Queuing Delay and Achievable Throughput in Random Access Wireless Ad Hoc Networks

Nabhendra Bisnik, Alhussein Abouzeid
Rensselaer Polytechnic Institute
Troy, NY 12180
bisnin@rpi.edu, abouzeid@ecse.rpi.edu

***Abstract*— In this paper we focus on characterizing the average end-to-end delay and maximum achievable per-node throughput in random access multihop wireless ad hoc networks with stationary nodes. We present an analytical model that takes into account the number of nodes, the random packet arrival process, the extent of locality of traffic, and the back off and collision avoidance mechanisms of random access MAC. We model random access multihop wireless networks as open G/G/1 queuing networks and use diffusion approximation to evaluate closed form expressions for the average end-to-end delay. The mean service time of nodes is derived and used to obtain the maximum achievable per-node throughput. The analytical results obtained here from the queuing network analysis are discussed with regard to similarities and differences from the well established information-theoretic results on throughput and delay scaling laws in ad hoc networks. We also investigate the extent of deviation of delay and achievable throughput in a real world network from the analytical results presented in this paper. We perform extensive simulations and verify that the analytical results closely match the results obtained from simulations.**

I. Introduction

A multihop wireless ad hoc network is a collection of nodes that communicate with each other without any established infrastructure or centralized control. The transmission power of a node is limited, thus the packets may have to be forwarded by several intermediate nodes before they reach their destinations. So each node may be a source, destination and relay. The wireless medium is shared and scarce, therefore ad hoc networks require an efficient MAC protocol [1]. Since ad hoc networks lack infrastructure and centralized control, the MAC protocols for ad hoc networks should be distributed, and thus random access MAC protocols, e.g. MACA [6] and MACAW [1], have been proposed. The delay and throughput of wireless ad hoc networks depend on the number of nodes, transmission range of the nodes, traffic pattern and the behavior of the MAC protocol [5], [4].

In this paper we investigate how the end-to-end delay and throughput in a random access MAC based multihop wireless network with stationary nodes depend on the number of nodes, transmission range and traffic pattern. We propose a queuing network model. The queuing network model proposed in this paper is unique in that it allows us to derive closed form expressions for the average end-to-end delay and maximum achievable throughput. The packet delay is defined as the time taken by a packet to reach its destination node after it is generated. The average end-to-end delay is the expectation of the packet delay over all packets and all possible network topologies. Our analysis takes into account the queuing delays at source and intermediate nodes. The packets are assumed to have a fixed size and random arrival process. Moreover we also characterize how the average end-to-end delay and maximum achievable throughput vary with the degree of locality of traffic. The primary purpose of this study is not to accurately predict the performance of standard protocols (like IEEE 802.11 MAC) but to gain insights into the queuing delays and achievable throughput in random access multihop wireless ad hoc networks.

Several studies have focused on finding the maximum achievable throughput and characterizing capacity-delay tradeoffs in wireless ad hoc networks [5], [8], [4], [9]. In [5] it is shown that for a wireless network with n stationary nodes, the per-node capacity scales as $\Theta(W/\sqrt{n \log n})$. In [8], the authors use simulations in order to study the dependence of per-node capacity on IEEE 802.11 MAC interactions and traffic pattern for various topologies like single cell, chain, uniform lattice and random network. An estimate of the expressions for one-hop capacity and upper bound of per-node throughput is obtained using the simulation results.

In [4], the authors characterize the delay-throughput tradeoffs in wireless networks with stationary and mobile nodes. It is shown that for a network with stationary nodes, the average delay and throughput are related by $D(n) = \Theta(nT(n))$, where $D(n)$ and $T(n)$ are the average end-to-end delay and throughput respectively. However the delay is defined as the time taken by a packet to reach the destination *after it has left the source*. Also, according to the network model, the packet size scales with throughput. Under these assumptions the delay is simply proportional to the average number of hops between a source destination pair. *i.e. in their model, there is no delay due to queuing*. If, more realistically, the packet size is assumed to be constant and the delay is defined as time taken by a packet to reach the destination *after its arrival/generation at the source*, there would be queuing delays at the source and intermediate nodes.

Several recent studies have proposed queuing models for performance evaluation of the IEEE 802.11 MAC. A finite queuing model is proposed and used in [14] for evaluating the packet blocking probability and MAC queuing delays in a Basic Service Set with N nodes. A queuing model for performance evaluation of IEEE 802.11 MAC based WLAN in the presence of HTTP traffic is proposed in [10]. In [11] the service time of a node, in IEEE 802.11 MAC based wireless ad hoc network, is modeled as a Markov modulated general arrival process. The resulting M/MMGI/1/K queuing model is used for delay analysis over a *single hop* in the network. An analytical model for evaluating closed form expression for the average queuing delay over *a single hop* in IEEE 802.11 based wireless networks is presented in [13]. In [12], the authors use a queuing theoretic approach in order to calculate the mean packet delay, maximum throughput and collision probability for an elementary four node network with hidden nodes and extend the results to *linear* wireless networks. It is worth noting that none of the prior works [14], [10], [11], [13], [12] extends to a general *two dimensional wireless network*.

