Mobile Networks and Applications 11, 75–90, 2006

DOI: 10.1007/s11036-005-4462-4

Analytical Models for Single-Hop and Multi-Hop Ad Hoc Networks

FARSHID ALIZADEH-SHABDIZ and SURESH SUBRAMANIAM
Department of Electrical and Computer Engineering, The George Washington University, Washington, DC 20005 USA,

Published online: 9 December 2005

Abstract. There is considerable interest in modeling the performance of ad hoc networks analytically. This paper presents approximate analytical models for the throughput performance of single-hop and multi-hop ad hoc networks. The inherent complexity of analysis of a multi-hop ad hoc network together with the fact that the behavior of a node is dependent not only on its neighbors' behavior, but also on the behavior of other unseen nodes makes multi-hop network analysis extremely difficult. However, our approach in this paper to analyze multi-hop networks offers an accurate approximation with moderate complexity. Our approach is based on characterizing the behavior of a node by its state and the state of the channel it sees. This approach is used to carry out an analysis of single-hop and multi-hop ad hoc networks in which different nodes may have different traffic loads. In order to validate the model, it is applied to IEEE 802.11-based networks, and it is shown through extensive simulations that the model is very accurate.

Keywords: ad hoc networks, analytical models, throughput, delay, IEEE 802.11, multi-hop networks

1. Introduction

There is considerable interest in modeling the performance of ad hoc networks analytically. The complexity of modeling the behavior of a node in a multi-hop network grows exponentially with the number of nodes in the network, and this makes exact analysis extremely difficult [5]. However, the approach in this paper to analyze multi-hop ad hoc networks offers a very accurate approximation with only a moderate amount of complexity. Our approach is based on characterizing the behavior of a node by its state and the state of the channel it sees. This approach is used to carry out an analysis of single-hop and multi-hop ad hoc networks in which different nodes may have different traffic loads. As an example, the model is applied to IEEE 802.11-based networks [4], which are the most studied ad hoc networks. Note that the MAC protocol is essentially the same for the various versions of 802.11, viz., 802.11 a, b, and g.

A review of the efforts to analytically model the MAC layer ofad hoc networks shows that there is no model for multi-hop ad hoc networks, and there are only models for saturated single-hop networks. There are reports of attempts to model multi-hop ad hoc networks, but none of them could successfully propose a complete model [9, 10].

A model for analyzing the binary exponential backoff mechanism of 802.11 Distributed Coordination Function (DCF) was introduced in [1], and it was used to compute the saturation throughput of single-hop 802.11-based networks. In [11] the saturation model proposed by Bianchi [1] is extended to IEEE 802.11e. A different approach is taken in [2] to analyze the saturation throughput. In that paper, the binary exponential backoff mechanism of 802.11 is approximated by a p-persistent mechanism. A simplified model for the 802.11 protocol is presented in [3]. That paper deviated from the 802.11 protocol and assumed that all nodes sharing a medium used the same Contention Window (CW), and the CW halved upon a successful transmission and doubled upon a collision. Analysis of single-hop 802.11 based networks in which different nodes may have different traffic loads is considered in [12], but the proposed model failed to capture some aspects of the standard, e.g., after successfully transmitting a packet the transmitter shall select a random backoff interval independent of the state of its queue or if a node receives a packet and the channel is busy it shall select a random backoff.

The rest of this paper is organized as follows. A list of notations is summarized in Table 1. The single-hop network model and assumptions are given in Section 2, and the behavior of a single node is analyzed using a Markov chain model in Section 3. Throughput and delay analyses of a single-hop network are presented in Section 4, and numerical results validating the model are given in Section 5. A throughput analysis of multi-hop networks is presented in Section 6, and simulation results validating the model are given in Section 7. The paper is concluded in Section 8.

2. Single-hop network model and assumptions

We consider the following model for the network: (1) The network consists of n stationary nodes sharing a common medium, and packets are transmitted from sources to destinations directly. (2) The nodes use the IEEE 802.11 DCF (either basic mode or RTS/CTS mode) as the MAC protocol. (3) Each node has an infinite buffer for storing packets. (4) Each node has one transceiver operating in half-duplex

Table 1
List of notations.

$E[P]$	Avg. duration of payload
$E[S_b]$	Backoff service time
$E[T_r]$	Avg. service time
$E[T_t]$	End-to-end delay
$E[T_x]$	Avg. transmission period
$E[T_w]$	Avg. waiting time
m	Max. number of retransmissions
m'	Max. contention window
n_a	Avg. nodes in the interference area with an idle channel
n_r	Avg. nodes in the Rx exclusive area with an idle channel
p	Probability of packet transmission failure
Pidle	Prob that a node and its channel are idle
Pnb	Prob. that a node marks the channel busy
Pns	Prob. that a node is not involved in communication
Ps	Prob. that a busy clot is successful
Ptr	Prob. of sensing a busy slot
q	Prob. of empty queue after processing a packet
S	Throughput
T_b	Waiting time for serving a packet
T_c	Collision slot duration
T_{c0}	Period of one collision
T_r	Avg. receiving period
T_s	Successful slot duration
T_{s0}	Period of one successful transmission
T_v	Vulnerable period
W_i	Contention window
W	Minimum contention window
Γ	Avg. Rx exclusive area
Γ_D	Avg. Rx exclusive area of a random node
λ	Avg. new packet arrival rate
τ	Probability of transmission in an idle slot
σ	Slot duration
$\bar{\sigma}$	Avg. slot duration

mode. (5) Packets arrive to a node according to a Poisson process with rate λ packets. (Even though a Poisson traffic model is assumed here, the analytical model can be generalized to other traffic models. We will assume varying packet arrival rates later in this paper.) (6) A packet transmission is considered to be successful if there are no other packet transmissions at the same time, i.e., the channel is assumed to be error-free. On the other hand, if two or more packet transmissions collide, then all of these packets are considered to have failed (i.e., no capture is assumed) and must be retransmitted.

3. A Markov Chain model for an IEEE 802.11 DCF node

We start by modeling the behavior of a node following the 802.11 MAC protocol. Readers are referred to [4] or [1] for details of the protocol's operation. Our modeling approach is as follows. The transmit and receive states of a node are modeled differently. The transmit states of a node are modeled using a Markov chain following the approach in [1] as shown in figure 1. The binary exponential backoff model in [1] is used in this paper to model the backoff mechanism of the 802.11 DCF protocol. However, when the traffic is non-saturating, a node is not in the backoff states all the time, and additional states are necessary, as shown in figure 1 and explained later in this section. The success or failure of receiving a packet is captured as a state of the channel. Thus, the behavior of a node is completely modeled with the node's transmit state diagram and the state of the channel seen by the node. (The channel state is the same for all nodes in a single-hop network.)

The state of the channel can be characterized by two probabilities—$p_{tr}(n)$, defined as the probability that at least one of n nodes transmits a packet in a random slot, and $p_s(n)$, the probability that there was a successful transmission, given that at least one node transmitted a packet. Let τ be defined as the probability that a node transmits a packet in a randomly chosen slot.

The node's state diagram is shown in figure 1. In order to simplify the model, it is assumed that the probability of failure given a node transmits a packet is independent of the state of the node, as in [1]. This probability is denoted by p. The fact that the collided nodes need to wait a timeout period before sensing the channel (according to the protocol) is ignored in the model [1]. Furthermore, let q denote the probability that the node's buffer is empty after the node finishes processing a packet in backoff (i.e., either successfully transmits the packet or drops it because it has been retransmitted the maximum number of times allowed).

The backoff states are denoted by either (i, w) or $(0', w)$ (the need for these states is explained below), where i is the index of the Contention Window $(CW) W_i = 2^i W$ (W denotes the minimum value of CW) for $0 \le i \le m'$, and $W_i = 2^{m'} W$ for $m' \le i \le m$. Here, m is the maximum number of retransmissions and m' defines the maximum value of W_i. The backoff counter value is denoted by w. (The backoff states, except $(0', w)$, are the same as in [1].) The *IDLE* state is the state in which a node does not have any packet to transmit. The *First-TX* state represents the first transmission of a packet after the *IDLE* state if the channel is sensed idle immediately after receiving a packet. If a packet transmission fails (which occurs with probability p), the node moves one level down (from *First-TX* to backoff level 0 or from level $i-1$ to level i) in the backoff states and chooses one of the numbers $0, \ldots, W_i - 1$ with equal probability $1/W_i$. The backoff counter then decrements once for every idle slot sensed. After a node finishes processing a packet (i.e., either successfully transmitted the packet or dropped it because it has been retransmitted the maximum number of times allowed), it resets i to zero and sets its backoff counter. After expiration of the counter, if the buffer is empty, the node moves to the *IDLE* state; otherwise, it transmits the next packet. In order to capture this process, two sets of backoff states are introduced corresponding to

