TUNABLE TRAFFIC CONTROL FOR MULTIHOP CSMA NETWORKS

Zhefu Shi, Cory Beard, Ken Mitchell
University of Missouri-Kansas City

ABSTRACT

A new analytical model for multihop CSMA networks is presented, using matrix exponential methods that model node queueing and backoff processes. This model also captures in a new way the collisions between nodes and then provides performance analysis for not just throughput, but also packet dropping, end-to-end delay, and jitter. All performance metrics are shown to be very accurate compared to simulation results. The model is then used to propose methods for controlling performance of CSMA multihop networks by using different traffic rates and backoff rates for different nodes or paths, depending on their network location.

Keywords: CSMA, QoS, matrix exponential

I. INTRODUCTION

The potential usefulness of multihop peer-to-peer wireless networks, whether they be ad hoc, sensor, personal area, or mesh networks, has been well established. These networks present especially promising opportunities in military and emergency civilian contexts, because of the high likelihood of a damaged infrastructure or even absence of an established infrastructure. Nodes in multihop networks take advantage of physical, MAC, and routing protocols to perform communication effectively in a given the context.

The focus of this paper is the MAC layer, specifically the use of Carrier Sense Multiple Access (CSMA). The paper’s accomplishments are to develop a new matrix exponential analytical model for CSMA performance in multihop networks and then to present new insights into how CSMA multihop network performance can be managed. The use of a CSMA approach is considered here, even through it is well understood that its performance can be quite limited in the presence of many nodes. However, the merits of a CSMA approach are still strong and worthy of consideration, since no explicit coordination protocols are needed. In CSMA, nodes simply sense the wireless medium, transmit if it is idle, or wait a random time and try again later if it is busy.

The CSMA multihop network is modeled in this paper as a continuous-time system, and backoff is accomplished through multiple levels of exponential random waits. This is similar to the IEEE 802.11 backoff process, but uses a continuous-time approach instead of a discrete-time approach and instead of using backoff windows it uses an exponential random variable. By changing the backoff rates of nodes dependent on where they are located in a network topology, this paper shows that throughput, delay, and jitter can be better controlled than if all nodes use the same backoff rates. This paper first introduces the novel analytical model and shows its accuracy in comparison to simulations, then several results are presented regarding the control of backoff rates and traffic rates per different paths.

We use matrix exponential distributions to model the complete binary exponential backoff and queueing process in [1]. Most work in this area has used much simplified models, but ours models the detailed queueing and backoff processes. Our model not only estimates throughput, but also other QoS metrics by using an M/ME/1 model for analysis. Boorstyn [2] used a continuous-time Markov model to develop a product form solution to analyze the throughput of arbitrary topology multihop packet radio networks that use the CSMA protocol with perfect capture and zero propagation delay. In this paper, we use the same assumptions. Kim [3] presented an M/ME/1 model by using an iterative method for the performance evaluation with the infinite backoffs and the same assumption as [2]. In [2], CSMA is modeled as a Markov process. We extend this to multiple backoff stages and exponentially changing backoff rates. In [4] and [5], the product form solution for a wider class of multiple access protocols than CSMA has been established. Kar [6] presented a similar model as [2] and focused on fairness issues regarding throughput with the same assumption as [2]. In [6], however, throughput was optimized based on changing arrival rates; we also consider backoff rates here. In [7], collision probabilities per backoff stage were considered comparing to the assumption collision probability is a constant in [8], like we do here, but for a discrete model.

II. MATRIX EXPONENTIAL ANALYTICAL MODEL

In Wireless Local Area Networks (WLANs) and wireless ad hoc networks, in order to guarantee that each node which contends for the channel can gain access, Carrier
Sense Multiple Access (CSMA) protocols are proposed. In this section, we present an analytic model combined with an iterative solution method.

![Markov chain with queueing of the multi-stage back off process.](image)

Fig. 1. Markov chain with queueing of the multi-stage back off process.

We specify the following notation first:

1) The frame arrival rate to the MAC layer is exponentially distributed with rate $\lambda_i$ at Node $i$.

