Theoretical Throughput Analysis of IEEE802.11 MAC DCF Multi-hop Ad Hoc Network

Yue Fang and A. Bruce McDonald
Northeastern University, Department of Electrical and Computer Engineering

Abstract Previous works on throughput analysis for IEEE 802.11 protocol are mainly focus on fully connected ad hoc network. However, with the rapid growth of large scale wireless ad hoc network, performance analysis for multi-hop network has gained more and more attention from the researchers. This paper extends the previous work to analyze the "Saturation Throughput" for multi-hop network. Simulation results validate the analysis and a routing protocol based on cross-layer information is proposed with the effort to achieve the theoretical throughput limit.

I. Introduction

Wireless ad hoc network, described by [1] as "art of networking without a network", is a set of nodes cooperate with each other to provide some basic networking functions such as routing, data forwarding any time, any where without the help of any fixed infrastructure. Along with the growing interest for ad hoc networks in recent years, IEEE 802.11 MAC layer protocol, especially the Distributed Coordination Function (DCF), originally designed for wireless local area communications, is now the most commonly used MAC layer protocol used in academic and industry work regarding to wireless ad hoc networks.

The theoretically throughput of IEEE 802.11 DCF has been studied in [2] and [3]. However, in these works, only fully connected ad hoc networks are considered, that is to say, in the network all the nodes are in each other's transmission range, so there is no hidden/exposed terminal problems. Those works provide a very good starting point, while the problem is that the assumption of fully connected network is not sufficient or realistic any more with the rapid growth of large scale wireless ad hoc network, such as, sensor network. In these networks, topology size is larger, number of nodes is bigger, and besides one hop communications between direct neighbors, there will be multi-hop communication between the nodes inside the network. In this paper, throughput analysis for IEEE 802.11 DCF in a multi-hop wireless ad hoc environment is presented. Simulation results validate the analysis and a new routing protocol using cross layer information [4] is proposed.

This paper is organized as follows. In Section-II. we briefly review IEEE802.11 DCF and the Markov model for backoff window size inherent from [2]. Definitions used throughout the paper are given in Section-III. and in Section-IV. theoretical throughput analysis is presented. Network model and the parameter used with model validation are provided in Section-V. Performance evaluation of the multi-hop wireless ad hoc network is given in Section-VI. Finally, conclusions and discussion of future work are presented in Section-VII.

II. IEEE 802.11 Distributed Coordination Function

This section consists two subsections. Subsection-A. This briefly summarizes the DCF as standardized by the 802.11 protocol. Subsection-B. reviews the Markov chain model for backoff window size introduced in [2] which we will used in our analysis.

A. IEEE 802.11 DCF

IEEE 802.11 MAC DCF employs both the two way handshaking DATA/ACK scheme which is called as basic access and the optional four way handshaking *request-to-send*/*clear-to-send*, abbreviates as RTS/CTS scheme. In this paper, we are particularly interested in the RTS/CTS scheme, so only this scheme is explained in detail here, while basic access scheme can be referred to [5] and [2].

When a station has packet to transmit, it first sense the channel. If the channel is idle for a period of time which equals to distributed inter frame space, abbreviated as DIFS, the station transmits its RTS to the intended receiver. Otherwise, if the channel is sensed busy, the station will continue to sense the channel until it is idle for DIFS, then the station will generate a random backoff interval before it can actually transmit in order to minimize the probability of collision with packets transmitted by other stations (as we will see in the future, there are four different kinds of collision). In addition, a station must perform randomly backoff between consecutive packets transmission to avoid channel capture.

A typical wireless communication employing IEEE 802.11 DCF can be described as follows: The sender station initials the communication by sending a short RTS frame to the intended receiver. The RTS indicates the length of time the sender would like to transmit. Upon correct reception of the RTS, the receiving station waits for a Short Inter frame Space (SIFS) to reduce the possibility of collision, and then transmits a CTS back to the source station. The source waits another SIFS before it is permitted to transmit the DATA frame. If the DATA frame is received correctly, the receiving station acknowledges the frame after a SIFS by sending an ACK frame to the source. Other stations utilize information in the duration field of RTS, CTS and DATA frame to adjust their Network Allocation Vector (NAV), which determines how long each station will defer any pending transmissions prior to sensing the medium.

B. Markov Chain Model for Backoff Window Size

DCF adopts an exponential backoff scheme when collision happens. Backoff time is slotted and stations are allowed to transmit at the beginning of the time slot. The slot time, σ, is set equal to the time needed for any station to detect the transmission from its neighbors. At each packet transmission, the backoff timer is random uniformly chosen from range (0, w-1). Where w is called contention window, contention window size starts with the minimum value CW_{min}, and will double at each failure of transmisson/retransmission until it reach a maximum value CW_{max}, the relation between CW_{max} and CW_{min} is $CW_{max} = 2^m CW_{min}$, where m is backoff stage.