We propose a detailed analytic model for multihop wireless ad hoc networks based on open G/G/1 queuing networks. We first evaluate the mean and second moment of service time over a single hop by taking into account the back off and collision avoidance mechanisms of IEEE 802.11 MAC. We then use the diffusion approximation for solving open queuing networks in order to derive closed form expression for the average end-to-end packet delay. Using the average service time of the nodes we obtain an expression for the achievable throughput. We present detailed discussions on how the maximum achievable throughput obtained from our queuing analysis compares with the per-node information theoretic capacity of Gupta-Kumar's model. The analytical results are verified against extensive simulations and numerical computations.

The rest of the paper is organized as follows. In Section II we briefly describe the well known diffusion approximation for solving open G/G/1 queuing networks. Detailed description of the network model is presented in Section III. The delay and throughput analysis of multihop wireless networks is presented in Section IV. Comments and discussions on the analytical results is presented in Section V. The comparison of the analytical and simulation results is presented in Section VI. Finally we present concluding remarks in Section VII.

II. Diffusion Approximation Method

The diffusion approximation [3] can be used for solving an open G/G/1 queuing network provided that all the nodes in the network are single server with first-come first-serve (FCFS) service strategy. The advantage of using the diffusion approximation in this work is that it allows us to obtain closed form expressions for the average end-to-end delay.

In this section we briefly describe how the diffusion approximation is used to solve an open G/G/1 queuing network. (Please see [3] for details). Suppose we have an open queuing network with n service stations, numbered from 1 to n. The external arrival of a packet is a renewal process with an average inter-arrival time of $1/\lambda_e$ and the coefficient of variance of inter-arrival time equals c_A^2. The mean and coefficient of variance of the service time at a station i are denoted by $1/\mu_i$ and c_{Bi}^2, respectively.

The *visit ratio* of a station in a queuing network is defined as the average number of times a packet is forwarded by (i.e. visits) the station. The visit ratio of station i, denoted by e_i, is given by

$$e_i = p_{0i}(n) + \sum_{j=1}^{j=n} p_{ji}(n) \cdot e_j \tag{1}$$

where p_{0i} denotes the probability that a packet enters the queuing network from station i and p_{ji} denotes the the probability that a packet is relayed to station i after completing its service at station j.

There are two sources of packet arrivals at a station: the packets that are generated at the station and the packets that are forwarded to the station by other neighboring stations. The resulting arrival rate is termed the *effective arrival rate* at a station. The effective arrival rate at station i, denoted by λ_i, is given by

$$\lambda_i = \lambda_e e_i \tag{2}$$

The *utilization factor* of station i, denoted by ρ_i, is given by

$$\rho_i = \lambda_i/\mu_i \tag{3}$$

The squared coefficient of variance of the inter-arrival time at a station i, denoted by c_{Ai}^2, is approximated using

$$c_{Ai}^2 = 1 + \sum_{j=0}^{n} (c_{Bj}^2 - 1) \cdot p_{ji}^2 \cdot e_j \cdot e_i^{-1} \tag{4}$$

where $c_{B0}^2 \triangleq c_A^2$ by convention.

According to the diffusion approximation, the approximate expression for the probability that the number of packets at station i equals k, denoted by $\hat{\pi}_i(k)$, is

$$\hat{\pi}_i(k) = \begin{cases} 1-\rho_i & k = 0 \\ \rho_i(1-\hat{\rho}_i)\hat{\rho}_i^{k-1} & k > 0 \end{cases} \tag{5}$$

where

$$\hat{\rho}_i = \exp\left(-\frac{2(1-\rho_i)}{c_{Ai}^2 \cdot \rho_i + c_{Bi}^2}\right) \tag{6}$$

The mean number of packets at a station i, denoted by $\overline{K_i}$, is

$$\overline{K_i} = \rho_i/(1-\hat{\rho}_i) \tag{7}$$

III. Queuing Network Model

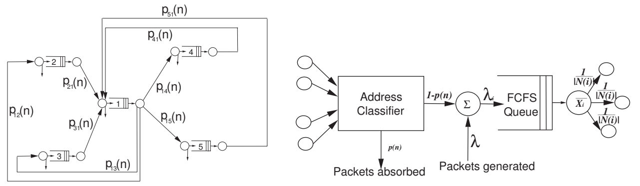
In this section we develop a queuing network model for multi-hop wireless networks and derive expressions for the parameters of the model.