2) The backoff rate is exponentially distributed with rate $b_i$ at the first stage, $b_i/2$ at the next stage, and so on, where $i$ is the node id. Random backoff in the MAC layer could be implemented according to any random variable, but it is advantageous here to use the exponential random variable to facilitate analysis.

3) The frame service rate for Node $i$ is exponential with rate $s_i$. In the simulation, we set all $s_i$ equal to each other.

4) The probability that the channel is busy when one node finishes the current backoff stage and tries to access the channel is defined as the collision probability. Different backoff stages have different collision probabilities which are denoted as $\delta_{i,g}$ for each node. Here $i$ is the node id, $g = 1 \ldots G$ is the backoff stage.

Now we describe the matrix exponential (ME) distributions we use. The reason for using matrix exponential (ME), is "Most important, many system performance measures which are normally ignored because of their computational and formualational difficulties can be dealt with easily in Linear Algebra Queueing Theory (LAQT)." [9]. The matrix $B$ is the service rate matrix of a matrix exponential distribution which is a family of distributions having rational Laplace-Stieltjes transforms [9]. The density function is $f(t) = \frac{pB}{t}e^{-Bt}\varepsilon$ and the corresponding distribution function is $F(t) = 1 - pe^{-Bt}\varepsilon$, where, $p$ is the starting vector and $\varepsilon$ is a summation operator (usually a column vector of all 1's).

Fig. 1 illustrates the Markov chain for one node using the CSMA process modeled here. Queue fill increases horizontally to the right, and backoff stages increase downward. We use matrix exponential methods to characterize the general service process at each queue fill (each column in Fig. 1), which makes the model M/ME/1. The operation of the model is illustrated, for example, by considering state “q,b,g”. In this state, $q$ frames are in the system, the node is in backoff mode, and it is in the $g$th stage of backoff. The node transfers to the sending state “q,s” if when it finishes backoff the channel is idle. The service matrix $B_i$ for a $G$ stage exponential backoff process for the $i^{th}$ node are

$$B_i = \begin{bmatrix}
    b_i & -b_i \delta_{i1} & 0 & \ldots & 0 \\
    0 & \frac{b_i}{2} & -\frac{b_i}{2} \delta_{i2} & \ldots & \frac{b_i}{2} (1 - \delta_{i2}) \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & \ldots & \frac{b_i}{2^{G-1}} & -\frac{b_i}{2^{G-1}} (1 - \delta_{iG}) & s \\
    0 & \ldots & \frac{b_i}{2^{G-1}} & \ldots & \frac{b_i}{2^{G-1}} 
\end{bmatrix},$$

(1)

In this paper, results are presented for multi-hop scenarios. Similar to [2] and [6], the propagation delay is assumed to be zero in our model. We assume the RTS and CTS are being used but the RTS and CTS transmission times are assumed included in the packet service time. Acknowledgements are obtained instantaneously. Transmission interference is negligible. When one node is sending a frame, it is not affected by another node. If another node finishes its backoff process and intends to send a frame, it only listens to the channel, but does not transmit, hence it does not interfere with the active transmission. In this case, the second node starts a new backoff stage or drops the current frame.

$$L_{i,s} = \begin{bmatrix}
    0 & 0 & \ldots & 0 & \ldots & 0 \\
    0 & 0 & \ldots & 0 & \ldots & 0 \\
    \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
    \frac{b_i}{2^{G-1}} & \frac{b_i}{2^{G-1}} & \ldots & \frac{b_i}{2^{G-1}} & 0 & \ldots \\
    0 & 0 & \ldots & 0 & 0 
\end{bmatrix}$$

(2)

A. Analytical queueing model and solution

We describe an analytical queueing model using an iterative solution method.