[2] proposed a Markov chain model for a single station's backoff window size. The combination of random variables backoff stage s(t) which corresponding to current contention window size and backoff counter b(t) which corresponding to the slots to be wait before the station can try to transmit again is modelled by a discrete-time Markov chain under the approximation that at each transmission attempt, each packet collides with constant and independent probability p regardless of the number of retransmission. Use the same notation as [2] [1], the Markov chain can be expressed as following:

$$\begin{cases} P\{i,k|i,k+1\} = 1 & k \in (0, W_i - 2) \quad i \in (0, m) \\ P\{0,k|i,0\} = (1-p)/W_0 & k \in (0, W_0 - 1) \quad i \in (0, m) \\ P\{i,k|i-1,0\} = p/W_i & k \in (0, W_i - 1) \quad i \in (1, m) \\ P\{m,k|m,0\} = p/W_m & k \in (0, W_m - 1) \end{cases}$$

Interested readers can refer to [2] for the detail description of this model.

III. Terminologies and Notations

Before presenting the throughput analysis, terminologies as well as the notations used later in this paper are listed in this section.

Table-1 is the list of the notations of variables used in the analysis of this paper. Notice that these variables are not independent and we will go through their relations in the rest of the paper when necessary.[2]

[1] $P\{i_1, k_1|i_0, k_0\} = P\{s(t+1) = i_1, b(t+1) = k_1, |s(t) = i_0, b(t) = k_0\}$

[2] transmission range *r* in Table-1is defined as the maximum euclidean distance between the communicating nodes so that any nodes with smaller euclidean distance can successfully communicate with each other. Both transmission power and receiving sensitivity can affect *r* .

Notations	Meaning
L	The radius of network
N	Number of nodes in the network
n_{avg}	Average number of neighbors of each node
ρ	Node density of the network
N_l	Number of links in the network
ρ_i	Link density of the network
W	CW_{min}
m	maximum backoff stage, $CW_{max} = 2^m W$
δ	propagation dealy
H	packet header, $H = PHY_{hdr} + MAC_{hdr}$
r	transmission range
$Num_{one-hop}$	number of non-overlapping one-hop set in the network at a given time

Table 1: Notation used through out the paper

Definition 1 Link
In a wireless network, when two nodes x, y are in each other's transmission range, it is said that there is a "link" between the nodes, which is represented by l(x,y)=1, otherwise l(x,y)=0. Link is not an actual entity, it represents the capability of communication between two nodes. In the scope of this paper, the link is regarded as bidirectional with same cost either direction.

Definition 2 Communications
c(x,y,t)=1: l(x,y)=1, and x,y is communicating with each other at time t .

c(x,y,t)=0: l(x,y)=1, and x,y is not communicating with each other at time t either because there is no traffic or x/y has to defer channel access because of other ongoing transmission.

Definition 3 Neighbor and non-active neighbor :
N_x: $N_x = \{y|\forall y \in V, l(x,y) = 1\}$; N_x is defined as the neighbor of node x.

$N'_x(t)$: if c(x,y,t)=1, then $N'_x(t) \cup y = N_x$, and $N'_x(t) \cap y = \phi$; $N'_x(t)$ is defined as the non_active neighbor of node x. The context of this set will vary with time.

Definition 4 Links and non-active links :
E_x: $E_x = \{l(x,y)|\forall y \in N_x\}$; E_x is defined as the set of links belongs to node x.

$E'_x(t)$: if c(x,y,t)=1, then $E'_x(t) = E_x - l(x,y)$;

$E'_x(t)$ is defined as the set of non-active links of node x. The access to these links has to be deferred at time t. The context of this set will vary with time too.

Definition 5 Ad Hoc Network
Ad Hoc network can be represented by graph **G(V,E)**, *where V is the set of nodes and E is the set of the links in the network.*

Definition 6 One-hop Set
One-hop Set O(x,y,t)is defined as the combination of the non-active neighbors of the communicating pair (x,y) at time t. The size of one-hop set S_o is defined as the total number of feasible

links inside it. At any given time, one-hop set (with regarding to link) won't overlap[3].

$$O(x, y, t) = E'x(t) \cup E'y(t)$$

The node size of one-hop set S_{o_node}*is the number of nodes inside it.*

Definition 7 Communication Set
Because all the links which are in two hops scope of the communication pair can not participate in any communication at same time[3]*, we define the communication pair and all the links affected by it to be a communication set* $C_{(x,y,t)}$*, where c(x,y,t)=1.*

$$C(x, y, t) = E_x \cup E_y \cup E_{N'_x} \cup E_{N'_y}$$

Different communication sets can overlap and

$$\cup_{\forall(x,y)} C_{(x,y,t)} = E.$$

The size of a communication set $S_c(x, y, t)$ *is defined as the total number of feasible links inside it.*

The term "feasible" is used here because communication set may overlap. Therefore, attention should be paid to avoid counting the same link twice in different communication set, for link that is already deferred by one communication can not be deferred again by *simultaneous* communication. Such as in Figure-1, hollow nodes represents nodes which are currently involved with ongoing communication, solid nodes are the one hop, while shadowed nodes are two hop neighbors. One-hop link is solid and the two hop link is dashed. Nodes (links) may have different characters in different communication set. For example, node H is one of the communication nodes in C(G,H,t) while it is also the two hops neighbor in C(O,P,t); link HK is the one hop link in C(G,H,t) and two hops link in C(O,P,t). Links between solid and hollow nodes or two hollow nodes must be deactivated while links between two shadowed nodes (not show in Figure-1) can be used for simultaneous transmission.