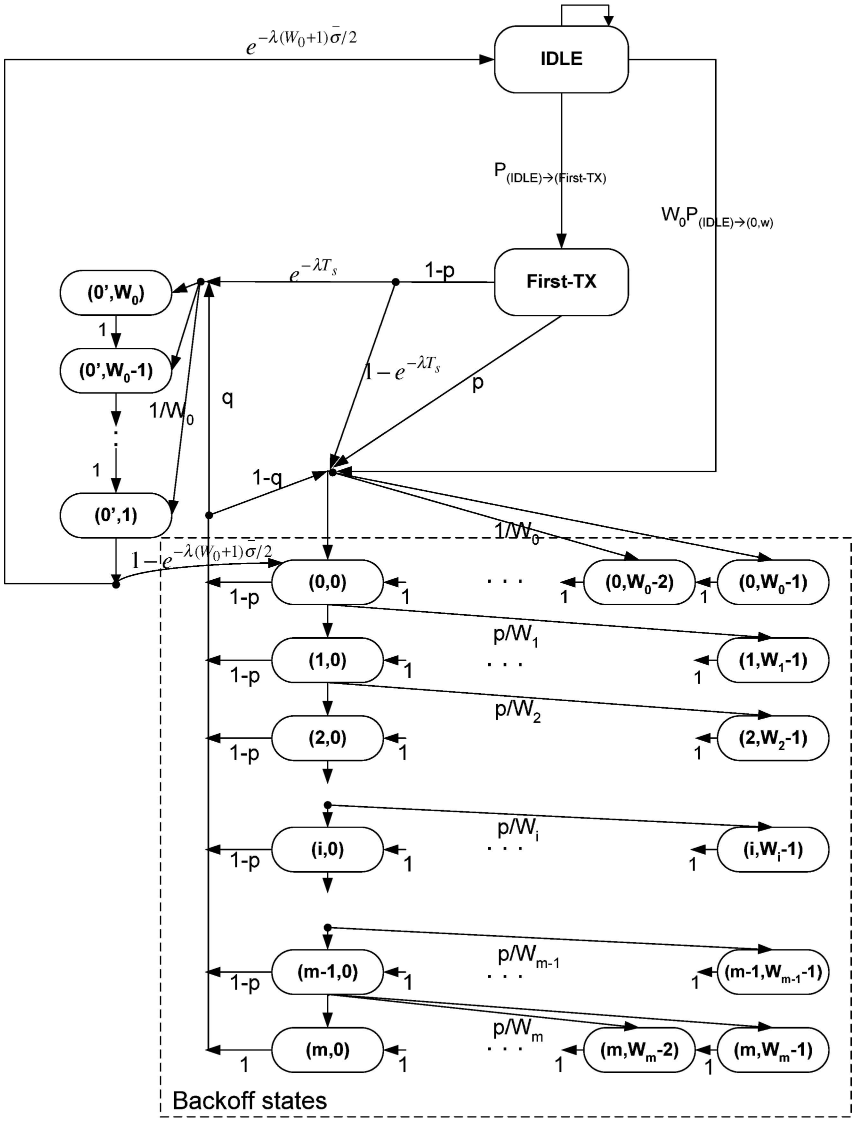


Figure 1. The state diagram for an 802.11 node.

$i = 0$, which are $(0, w)$ and $(0', w)$. A node chooses one of the $(0', w)$ states, if at the time of backoff counter setting, the transmit queue is empty. If, however, the transmit queue is not empty, the node starts backoff procedure by choosing one of the $(0, w)$ states. If a node selects one of the $(0', w)$ states, it checks the queue after backoff counter expiration. If the queue is empty, the node goes to the *IDLE* state, otherwise it transmits the packet (as shown by the transition to state $(0, 0)$). The standard specifies that the backoff counter is chosen from the range $[0, W_0 - 1]$. We deviate from the standard slightly and choose the backoff counter from the range $[1, W_0]$ because it simplifies the calculation of the state probabilities later.

Transitions from state to state occur at the end of *channel slots*. Three types of channel slots are defined, each of different duration: *idle*, *fail*, or *success*, depending on whether a slot on the channel is idle, a collision between two or more transmissions happened during the slot, or whether a packet was successfully transmitted during the slot, respectively. The duration of a channel slot is the period of time that the channel stays in one state: idle, fail, or success. The average lengths of idle, successful and failed slots are denoted by σ, T_s, and T_c, respectively.

The described state diagram is an embedded Markov chain, since the future state of a node given the present state is independent of the past, and the state occupancy time depends on the channel slot time. Next, the transition probabilities for the state diagram are derived. The transition probability from state 'a' to state 'b' is denoted by $P_{a \to b}$. Then

$$P_{(\mathrm{IDLE}) \to (\mathrm{First-TX})} = (1 - p_{tr}(n-1))(1 - e^{-\lambda\sigma}),$$
$$P_{(\mathrm{First-TX}) \to (0'w)} = (1 - p)e^{-\lambda T_s}/W_0, \quad 1 \le w \le W_0,$$
$$P_{(\mathrm{First-TX}) \to (0'w)} = [(1 - p)(1 - e^{-\lambda T_s}) + p]/$$

$$
\begin{aligned}
&W_0, 0 \le w \le (W_0 - 1),\\
P_{(\mathrm{IDLE})\to(0,\mathrm{w})} &= [p_{tr}(n-1)p_s(n-1)(1-e^{\lambda T_s})\\
&\quad + pt_r(n-1)(1-p_s(n-1))\\
&\quad (1-e^{-\lambda T_s})]/W_0, 0 \le w \le W_0 - 1.
\end{aligned} \tag{1}
$$

Here, $1 - p_{tr}(n-1)$ is the probability that an idle channel slot occurs given the node in question is idle. When the node in question is idle, a successful channel slot occurs with probability $p_{tr}(n-1)p_s(n-1)$, and a collision slot appears on the channel with probability $p_{tr}(n-1)(1-p_s(n-1))$.

The transition probabilities in the backoff states are given as follows:

$$
\begin{aligned}
P_{(i,w+1)\to(i,w)} &= 1, \quad 0 \le w \le W_i - 2, 0 \le i \le m\\
P_{(i-1,0)\to(i,w)} &= p/W_i, \quad 0 \le w \le W_i - 1, 1 \le i \le m\\
P_{(i,0)\to(0,w)} &= (1-p)(1-q)/W_0, \quad 0 \le w \le W_0 - 1,\\
&\quad 0 \le i \le m-1\\
P_{(m,0)\to(0,w)} &= (1-q)/W_0, \quad 0 \le w \le W_0 - 1\\
P_{(i,0)\to(0',w)} &= q(1-p)/W_0, \quad 1 \le w \le W_0,\\
&\quad 0 \le i \le m-1\\
P_{(m,0)\to(0'w)} &= q/W_0, \quad 1 \le w \le W_0 .
\end{aligned} \tag{2}
$$

The steady-state distribution of the Markov chain is now obtained. Let $B(i, j)$ denote the probability of being in backoff state (i, j), and let *B(First-TX)* and *B(IDLE)* be the probabilities of the *First-TX* and *IDLE* states, respectively. From the balance equations, one can obtain

$$
B(i,w) = \begin{cases}
(1-p)\sum_{j=0}^{m-1} B(j,0) + B(m,0), \\
\quad \text{if } i = 0, w = 0\\
\dfrac{w_i - w}{W_i} B(i,0), \quad \text{if } 0 \le i \le m, 0 \le w < W_i\\
\dfrac{W_0 - w}{W_0}\big[B(0,0) - B(0',1)\\
\quad \big(1 - e^{-\lambda\bar{\sigma}\,\frac{W_0+1}{2}}\big)\big], \quad \text{if } i = 0, 0 < w < W_i\\
\dfrac{W_0 + 1 - w}{W_0} B(0'1),\\
\quad \text{if } i = 0', 1 \le w \le W_0,
\end{cases}
$$

$$
B(\mathrm{First-TX}) = B(\mathrm{IDLE})(1 - P_{tr}(n-1))(1 - e^{-\lambda\sigma}) , \tag{3}
$$

and

$$
\begin{aligned}
&B(\mathrm{IDLE})[(1 - p_{tr}(n-1))(1-e^{-\lambda\sigma}) + p_{tr}(n-1)\\
&\quad p_s(n-1)(1-e^{-\lambda T_s})\\
&\quad + p_{tr}(n-1)(1-p_s(n-1))(1-e^{-\lambda T_c})]\\
&\quad = [qB(0,0) + B(\mathrm{First-TX})\\
&\quad \times (1-p)e^{-\lambda T_s}]e^{-\lambda\bar{\sigma}\,\frac{W_0+1}{2}} .
\end{aligned} \tag{4}
$$

The average time between successive backoff counter decrements is denoted by $\bar{\sigma}$. By conditioning on the state of the channel in each slot, given that the node in question is in backoff, $\bar{\sigma}$ can be obtained as

$$
\begin{aligned}
\bar{\sigma} &= (1 - p_{tr}(n-1))\sigma + p_{tr}(n-1)p_s(n-1)(T_s + \sigma)\\
&\quad + p_{tr}(n-1)(1 - p_s(n-1))(T_c + \sigma) .
\end{aligned} \tag{5}
$$

Finally, $B(0, 0)$ can be found by noting that the steady-state probabilities sum to one. In order to determine state probabilities, the value of q needs to be calculated. Recall that q is the probability that the transmit queue is empty when a node finishes processing a packet in backoff. In other words, q is the probability that there is no other packet to be processed after processing a packet in backoff states. This happens when the node enters the backoff states with exactly one packet to transmit, and no new packets arrive until the packet's processing is finished. Note that entering the backoff mode with exactly one packet and not receiving a new packet while processing a packet are statistically independent events. If the time it takes to process a packet in backoff mode is denoted by S_b, then q is (approximately) given by

$$
q = e^{-\lambda E[S_b]}\,\text{Prob(One pkt. in buffer when entering backoff)}. \tag{6}
$$

As a further approximation, it is assumed that the probability of entering backoff with one packet is one. For light loads, this is a reasonable assumption because backoff states are rarely visited, and the probability of entering backoff with more than one packet is likely to be very small. For heavy loads, q is almost zero because of large λ and large S_b, and is therefore insensitive to the probability of entering backoff with exactly one packet. The backoff service time, S_b, is the duration of time that is spent in backoff states before a packet is transmitted successfully or dropped due to the maximum retransmission constraint, given that backoff mode is entered by a node for this packet's transmission. One can obtain S_b by determining the time spent in backoff states conditioned on the event that a packet is successfully transmitted after i collisions, $0 \le i < m$, or dropped after m collisions, once backoff is first entered. Thus,

$$
\begin{aligned}
E[S_b] &= \sum_{i=0}^{m-1} p^i(1-p)\left(T_s + iT_c + \sum_{j=0}^{i}\frac{w_j - 1}{2}\bar{\sigma}\right)\\
&\quad + p^m\left[T_s(1-p) + pTc + mT_c + \sum_{j=0}^{m}\frac{w_j - 1}{2}\bar{\sigma}\right].
\end{aligned} \tag{7}
$$