1) The standard Matrix Exponential M/ME/1 solution for the steady state probability at queue fill $q$ for Node $i$ is

$$\pi_{i,q} = \pi_{i,0} p (\lambda_i (\lambda_i I + B_i - \lambda_i \varepsilon p)^{-1} y)^q,$$

(3)

$$p = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^T.$$

2) Define $\pi_{i,s}$, the probability that Node $i$ is in a sending state, as that node’s channel utilization,

$$\pi_{i,s} = \sum_{l=1}^{Q_i} \pi_{i,l} \varepsilon_1, \quad \varepsilon_1 = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}^T.$$

(4)
This is the steady state probability of being in a sending state for the \(i^{th}\) node. Assume all \(s_i\) are equal.

\[
\delta_{i,1} = \frac{\pi_{i,s}}{\pi_{i,0} + \pi_{i,1} + \pi_{i,s}}.
\]  

(5)

Here, \(g_1\) means the idle or \(g^{th}\) backoff stage at node 1, and \(x_{1}\) means any status, (idle, \(g^{th}\) backoff stage or sending state) at node 1, etc.

3) Use step 2 to define \(\delta_{i,g}\) for every node so the collision probability for every node can be evaluated.

4) The \(\delta_{i,g}\) for each node is dependent on the \(\pi_{j,s}\) of other nodes. This creates a set of mutually dependent equations. These equations can be solved exactly for a few nodes or with numerical methods for more nodes.

By reconstructing the \(B_i\) matrix each time using the collision probability evaluated from the previous iteration step using the above equations, the queue fill distribution can be evaluated using the following equations.

\[
\pi_{i,q} = \pi_{i,0}p(\lambda_i(I + B_i - \lambda_i\epsilon p)^{-1})^q, \quad 0 < q < Q_i, \quad (6)
\]

\[
\pi_{Q_i} = \pi_{i,0}p\lambda_i B_i^{-1}(\lambda_i(I + B_i - \lambda_i\epsilon p)^{-1})^{Q_i-1}, \quad (7)
\]

\[
\pi_{i,0} + \sum_{q=1}^{Q_i} \pi_{i,q} \epsilon = 1. \quad (8)
\]

Fig. 2. Markov chain with queueing of the multi-stage backoff process.

Here we present a new method for collision probability calculation. Fig. 2 is a Markov chain for the sending or backoff status of each node in a 3-node tandem topology, of each node there is a single backoff stage. If a state has a “0” for node \(i\), it is idle; if it is in the first backoff stage it will have a “1”, and if it is sending it will have an “s”.

The third node is the sink, so its status is not shown here. \(1 - \zeta_t\) means the probability node \(i\) goes to the idle state when node \(i\) finishes the current sending, and \(\zeta_t\) means the probability that node \(i\) goes to the first backoff state when it finishes sending. Also \(1 - \gamma_i\) means the probability that node \(i\) goes to the idle state when node \(i\) finishes its last backoff stage. Notice that the state \((s, s)\) doesn’t exist because node 1 and 2 are in the interference range of each other. In the wireless network, the situation doesn’t exist that any pair of nodes in the interference range of each other both can be in a sending status at the same time.

In this case, the collision probability for node 1 at stage 1 is: \(\delta_{1,1} = \frac{\pi_{1,x}}{\pi_{1,0} + \pi_{1,1} + \pi_{1,s}}\).

In general, assume there are \(N\) nodes in the interference range, each has \(G\) backoff stages, node \(i\)’s collision probability at stage \(g\) is:

\[
\delta_{i,g} = \frac{\sum \pi_{g_1,g_2,...,g_i,...,g_N}}{\sum \pi_{x_1,x_2,...,g_i,...,x_N}}. \quad (5)
\]

Here, \(g_1\) means the idle or \(g^{th}\) backoff stage at node 1, and \(x_{1}\) means any status, (idle, \(g^{th}\) backoff stage or sending state) at node 1, etc.

Fig. 3. 7 nodes, 3 paths, traffic splitting.

Fig. 4. 6 nodes, 3 paths, traffic merging.