If $S_c(O, P, t)$ is calculated first, link OK, ON, HK, HN and MN are counted, so we have:

$$S_c(O, P, t) = 13$$

Then for $S_c(G, H, t)$, link OK, ON, HK, HN and MN should not be counted,thus we got :

$$S_c(G, H, t) = 12$$

Lemma-1 summarizes the channel access status of the links in the network.

Lemma 1 *When c(x,y)=1, then all the nodes which are two hops or less away from x, y have to defer their channel access. Or to say:*

when c(x,y)=1, then c(m,)=0,* $\forall m, m \in N'_x$ *or* $m \in N'_y$*, c(n,*)=0,* $\forall n, n \in N'_m$.

[3]In some extreme case, such as perfect scheduling, fixed data length, one-hop set may overlap, however, those cases are already exceed our interest

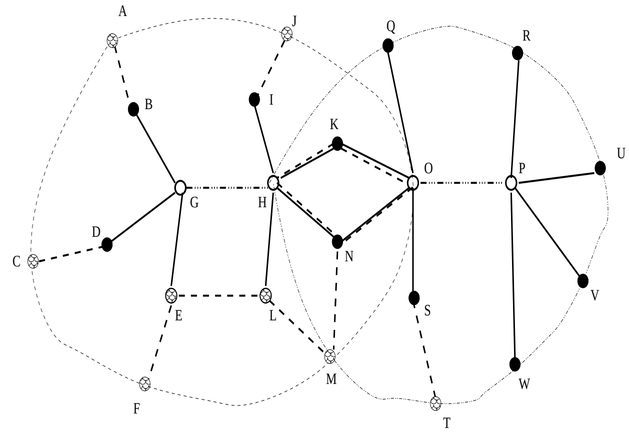

Figure 1: A example of how to calculate the size of communication set

IV. Throughput Analysis

As in [2] and [3], "saturation throughput" will be used as the fundamental performance figure in this paper. Here "saturation throughput" is the limit reached by the system throughput as the offered load increases, and represent the maximum goodput that the system can carry *in stable conditions*. In other words, for each node in the network, there will always be packets ready to be transmitted. Note that in [2] assumptions are made that there is no hidden terminals problem and all the nodes are in each other's transmission range, so at any time there will be only one valid transmission in the system, while in multi-hop system, there will be multiple simultaneous transmissions in the networks, therefore "saturation throughput" is the total goodput achieved by the network.

There are three subsections in section-IV. In subsection-A., important result from previous work [2] is repeated, serving as the start point of our work. Subsection-B. lists the preliminary result about the characteristics of multi-hop network. While in subsection-C. the detail analysis of the theoretical throughput limit of multi-hop wireless ad hoc network is provided.

A. Previous Result

In [2], assumption are made that at each transmission attempt, regardless of the number of retransmission suffered, each packet collides with constant and independent probability p. Also a Markov chain model for the backoff window size is proposed, the brief description of this model is shown in section-II.-B. Based on those two important assumptions, the probability that a station transmits in a randomly chosen slot time τ is derived as: (see details in [2])

$$\tau = \frac{2(1-2p)}{(1-2p)(W+1) + pW(1-(2p)^m)} \quad (1)$$

p in (1) is the probability that any transmitted frame will result in a collision. Note that in [2] τ is derived without any special assumption such as no hidden terminal problem/ all nodes are in each other's transmission range. However, fully connected network implies that all the nodes in the network are completely synchronized, given the synchronization condition, the time axis can to divided into a sequence of non-overlapping SLOTs (Φ)[4]. Φ is defined as the time elapsed between two successive successful transmission. Note in "saturation condition" that a node always has packet to transmit, apparently each node will be either in backoff state or in transmit/receive state, fully connected network guaranteed that the vision of the network seen from every node will be exactly same. In other words, nodes in the network are synchronized in Φ.

In multi-hop environment, generally, fully connected synchronization (at Φ level) can not be achieved any more. However, slot (σ) level synchronization can be maintained under "saturation" condition. The reason is that as previously stated, node is either in backoff or transmit/receive state, thus neighboring nodes will be synchronized by Φ, and nodes two or more hops away will be synchronized by σ through the intermediate node. In this case, (1) can also be used in "saturated" multi-hop environment without any modification. Notice here in (1), the probability of retransmission is already counted, which has no difference with "first transmission" in the information transmitted point of view.