A. The network model

The network consists of $n+1$ nodes, numbered 1 to $n+1$, that are distributed uniformly and independently over a torus of unit area. We assume a torus area in order to avoid complications in the analysis caused by edge effects. Each node is assumed to have an equal transmission range, denoted by $r(n)$. Let r_{ij} denote the distance between nodes i and j. Nodes i and j are said to be *neighbors* if they can directly communicate with each other, i.e. if $r_{ij} \leq r(n)$. Let $N(i)$ denote the set of nodes that are neighbors of node i. All the neighbors of a node lie on a disc of area $A(n) = \pi r(n)^2$ centered at the node. The area $A(n)$ is termed the *"communication area"* of a node. The communication area is chosen such that the network is connected which ensures that $N(i) \neq \phi \ \forall \ i$. The transmission rate of each node equals W bits/second.

We use a special case of the Protocol Model of interference described in [5]. If node i transmits to node j then the transmission will be successful only if (i) $r_{ij} \leq r(n)$ and (ii) $r_{kj} > r(n)$ for every node $k \neq i, j$ that transmits simultaneously with node i. In other words, node i can successfully transmit a packet to node j only if i is a neighbor of j and no other neighbor of j is transmitting concurrently with i. (This is equivalent to setting Δ = 0 in the Protocol Model in [5]).

The traffic model for the network may be described as follows. Each node in the network could be a source, destination and/or relay of packets. Each node generates packets with rate λ packets/sec. The delay analysis is possible for any packet generation process as long as the mean and SCV of packet inter arrival time is known. For the sake of simplicity, we assume in our model that the packet generation process at each node is an i.i.d. Poisson process. The size of each packet is constant and equals L bits. When a node receives a packet from any of its



(a) Representation of multihop wireless ad hoc network as a queuing network.

(b) Representation of a node of multihop wireless ad hoc network as a station in the queuing network.

Fig. 1. Queuing network model for multihop wireless ad hoc network.

neighbors, it either forwards the packet to its neighbors with probability $(1-p(n))$ or absorbs the packet with probability $p(n)$. The probability $p(n)$ is referred to as "absorption probability". In other words, the absorption probability is the probability that a node is the destination of a packet given that the node has received the packet from its neighbors. When a node forwards a packet, each of its neighbors is equally likely to receive the packet. The advantage of such a model is that we can control the locality of the traffic by varying the parameter $p(n)$. The traffic is highly localized if $p(n)$ is large while a small value of $p(n)$ implies unlocalized traffic. This helps in characterizing the effect of the locality of the traffic on the average delay and maximum achievable throughput.

We assume that each node in the network has infinite buffers which means that no packets are dropped in the network. The packets are served by the nodes on first come first serve basis.

Multihop wireless ad hoc networks can be modeled as a queuing network as shown in Figure 1(a). The stations of the queuing network correspond to the nodes of the wireless network. The forwarding probabilities in the queuing network, denoted by p_{ij}, correspond to the probability that a packet that is transmitted by node i enters node j's queue. Figure 1(b) shows a representation of a node in the ad hoc network as a station in the queuing network.

The end-to-end delay for a packet equals the sum of the queuing and transmission delays at the source and at the intermediate nodes. We will use the queuing network model shown in Figures 1(a) and 1(b) in order to mathematically analyze the end-to-end delay.

B. Parameters of the queuing network model

In this section we derive expressions for the parameters of the queuing network model of multihop wireless networks. The detailed proofs of the Lemmas stated below may be found in [2].

Lemma 1: The expected[1] probability that a packet is forwarded from node i to node j, denoted by $\overline{p_{ij}}(n)$, is

$$\overline{p_{ij}}(n) = \begin{cases} \frac{1-p(n)}{n}(1-(1-A(n))^n) & i \neq j \\ 0 & i = j \end{cases} \tag{8}$$

Lemma 2: The expected visit ratio of node i, denoted by $\overline{e_i}$, is given by

$$\overline{e_i} = \frac{1}{(n+1)p(n)} \;\; \forall \; i \tag{9}$$

when $(1-A(n))^n \to 0$, i.e. when $A(n)$ is chosen such that the network is connected with high probability.

Lemmas 1 and 2 (equations 8-9) indicate that the nodes visit ratio and the forwarding probabilities averaged over all possible instances of the topologies are similar to the visit ratios and forwarding probabilities of an average topology where each node has a number of neighbors equal to the average case. Thus, as a result of these two lemmas, one may derive the remaining set of model parameters (effective packet arrival rate and number of packets traversed) by considering the average case topology. Applying these results in the diffusion model will provide expressions for the average end to end delay, defined as the expectation of the packet delay over all packets and all possible network topology instances.