After some algebra, we get:

$$
\begin{aligned}
E[S_b] &= (1 - p^{m+1})\left(T_s + \frac{T_{cp} - \bar{\sigma}/2}{1-p}\right)\\
&\quad + \bar{\sigma}\left(\frac{\sum_{i=0}^{m} W_i p^i}{2}\right),
\end{aligned} \tag{8}
$$

where

$$\sum_{i=0}^{m} W_i P^i = \begin{cases} \dfrac{W(1-(2p)^{m+1})}{1-2p}, & m \le m' \\ \dfrac{W(1-(2p)^{m'+1})}{1-2p} \\ \quad + \dfrac{Wp(2p)^{m'}(1-p^{m-m'})}{1-p}, & m > m'. \end{cases} \tag{9}$$

Now all the state probabilities have been found as functions of p, $p_{tr}(n-1)$, and $p_s(n-1)$. The probability, τ, that a station transmits in a randomly chosen idle slot can be found as $\sum_{i=0}^{m} B(i,0) + B(Fist - TX)$.

Note that a node's state diagram depends only on the channel states and the packet arrival rate at that node and not on the packet arrival rates at other nodes. Therefore, the steady-state probabilities and the value of τ can be found independently for each node as a function of channel state probabilities. The next section shows how the channel state probabilities can, in turn, be expressed in terms of each node's τ.

4. Single-Hop network analysis

We are now ready to present the throughput and delay analyses for a single-hop network.

4.1. Throughput analysis

Now, consider the n-node single-hop ad hoc network with node i having a packet arrival rate of λ_i packets per second, $1 \le i \le n$. Since the node probabilities such as p and τ depend on a node's packet arrival rate, the probabilities for node i are denoted by adding an index i. Let S denote the normalized channel throughput, i.e., the fraction of time that the channel is used to successfully transmit user payload.

First, the equations for τ_i (the probability τ for node i), $1 \le i \le n$, are written in terms of p_i, $p_{tr}(n-1)$ and $p_s(n-1)$. Now, the probability of failure p_i given that node i transmits a packet is equal to the probability that at least one of the other $(n-1)$ nodes also transmits in the same slot, and can be written as follows:

$$p_i = 1 - \prod_{j=1, j\neq 1}^{n} (1-\tau_j). \tag{10}$$

The probability that at least one node transmits in an idle slot is

$$p_{tr}(n) = 1 - \prod_{j=1}^{n} (1-\tau_j), \tag{11}$$

and the probability that a transmitted packet is successful is

$$p_s(n) = \frac{\sum_{i=1}^{n} \left(\tau_i \prod_{j=1, j\neq i}^{n} (1-\tau_j)\right)}{P_{tr}(n)}. \tag{12}$$

Now, the equations for τ_i, p_i, $p_{tr}(n)$, and $p_s(n)$ are solved simultaneously.

Denoting the average duration of payload by $E[P]$,

$$S = \frac{p_{tr}(n) p_s(n) E[p]}{(1-p_{tr}(n))\sigma + p_{tr}(n) p_s(n) T_s + p_{tr}(n)(1-p_s(n)) T_c}, \tag{13}$$

where the numerator is the average payload size per slot and the denominator is the average slot duration. This completes the throughput analysis.

Note that the model here is applicable to both the basic and RTS/CTS modes [1]. The only difference is in the expressions for T_s and T_c in the two modes. The model also covers the case of multiple packet transmissions without interruption.

4.2. Delay analysis

In this section, the analytical model is used to find the steady-state expected packet delay at the MAC layer. For simplicity, it is assumed here that packets arrive at the same rate at all nodes, but the approach can be generalized to the case of different arrival rates for the nodes.

In the non-saturation case, the goal is finding the end-to-end delay, which is defined as the average time from when a packet enters the MAC layer queue until it is successfully transmitted. End-to-end delay is denoted by $E[T_t]$. In the case of saturation, the period of time that a packet waits in the queue is not defined, so only transmission delay will be calculated. Transmission delay is defined as the time from when a packet becomes eligible to be transmitted until it is transmitted successfully. The transmission delay is denoted by $E[T_x]$. The saturation transmission delay is equal to the period of time that a packet (that is ultimately successfully transmitted) spends in backoff. Recall that nodes stay in backoff all the time under saturation traffic. So, the transmission delay is equal to the backoff service time ($E[S_b]$, calculated in (8)) excluding the term corresponding to dropping a packet, i.e.,

$$E[T_x] = E[S_b] - p^{m+1}\left[(m+1)T_c + \bar{\sigma}\sum_{i=0}^{m}\frac{W_i - 1}{2}\right]. \tag{14}$$

The remainder of this section discusses how to find the non-saturation end-to-end delay. The end-to-end delay consists of two components: waiting time and service time. Waiting time is the period of time that a packet waits in the queue for other packets to be transmitted. Service time is defined as the time from when a packet begins to be processed until it is successfully transmitted. Service time includes the actual transmission period and the time that the channel is sensed to be idle and the backoff counter decremented. Note that the time the channel is sensed to be busy and backoff counter is frozen is considered as part of the waiting time. Let us denote the average waiting and service times by $E[T_w]$ and $E[T_r]$, respectively.

To find the non-saturation service time, note that when a packet is generated, it is transmitted without delay if the node and the channel are both idle. Otherwise, the packet undergoes the backoff process. If the packet undergoes the backoff process, the service time is exactly equal to the backoff service time $E[S_b]$, but with the average slot time $\bar{\sigma}$ replaced by the exact slot time σ. Let us denote this as $E[S_b]|_{(\bar{\sigma}=\sigma)}$. This is because the time that the medium is busy while the packet is in backoff states is captured as part of the waiting time. If the packet was generated while the node and the channel were idle, the service time would be T_s in the case of success and $T_c + E[S_b]|_{(\bar{\sigma}=\sigma)}$ in the case of collision. The probability of a node being idle is $B(\text{IDLE})$, and the probability of the channel being idle is the proportion of time that the channel is sensed to be idle given that the node is idle. So, the node and the medium are idle with the probability

$$p_{\text{idle}} = B(\text{IDLE})\frac{(1 - \mathrm{p_t}\,(\mathrm{n} - 1))\sigma}{\bar{\sigma}}\ . \tag{15}$$

Hence, the non-saturation service time can be written as follows:

$$E[T_r] = p_{\text{idle}}\big[(1-p)T_s + p\big(T_c + E[S_b]|_{(\bar{\sigma}=\sigma)}\big)\big] \\ +(1 - p_{\text{idle}})E[S_b]|_{(\bar{\sigma}=\sigma)}\ . \tag{16}$$

The average waiting time is the sum of the times it takes to serve each preceding packet. In order to find the average waiting time, the whole network is modeled as a virtual queue to which packets arrive and receive service from the channel (server). Thus, the network is modeled as an M/G/1 queue with an arrival rate equal to the total rate of arrivals to the network ($n\lambda$). Since the service order is independent of the packet transmission time, the Pollaczek-Khinchin mean value formula can be used to find the average waiting time as follows:

$$E[T_w] = \frac{\lambda n E\big[T_b^2\big]}{2(1 - \lambda n E[T_b])}. \tag{17}$$

Here, T_b represents the time that a given packet waits for another packet to be served. This time consists of the random backoff period, the collision period, and the successful transmission period. The distribution of T_b, can be found by solving the state diagram of a node and the fact that the duration of time that a node resides in each state is known. Due to the lake of space the details are not presented. In order to find the first two moments of T_b, its distribution has to be obtained. The distribution of T_b can be written by conditioning T_b on two events: (a) the transmitted packet arrived when the node and the medium are both idle, and (b) either the node is not idle or the medium is busy when the packet to be transmitted arrived.

The conditional distribution of T_b, given the packet arrived when the node and the medium are idle, is as follows:

$$\text{Prob}(T_b = T_s) = (1-p),$$

$$\text{Prob}\left(T_b = T_s + \frac{iT_c}{2} + \frac{\sigma}{2}\sum_{j=0}^{i-1}\frac{W_j - 1}{2}\right) = p^i(1-p), \\ 0 < i \le m,$$

$$\text{Prob}\left(T_b = (m+1)\frac{T_c}{2} + \frac{\sigma}{2}\sum_{j=0}^{m-1}\frac{W_j - 1}{2}\right) = p^{m+1}. \tag{18}$$

Otherwise, the distribution of T_b is as follows:

$$\text{Prob}\left(T_b = T_s + \frac{iT_c}{2} + \frac{\sigma}{2}\sum_{j=0}^{i-1}\frac{W_j - 1}{2}\right) = p^i(1-p), \\ 0 < i \le m,$$

$$\text{Prob}\left(T_b = (m+1)\frac{T_c}{2} + \frac{\sigma}{2}\sum_{j=0}^{m-1}\frac{W_j - 1}{2}\right) = p^{m+1}. \tag{19}$$

Using the conditional distributions of T_b and the probability p_{idle} the first two moments of T_b are obtained. After some algebra, the first moment of T_b can be written compactly as follows:

$$E[T_b] = \frac{\sigma}{4}(1 - p_{\text{idle}})\left[\sum_{i=0}^{m} W_i p^i - \frac{1 - p^{(m+1)}}{1-p}\right] \\ +\frac{\sigma}{4}p_{\text{idle}}\left[p\sum_{i=0}^{m=1} W_i P^i - \frac{p(1-p^m)}{1-p}\right] \\ +(1 - p^{m+1})\left(T_s + \frac{p_c^T/2}{(1-p)}\right). \tag{20}$$

Note that when two or more packets collide, the collision period and the following backoff period durations must not be included more than once in computing T_b for a packet. Therefore, these durations are divided by the number of colliding packets, and thus are shared between the colliding packets. In order to simplify the analysis, it is assumed that every collision involves exactly two packets. This is the reason that the factor 1/2 appears in the conditional distributions for T_b. Having both $E[T_w]$ and $E[T_r]$ allows us to compute the end-to-end non-saturation average packet delay as:

$$E[T_t] = E[T_w] + E[T_r]. \tag{21}$$

Note that the delay analysis was performed without considering the operational mode of the protocol, RTS/CTS or basic mode.