B. Pre-work

Fundamental assumptions in the throughput analysis are:

- There are N nodes uniformly distributed in a circular space with radius L, each node will have n_{avg} neighbors in average.
- The mobility of the nodes will not change the communication set during the period of communication. Or in other words, the network is relatively stable during the transmission of data packet or control message.
- Time is slotted and slots (σ) are synchronized in the system.
- To simplify the analysis, the data packet length is fixed, this condition can be loosed using the expectation of packet length to replace the fixed data packet length in the course of derivation.
- Each packet will be received correctly if there is no collision, under this assumption, the probability that an ACK packet collided with other packets is negligible.

The network will be characterized by either N and n_{avg} or N and ρ. Because the nodes are uniformly distributed in the network, the following relation is true:

$$\rho = \frac{n_{avg}+1}{\pi r^2}$$

[4] Φ (variable) refers to time interval between two consecutive backoff time counter decrements, which is different from σ, (constant), the slot time discussed in section-B.

Next, given the N and n_{avg}, the radius of the topology L and number of links in the network N_l can be calculated as:

$$\begin{aligned} \frac{\pi r^2}{n_{avg}+1} &= \frac{\pi L^2}{N} \\ L &= \sqrt{\frac{N}{n_{avg}+1}} \cdot r \\ N_l &= \frac{N \cdot n_{avg}}{2} \end{aligned}$$

Thus the link density is :

$$\begin{aligned} \rho_l &= \frac{N_l}{\pi L^2} \\ &= \frac{n_{avg}(n_{avg}+1)}{2\pi r^2} \end{aligned}$$

Other important parameters includes the average area covered the one-hop set and communication set ($\overline{S_{onehop}}$ and $\overline{S_{comm}}$) and the expectation of the size of one-hop set and communication set $E[S_o]$ ($\overline{S_o}$) and $E[S_c]$ ($\overline{S_c}$). The values of those parameters are derived as follows:

1. Area covered by one hop set - $\overline{S_{onehop}}$
   (a) *worst case*: suppose the euclidean distance between the communicating nodes is l. then:
   $$S_{one-hop}(l) = \frac{4}{3}\pi l^2 + \frac{\sqrt{3}}{2} l^2$$
   (b) *expectation*: the expectation of the $S_{one-hop}$ can be calculated by:
   $$\overline{S_{one-hop}} = \int_0^r S_{one-hop}(l) f(l) dl$$
   Under realistic assumption that the euclidean distance between the communicating pair is uniformly distributed between (0,r) and use the area of rectangular to approximate the area of the one-hop set, we will have
   $$\begin{aligned} \overline{S_{one-hop}} &= \int_0^r \frac{1}{r} \cdot (2r+l) \cdot 2r dl \\ &= 5r^2 \end{aligned} \quad (2)$$
2. Expectation of size of one hop set - $E[S_o]$ ($\overline{S_o}$)
   Given the nodes are uniformly distributed in the topology and the average area covered by one-hop set, it is obvious that
   $$\begin{aligned} \overline{S_o} &= \rho_l \cdot \overline{S_{onehop}} \\ &= 5\rho r^2 \\ &= \frac{5 \cdot n_{avg}(n_{avg}+1)}{2\pi} \end{aligned} \quad (3)$$
   $$\begin{aligned} \overline{S_{o\_node}} &= \rho \cdot \overline{S_{onehop}} \\ &= 5\rho r^2 \\ &= \frac{5(n_{avg}+1)}{2\pi} \end{aligned} \quad (4)$$

3. Area covered by communication set - $\overline{S_{comm}}$

   (a) *typical case:*
   suppose the euclidean distance between ongoing communication (A,B) is l. Use the similar approach as $\overline{S_{one-hop}}$, S_{comm} is regarded as the sum of area covered by circles with A/B as centers with radius 2r. Use the area of rectangular to approximate the area covered by two overlapping circle. we got:

$$S_{comm}(l) \quad = \quad (4r + l) \cdot 4r$$

   Note that although not all nodes in the area of those circle areas will be the two hop neighbors of A/B, however, A bias factor $0 < f_{bias} \leq 1$ is needed to account for the nodes (links) that are not equal or less than two hops away from the communication nodes in the calculated "area" of communication set. Thus the actual area of communication set is less than the total area of 2 circles discussed above. Intuitively , f_{bias} would be related to ρ, and will decrease with the growth of ρ. In this paper, f_{bias} is set to be 0.34 based on simulation results. So we have:

$$S_{comm}(l) \quad = \quad f_{bias} \cdot (4r + l) \cdot 4r$$

   (b) *expectation:* the expectation of S_{comm} can be calculated by:

$$\begin{aligned} \overline{S_{comm}} &= f_{bias} \cdot \int_0^r S_{comm}(l) f(l) dl \\ &= f_{bias} \cdot \int_0^r \frac{1}{r} \cdot (4r + l) \cdot 4r dl \\ &= f_{bias} \cdot 18 \cdot r^2 \end{aligned} \tag{5}$$

4. Expectation of size of communication set - $\overline{S_c}$

   Given the nodes are uniformly distributed in the topology and the average area covered by communication set, it is obvious that

$$\begin{aligned} \overline{S_c} &= f_{bias} \cdot \rho_l \cdot \overline{S_{comm}} \\ &= 18 \cdot f_{bias} \rho_l r^2 \\ &= \frac{9 \cdot f_{bias} \cdot n_{avg} \cdot (n_{avg} + 1)}{\pi} \end{aligned} \tag{6}$$

Notice that the "Saturation throughput" can only be achieved when all the possible links that can transmit actually do transmit. In other words, it's an ideal upper limit that the system could achieve by carefully pick the nodes to transmit so that there would be maximum number of simultaneous transmissions in the network. The problem is whether this ideal condition is achievable and how to do that? The following algorithm gives one of the possible ways.

