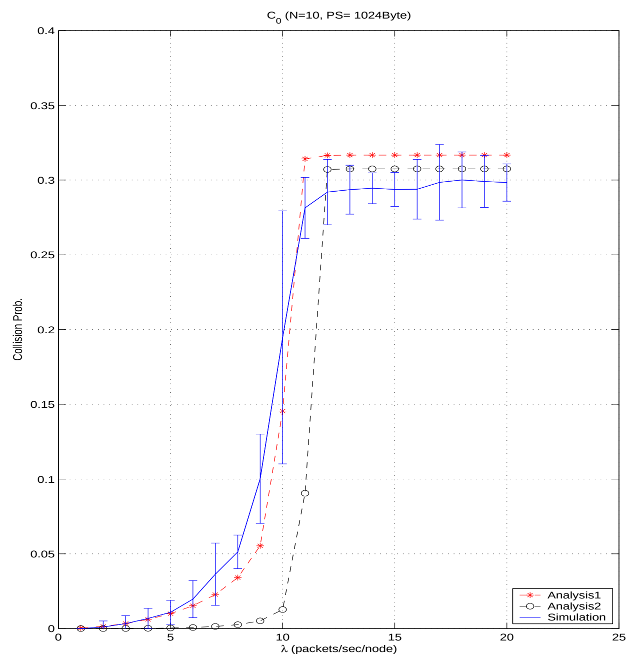

Fig. 9. Collision Probabilities vs Load

off algorithm. In Figure-9 95% confidence intervals are also shown. It can be observed from figures that the first analysis model performs better than the second analysis model in the transition part while the second is better than the first in the saturation part of the traffic load. Both analytical results match the simulation result very closely. Since much less research has been done on the transition part when compared to the saturation part, the first analytical model gives opportunity to delve into the transition part. Hence, the first model is used in the following sensitivity analysis section.

VI. Sensitivity Analysis

The objective of this study requires the development of a sensitivity analysis which gives a deeper insight into the inherent characteristics of IEEE 802.11 MAC protocol.

The first sensitivity analysis is done by changing the number of nodes in the system. The results for all performance metrics are shown in Figure-10 and Figure-11. As the number of nodes increases, all performance metrics deteriorate in a very clear way. That is because more nodes lead to more collision in the network. By restricting the number of nodes accepted in the WLAN, the desired performance can be obtained for a specific QoS application.

Next, sensitivity analysis results for all performance metrics are shown in Figure-12 and Figure-13 while varying the queue length K in terms of how many packet it can accommodate. As the queue length increases, the end-to-end delay increases, on the contrary, both the MAC delay and jitter remain same after K is 10. In the transition part, the waiting time variance jumps up with bigger queue length. Furthermore, the less throughput and bigger dropping and blocking probabilities are observed with less queue length in the transition part. Looking at these figures, the appropriate queue length can be easily chosen for a specific QoS application.

The relation between M and the maximum back-off stage m is given in Equation-17. The sensitivity analysis is performed by changing the value of M and the results are shown in Figure-14 and Figure-15 given that m is 5. It is seen that as M increases, all metrics get better except for the variance of waiting time.

The performance results while changing m are shown in Figure-16 and Figure-17. All throughput metrics are approximately close to each other. At saturation $m = 3$ gives the best result. However, better results for delay metrics are observed with increasing m, shown in Figure-16. There is a tradeoff in terms of balancing better delay or throughput for a particular QoS application. When compared the results with the ones in M, they are not as distinguishable as them.