5. Single-hop network performance results

We now present numerical results that validate the node model and the throughput and delay analyses. The model is validated by comparing the predicted results of the model with simulations. The simulations were done using the Frequency Hopping Spread Spectrum (FHSS) system parameters. The minimum backoff window size (W) is 32 and m' is equal to 5. A summary of FHSS system parameters is shown in Table 2

The value of m was set to the short messages maximum retransmission (*Short Retury Limit*) for RTS/CTS mode, and

Table 2
FHSS system parameters.

MAC header	272 bits
Physical header (PHY)	128 bits
ACK	112 bits + PHY header
RTS	160 bits + PHY header
CTS	112 bits + PHY header
SIFS	28 μs
DIFS	128 μs
m (Short Retry Limit)	7
m (Long Retry Limit)	4
Slot duration (σ)	50 μs
Propagation delay (δ)	1 μs
Channel bit rate	1 Mbps
Timeout	300 μs

to the long messages maximum retransmission (*Long Retry Limit*) for basic mode. The buffer size was set to a large number (forty packets in the simulations) in order to approximate an infinite buffer. The simulation was done for a fixed packet length (8184 bits).

The throughput simulation results are collected after 80 seconds of operation of the network. The throughput versus the offered load for RTS/CTS mode is shown in figure 2. In order to confirm the validity of the model for the case when the traffic at different nodes is not identical, we show results for a network in which packets arrive to each user at a different rate in figures 3 and 4. In these experiments, the packet arrival rate for each node is selected from a uniform distribution within the interval shown next to the curves. As seen in these figures, the capacity in RTS/CTS mode is not sensitive to the network size, but the capacity of the network in basic mode decreases with increasing network size. This is because basic mode throughput is very sensitive to collisions and increasing the network size increases the probability of collision. Note that the maximum capacity of basic mode and RTS/CTS mode are almost the same. Also, observe the excellent match between analytical and simulation results.

Delay results are shown in figures 5 and 6. Figure 5 shows the saturation transmission delay versus the network size for both RTS/CTS and basic modes. Figure 6 shows end-to-end delay results using analysis and simulation for RTS/CTS mode. End-to-end delay increases rapidly with network size. Once again, we notice that the model and simulation results are very close.

6. Multi-hop network throughput analysis

The network model and assumptions for the multi-hop network analysis are the same as for a single-hop network, except that in a multi-hop network the final destination of a packet might not be reached directly and the other nodes can be used to route the packet to the final destination. In the case of a multi-hop network, the packets arriving to a node are composed of newly generated packets and transit packets routed through the node. It is assumed that the total packet arrival rate at a node (sum of new and transit packet arrival rates) is known. The transit packet arrival rates can be obtained from the arrival rate of packets to the network, the traffic distri-

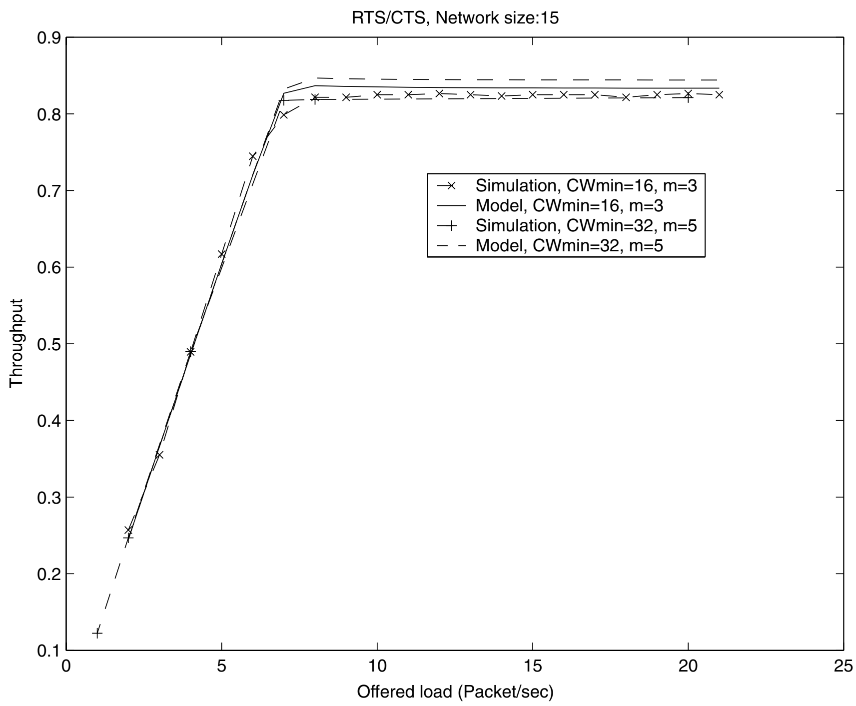


Figure 2. Throughput versus offered load per node; RTS/CTS mode and $n = 15$.

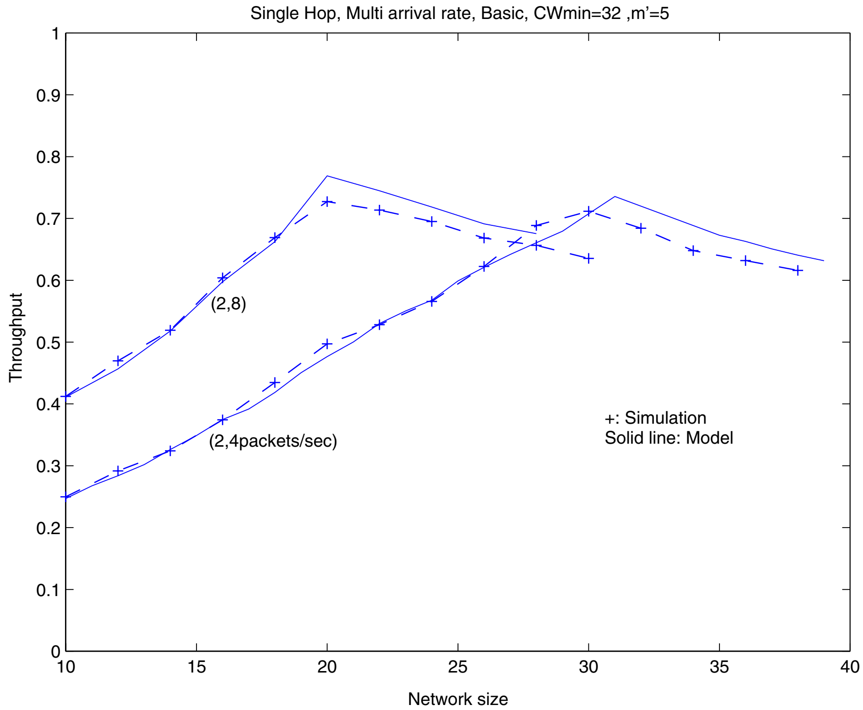


Figure 3. Throughput vs. network size for different traffic loads: basic mode.

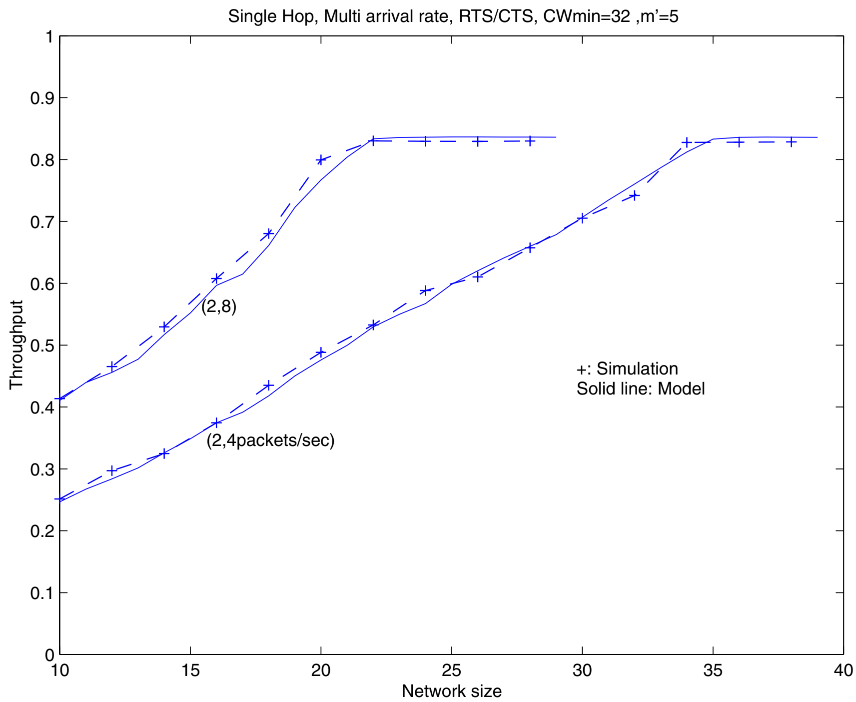


Figure 4. Throughput vs. network size for different traffic loads: RTS/CTS mode.

bution, and the routing algorithm. It is also assumed that the total number of nodes is large enough to ignore edge effects over the service area. We present the analysis for the throughput assuming that each node's neighbors are known, and the packet arrival rates for every node are given. The equations are written for a given node in the network with n neighbors (i.e., number of nodes with which the given node can interfere).

