The 14th IEEE 2003 International Symposium on Personal,Indoor and Mobile Radio Communication Proceedings

Performance of Wireless LANs Based on IEEE 802.11 MAC Protocols

Hongqiang Zhai and Yuguang Fang
Department of Electrical and Computer Engineering
University of Florida. Gainesville, Florida
zhai@ecel.ufl.edu, fang@ece.ufl.edu

***Abstract*-IEEE 802.11 MAC protocol has been the standard for Wireless LANs, and also adopted in many network simulation packages for wireless multi-hop ad hoc networks. However, it is well known that, as the traffic goes up, the performance of IEEE 802.11 MAC drops dramatically in terms of delay and throughput, especially when the station approaches its saturation state. To explore the inherent problems in this protocol, it is important to characterize the probability distribution of the packet service time at the MAC layer. In this paper, by modeling the exponential backoff process as a Markov chain, we can use the signal transfer function of the generalized state transition diagram to derive the probability distribution of the MAC layer service time. We then present the discrete probability distribution for MAC layer packet service time, which is shown to accurately match the simulation data from network simulations. Finally, based on the probability model for the MAC layer service time, we are able to analyze a few performance metrics of the wireless LAN and give better explanation to the performance degradation in delay and throughput at various traffic load.**

Keywords-Performance Analysis; 802.11 MAC; Queue Dynamics

I. INTRODUCTION

The Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol, i.e., IEEE 802.11 MAC protocol, has been proposed as the standard protocol for wireless local area networks (LANs), which is also widely implemented in wireless testbeds and simulation packages for wireless multi-hop ad hoc networks.

However, there are many problems encountered in the upper protocol layers in IEEE 802.11 wireless networks. The packet delay greatly increases when there are serious collisions due to the heavy traffic. Packets may be dropped either by the buffer overflow or by the MAC layer contentions. Such packet losses may affect high layer networking schemes such as the TCP window adaptation and networking routing maintenance. The routing simulations [1] [2] over ad hoc networks indicate that network capacity is poorly utilized in terms of throughput and packet delay when the 802.11 MAC protocol is integrated with routing algorithms. TCP in the wireless ad hoc networks is unstable and has poor throughput due to TCP's inability to recognize the difference between the link failure and the congestion. Besides, one TCP connection from one-hop neighbors will capture the entire bandwidth, leading to the one-hop unfairness problem [3] [4].

Performance analysis of the IEEE 802.11 MAC protocol could help to discover the inherent cause of the above problems and suggest possible solutions. Many papers on this topic have been published. Cali [5] derived the protocol capacity of IEEE 802.11 MAC protocol and presented an adaptive backoff mechanism instead of the exponential backoff mechanism. Bianchi [6] proposed a Markov chain model for the binary exponential backoff procedure to analyze and compute the IEEE 802.11 DCF saturated throughput. The performance evaluation in [6] assumes the saturated scenario where all stations always have data to transmit. Based upon Bianchi's model, Foh and Zuckerman presented the delay performance of 802.11 MAC in [7]. However, to the authors' knowledge, there is no study on the queue dynamics of the IEEE 802.11 wireless LANs.

In this paper, we first characterize the probability distribution of the MAC layer packet service time (i.e., the time interval between the time instant a packet starts for transmission and the time instant that the packet either is acknowledged for correct reception by the intended receiver or is dropped). Based on the probability distribution model of the MAC layer packet service time, we then study the queueing performance of the wireless LANs based on the 802.11 protocols.

II. BACKGROUND

A. Distributed Coordination Function (DCF)

Before we present our analysis for 802.11 MAC, we first briefly describe the main procedures in the DCF of 802.11 MAC protocol. In the DCF protocol, a station shall ensure that the medium is idle before attempting to transmit. It selects a random backoff interval less than the current contention window (CW) based on the uniform distribution, and then decreases the backoff timer by one at each slot when the medium is idle (may wait for DIFS followed a successful transmission). If the medium is determined to be busy, the station defers until the end of the current transmission. Transmission shall commence whenever the backoff timer reaches zero. When there are conflicts during the transmission or when the transmission fails, the station invokes the backoff procedure. To begin the backoff procedure, the contention window size CW, which takes an initial value of Cwmin, doubles its value before it reaches a maximum upper limit CWmax, and remains the value CWmax when it is reached until it is reset. Then, the station sets its backoff timer to a random number uniformly distributed over the interval [0, CW) and attempts to retransmit when the backoff timer is decreased to zero again. If the maximum transmission failure limit is reached, the retransmission shall cease, and the packet shall be discarded [10].

This work was supported in part by the US Office of Naval Research under grant N000140210464 (ONR Young Investigator Award) and N000140210554.