Algorithm

- **Step 1:** List the size of all the possible communication set in a ascending order. Basically, each link in the network will generate a possible communication set.

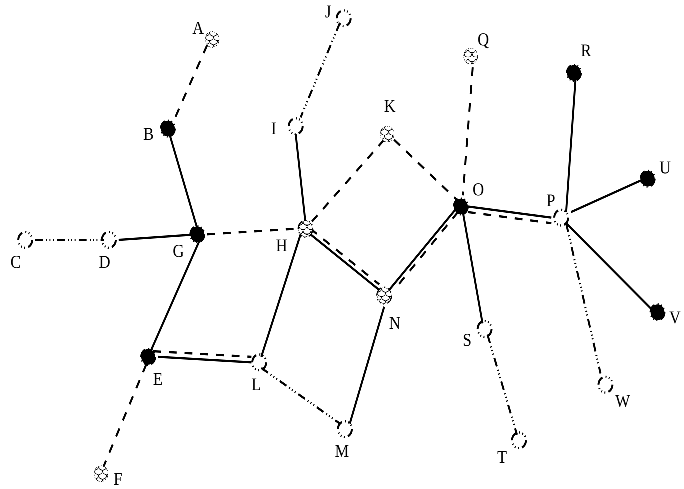

Figure 2: Flow assignation of Figure-1 using the proposed algorithm

- **Step 2:** Pick the first communication set with least size, transmission along the corresponding link can be granted, therefore there will be least links affected by the addition of the flow.
- **Step 3:** Update the size of all the remaining possible communication sets.
- **Step 4:** If two communication sets have the same size, the tie is broken by pick the communication set in which the distance between two communicating nodes is closer.
- **Step 5:** Repeat steps 1-4 until there is no possible communication set left.

The metric used in the algorithm is $\overline{S_c}$ as we defined in Definition-7. Using the algorithm, as shown in Figure-2, there will be five simultaneous transmission in the network to achieve "saturation condition".

Figure-3 depicts the ideal load assignment of a medium size ad hoc network which is in "saturation condition". In which solid links represent ongoing communication, while dashed and darkened links are the one-hop and two-hop away links respectively. Most links in Figure-3 are either solid or dashed, the percentage of darkened links (two hop links communication) will be low, otherwise there still are some links can bear more traffic at the given time. Table-2 shows the actual percentage of two hop links by running the proposed algorithm on networks with different size and density. Most of time the percentage of two hop link is less than 4%, thus the network has been partitioned into several non-overlapping one hop sets. Given the number of non-overlap one-hop sets and "saturation throughput" of a single communication set, it is straightforward that "saturation throughput" of multi-hop network will be the product of the two.

Because all the nodes are uniformly distributed in the network, the number of non-overlap one-hop sets can be easily

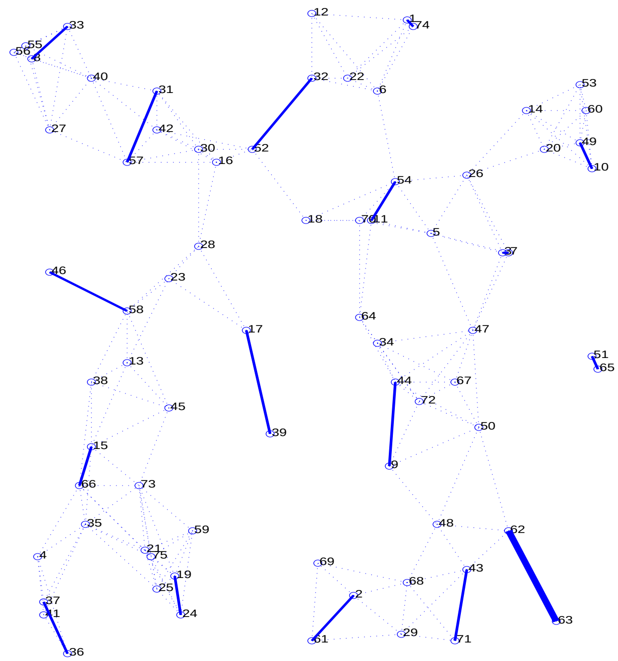

Figure 3: Flow assignation of a larger network using the proposed algorithm

n_{avg}	two hop links (%)	n_{avg}	two hop links (%)
2	3.8975	25	0
3	2.5075	30	0
4	1.9325	35	0.24
5	0.81	50	0
6	1.875	60	0.43
7	0.3225	75	0
8	0.2775	85	0
9	0.9825	100	0
10	0.025	110	0
20	0	125	0