Lemma 3: The effective packet arrival rate at a node i, denoted by λ_i, is

$$\lambda_i = \lambda / p(n) \tag{10}$$

Lemma 4: The expected number of hops traversed by a packet between its source and destination, denoted by $\overline{s}$, is

$$\overline{s} = 1/p(n) \tag{11}$$

Notice that the average queuing delay depends on the service time distribution of the nodes, which in turn depends on the MAC protocol used by the nodes. This is the focus of the following section.

IV. Queuing Analysis

In this section we first present a model for a random access MAC that accounts for the back off and collision avoidance mechanisms of IEEE 802.11 MAC. We then present the delay analysis of multihop wireless ad hoc networks by integrating the MAC model with the queuing network model developed in Section III.

A. The MAC model

1) Interfering neighbors: Two nodes are said to be *interfering neighbors* if they lie within a distance of $2r(n)$ of each other. The transmission of a node would be successful if none of the interfering neighbors of the node transmits concurrently. Also two nodes may successfully transmit at the same time if they are not interfering neighbors of each other. The definition of interfering neighbors is similar to the definition given in [5].

2) The random access MAC model: Before transmitting each packet the nodes count down a random back-off timer. The duration of the timer is exponentially distributed with mean $1/\xi$. As in IEEE 802.11, the timer of a node freezes each time an interfering neighbor starts transmitting a packet. When the back off timer of a node expires, it starts transmitting the packet and the back off timers of all its interfering neighbors are immediately frozen. The timers of the interfering neighbors are resumed as soon as the transmission of the packet is complete. The time required to transmit a packet from a node to its neighbor is $L/W+T_o$, where T_o is the time required for the exchange of RTS, CTS and ACK packets. We assume that T_o is negligible compared to L/W, so in our analysis we assume that the time required to transmit a packet is L/W. The model is mathematically tractable and at the same time captures the behavior of IEEE 802.11 MAC protocol.

B. Delay analysis

With the help of the following three Lemmas we determine the mean and second moments of the service time of nodes using the

[1] All expected values in this paper are the expectation over all packets and all possible network topologies.

random access MAC model. The proofs of the Lemmas may be found in [2]. We then present the result for end-to-end delay in multihop wireless networks.

Lemma 5: Let H_i denote the number of interfering neighbors of a node i. Then

$$E[H_i] = 4nA(n) \tag{12}$$

$$E[H_i^2] = 4nA(n)(1 + 4(n-1)A(n)) \tag{13}$$

where $A(n) = \pi \cdot r(n)^2$.

Lemma 6: Let M_i denote the number of interfering neighbors of a node i that have at least one packet to transmit. Then under steady state,

$$E[M_i] = \rho 4nA(n) \tag{14}$$

$$E[M_i^2] = \rho^2 \cdot 4nA(n)(1 + 4(n-1)A(n)) + (1-\rho)\rho 4nA(n) \tag{15}$$

where ρ is the utilization factor of the nodes.

Lemma 7: Let Z_i denote the number of times the timer of a node i is frozen before its expiration. Then

$$E[Z_i] = 4 \cdot \rho nA(n) \tag{16}$$

Theorem 1: Let $\overline{X_i}$ and $\overline{X_i^2}$ denote the mean and second moment of service time of a packet by a node i. Then

$$\overline{X_i} = E[X_i] = \frac{\frac{1}{\xi} + \frac{L}{W}}{1 - 4nA(n)\lambda_i \frac{L}{W}} \tag{17}$$

$$\overline{X_i^2} = E[X_i^2](1 + 3\overline{m} + 2\overline{m^2})\frac{L^2}{W^2} + 2(2\overline{m}+1)\frac{L}{W}\frac{1}{\xi} + \frac{2}{\xi^2} \tag{18}$$

where $\overline{m} = E[M_i]$ (in (14)) and $\overline{m^2} = E[M_i^2]$ (in (15)).

Proof: The time taken by node i to serve a packet, denoted by X_i, is the sum of three terms: (i) the duration of the random back off timer (t_i), (ii) the duration for which the timer remains frozen ($Z_i L/W$), and (iii) the transmission time (L/W). Thus

$$X_i = t_i + Z_i \frac{L}{W} + \frac{L}{W} \tag{19}$$

Taking expectation of both sides we get,

$$E[X_i] = E[t_i] + E[Z_i] \cdot \frac{L}{W} + \frac{L}{W} = \frac{1}{\xi} + 4\rho nA(n)\frac{L}{W} + \frac{L}{W} \tag{20}$$

Substituting $\rho = \lambda_i \overline{X_i}$ and by rearranging, we get (17).