B. The Markov chain model for the exponential backoff transmission procedure

Bianchi [6] presented a two-dimensional Markov chain model for the backoff mechanism for 802.11 MAC DCF. By deriving the steady state probability of the Markov chain, he carried out an accurate analysis for the throughput at the saturated state. In this paper, we will extend this model and characterize the probability distribution of the MAC layer packet service time (or MAC layer service time for short).

To model the backoff mechanism as a Markov chain, there is a key approximation in the model when the network enters the steady state. At each transmission attempt, regardless of the number of retransmissions occurred, we assume that each packet collides with constant and independent probability p_c. This assumption is reasonable in steady state since p_c mainly depends on the overall network traffic.

III. THE PROBABILITY DISTRIBUTION OF THE MAC LAYER SERVICE TIME

A. MAC Layer Service Time

As described in section II, there are three basic processes when the MAC layer transmits a packet: the decrement process of the backoff timer, the successful packet transmission process that takes a time period of T_{suc} and the collision process for the packet transmission that takes a time period of T_{col}. Here, T_{suc} is the average time the medium is sensed busy because of a successful transmission, and T_{col} is the average time the medium is sensed busy by each station due to collisions.

The MAC layer service time is the time interval from the time instant that a packet becomes the head of the queue and ready to transmit to the time instant that either the packet is acknowledged for a successful transmission or the packet is dropped. This time is important when we examine the performance in the high layer. Apparently, the distribution of the MAC layer service time is a discrete probability distribution because the smallest time unit of the backoff timer is a time slot. T_{suc} and T_{col} depend on the transmission rate, the length of the packet and overhead, and the specific transmission scheme (the basic access DATA/ACK scheme or the RTS/CTS scheme) [6].

To obtain the distribution of the MAC layer service time denoted by random variable T_S, we apply the generalized state transition diagram, where we mark the transition time on each branch along with the transition probability in the state transition diagram (the Markov chain). The transition time, which is the duration for the state transition to take place, is expressed as an exponent of Z variable in each branch. Thus, the probability generating function of total transition time can be obtained from the signal transfer function of the generalized state transition diagram using the well-known Mason formula [8].

B. Decrement Process of Backoff Timer

Let us first consider the decrement process of backoff timer because the successful transmission process and collision process are simply determined by the value of T_{suc} and T_{col} respectively. In the backoff process, if the medium is idle, the backoff timer will decrease by one for every idle slot detected. When detecting an ongoing successful transmission, the backoff timer will be suspended and defer a time period of T_{suc}, while if there are collisions among the nodes, the deferring time will be T_{col}.

As mentioned in section II, p_c is the probability of a collision seen by a packet being transmitted on the medium. Assuming that there are n stations in the wireless LAN we are considering, we observe that p_c is also the probability that there is at least one packet transmission in the medium among other (n-1) stations in the interference range of the station under consideration. This yields

$$p_c = 1-[1-(1-p_0)\tau)]^{n-1}, \tag{1}$$

where p_0 is the probability that there are no packets ready to transmit at the MAC layer in the wireless station under consideration, and τ is the packet transmission probability that the station transmits in a randomly chosen slot time given that the station has packets to transmit at the MAC layer. Let W be the minimum value of contention window size. Since the contention window is reset after the maximum α times of retransmissions as defined in the protocols [10], following similar procedure used in [6] and [9], we can obtain

$$\tau = \begin{cases} \dfrac{2(1-p_c^{\alpha+1})}{1-p_c^{\alpha+1}+(1-p_c)W(\sum_{i=0}^{\alpha}(2p_c)^i)} & ,\alpha \le m \\ \dfrac{2(1-p_c^{\alpha+1})}{1-p_c^{\alpha+1}+p_cW\sum_{i=0}^{m-1}(2p_c)^i+W(1-2^m p_c^{\alpha+1})} & ,\alpha > m \end{cases}, \tag{2}$$

where m is the maximum number of the stages allowed in the exponential backoff procedure (definition is clarified below).

Let P_{suc} be the probability that there is one successful transmission among other (n-1) stations in the considered slot time given that the current station does not transmit. Then,

$$\begin{aligned} P_{suc} &= C_1^{n-1}(1-p_0)\tau(1-(1-p_0)\tau)^{(n-2)} \\ &= (n-1)((1-p_c)^{(n-2)/(n-1)}+p_c-1) \end{aligned}. \tag{3}$$

Then $p_c - P_{suc}$ is the probability that there are collisions among other (n-1) stations (or neighbors).