Table 2: Percentage of blue link in the network

calculated by

$$Num_{one-hop} = \lfloor \frac{\pi L^2}{S_{one-hop}} \rfloor = \lfloor \frac{\frac{\pi r^2 \cdot N}{n_{avg}+1}}{5r^2} \rfloor = \lfloor \frac{\pi \cdot N}{5(n_{avg}+1)} \rfloor \quad (7)$$

where $\lfloor x \rfloor$ means the maximum integer which is less than x.

From (7) we can see that the number of non-overlapping one-hop set is only determined by N and n_{avg}

C. Throughput

We revisit (1) from [2] here: The probability of a station to transmit in a random selected slot is derived under the Markov chain backoff window model:

$$\tau = \frac{2(1-2p)}{(1-2p)(W+1)+pW(1-(2p)^m)}$$

where p is the conditional collision probability and W is the minimum backoff window size, $W = CW_{min}$.

In general, these probabilities depend on the conditional collision probability p, which is still unknown. To find the value of p, it is sufficient to note that the probability p that a transmitted packet encounters a collision, is the probability that, there are more than one link try to transmit at same time inside the communication set. This yields

$$p = 1-(1-\tau)^{\overline{S_o}} \quad (8)$$

Equations (1) and (8) represent a nonlinear system in the two unknowns τ and p, which can be solved using numerical techniques. It is easy to prove that this system has a unique solution. In fact, inverting (8), we obtain

$$\tau(p) = 1-(1-p)^{\frac{1}{\overline{S_o}}}$$

This is a continuous and monotone increasing function in the range $p \in (0,1)$, that starts from $\tau(0) = 0$ and grows up to $\tau(1) = 1$. Equation $\tau(p)$ defined by (1) is also continuous in the range $p \in (0,1)$. [2]

Normalized system throughput S can be expressed by:

$$S = \frac{\text{E[payload information transmitted in } \Phi \text{]}}{\text{E}[\Phi]} \text{ [2]}$$

where

$$E[\text{payload info in } \Phi] = Num_{one\_hop} P_{success} E(P)$$
$$E[\Phi] = P_{idle}\sigma + P_{tr}P_{success}T_s + P_{tr}P_{RTS}T_{RTS} + P_{tr}P_{CTS}T_{CTS}$$

$P_{success}$ is the probability that a packet is transmitted successfully, and P_{idle} is the probability all the links inside the communication set are idle, whereas P_{RTS} and P_{CTS} are the probabilities of the transmission collided during RTS time or during CTS time, respectively. σ is the duration of an empty slot time, T_s is the average time needed for a successful transmission, while T_{RTS}, T_{CTS} are the period of time during which the channel is sensed busy by *non-colliding stations* when two RTS frames collide or RTS frame collides with CTS frame.

Let node A be the node initial the communication, node B be the node which node A is communicating with and N'(x) be the number of non-active neighbors of x, the above probability and time can be expressed as:

$$
\begin{aligned}
P_{idle} &= (1-\tau)^{S_o} & (9)\\
P_{tr} &= 1 - P_{idle}\\
&= 1-(1-\tau)^{S_o} & (10)\\
P_{RTS} &= P_{collision\_at\_RTS/not\_idle}\\
&= \frac{1-(1-\tau)^{N'(B)}}{1-(1-\tau)^{S_o}}\\
&= \frac{1-(1-\tau)^{(n_{avg}-1)}}{1-(1-\tau)^{S_o}} & (11)\\
P_{CTS} &= P_{no\_collision\_at\_RTS,not\_idle}\cdot\\
&\quad P_{collision\_at\_CTS/not\_idle,no\_coll\_at\_RTS}\\
&= \frac{(1-\tau)^{(n_{avg}-1)}}{1-(1-\tau)^{S_o}}\cdot(1-(1-\tau)^{N'(A)})\\
&= \frac{(1-\tau)^{(n_{avg}-1)}}{1-(1-\tau)^{S_o}}\cdot(1-(1-\tau)^{(n_{avg}-1)})\\
&= \frac{(1-\tau)^{(n_{avg}-1)}\cdot(1-(1-\tau)^{(n_{avg}-1)})}{1-(1-\tau)^{S_o}} & (12)\\
P_{success} &= P_{not\_idle}\cdot P_{no\_collision\_at\_RTS/not\_idle}\cdot\\
&\quad P_{no\_collision\_at\_CTS/not\_idle,no\_coll\_at\_RTS}\\
&= P_{tr} - P_{tr}P_{RTS} - P_{tr}P_{CTS}\\
&= (1-(1-\tau)^{S_o})\cdot(1-(1-(1-\tau)^{n_{avg}}) -\\
&\quad (1-\tau)^{n_{avg}}\cdot(1-(1-\tau)^{n_{avg}}))\\
&= (1-(1-\tau)^{S_o})\cdot(1-(1-\tau)^{2n_{avg}}) & (13)\\
T_s &= RTS+SIFS+\delta+CTS+SIFS+\delta+\\
&\quad H+E[P]+SIFS+\delta+ACK\\
&\quad +DIFS+\delta & (14)\\
T_{RTS} &= RTS+SIFS+\delta & (15)\\
T_{CTS} &= RTS+SIFS+\delta+CTS+\\
&\quad SIFS+\sigma+DIFS & (16)
\end{aligned}
$$