Fig. 5. 5 node tandem topology

Fig. 3 through Fig. 5 are the topologies we test for our analytical model, which cover all the basic topologies for traffic splitting, traffic merging, and tandem. Fig. 6 through Fig. 9 show the channel utilization, mean end-to-end delay evaluated by the theoretical model and simulation for the
case of 7 nodes, 3 paths with the traffic splitting topology and a 2-stage exponential backoff process with changes in arrival rate and backoff rate. Next, Fig. 10 through Fig. 13 show the channel utilization, mean end-to-end delay evaluated by the theoretical model and simulation for the case of 6 nodes, 3 paths with a traffic merging topology and a 2-stage exponential backoff process with changes in arrival rate and backoff rate. Simulation was performed using a custom built model in CSIM using the same assumptions as the analytical model. For the case of changing arrival rates, the arrival rates of all the nodes, $\lambda_i$, are from 10 to 200 frames/sec, the backoff rates $b = 2000$ attempts/sec, the sending rates $s = 200$ frames/sec (same for all nodes). For the case of changing backoff rates, the arrival rates are $\lambda_1 = 100$, $\lambda_2 = 15$ and $\lambda_3 = 25$, $b_i$ is from 200 to 6000 occurrences/sec, and the sending rates are $s = 200$ frames/sec.

The simulation results match the analytical results very closely for all performance metrics. A key benefit of this model is also seen in that the analytical results match with the simulation results for all ranges of traffic loads. This is an improvement to the popular throughput model in [8] which only addresses constant backlog conditions (i.e., high loads).

Now we can see our first important application of this model. Fig. 14 that shows the effect of changing backoff rate for all 5 nodes in the 5 node tandem topology, for the end-to-end throughput over the whole path. Notice that at this traffic load increasing backoff rates does not always help to increase end-to-end throughput. For example, when $\lambda = 30$, increasing $b$ when $b > 100$ makes the end-to-end throughput actually decreases. Development of a search algorithm is critical for optimizing the value of the backoff rate that is needed to maximize the end-to-end throughput for a particular topology and load. Fig. 15 through 16 show the effect on end-to-end delay and variance of end-to-end delay from only changing the backoff rate at node 1, $b_1$, in the 5 node tandem topology. Also Fig. 17 through 18 show the effect on the number of blocked frames (due to full queues) and dropped frames (due to collisions at all backoff stages) from changing arrival rate at node 1, $\lambda_1$, in the 5 node tandem. Notice for node 1, the number of blocking and dropped frames increases as $\lambda_1$ increases. For the other downstream nodes,
the number of dropped frames increases as $\lambda_1$ increases, but the change of the number of blocked frames can be ignored, which means at this traffic load, dropping is the only reason for changes in the end-to-end throughput.

III. ANALYSIS OF THROUGHPUT AND QUALITY OF SERVICE

First define the channel utilization for node $i$:

$$\pi_{i,s} = \frac{\lambda_i (1 - \pi_{i,Q})(1 - \prod_{g=1}^{G} \delta_{i,g})}{s_i}.$$  \hspace{1cm} (9)

The term $(1 - \pi_{i,Q})$ means the ratio of the frames left after dropping because the queue is full. The term $(1 - \prod_{g=1}^{G} \delta_{i,g})$ means the ratio of the frames left after collision. This result is applicable for cases of 1) infinite queue and 2) finite queue because the frame loss comes from blocking when queues are full and from dropping after collisions through all backoff stages. The throughput for the whole path is throughput of the last node in the path.

Secondly define the mean service time $\mu_{i,srv}$ and $\rho_i$ and $\rho$ as follows

$$\mu_{i,srv} = \frac{1}{\lambda_i} = \frac{pB_i^{-1}}{\epsilon}, \quad \rho_i = \frac{\lambda_i}{\mu_i},$$  \hspace{1cm} (10)

From [9], the mean residual vector for a frame in service at Node $i$, $\rho_{i,srv}$, and the mean remaining time for a frame in service, $\mu_{i,r}$, are

$$\rho_{i,srv} = \frac{pB_i^{-1}}{\epsilon}, \quad \mu_{i,r} = pB_i^{-1}\epsilon.$$  \hspace{1cm} (11)