The sensitivity analysis results are displayed in Figure-

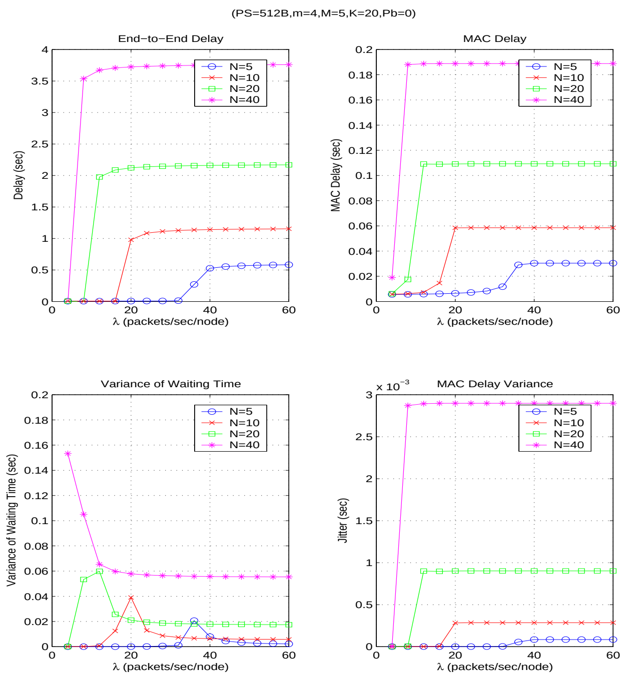

Fig. 10. Sensitivity analysis for delay metrics while varying the number of nodes N

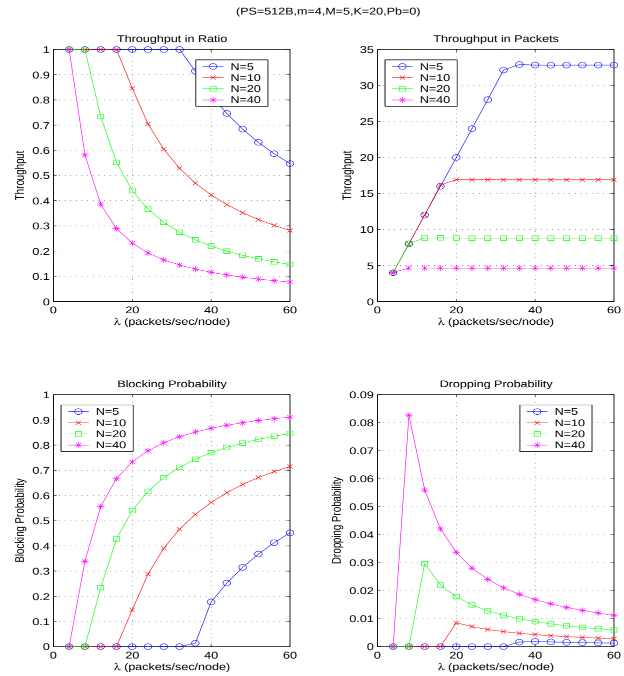

Fig. 11. Sensitivity analysis for throughput metrics while varying the number of nodes N

18 and Figure-19 when altering the initial back-off window size W. Having smaller initial back-off window size leads to not only better delay metrics except for the variance of waiting time but also worse throughput metrics except for the blocking probability. Like m, W causes a tradeoff between delay and throughput. Therefore, the special W must be chosen for a specific application.

Finally, the effect of bit error rate in the channel on performance metrics is shown in the Figure-20 and Figure-21. Keep in mind that all previous figures were obtained considering the ideal channel. Obviously, not only all delay metrics in Figure-20 increase but also all throughput metrics in Figure-21 decrease with enlarging bit error rates. Especially, if P_b is bigger than 10^{-3}, the system collapses at saturation for this specific scenario. As also seen from the figures, the best results like in the ideal channel is obtained if P_b is smaller than 10^{-6}. Furthermore, the other important result is that as the bit error rate increases in the channel, the transition starts earlier which makes more variance and less degradation in the metrics.

VII. Discussion

The analysis presented in this paper has shown that the analytical model is capable of providing a good match to IEEE 802.11 MAC protocol. The following discussion presents possible enhancements to an admission controller as an application that suggest how the controller might benefit from the results presented in this paper. While these suggestions are based on knowledge gained through the results presented in this paper, further study is required to verify the extent of their potential benefits to the controller or other applications such as scheduler or routing protocol. Extension of the simulation model to an admission controller is planned as future work.