The analysis can be applied to different modes of 802.11 DCF MAC protocol as well as any other ad hoc MAC layer protocol which satisfies the assumptions of this section. The analysis is explained for a general ad hoc MAC layer protocol with known T_s, T_c, and *vulnerable period* (T_v). The vulnerable period is the period of communication that is vulnerable to interference. If no interference occurs during this

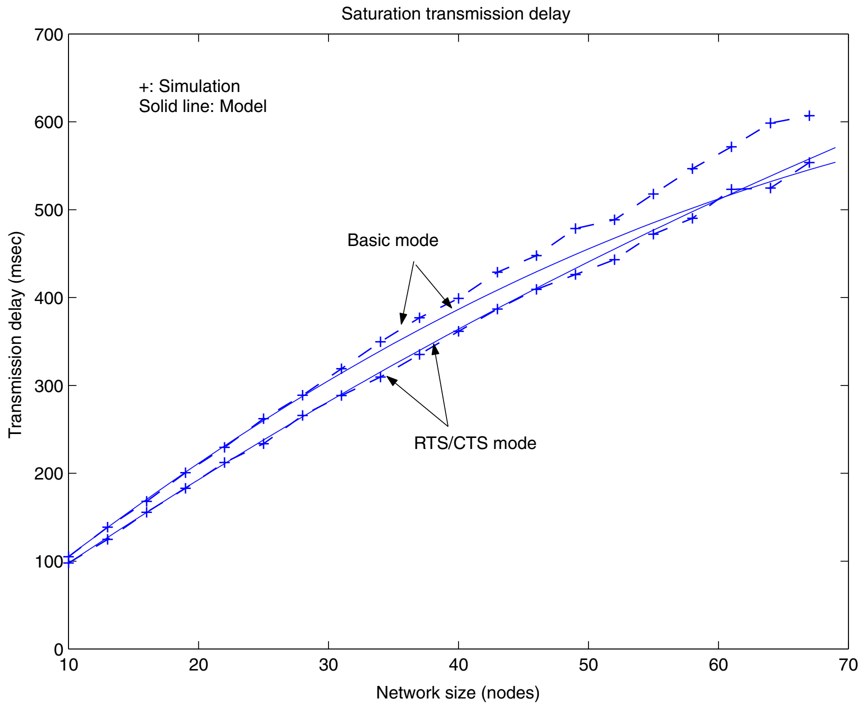


Figure 5. Saturation transmission delay vs. the network size: RTS/CTS and basic modes.

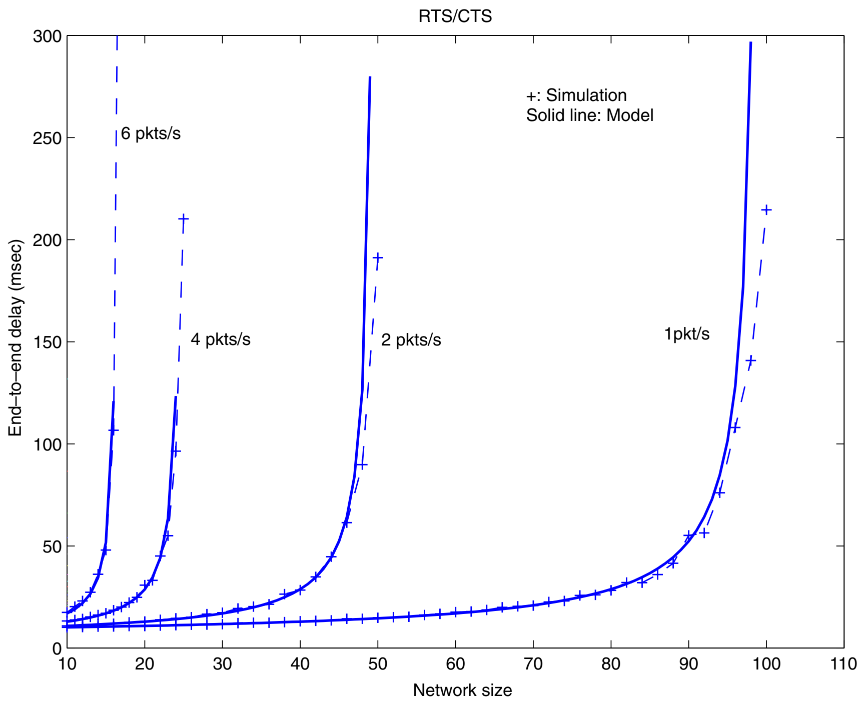


Figure 6. End-to-end delay vs. the network size for different offered loads: RTS/CTS mode.

period, a successful packet transmission occurs. Destination here refers to the next hop destination of the packet. The analysis is simplified by assuming that a collision can only happen at a destination. The first T_v duration of the first packet from a source to the destination is vulnerable to interference.

Let us give some examples of MAC protocols which fit in the above general framework. The basic mode of IEEE 802.11 DCF if collision during ACK transmission is ignored is an example. The vulnerable period, T_v, is equal to the data packet duration. Another example is RTS/CTS mode of 802.11 DCF with busy-tone option. Busy-Tone Multiple Access (BTMA) is a classical method introduced in [6] as a natural extension of Carrier Sense Multiple Access (CSMA) to eliminate the hidden terminal problem completely by marking the channel busy during the entire time of communication [7].

To generalize the multi-hop network analysis to any MAC protocol within the above framework, let us refer to the first message which is transmitted by a source to initiate communication as MSG1, and the first message by the destination to confirm the establishment of communication as MSG2. MSG1 and MSG2 correspond to RTS and CTS, respectively, in RTS/CTS mode, and data packet and ACK, respectively, in basic mode.

Given each node's neighbors and their packet arrival rates, the equations are written for a given node in the network. Let us refer to this node as node S. Let us assume that node S has n neighbors, where a neighbor is a node with which S can interfere, and vice versa. The corresponding parameters of each neighbor are denoted by adding an index i, $1 \leq i \leq n$.

In the multi-hop network model, transmissions from a node are assumed to cause interference at other nodes within a circle of radius r centered on the node. The transmission range of a node is no larger than the interference range and is given by αr, where $0 < \alpha \leq 1$, i.e., a node's transmission may be addressed to any node within a circle of radius αr centered on the node. Here, α is a parameter that is dependent on the physical channel. Note that the transmission range only impacts the selection of the next node on the route to the final destination, and it is the interference range that affects channel and packet transmission status. (Therefore, α does not occur in the various probabilities calculated later.)

The area in which the transmission causes interference is referred to as the *interference area* of a node, and the area with acceptable SNR reception is called the *receiving area*.

In multi-hop networks, source and destination do not cover the same area as shown in figure 7, unlike the identical coverage in single-hop networks. In this paper, the part of a destination interference area which is not exposed to the source is called the "*Rx exclusive area*," and the intersection of the source and destination interference areas is denoted by C. These areas for a source (node S) and the destination (node D) are shown in figure 7. Nodes in the Rx exclusive area do not detect the initiation of communication by the source until the destination reacts to it. Let us denote the ratio of the average number of nodes in the Rx exclusive area to the number of nodes in the interference area by Γ. If the number of nodes in the Rx exclusive area of node S and node D is denoted by $N(S, D)$, Γ can be written as follows:

$$\Gamma = \frac{E_i[N(S, i)]}{n}. \tag{22}$$

Here, the probability of $N(S, i)$ is equal to the probability of node S transmitting an arbitrary packet to a neighboring node i, and is a function of both the traffic and the routing algorithm.

If the next hop destination node is selected from the neighbors with equal probability and $\alpha = 1$, the value of Γ is denoted by Γ_0 and can be found as:

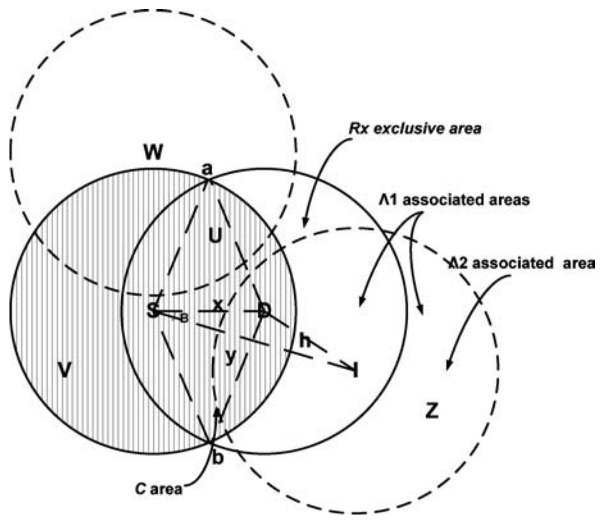


Figure 7. Example of a transmitter and a receiver in a multi-hop network.

$$\Gamma_0 = \frac{\frac{1}{n}\sum_{i=1}^{n} N(s, i)}{n} = \frac{avg_i[N(S, i)]}{n}, \tag{23}$$

in which, $\frac{1}{n}\sum_{i=1}^{n} f(i)$ is denoted by $avg_i[f(i)]$. This notation is also used later to simplify the equations.

As mentioned earlier, the analysis consists of finding p, $p_{tr}(n)$, and $p_s(n)$ for the arbitrary transmitting node S which has n neighbors. The computation of these probabilities is discussed in the next three sub-sections. Following this, the evaluation of the throughput will be discussed.