Thus, the backoff timer has the probability of 1- p_c to decrement by 1 after an empty slot time σ, the probability of P_{suc} to stay at the original state after T_{suc}, and the probability of $p_c - P_{suc}$ to stay at the original state after T_{col}. So the decrement process of backoff timer is a Markov process. The signal transfer function of its generalized state transition diagram is

$$H_d(Z) = (1-p_c)Z^{\sigma}/[1-p_{suc}Z^{T_{suc}}-(p_c-p_{suc})Z^{T_{col}}]. \tag{4}$$

From above formula, we observe that $H_d(Z)$ is a function of p_c, the number of stations n and the dummy variable Z.

C. Generalized State Transition Diagram

Now, it is possible to draw the generalized state transition diagram for the packet transmission process as shown in Fig. 1.

In Fig.1, {s(t), b(t)} is the state of the bi-dimensional discrete-time Markov chain, where b(t) is the stochastic process representing the backoff timer count for a given station, and s(t) is the stochastic process representing the backoff stage with values (0, ... α) for the station at time t. Let m be the "maximum backoff stage" at which the contention window size takes the maximum value, i.e. CWmax = 2^mCWmin. At different "backoff stage" i ∈ (0, α), the contention window size $W_i = 2^i CWmin$ if $0 \le i \le m$, and W_i= CWmax if $m \le i \le \alpha$.

As we defined before, the random variable T_S is the duration of time taken for a state transition from the start state (beginning to be served) to the end state (being transmitted successfully or discarded after maximum α times retransmission failures). Thus, its Probability Generating Function (PGF), denoted as *B(Z)* that is the function of p_c, n and Z, is simply the signal transfer function from the start state to the end state given by:

$$HW_i(Z) = \begin{cases} \sum_{j=0}^{2^iW-1} H_d{}^j(Z)/(2^iW), & (0 \le i \le m) \\ HW_m(Z), & (m < i \le \alpha) \end{cases}$$

$$H_i(Z) = \prod_{j=0}^{i} HW_j(Z), (0 \le i \le \alpha) \qquad .(5)$$

$$B(Z) = (1-p_c)Z^{T_{suc}} \sum_{i=0}^{\alpha} (p_c Z^{T_{col}})^i H_i(Z) + (p_c Z^{T_{col}})^{\alpha+1} H_\alpha(Z)$$

Since *B(Z)* can be expanded in power series, i.e.

$$B(Z) = \sum_{i=0}^{\infty} P(T_s = i)Z^i \;, \qquad (6)$$

we can obtain the arbitrary n^{th} moment of T_S by differentiation (hence the mean value and the variance). Here the unit of T_S is slot.

D. Probability Distribution Modeling

Fig. 2 shows the MAC layer service time corresponding to the various combinations of T_{suc}, T_{col} and σ. From the probability generation function (PGF) of the MAC layer service time, we can easily obtain the discrete probability distribution. Fig. 2 shows the probability distribution of the MAC service time at each discrete value. This example uses RTS/CTS mechanisms. Packet length is 1000bytes and data transmission rate is 2 Mbps. We can see that there may be one or more successful transmissions during the decrement processes of the backoff timer and T_{suc} is much larger than T_{col} and σ in the RTS/CTS mechanism.

We also notice that its envelope is similar to an exponential distribution. If we can use some continuous distribution to approximate the discrete one, it will give great convenience to analyze the queueing characteristics. Fig.2 shows the approximate probability density distribution (PDF) of T_S and several continuous PDF including Gamma distribution, log-normal distribution, exponential distribution and Erlang-2 distribution. We can see log-normal distribution is a good approximation not only at the high collision probability case but also at the low collision probability case, and it also has a very close tail distribution with that of T_S. In addition, the exponential distribution seems to be a reasonably good approximation except at very low collision probability. Here, the PDF of T_S is obtained by assuming that the probability density is uniform in a small interval and is represented by a histogram while other continuous PDF is determined by the value of mean and/or variance of T_S.