- The admission controller decides whether a new connection can be granted or not, depending on the status of the network resources and the level of service called for by the new request. The purpose of any admission control is to ensure that admittance of a new data flow into a resource-limited network does not degrade QoS committed by the network to the admitted data flows while optimizing network resource usage. Since the network resource is changing dynamically depending on the number of nodes, the packet size or system parameters, the optimum resource usage should be decided. Seen from the sensitivity results, if the network is in the saturation, there is no need to grant any new request. It is suggested that each node in the network measures the traffic condition (traffic load) on the wireless link. When the traffic load is

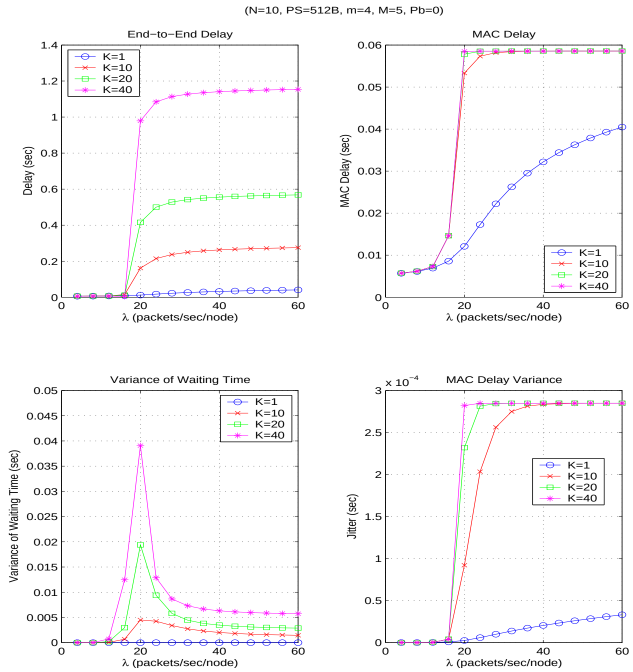

Fig. 12. Sensitivity analysis for delay metrics while varying the queue length K

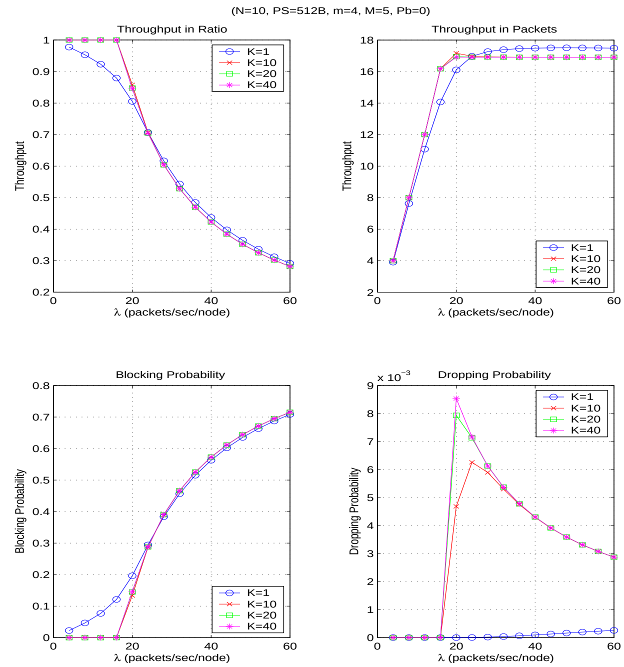

Fig. 13. Sensitivity analysis for throughput metrics while varying the queue length K

greater than a threshold, this means that the 802.11 wireless network is experiencing the overload, long medium access delay and possible the degradation of throughput. For instance, this threshold could be the MAC delay variance since it is easy to obtain the MAC delay for each node.

- The queue in the MAC layer should be in the right length for each admitted request in the admission controller. After a threshold queue length, even though the MAC delay remains same, the end-to-end delay increases steadily. Using the analytical model it is easy to find this threshold queue length.
- In any admission controller once a request is admitted, the service provided to it must not be affected by requests of other nodes. When the number of nodes in the WLAN gets more, the performance of admitted services scales down because of the more packet collisions. Hence, the channel access time for the ordinary service (best-effort) must be increased depending on the channel conditions. From sensitivity analysis results changing the system parameters, m, M, and/or W dynamically will provide a scalable admission controller.