6.1. The probability of failure (p)

Let us now proceed with finding the probability of failure, p, given that the node transmits a packet. The reader is referred to 7. Given that node S initiated communication by transmitting MSG1 in a given slot, the transmission will be successful if and only if all of the following events happen: (1) No node in the Rx exclusive area is involved in communication; otherwise, the destination's channel would be marked as busy. (2) No MSG1 transmission is in area C in the given slot. (3) No MSG2 transmission is in area C in the given slot, given that there was no MSG1 transmission in the source's interference area. (4) No MSG1 transmission is in the Rx exclusive area during the vulnerable period (of duration T_v). (5) No MSG2 transmission is in the Rx exclusive area during the vulnerable period given that there was no MSG1 transmission in that area.

Let us refer to the above events as *event1* to *event5*, and their corresponding probabilities as *p-event1* to *p-event5*, respectively. *p-event1* through *p-event5* are found next.

p-event1: *p-event1* is equal to the probability that all the nodes in the Rx exclusive area are silent (i.e., not involved in communication). Let us denote the probability of node i being silent by p_{nsi}. Not all the nodes in the Rx exclusive area have opportunity to get involved in communication and not all of them sense an idle channel at the same time as the destination. There is a percentage of nodes in that area which marked their channel busy because of activities outside the destination interference area. Let us denote the average number of nodes in the Rx exclusive area which sense an idle channel at the same time as the destination does by n_r.n_r is determined later. (Here and in the rest of the analysis, the averaging is done over the next hop destination by assuming, for simplifying reasons, that each neighbor is equally likely to be the next node selected.) In summary, *p-event1* is the probability that n_r nodes in the Rx exclusive area are silent. , and is written as follows:

$$p\text{-}event1 = (avg_i[p_{nsi}])^{n_r} \ . \tag{24}$$

p-event2 and p-event3: **p-event2** is equal to the probability that no node in area C transmits MSG1 in the given slot. The probability of no MSG1 transmission by a node i in area C in the given slot is $(1 - \tau_i)$.

Let us elaborate on *p-event3*, the probability of no MSG2 transmission in the given slot, given that there was no MSG1 transmission in area C. An MSG2 transmission in area C cannot be a response to an MSG1 transmission that was originated in area C, otherwise the channel would be busy and *event2* would not have occurred. We are interested in the cases where a node in area C (such as U in figure 7) responds to MSG1 (from a node such as W), which was not sensed by node S. Therefore, node S has to be in the Rx exclusive area of the node that is transmitting MSG2 (node U) in the same slot as node S. If we consider all the MSG2 transmissions in the interference area of node S, only a fraction Γ of nodes on average meet this criterion. The probability of an MSG2 transmission is equal to the probability of a successful MSG1 transmission, which is $\tau_i(1 - p_i)$. Now, the probability of MSG1 transmission by neighbors of node U (such as node W) is approximated by the probability of transmission of node U itself. Therefore, the probability of no MSG2 transmission by a node i in area C in the same slot that node S is transmitting MSG1 is equal to $(1 - \tau_i(1 - p_i))^{\Gamma_i}$.

The last step in the calculation of *p-event2* and *p-event3* is finding the number of nodes with idle channels in area C which participate in the above calculation. The number of nodes with an idle channel in area C is approximated as the total number of nodes in area C, $(1-\Gamma)n$. This approximation is based on the fact that area C falls inside the source and the destination interference areas, and no node in these areas can be involved in communication. So, a large area surrounding the area C is idle, and as a result the probability that a node in area C senses a busy channel, because of activities outside area C, is negligible. Therefore, *p-event2* and *p-event3* are written as follows:

$$p\text{-}event2 = (avg_i[(1 - \tau_i)])^{(1-\Gamma)n},$$
$$p\text{-}event3 = (avg_i[(1 - \tau_i + \tau_i p_i)^{\Gamma_i}])^{(1-\Gamma)n} \ . \tag{25}$$

p-event4 and p-event5: Lastly, *p-event4* and *p-event5* are obtained in a similar way as *p-event2* and *p-event3*. The probability of having no MSG1 transmission by a node i in the Rx exclusive area in a slot (of duration σ) is $(1 - \tau_i)$. Thus, the probability that no node in the Rx exclusive area transmits MSG1 over a duration T_v is written as:

$$p\text{-}event4 = (avg_i[(1 - \tau_i)])^{n_r \frac{T_v}{\sigma}} . \tag{26}$$

Let us now find *p-event5*, the probability of having no MSG2 transmission by a node i in the Rx exclusive area over a duration T_v, given that there was no MSG1 transmission in that area. An MSG2 transmission in the Rx exclusive area cannot be a response to an MSG1 transmission that originated in the Rx exclusive area, otherwise the destination channel would be busy. Therefore, using the same analogy as in the calculation of *p-event2* and *p-event3*, the fraction Γ of nodes with an idle channel in the Rx exclusive area has to be considered. Not all the nodes in the Rx exclusive area sense an idle channel at the same time as the source and destination nodes. Recall that the average number of nodes in the Rx exclusive area that sense an idle channel at the same time as the node S is denoted by n_r. Thus,

$$p\text{-}event5 = (avg_i[(1 - \tau_i + \tau_i p_i)^{\Gamma_i})^{n_r \frac{T_v}{\sigma}} \ . \tag{27}$$

The transmission is successful if all of the five events happen. Therefore, the probability of failure is:

$$\begin{aligned} p = 1 - &(avg_i[(1 - \tau_i)(1 - \tau_i + \tau_i p_i)^{\Gamma_i}])^{(1-\Gamma)n} \\ &\times \left(avg_i[(1 - \tau_i)(1 - \tau_i + \tau_i p_i)^{\Gamma_i}]\right)^{n_r \frac{T_v}{\sigma}} \\ &\times (avg_i[p_{nsi}])^{n_r} \end{aligned} \tag{28}$$

The underlying independence assumptions used to calculate p imply that the nodes in area C are observing the same channel condition, and n_r nodes in the Rx exclusive area are also observing the same channel condition. These are valid assumptions as discussed before. It is also assumed that each of the n_r nodes in the Rx exclusive area being silent is independent of the others.

6.2. The probabilities p_{tr} and p_s

The probability p_{tr}: The probability that node S sees a busy slot, p_{tr}, is one minus the probability that node S sees an idle slot. Node S sees an idle slot if both of the following events happen: no node with an idle channel in the interference area transmits MSG1 (including node S itself), and no node with an idle channel transmits MSG2 given that the corresponding MSG1 has not been seen by node S. Otherwise, the channel would be busy because of the MSG1 transmission. In the second event, the fact that node S is exposed to MSG2 without sensing the corresponding MSG1 means that node S has to be in the Rx exclusive area of the communication, and therefore, the fraction Γ of nodes has to be considered. The probability of no MSG2 transmission by a neighbor (such as U) given that the corresponding MSG1 (by a node such as W) has not been seen by node S is $(1-\tau_i(1-p_i))^{\Gamma_i}$, as in the calculation of *p-event2*. The probability of no MSG1 transmission by a node i is $(1-\tau_i)$. If node S senses an idle channel, not all the nodes in its interference area sense an idle channel at the same time. Let us denote by n_a the average number of nodes in the interference area of node S that sense an idle channel at the same time as node S. n_a will be found later. Hence, p_{tr} can be written as follows:

$$p_{\text{tr}} = 1-(1-\tau)\big(avg_i[(1-\tau_i)(1-\tau_i+\tau_i p_i)^{\Gamma_i}\big)^{n_a} \quad (29)$$

The independence assumption used to calculate p_{tr} is based on the fact that n_a nodes in the interference area observe the same channel condition. Note that only MSG1 transmission by node S has been considered in p_{tr} calculation and not MSG2 transmission. This is because the MSG2 transmission by node S has to be a response to an MSG1 which is originated in the receiving area of node S, and the corresponding MSG1 would have made the channel busy.

The probability p_s: Let us now find the probability that node S observes a successful slot, given that it observes a busy slot. A slot is successful when the node is exposed to at least one successful transmission by a source, or when the node is not exposed to a source but it is exposed to at least one successful reception by a destination, or when the node itself transmits a packet successfully. Let p_{s1} be the probability that there is at least one successful transmission in the node's interference area, given that at least one node in the interference area is involved in communication. Let p_{s2} be the probability that there is at least one successful transmission in the node's interference area, given that at least one node in the interference area is involved in communication. Let p_{s2} be the probability that there is at least one successful reception in the node's interference area, given that the node is not exposed to a source and it senses a busy slot. The probability that a node i successfully transmits in a slot is $\tau_i(1-p_i)$, and since the node itself and on average n_a neighbors participate in generating a busy slot,

$$P_{s1} = \frac{1-(1-\tau(1-p))(avg_i[(1-\tau_i(1-p_i))])^{n_a}}{p_{\text{tr}}} \quad (30)$$

The probability that a node i successfully receives in a slot is $\tau_i(1-p_i)$. In order to eliminate the cases that node S is exposed to both receiver and transmitter, only those cases in which node S is in the Rx exclusive area of a communication have to be considered; thus only the fraction Γ of successful receptions must be considered. Thus,

$$p_{s2} = \Gamma\frac{1-(avg_i[(1-\tau_i(1-p_i))])^{n_a}}{p_{\text{tr}}} \quad (31)$$

Note that successful transmission of node S has been considered in the calculation of p_{s1}, but successful reception of node S has not been considered in the calculation of p_{s2}, since successful reception of node S corresponds to the case that a node in the receiving area of node S is successfully transmitting a packet.