P_{RTS} and P_{CTS} is derived based on the fact that collision of RTS frames happens when least one of non-active neighbors of node B transmits simultaneously with node A and collision of CTS frames happens when no collision happens at RTS frame but at least one of the non-active neighbor of node A transmits simultaneously with node B when it reply CTS frame to node A.

Note here we neglect the fact that the two or more colliding stations, before sensing the channel again, need to wait an ACK Timeout, and thus T_{RTS} and T_{CTS} for these colliding stations is greater than that considered here. Retransmission is taken into account in the expression of probability a station transmit at any given time slot Φ, so it is not included again in T_{RTS} and T_{CTS}.

V. Model Validation

The topology considered in this paper consists of a finite number of stations uniformly distributed in a circular space with radius L. The following justified assumptions are made to improve analytical tractability without loss of generality:

- The effect of propagation delay is insignificant with respect to frame transmission time and media access delay; hence, it is negligible in the analysis. This is a very realistic assumption for transmission ranges ≤ 250m.
- Only a single ad hoc network consisting of one BSS is considered. Hence, it is assumed there are no interfering BSSs—the DSSS spreading sequence is unique.
- Collisions caused by simultaneous transmission are assumed to be the most significant cause of packet corruption.
- Node mobility is assumed negligible with respect to frame transmission time. Hence, transmissions always complete before a mobile receiver moves out of range.

In order to make sure the "saturation throughput" is achieved, enough load has to be generated, so for every link, on each direction, up to half of the link capacity of load has been injected at the beginning of simulation. To validate the analytical result given in Section-IV., throughput is measured by the fraction of time spent on successfully transmission versus the whole simulation time (exclude the transit time needed for the network to reach steady state).

Network simulator (*ns2*) is used in our simulation. The system parameter for analysis and simulation are the default value used by *ns-2*, whose values are listed in Table-3.

Parameters	Value	Parameters	Value
W_{min}	32	DIFS	50 μs
W_{max}	1024	SIFS	10 μs
m	5	rate	2Mb/s
σ	20 μs	RTS	44 bytes
H	6 bytes	CTS	38 bytes
r	250 m	ACK	39 bytes

Table 3: System parameter values.

To validate the model, a single one-hop set is picked up from the network regarding to one specific communication, and measure the throughput of this specific one-hop set. The comparison between the analytical and simulation result is shown in Figure-4. To elminate the effect of routing, communication are limited between neighbors. Each simulation point is run 10 times and the average of the result from the top 2 run is used to guarantee the "saturation condition".

From Figure-4 we can see that the analysis relatively accurate for networks with low and mediate density thus the model is validated. For network with average number of neighbors up to 10 the simulation result is quite close to the analytical one, with the density increase, the discrepancy begins to grow.