VIII. Conclusion

This paper represents an important advance in the analysis and design of effective ad hoc networks. It also provides an significant contribution to the performance analysis of IEEE 802.11 wireless LANs. A novel queuing theoretic model based on the M/MMGI/1/K queue and parametric service model for IEEE 802.11 DCF using RTS/CTS was solved. The results are based on the single-hop case, but they represent an extensible and flexible approach. It is investigated using the sensitivity analysis results what causes the deterioration of performance measures in a wireless network. The applications for the type of queuing model developed in this work include routing, admission control and scheduling. To the best of the authors' knowledge this is the first comprehensive queuing analysis of IEEE 802.11 DCF using RTS/CTS that has been well-validated and published in the literature.

References

[1] Law A. and Kelton W. D., *Simulation modeling and analysis*, 2nd ed., McGraw-Hill, 1991.

[2] A. Aguiar and J. Gross, *Wireless channel models*, Technical Report TKN-03-007 (Telecommunication Networks Group,Technische Universitt Berlin), April 2003.

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

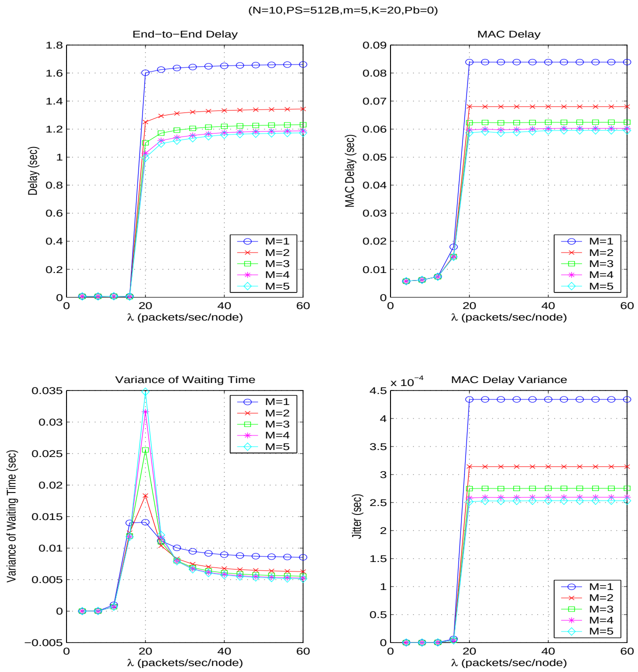

Fig. 14. Sensitivity analysis for delay metrics while varying M

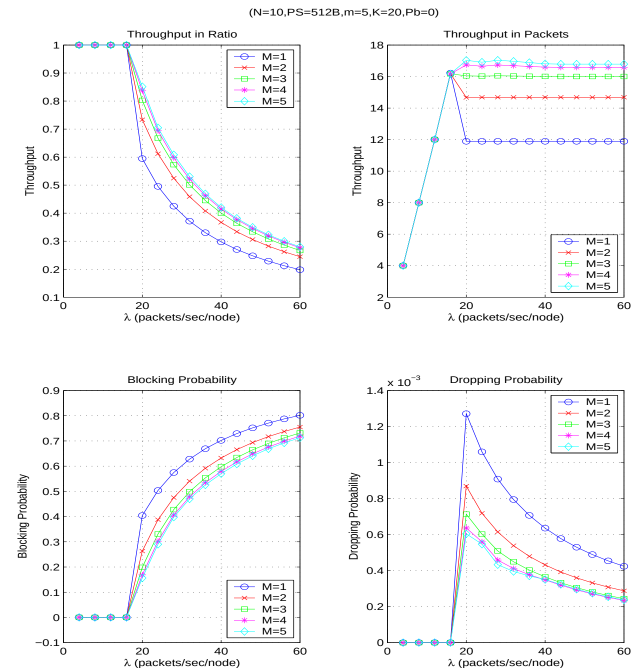

Fig. 15. Sensitivity analysis for throughput metrics while varying M

[4] IEEE Standards Board, *Ieee standard 802.11 - wireless lan medium access control (mac) and physical layer (phy) specifications*, The Institute of Electrical and Electronics Engineers, Inc, 345 East 47th Street, New York, NY 10017-2394, USA, June 1997.