Customers only care about the delay for frames that are successfully transmitted. Since dropped frames need to go through all the backoff stages, we define $\overline{T_{i,srv,d}}$ as the average delay time for a convolution of $G$ backoff stages where the rate at each stage is $\frac{b_g}{b_i}, g = 1 \cdots G$. For example, $\overline{T_{i,srv,d}} = \frac{1}{b_i}$ for the single stage case. Considering $\overline{T_{i,srv}}$ has two parts, 1) delay of the dropped frames, $\overline{T_{i,srv,d}}$, and 2) delay of the successful transmission, $\overline{T_{i,srv,s}}$, we have

$$\overline{T_{i,srv}} = \overline{T_{i,srv,d}} + \frac{\prod_{g=1}^{G} \delta_{i,g}}{1 - \prod_{g=1}^{G} \delta_{i,g}}.$$  \hspace{1cm} (12)
The density function of the successful transmission time is
\[ f(t) = p e^{-B_i t} L_{i,s} \epsilon. \]
Therefore, the mean service and system time for successful transmission is
\[ T_{i,srv,s} = \frac{pB_i^{-2} L_{i,s} \epsilon}{pB_i^{-1} L_{i,s} \epsilon}. \]  
(13)

The mean end-to-end delay and jitter can also be expressed with the Laplace transform by taking advantage of changing the time domain convolution of delay at each node to a product in the frequency domain. The Laplace expression for
\[ f(t) = p e^{-B_i t} L_{i,s} \epsilon \]  

is
\[ F(s) = p(B + sI)^{-1} L_{i,s} \epsilon. \]

Since we assume the service process in each node is independent, the Laplace transform of end-to-end delay for a specific path, the convolution, can be expressed as
\[ \prod_{i=1}^{N} p_i(B_i + sI)^{-1} L_i \epsilon. \]

The first and second moment of the convolution can be calculated by 1) taking the first and second derivative, respectively, 2) set \( s = 0 \). Then the mean end-to-end delay and jitter can be calculated easily.

Fig. 19 and Fig. 20 show our second and third important results concerning control of a multihop CSMA network. First of all in Fig. 19, the arrival rate to path 1 in the 7 node, 3 path topology, \( \lambda_1 \) decreases, which makes the arrival rate to path 2, \( \lambda_2 \), increase, since their sum is kept constant. This results in increased the end-to-end throughput for path 2, but the end-to-end throughput of path 3 decreases. Loads in one part of a topology have a profound effect on other parts of the topology and we will continue to study how throughput can be improved through traffic engineering.

The third important result is seen in Fig. 20. When \( b \leq 520 \) from Fig. 20, end-to-end throughput of paths 1 and 3 increase, and end-to-end throughput of path 2 decreases. When \( b > 520 \), the opposite occurs for all paths. This will continue to be studied.

IV. CONCLUSION

This paper provides a new analytical model for continuous-time CSMA medium access control that applies for all ranges of network loads for multihop network, which
is in contrast with the majority of recent CSMA modeling which assumes continuously backlogged queues (i.e., high loads). It presents an iterative solution methodology where stage-dependent collision probabilities of different nodes are related to each other through each others’ channel utilizations. Backoff processes are modeled as exponential random variables with multiple backoff stages. From this model, dependencies were found with respect to arrival rates to the MAC layer ($\lambda_i$) and backoff rates ($b_i$) for end-to-end throughput, delay, and jitter.

With the help of the analytical model, we found several important applications. First of all, as frames progress through a network, dropping is the main reason for loss, instead of blocking. Secondly, increasing backoff rate, $b$, will not always increase the end-to-end throughput. In some traffic loads, increasing backoff rates will decrease the end-to-end throughput. And finally, optimized choice of traffic load paths is another method to improve the end-to-end quality of service. The study of these three applications is important because of their potential to improve end-to-end quality of service, especially in urgent emergency or military scenarios. This work now also provides a basis for future work to devise more advanced strategies to improve end-to-end quality of service, increase overall system utilization, incorporate continuous-time MAC protocols in future wireless networks, or prioritize access to channels for important uses, like emergency applications.

REFERENCES