The proof of (18) may be found in [2]. ■

Corollary 1: The standard deviation of service time of a node i, denoted by $\sigma^2_{X_i}$, is given by

$$\sigma^2_{X_i} = \frac{L^2}{W^2}(\overline{m} + \overline{m^2} + \sigma^2_m) + 2(2\overline{m}+1)\frac{L}{W}\frac{1}{\xi} + \frac{1}{\xi^2} \tag{21}$$

where $\sigma^2_m = \overline{m^2} - \overline{m}^2$.

The squared coefficient of variance of the service time at a node i, denoted by c^2_{Bi} is given by $\sigma^2_{X_i}/\overline{X_i}^2$. Using (4), the squared coefficient of variance of the inter arrival time at a node i, denoted by c^2_{Ai}, is given by

$$c^2_{Ai} = 1 + \sum_{j=1, j\neq i}^{n+1} (c^2_{Bi} - 1)\frac{1-p(n)}{n} = 1 + (c^2_{Bi} - 1)(1-p(n))$$

With the knowledge of c^2_{Ai}, c^2_{Bi} and ρ, we can find the parameter $\hat{\rho}$ as given in (6).

Theorem 2: For the random access MAC model the average end-to-end delay in a multihop wireless network, denoted by $D(n)$, is given by

$$D(n) = \frac{\rho_i}{\lambda \cdot (1-\hat{\rho})} \tag{22}$$

Proof: Let $\overline{D_i}$ denote the average delay at a node i. According to Little's Law, $\overline{D_i} = \overline{K_i}/\lambda_i$, where $\overline{K_i}$ is the average number of packets in the queue of node i. Substituting $\overline{K_i}$ from (7) we get

$$\overline{D_i} = \overline{K_i}/\lambda_i = \rho/(\lambda_i(1-\hat{\rho}))$$

By Lemmas 1 and 2 the average delay at each node is the same. Thus the average end-to-end delay equals the product of the average number of hops traversed by a packet and the average delay at each node. Hence $D(n) = \overline{s} \cdot \overline{D_i}$ which leads to (22). ■

C. Maximum achievable throughput

The *maximum achievable throughput*, denoted by λ_{max}, is defined as the maximum packet arrival rate at each node for which the average end-to-end delay remains finite. If the packet arrival rate exceeds λ_{max}, the delay will become unbounded. The following corollary, that follows from Theorem 1, yields a relationship between the maximum achievable throughput and the network parameters.

Corollary 2: For a multihop wireless network the maximum achievable throughput λ_{max} is

$$\lambda_{max} = \frac{p(n)}{\frac{1}{\xi} + \frac{L}{W} + 4nA(n)\frac{L}{W}} \tag{23}$$

Also from (23), $\lambda_{max} = o\left(1/\overline{s}nA(n)\right)$.

Corollary 2 directly follows from $\rho_i = \lambda_i \overline{X_i} < 1$.

V. Discussions

In this section we discuss the implications of the analytical results derived in the last section. We first present a brief intuitive interpretation of the mean service time followed by a discussion on the maximum achievable throughput evaluated in Corollary 2 and how it compares with well known information theoretic results [5]. We also discuss how our analytical results vary from those obtained for a more pragmatic network model.

A. Interpretation of mean service time

We now present a mathematically non-rigorous, but intuitive, derivation of mean service time of a node for the random access MAC model. This derivation further elucidates the result on service time. Consider a hypothetical *two node network* where one of the nodes transmits packets to the other node. Both nodes use the random access MAC model described in IV-A. In this scenario there is no contention for the channel and the average service time of the transmitter would be $\frac{1}{\xi} + \frac{L}{W}$. We refer to $\frac{1}{\xi} + \frac{L}{W}$ as the *uncontended service time*.

Now consider a node (say node 0) with m interfering neighbors, numbered 1 through m. The node and its interfering neighbors use the random access MAC model for collision avoidance. Packets of size L_j bits arrive at a rate of α_j packets/second at neighbor j. From the point of view of node 0, the channel is available when no other interfering neighbor is transmitting. Under steady state, the fraction of time that the channel is available to node 0 is $1 - \sum_{k=1}^{m} \alpha_k (L_k/W)$. So the service time of node 0 would be the uncontended service time scaled by the fraction of time the channel is available to node 0. Hence the service time of node 0 equals $\frac{1/\xi + L_0/W}{1 - \sum_{k=1}^{m} \alpha_k (L_k/W)}$. We refer to $\sum_{k=1}^{m} \alpha_k (L_k/W)$ as the *contention term*.

In a multihop wireless network, m is analogous to the number of interfering neighbors and $\alpha_j = \lambda_i, L_j = L \ \forall\ j$. The expected value of the contention term (or the fraction of time the channel is not available to a node) is $4nA(n)\lambda_i \frac{L}{W}$ and therefore the service time of a node equals $\frac{1/\xi + L/W}{1 - 4nA(n)\lambda_i (L/W)}$.