[5] R. B. Cooper, *Introduction to queueing theory*, 1st ed., North-Holland Publishing Company, 1981.

[6] H. Kassab E. Koksal and H. Balakrishnan, *An analysis of short-term fairness in wireless media access protocols*, Extended version of short paper in Proc. ACM SIGMETRICS, June 2000.

[7] Farshad Eshghi and Ahmed K. Elhakeem, *Performance analysis of ad hoc wireless lans for real-time traffic*, IEEE Journal on selected areas in communications **vol. 21** (2003), no. 2, pp. 204–215.

[8] Y. Fang and A. B. Mcdonald, *Theoretical channel capacity in multihop ad hoc networks*, IEEE LANMAN'04 (San Francisco), April 2004.

[9] Marco Conti Frederico Cali and Enrico Gregori, *Dynamic tuning of the ieee 802.11 protocol to achieve a theoretical throughput limit*, IEEE/ACM Transactions on Networking **vol. 8** (2000), no. 6, pp. 785–799.

[10] E. Gilbert, *Capacity of a burst noise channel*, Bell Syst. Tech J. **39** (1960), 1253–66.

[11] Shiduan Cheng Haitao Wu and Jian Ma, *Performance of reliable transport protocol over ieee 802.11 wireless lan: Analysis and enhancement*, Proceedings of the IEEE INFOCOM'02 (New York), June 2002.

[12] Z. Yang W. Cheng J. He, Z. Tang and C. T. Chou, *Performance evaluation of distributed access scheme in error-prone channel*, IEEE TENCON02, 2002.

[13] N. Cocker M. Veeraraghavan and T. Moors, *Support of voice services in ieee 802.11 wireless lans*, IEEE INFOCOM, 2001.

[14] A. C. Boucouvalas P. Chatzimisios and V. Vitsas, *Influence of channel ber on ieee 802.11 dcf*, ELECTRONICS LETTERS, vol. vol. 39, November 2003.

[15] P. Kirstein T. Pagtzis and S. Hailes, *Operational and fairness issues with connection-less traffic over ieee802.11b*, Proceedings of IEEE International Conference on Communications (ICC) (Helsinki, Finland), June 2001.

[16] Hideaki Takagi, *Queueing analysis a foundation of performance evaluation volume 2: Finite systems*, 1st ed., North-Holland Publishing Company, 1993.

[17] Y.C. Tay and K.C. Chua, *A capacity analysis for the ieee 802.11 mac protocol*, ACM/Baltzer Wireless Networks **vol. 7** (2001), no. 2, pp. 159–171.

[18] O. Tickoo and B. Sikdar, *Queueing analysis and delay mitigation in ieee 802.11 random access mac based wireless networks*, Proceedings of the IEEE INFOCOM, 2004.

[19] P.D. Welch, *On the problem of the initial transient in steady-state simulation*, Yorktown Heights, 1981.

[20] G. Xylomenos and G. C. Polyzos, *Tcp and udp performance over a wireless lan*, Proceedings of the IEEE INFOCOM, 1999.