6.3. Durations of success and collision periods (T_s and T_c)

There might be two or more neighbors independently initiating a communication, and their transmissions might overlap in time. Therefore, the lengths of successful (T_s) and collision slots (T_c) are not necessarily the same as a single successful or collision transmission. Any transmission or reception in the interference area is sensed by a fraction $(1-\Gamma_0)$ of neighbors on average, so the other Γ_0 fraction of neighbors might initiate communication and create an overlapping situation. The reader is referred to figure 8. For node S that has a neighbor that is involved in communication (e.g., node A), there are neighbors (e.g., node K) who do not get exposed to communication and can start independent communication during the first transmission. Let us denote the duration of a single successful and collision transmission by T_{s0} and T_{c0}, respectively. A successful slot is defined as the duration of time that the channel is busy and at least one successful transmission occurred in that period. A collision slot is defined as a busy period caused by one or more collisions with no successful transmission during that period. In order to calculate the lengths of the busy slots, let us ignore overlapping of more than two transmissions. This is a valid assumption for most network configurations, since the activity of one neighbor marks the channels of a fraction $1-\Gamma_0$ of nodes in the interference area as busy, on average. For example, the average value of Γ_0 can be calculated to be 0.4135 (the value of Γ_0 is found from equation (23)), if neighbors are randomly positioned in the interference area. From a given node's perspective the duration of successful or collision slot might be extended because of another successful or collision slot, and therefore four different scenarios are distinguished, as shown in figure 8. The scenario in which a collision slot is extended because of a second collision slot needs to be considered in the T_c calculation, and the other three scenarios are part of the T_s calculation. If overlapping of more than two transmissions is ignored, the average success and collision slot durations can be calculated as the sum of one transmission's duration and the average extension due to the second transmission. Thus,

$$\begin{aligned}
T_s &= \frac{T_{s0}}{2}[1-(avg_i[(1-\tau_i(1-p_i))])^{\Gamma_0 n}]\frac{T_{s0}}{\bar{\sigma}} \\
&\quad + \frac{T_{c0}}{2}[1-(avg_i[(1-\tau_i(1-p_i))])^{\Gamma_0 n}]\frac{T_{s0}}{\bar{\sigma}} \\
&\quad + \frac{T_{c0}}{2}[1-(avg_i[(1-\tau_i p_i)])^{\Gamma_0 n}]\frac{T_{c0}}{\bar{\sigma}}T_{s0}, \\
T_c &= \frac{T_{c0}}{2}[1-(avg_i[(1-\tau_i p_i)])^{\Gamma_{0n}}]\frac{T_{c0}}{\bar{\sigma}} + T_{c0} \qquad (32)
\end{aligned}$$

Here, the average extension is equal to half the length of the second transmission multiplied by the probability of its occurrence. The probability of the occurrence of a second successful transmission during a successful slot is equal to the probability that the other Γ_0 fraction of neighbors has at least one successful transmission during the T_{s0} period. A

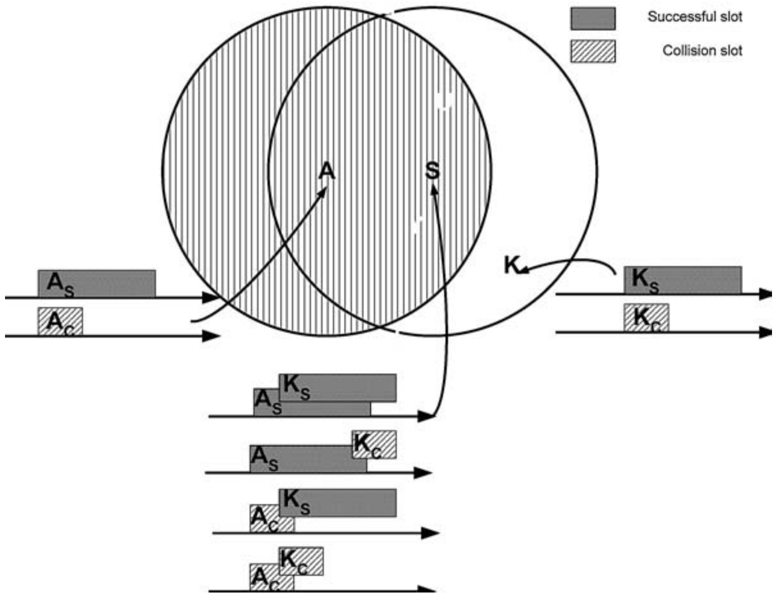


Figure 8. Four different scenarios of overlapping successful and collision slots in a multi-hop network.

collision slot followed by a successful slot is captured in the second term of (32), which is the average extension multiplied by the probability that a Γ_0 fraction of neighbors has at least one successful transmission during the first collision slot. Extension of a successful slot due to a second collision is calculated in the third term of (32). The second collision has to occur during the last T_{c0} period of the first successful slot in order to affect the length of the busy slot, and the probability of a collision slot is τ_p. Extension of a collision slot due to a second collision is found the same way as explained above.

6.4. Throughput calculation

The nodal throughput can be found in the same way as in a single hop network. Hence, the node throughput is

$$S_{\text{node}} = \frac{\tau(1-p)E[P]}{p_{\text{tr}}[p_{s1}T_s + p_{s2}T_r + (1-p_{s1}-p_{s2})T_c] + (1-p_{\text{tr}})\sigma}. \quad (33)$$

The denominator is the average slot time, $\overline{\sigma}$, from node S's point of view. Here, T_r is the duration that the channel is marked busy by a destination node while receiving a packet. The destination marks the channel busy a time T_v after the initiation of communication. Hence, $T_r = T_s - T_v$.

During the multi-hop throughput analysis description, three parameters, viz., n_a, n_r, and p_{ns}, were introduced. These are obtained now.

The probability p_{ns}: The first step in calculating p_{ns} is finding the probability that a neighbor marks the channel busy, denoted by p_{nb}. The probability p_{nb} can be expressed as the ratio

$$p_{\text{nb}} = \frac{\text{The average busy period made by a neighbor}}{\text{The average duration of a slot}} \quad (34)$$

The numerator is the sum of the average period that a neighbor marks the channel busy as a source, and the period that a neighbor marks the channel busy as a destination, given that the corresponding source was not in the neighborhood. A node i transmits successfully with the probability $\tau_i(1-p_i)$, transmits and experiences collision with the probability $\tau_i p_i$, and receives successfully with the probability $\tau_i(1-p_i)$. The duration of time that the channel is marked busy during a successful reception is $T_{r0} = T_{s0} - T_v$. Note that a fraction Γ of receptions has to be considered on average in order to exclude the cases that a source and the destination are both in the interference area. Hence, (34) becomes

$$p_{\text{nb}} = \frac{\tau(1-p)T_{s0} + \tau p T_{c0} + \Gamma_\tau(1-p)T_{\tau 0}}{\overline{\sigma}} \quad (35)$$

Let us define Λ_1 as the *average interference area of a node* (within the Rx exclusive area) excluding the source's interference area, normalized to the node's interference area. Λ_1 is visualized for node I in figure 7. We multiply p_{nb} by Λ_1 in order to eliminate cases of communication with nodes within the interference area of node S, and write p_{ns} is as follows:

$$p_{\text{ns}} = 1 - \Lambda_{1pnb}. \quad (36)$$

The number of nodes n_a: In a multi-hop network, a node and its neighbor may or may not detect an idle channel at the same time. Recall that n_a is defined as the average number of nodes in the interference area of a node S that sense an idle channel at the same time as S. n_a is written as follows:

$$n_a = (n-1)(avg_i[(p_{\text{nsi}})])^{\Gamma_0 n}, \quad (37)$$

where $n - 1$ is used to exclude the node itself.

The number of nodes n_r: We finally calculate n_r, which is the average number of nodes in the Rx exclusive area which sense an idle channel, given that the source and the destination channels are idle. Let us define Λ_2 as the average interference area of a node within the Rx exclusive area excluding the source's and the destination's interference areas, normalized to a node's interference area. Λ_2 is visualized for node I in figure 7. Then, n_r can be written as follows:

$$n_r = \Gamma n(avg_i[(p_{\mathrm{nsi}})])^{\Lambda 2n}. \tag{38}$$

The multi-hop network analysis is completed after finding Λ_1, Λ_2, Γ, and Γ_0. The value of Λ_1, Λ_2, Γ, and Γ_0 are found as functions of the interference area, the receiving area, and the spatial distribution of destination nodes in the receiving area. If the distribution of destination nodes in the receiving area can be expressed analytically, these parameters can be found analytically as well. Otherwise, they have to be found empirically.

7. Multi-hop network performance results

We now present numerical results to validate the multi-hop network analytical model for throughput. Throughput here refers to the MAC layer throughput. Network layer throughput may be obtained from the MAC layer throughput if the average number of hops is known. All the simulation parameters are the same as for the single hop network, except parameters explicitly mentioned here, which are multi-hop specific. RTS/CTS mode of 802.11 DCF with busy-tone is used as the MAC layer protocol. The interference area and receiving area are assumed to be the same.

The multi-hop model validation has been done for two different routing algorithms, Most Forward within radius R (MFR) [8] and Random Forward within radius R (RFR) routing algorithm. In the MFR routing algorithm, the next hop is selected in such a way that progress toward the final destination is maximized. Progress here is defined as the length of the projection of the vector connecting the source to the next hop node onto the vector connecting the source to the final destination. This algorithm results in the shortest-path (in terms of number of hops) in most cases, which is the basis for the most suited routing algorithms for mobile ad-hoc networks such as AODV and DSR. In the RFR routing algorithm, if the destination is not an immediate neighbor of the source, then the next hop node is selected randomly between the nodes with positive forward progress.