B. Implications of maximum achievable throughput result

The result of Corollary 2 re-emphasizes the importance of carefully choosing the transmission ranges of nodes. λ_{max} increases with decrease in $r(n)$. However if $r(n)$ is too small then the network would become disconnected. According to [7], the network is asymptotically connected for $r(n) = \omega(\sqrt{\log n/n})$. So for a connected network $A(n) = \omega(\log n/n)$ and $\lambda_{max} = o(\frac{p(n)}{c+4\log(n)(L/W)})$.

Another interesting observation is the dependence of λ_{max} on the traffic pattern. λ_{max} is directly proportional to $p(n)$. From (11), the expected number of hops traversed by a packet equals $1/p(n)$. Thus another way of interpreting the result is that λ_{max} is inversely proportional to the expected number of hops between a source-destination pair.

We further investigate how our result on the maximum achievable throughput compares with the result by Gupta-Kumar on throughput capacity. According to the Gupta-Kumar model, the nodes are distributed uniformly and independently over a sphere of unit surface area and each source chooses a random destination. Therefore the expected distance between a source and the corresponding destination equals the expected distance between two points uniformly and independently distributed on a sphere. Thus the expected distance between a source destination pair in Gupta-Kumar's model is a constant (i.e. does not vary with n), say $\overline{s_{GK}}$. The transmission range in their model is $\omega(\sqrt{\log n/n})$. Thus the expected number of hops between a source-destination pair in Gupta-Kumar model is $o(\sqrt{n/\log n})$.

In order to compare our results with Gupta-Kumar's results we choose our model parameters such that we have comparable average number of hops between a source-destination pair and comparable transmission range. In our model if we choose $p(n) = \sqrt{\log n/n}$, then the expected number of hops between a source and destination node is $\overline{s} = 1/p(n) = \sqrt{n/\log n}$, which is comparable to the Gupta-Kumar model. Also $r(n) = \sqrt{\log n/n}$ or $A(n) = \pi \log n/n$ makes the transmission range of our model comparable to that of the Gupta-Kumar model. So for the model parameters that are comparable to the Gupta-Kumar model, the maximum achievable throughput is

$$\lambda_{max} = \frac{\frac{1}{4\pi}\frac{W}{\sqrt{n\log n}L}}{1+\frac{c}{4\pi\log n(L/W)}} \tag{24}$$

or $\lambda_{max} = o(W/\sqrt{n\log n})$.

The above discussion implies that for the similar values of parameters of the network model we get a bound similar to the Gupta-Kumar's bound on throughput capacity, but for our model the bound is not achievable. The reason for the bound not being achievable is that in our model we consider a random access MAC protocol rather than a perfect deterministic scheduling. Thus the bound is not achievable because some amount of channel capacity is wasted by the nodes during contention for the channel.

C. Comparison with delay and throughput in real networks

The analytical model in this paper is kept reasonably simple so that it is possible to obtain closed form expressions for delay and throughput. In particular our MAC model does not take into account packet collisions and our routing model is random walk of packets over the network. Thus our model deviates from the real world scenarios where the packets collide due to random access MAC and the packets are routed along fixed paths dictated by routing protocols. In this subsection we discuss how much the delay and maximum achievable throughput in real world networks deviate from our analytical results.

1) Effect of packet collisions: Consider a more practical MAC model where a node transmits as soon as its transmit timer expires and the interfering neighbors freeze their timers only when they sense the transmission. For such a MAC, the transmission of node i may collide with the transmission of an interfering neighbor if the difference between the time instances when the transmit timers of node i and that of the interfering neighbor expire is less than the propagation delay between the nodes. Let d denote the propagation delay between node i and its interfering neighbor that has a packet to send, then the probability that the transmission of i does not collide with that of the interfering neighbor equals $e^{-2\xi d}$. Since the interfering neighbors are located within two hops of node i, $d \leq 2r/c = \delta$, where c is velocity of electromagnetic waves. Thus the probability that the transmission of node i does not collide with an interfering node's transmission is greater than $e^{-2\xi\delta}$. So if node i has I interfering neighbors, then the probability that a transmission is a success is bounded by

$$P[\text{Success}] \geq e^{-2\xi\delta I} \tag{25}$$

Let P_s denote the expected probability of success, averaged over all possible topologies, then

$$P_s \geq \left(1 - \left(1 - e^{-2\xi\delta}\right) 4A(n)\right)^n = P_s^{(L)} \tag{26}$$