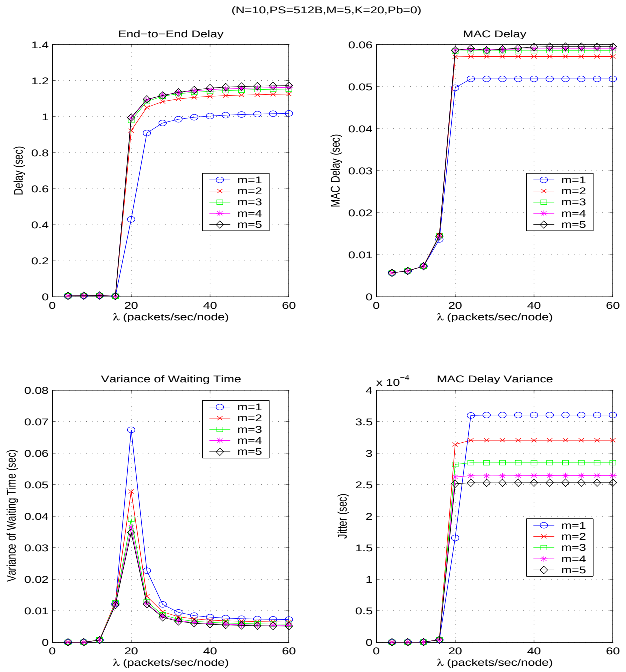

Fig. 16. Sensitivity analysis for delay metrics while varying the maximum back-off stage m

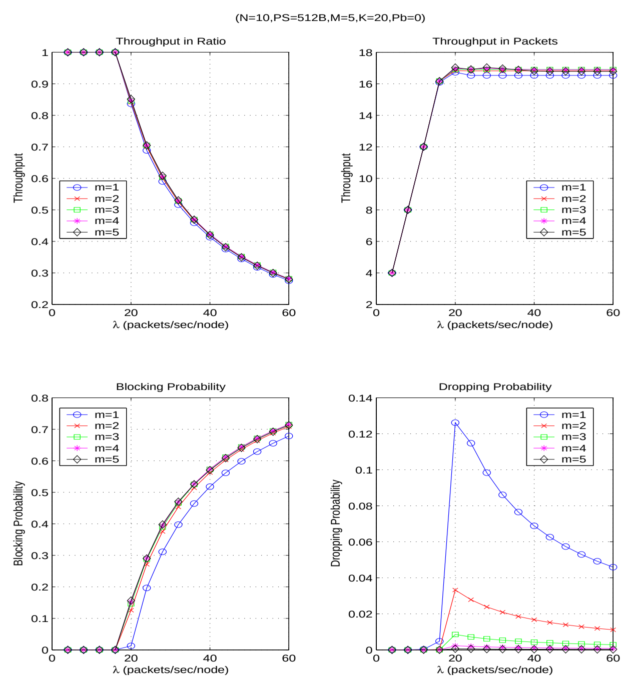

Fig. 17. Sensitivity analysis for throughput metrics while varying the maximum back-off stage m

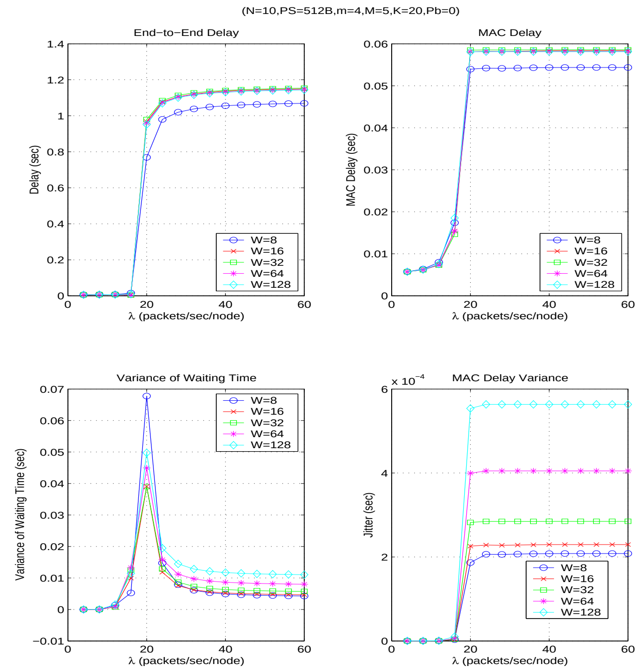

Fig. 18. Sensitivity analysis for delay metrics while varying initial back-off window size W