In order to validate the model, a 40 node network was considered, in which the nodes were randomly positioned in a square coverage area with a given transmission radius. The simulation was run for 20 seconds. Given the node positions and the routing algorithm, the arrival rate of transit packets, and hence, the total packet arrival rate can be found. The model can now be applied because we are given the neighbors of each node and the packet arrival rate at each node. The results are shown in figures 9 and 10. The numbers next to the curves indicate the (measured) average neighborhood

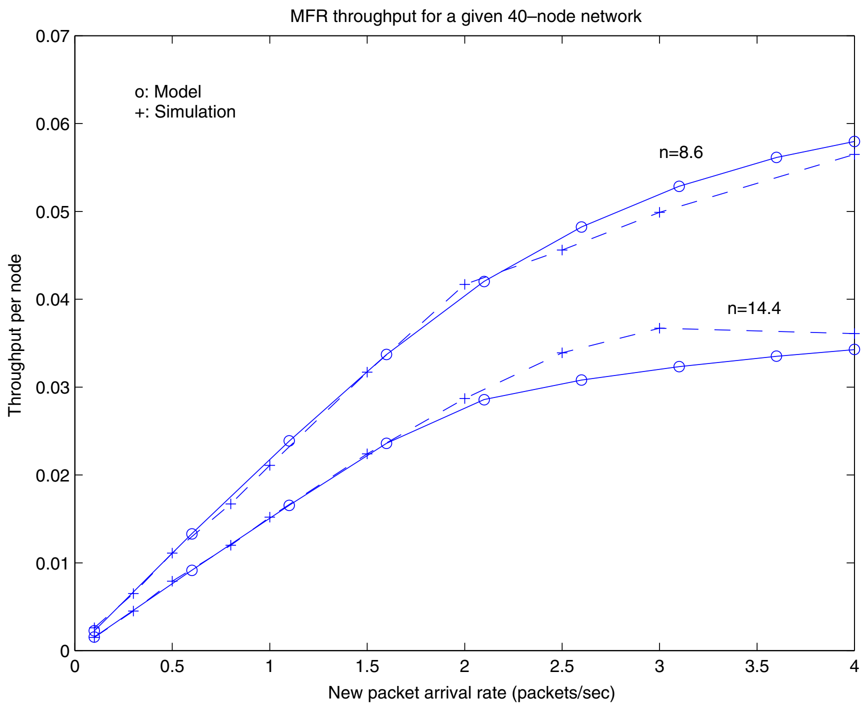


Figure 9. Throughput per node vs. new packet arrival rate: MFR routing.

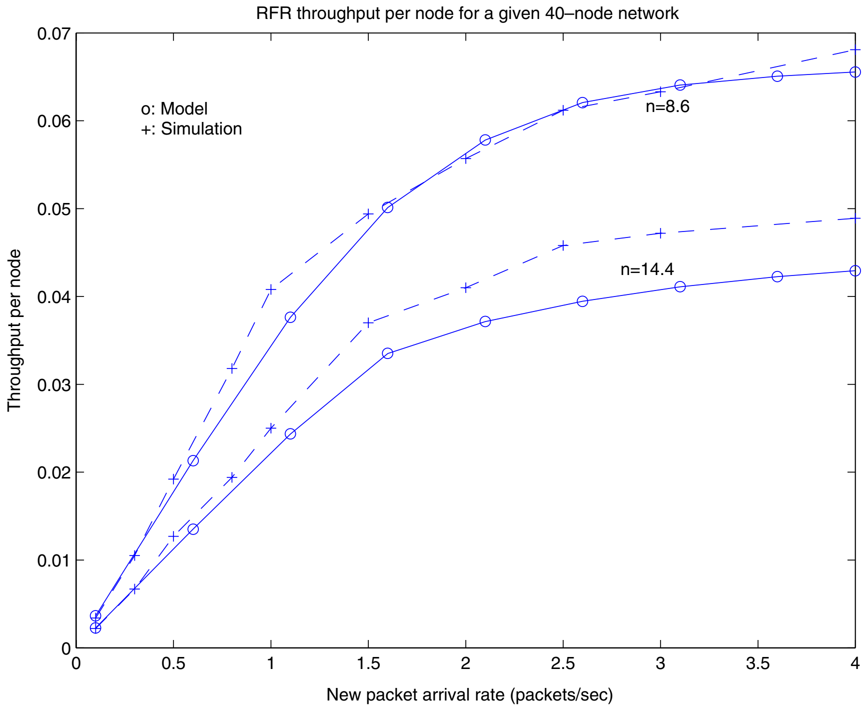


Figure 10. Throughput per node vs. new packet arrival rate: RFR routing.

size. We note that the analytical and simulation results are reasonably close to each other, and that the analytical model is able to accurately predict throughput trends.

Note that we have presented the analysis for a network for which each node's neighbors and every node's arrival rates are exactly specified. As remarked earlier, one may perform an average-case analysis for a network in which the nodes are randomly distributed over a given area, for example, uniformly distributed. In such a case, the quantities Γ and Γ_0 may be found analytically, rather than empirically. We have also applied our model to conduct such an average-case analysis and have verified the results against simulation results. Those are not presented here.

8. Conclusions

Analytically modeling the performance of ad hoc networks is of great current interest. There exists some work on analytically computing the throughput of saturated single-hop networks. Ad hoc networks are not expected to operate near saturation, and it is important to obtain the performance under given, arbitrary traffic conditions. An example of when this would be useful is in the performance comparison of two different routing algorithms that result in different traffic conditions. We are unaware of a complete analytical model that is applicable under non-saturating and varying traffic loads, and to multi-hop networks based on CSMA/CA protocols.

In this paper, an analytical model is presented for single-hop and multi-hop ad hoc networks. The model can be applied to any multi-hop and single-hop ad hoc network as well as different modes of the 802.11 MAC protocol with some reasonable assumptions. Using the model, we evaluate the throughput for single-hop and multi-hop networks, and delay for single-hop networks analytically. Comparisons with simulation results demonstrate the model's accuracy.

Acknowledgments

This work was supported in part by NSF grants ANI-9973111 and CNS-0434956.

References

[1] G. Bianchi, Performance analysis of the IEEE 802.11 distributed coordination function, IEEE JSAC 18(3) (2000) 535–547.

[2] F. Cali, M. Conti and E. Gregori, Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit, IEEE/ACM Trans. on Networking 8(6) (2000) 785–799.

[3] F. Eshghi and A.K. Elhakeem, Performance analysis of ad hoc wireless LAN for real time traffic, IEEE JSAC 21(2) (2003) 204–215.

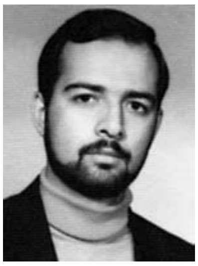
[4] IEEE standard, Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, IEEE standard 802.11, 1997 (1999).

[5] F. Tobagi, Modeling and performance analysis of multihop packet radio networks, Proc. of the IEEE 75(1) (1987) 135–155.

[6] F. Tobagi and L. Kleinrock, Packet switching in radio broadcast channels: Part II-hidden terminal problem in carrier sense multiple access and the busy tone solutions, IEEE Trans. on Commun 23(12) (1975) 1417–1433.

[7] Z. Haas and J. Deng, Dual busy tone multiple access (DBTMA) - A multiple access control scheme for Ad hoc networks, IEEE Trans. on Commun. 50(6) (2002) 975–985.

[8] L. Kleinrock and J. Silvester, Optimum transmission radii for packet radio networks, IEEE National Telecommunication Conference (1978) 4.3.1–4.3.5.
[9] L. Wu and P. Varshney, Performance analysis of CSMA and BTMA protocols in multi-hop networks(I). Single channel case, Information Sciences, Elsevier Sciences Inc. 120 (1999) 159–177.
[10] Y. Wang and J.J. Garcia-Luna-Aceves, Collision avoidance in multi-hop ad hoc networks, IEEE Int'l Symp. on MASCOTS'02 (Oct. 2002) pp. 145–154.
[11] Z. Kong, D.H.K. Tsang, B. Bensaou and D. Gao, Performance analysis of IEEE 802.11e contention-based channel access, IEEE JSAC, 22(10) (2004) 2095–2106.
[12] E. Ziouva and T. Antonakopoulos, The IEEE 802.11 distributed coordination function in small scale ad-hoc wireless LANs, International Journal of Wireless Information Networks 10(1) (2003) 1–15.



Farshid Alizadeh-Shabdiz received his B.Sc. in 1989 at University of Science and Technology, M.Sc. in 1991 at Tehran University, Iran, and D.Sc. in 2004 at the George Washington University. He is a senior research engineer in Advanced Solution Group, part of Cross Country Automotive Services, and he is also a part time faculty member at Boston University.

Dr. Alizadeh-Shabdiz was part of the design and implementation team of the three first satellite-based mobile networks: ICO global medium orbit satellite network voice and data services, Thuraya GEO satellite network, and the first phase of Inmarsat high speed data network. His research interests include MAC layer, physical layer and network layer of wireless and satellite networks.

E-mail: alizadeh@gwu.edu

Suresh Subramaniam received the Ph.D. degree in electrical engineering from the University of Washington, Seattle, in 1997. He is an Associate Professor in the Department of Electrical and Computer Engineering at the George Washington University, Washington, DC. His research interests are in the architectural, algorithmic, and performance aspects of communication networks, with particular emphasis on optical and wireless ad hoc networks.

Dr. Subramaniam is a co-editor of the books "Optical WDM Networks – Principles and Practice" and "Emerging Optical Network Technologies: Architectures, Protocols, and Performance". He has been on the program committees of several conferences including Infocom, ICC, Globecom, and Broadnets, and served as TPC Co-Chair for the 2004 Broadband Optical Networking Symposium. He currently serves on the editorial boards of Journal of Communications and Networks and IEEE Communications Surveys and Tutorials. He is a co-recipient of the Best Paper Award at the 1997 SPIE Conference on All-Optical Communication Systems.

E-mail: suresh@gwu.edu