The expected number of times a node transmits a packet before it is received successfully by its neighbor equals $1/P_s$. It is easy to see that the RHS of eqn. (20) is scaled by a factor of $1/P_s$ and the mean service time may be evaluated by rearranging the resulting equation. So for the more practical MAC model, that allows packet collisions, the mean service time is bounded by

$$\frac{\frac{1}{\xi}+\frac{L}{W}}{1-4nA(n)\lambda_i L/W} \leq \overline{X_i} \leq \frac{\frac{1}{\xi}+\frac{L}{W}}{P_s^{(L)} - 4nA(n)\lambda_i L/W} \tag{27}$$

The maximum achievable throughput, evaluated using $\lambda_i\overline{X_i} < 1$, is bounded by

$$\lambda_{max}^{(L)} = \frac{P_s^{(L)} p(n)}{\frac{1}{\xi}+\frac{L}{W}+4nA(n)\frac{L}{W}} \leq \lambda_{max} \leq \frac{p(n)}{\frac{1}{\xi}+\frac{L}{W}+4nA(n)\frac{L}{W}} = \lambda_{max}^{(U)} \tag{28}$$

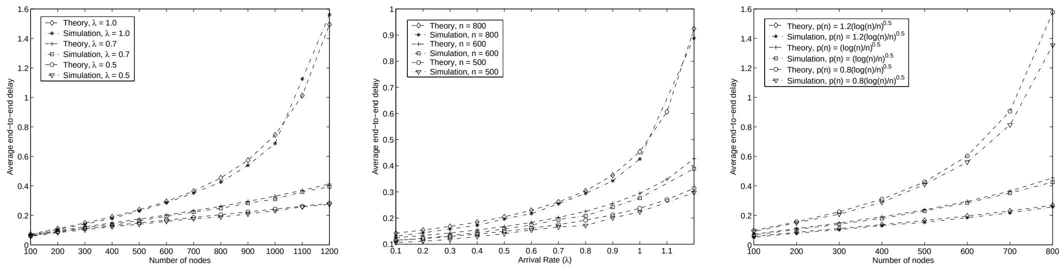
The dependence of $\lambda_{max}^{(L)}$, the lower bound of λ_{max}, on the rate of transmit timer, ξ, is particularly interesting. As ξ increases, both $P_s^{(L)}$ and $1/\xi$ terms in the denominator decrease. Thus there is a tradeoff in choosing the rate of the transmit timer - a high ξ leads to lower waiting time before transmission but leads to higher probability of packet collision. Let $\xi^\star$ be the optimal value of ξ that maximizes the lower bound of λ_{max}. Equating $d\lambda_{max}^{(L)}/d\xi$ to 0 yields that $\xi^\star$ satisfies the following relation

$$\frac{(b(n)\xi^{\star 2}+\xi^\star)e^{-2\xi^\star\delta}}{(1-4A(n)(1-e^{-2\xi^\star\delta}))} = \frac{1}{8nA(n)\delta} \tag{29}$$

where $b(n) = L/W + 4nA(n)L/W$. Closed form expression for $\xi^\star$ cannot be evaluated from the above relation. However by approximating $e^{-2\xi^\star\delta} \approx 1$ (high probability of success) and solving the resulting quadratic equation we get

$$\xi^\star \approx \frac{1}{2L/W}\frac{1}{1+4nA(n)}\left(\sqrt{1+\frac{(1+4nA(n))L}{2nA(n)W\delta}}-1\right) \tag{30}$$

As expected, $\xi^\star$ decreases with increase in the expected number of interfering neighbors, packet transmission time and propagation delay.



(a) Average end-to-end delay vs. number of nodes.
(b) Average end-to-end delay vs. arrival rate.
(c) Average end-to-end delay vs. number of nodes.

Fig. 2. Comparison of the analytical results with simulation results.

2) Effect of deterministic routing: In the routing model used in this paper, a node forwards a packet to any of its neighbor with equal probability which spreads the traffic evenly throughout the network. On the other hand, a deterministic routing protocol routes each packet belonging to a particular flow (typically identified by a source-destination pair) along a deterministic path, determined using some goodness metric. This may lead to the unfortunate situation where large number of flows pass through a few nodes that are perceived by the protocol to have good paths to many destinations. This leads to creation of routing bottlenecks leading to large queuing delays at intermediate nodes and higher end-to-end delays. (Several protocols have been designed to particulary avoid this by routing around congested areas and hence achieving load balancing.) In this case, (22) can be viewed as a lower bound on the average end-to-end delay in networks with deterministic routing.

VI. SIMULATIONS

In this section we compare the simulation results with the analytical results. The aim of the comparison is to verify the validity of the assumptions made in our analysis and the accuracy of the diffusion approximation method as applied in modeling stationary multihop ad hoc networks.