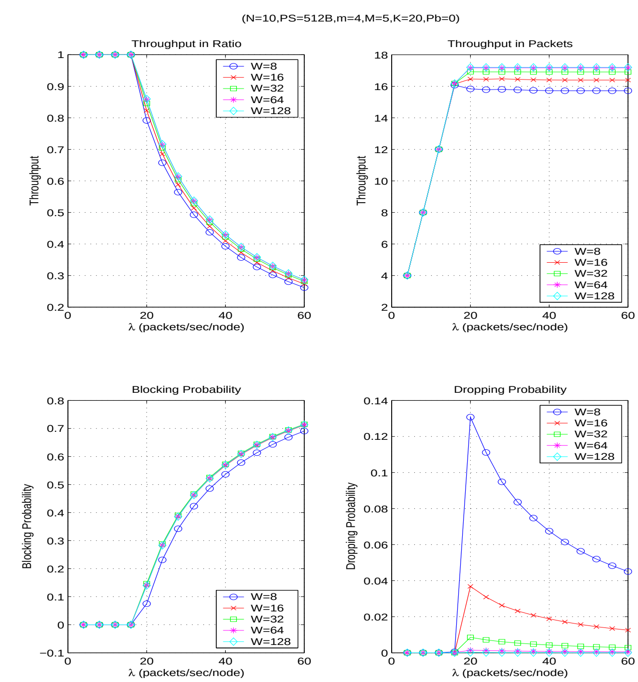

Fig. 19. Sensitivity analysis for throughput metrics while varying initial back-off window size W

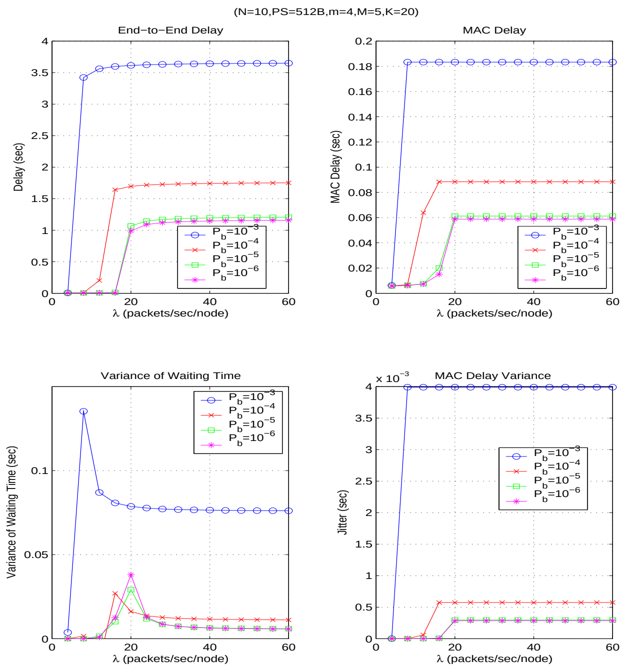

Fig. 20. Sensitivity analysis for delay metrics while varying the the bit error rate P_b

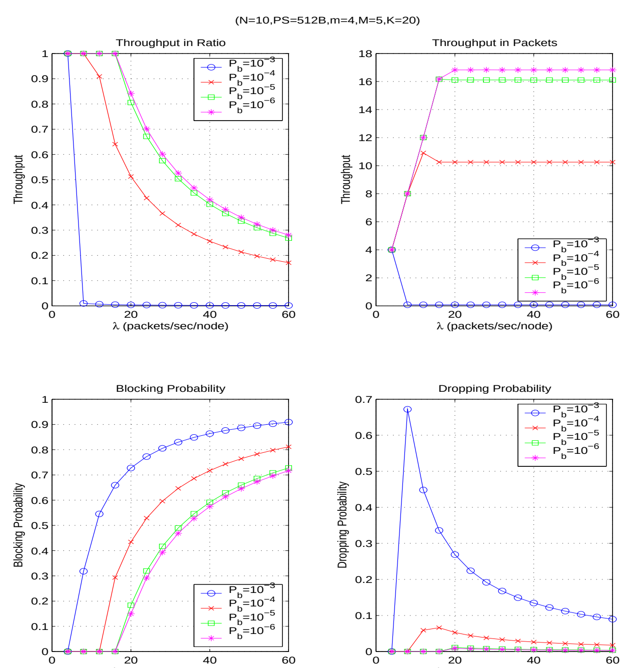

Fig. 21. Sensitivity analysis for throughput metrics while varying the bit error rate P_b