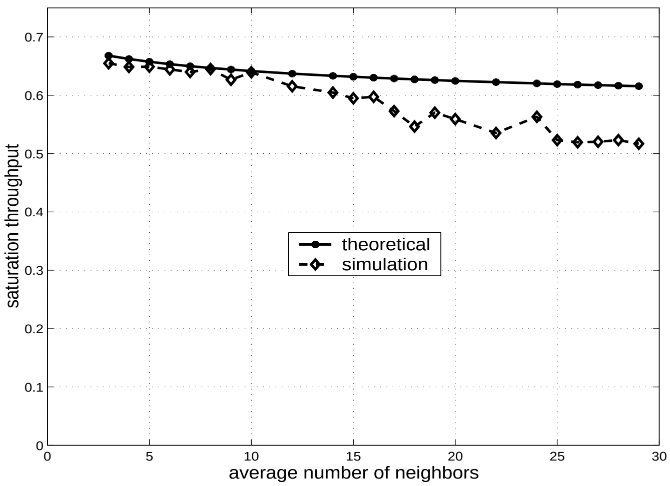

Figure 4: Model Validation

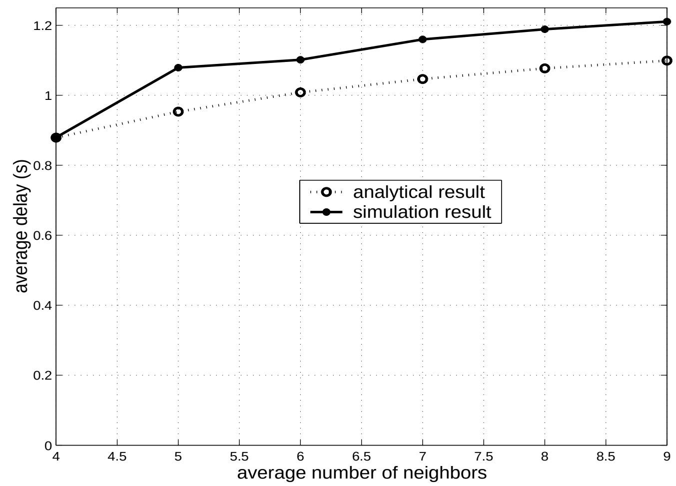

Figure 5: Average Delay experienced

The possible reason for this is as the density grows, backoff becomes more frequent, thus the probability of transmission will be less than the analysis expected.

VI. Performance Evaluation

Figure-5 depicts the delay experienced at a randomly selected node in one-hop set, the maximum queue length used in simulation is set to be 1000. Under ”saturation condition”, it is reasonable to assume the queue is always full, the delay one packet experiences will be the sum of queueing delay and the processing delay, which in turn could be approximated by the time difference between packet enter the queue and left the queue. Once a packet becomes the head of the queue, it will remain there until it is successfully received by next hop or the retransmit limit is reached, which one comes first. Based on the above assumption, theoretically the delay (D) can be approximated by:

$$D = \text{Maximum queue length} \cdot \text{Delay of the 1st packet in queue} = L_{queue} \cdot D_{packet}$$

where

$$D_{packet} \cong P_s ucc \cdot T_s + (1 - P_s ucc) \cdot \Phi$$

Figure-5 shows basically simulation result agrees with theoretical analysis and they have same growing tendency . How ever, because in reality the node will have to deal with some queuing issues even the communication is between neighbors, so the simulation result is slightly larger than theoretical result ($< 15\%$).

It has be proposed that the ”saturation throughput” depends on the both the ”saturation throughput” and $Num_{one-hop}$ and the overall network throughput should be the product of them. However, simulation result shows great discrepancy between the theory and practice in Figure-6.

The reason of the discrepancy lies in several aspects. The most important one is that the analytical ”overall throughput” is derived under the assumption of ideal condition, or, in other words, all the links that may transmit actually do transmit, in practice, this condition is seldom satisfied under current routing/scheduling scheme.

Here is the place to put in the idea of the cross-layer routing based on MAC layer delay.

VII. Conclusion

In this paper, the ”saturation throughput” of multi-hop wireless ad hoc network is studied. Simulation results shows that the analytical model for one-hop set throughput is accurate, while there is enough space left for the improvement of overall network throughput using interactive routing algorithms.

References

[1] Magnus Frodigh, Per Johansson and Peter Lasson, ”Wireless ad hoc network—The art of networking without a network”, Ericsson Review No. 4 2000.

[2] Giuseppe Bianchi ”Performance Analysis of the IEEE 802.11 Distributed Coordination Function”, IEEE Journal on Selected Areas in Communications, Vol. 18, No. 3, March 2000.

[3] Frederico Cali, Marco Conti and Enrico Gregori ”Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit”, IEEE/ACM Transactions on Networking, Vol. 8, No. 6, December 2000.

[4] Yue Fang and A.Bruce Mcdonald, ” Cross-Layer Performance Effects of Path Coupling in Wireless Ad Hoc

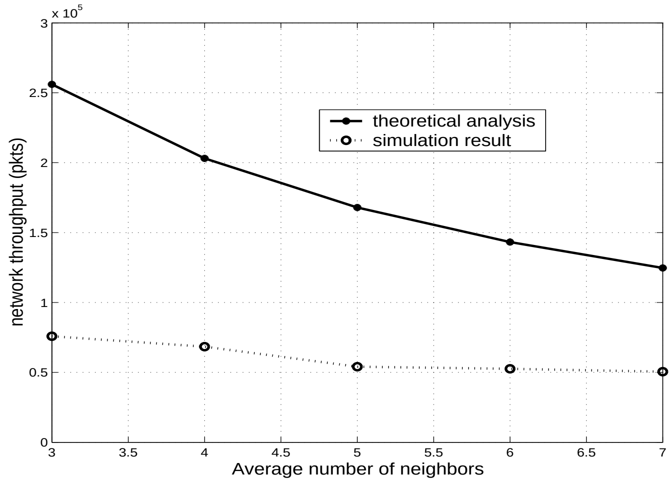

Figure 6: Overall network throughput— analytical vs. simulation result

Networks: Power and Throughput Implications of IEEE 802.11 MAC".

[5] IEEE Standard for Wireless LAN Medium Access Control (MAC) and Phyiscal Layer (PHY) Specifications. Nov, 1997, P802.11

[6] Brain P. Crow, Indra Widjaja, Jeong Geun Kim and Prescott T. Sakai, "IEEE 802.11 Wireless Local Area Networks" In *IEEE Communications Magazine*, pp.116-126, Sep, 1997