The simulation setting is the following. The network topology for the simulations consists of n nodes scattered randomly over a torus of unit surface area. Each node can communicate with nodes within a distance $r(n) = \sqrt{\log n/n}$. The random access MAC protocol used by the nodes is the same as described in $IV - B$. Each node produces packets of size $L = 1$ Kbits at the rate of λ packets/sec. The transmission rate of each node is $W = 10^6$ bits/sec. The probabilistic routing described in Section III is used for the simulations. The average delay for a particular topology is obtained by averaging the end-to-end delay of all packets produced during the simulation. In order to average out the effect of topology, each simulation is repeated over several topologies. The average end-to-end delay is obtained by averaging the delay over all topologies.

Figure 2(a) shows how the average end to end delay, as obtained from the simulations, varies with the number of nodes for $\lambda = 0.5$, 0.7 and 1.0 with $p(n) = \sqrt{\log n/n}$. Figure 2(b) shows how the average end to end delay varies with the arrival rate (λ) for $n = 500$, 600 and 800 with $p(n) = \sqrt{\log n/n}$. Figure 2(c) shows how the average end to end delay varies with the number of nodes for various values of absorption probability with $\lambda = 1$ packets/sec. The theoretical values of the average end-to-end delay as obtained from the analytical results are plotted alongside the simulation results in Figures 2(a), 2(b) and 2(c). It is observed that the simulation results agree closely with the theoretical values.

VII. CONCLUSION AND FUTURE WORK

The characterization of capacity and delay in ad hoc networks has been the focus of considerable research. However capacity and delay of networks based on random access MAC, like IEEE 802.11, have not been substantially studied. In this paper we presented delay analysis of random access MAC multihop wireless ad hoc networks. We derived closed form expressions for the average end-to-end delay and maximum achievable throughput. We showed that, for comparable network parameters, the upper bound on maximum achievable throughput is of the same order as the Gupta-Kumar's bound. However for the random access MAC the bound is not achievable. The analytical results are verified using simulations.

The results and framework presented in this paper leads to several venues for future research. Our current directions include the delay analysis and characterization of the maximum achievable throughput for hierarchical networks, many to one communication scenarios, wireless networks with sleeping nodes and wireless networks with other medium access control algorithms.

REFERENCES

[1] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. MACAW: A media access protocol for wireless LANs. In *SIGCOMM*, pages 212–225. ACM Press, 1994.

[2] N. Bisnik and A. Abouzeid. Queuing network models for delay analysis of multihop wireless ad hoc networks. Technical report, ECSE, RPI. `http://www.ecse.rpi.edu/homepages/abouzeid/qmodels.pdf`.

[3] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi. *Queuing Networks and Markov Chains*, chapter 10, pages 423–430. John Wiley and Sons, 1998.

[4] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proceedings of IEEE INFOCOM*. IEEE, March 2004.

[5] P. Gupta and P. R. Kumar. Capacity of wireless networks. *IEEE Trans. on Information Theory*, pages 388–404, March 2000.

[6] P. Karn. MACA: a new channel access methos for packet radio. In *Proceedings of the 9th Comouter Networking Conference*, pages 134–140, September 1990.

[7] A. Kumar, D. Manjunath, and J. Kuri. *Communication Networking An Analytical Approach*, chapter 8, pages 456–476. Morgan Kaufman Publishers, 2004.

[8] J. Li, C. Blake, D. S. D. Couto, H. I. Lee, and R. Morris. Capacity of ad hoc wireless networks. In *MobiCom '01*, pages 61–69, New York, NY, USA, 2001. ACM Press.

[9] X. Lin and N. B. Sheroff. On the fundamental relationship between the achievable capacity and delay in mobile wireless networks. In B. K. Szymanski and B. Yener, editors, *Advances in Pervasive Computing and Networking*, pages 17–55. Springer Science, New York, NY, 2004.

[10] D. Miorandi, A. A. Kherani, and E. Altman. A queueing model for HTTP traffic over IEEE 802.11 WLANs. In *Proceedings of 16th ITC Specialist Seminar*, August 2004.

[11] M. Ozdemir and A. B. McDonald. An M/MMGI/1/K queuing model for IEEE 802.11 ad hoc networks. In *Proceedings of PE-WASUN'05*, pages 107–111. ACM Press, 2004.

[12] S. Ray, D. Starobinski, and J. B. Carruthers. Performance of wireless networks with hidden nodes: A queuing-theoretic analysis. *To appear in Journal of Computer Communications*.

[13] O. Tickoo and B. Sikdar. A queueing model for finite load IEEE 802.11 random access MAC. In *To appear in the proceedings of IEEE ICC, Paris, France*, June 2004.

[14] G. Zeng, H. Zhu, and I. Chlamtac. A novel queueing model for 802.11 wireless LANs. In *Proceedings of WNCG Wireless Networking Symposium*, 2003.