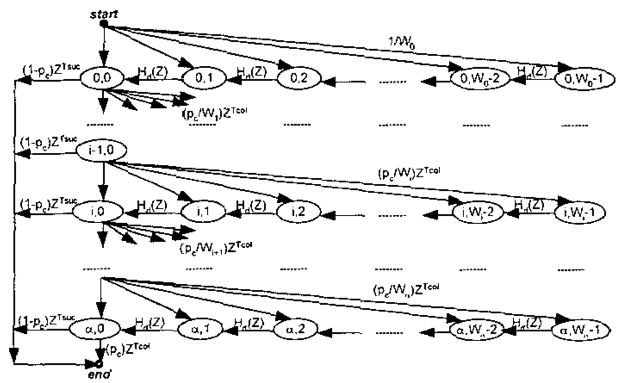

Figure 1. Generalized state transition diagram for transmission process

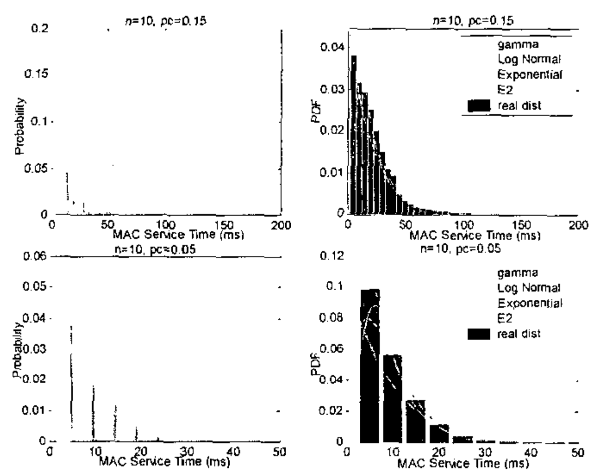

Figure 2. Probability Distribution of MAC Layer Service Time

E. Simulation Results

In our simulation study, we use ns-2. The wireless channel capacity is set to 2Mbps. Packet length is 1000 bytes, and the maximum queue length is 50. We use different number of mobile stations in a rectangular grid with dimension 200m x 200m to simulate the Wireless LAN.

The mobile stations are randomly moving and have the same rate of packet input. Our simulation indicates the distribution of MAC layer service time is independent of the packet input distribution whether it is deterministic or Poisson distributed. It mainly depends on the total traffic in the network before saturation and on the number of mobile stations after saturation.

Fig.3 shows the simulation results of the MAC layer service time in the network with 17 mobile stations and total traffic of 0.8 and 1.6 Mbps respectively. It displays the same characteristics with that obtained from the theoretical result.

IV. Queueing Modeling and Analysis

A. Problem formulation:

Many applications are sensitive to end-to-end delay and queue characteristics such as average queue length, waiting time, queue blocking probability, service time, and goodput. Thus, it is necessary to investigate the queueing modeling and analysis.

A queue model can be characterized by the service time distribution and the arrival process in addition to the service discipline. We have characterized the service time distribution in the previous section. In this paper, we assume that the packet arrivals at each mobile station follow the Poisson process or a deterministic distribution with average arrival rate λ. The packet transmission process at each station can be modeled as a general single "server". The buffer size at each station is K. Thus the queueing model for each station can be modeled as an M/G/1/K when Poisson arrivals of packets are assumed.

B. The steady-state probability of the M/G/1/K queue

Let p_n represent the steady-state probability of n packets in the queueing system, and let π_n represent the probability of n packets in the queueing system upon a departure at the steady state, and let $\boldsymbol{P}=\{p_{ij}\}$ represent the queue transition probability matrix:

$$p_{ij} = \Pr\{X_{n+1} = j \mid X_n = i\}, \tag{7}$$

where X_n is the number of packets seen upon the nth departure.

To obtain p_{ij}, we define

$$k_n = \Pr\{n \text{ arrivals during service time } S=t\} = \sum_{i=0}^{L} \frac{e^{-\lambda' i}(\lambda' i)^n}{n!} P\{Ts = i\}, \tag{8}$$

where λ' is the average arrival rate. We can easily obtain

$$\mathbf{P} = \{p_{ij}\} = \begin{bmatrix} k_0 & k_1 & k_2 & \cdots & k_{K-2} & 1-\sum_{n=0}^{K-2} k_n \\ k_0 & k_1 & k_2 & \cdots & k_{K-2} & 1-\sum_{n=0}^{K-2} k_n \\ 0 & k_0 & k_1 & \cdots & k_{K-3} & 1-\sum_{n=0}^{K-3} k_n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & k_0 & 1-k_0 \end{bmatrix} \tag{9}$$

Moreover, we notice that

$$k_0 = B(e^{-\lambda'}), \; k_n = \frac{\lambda'^n}{(-1)^n n!} \frac{\partial^n B(e^{-\lambda'})}{\partial \lambda'^n} \tag{10}$$

According to the balance equation:

$$\pi P = \pi, \tag{11}$$

where $\pi = \{\pi_n\}$, we can compute the π. For the finite system size K with Poisson input [12], we have

$$p_0 = \frac{\pi_0}{\pi_0+\rho}, \; p_n = \frac{\pi_n}{\pi_0+\rho} (0 \le n \le K-1), \; p_K = 1-\frac{1}{\pi_0+\rho} \tag{12}$$

where ρ is the traffic intensity $\rho = \lambda' \overline{T_S}$.

C. Conditional Collision Probability p_c

From (12), we know that p_n is a function of p_c, λ', n. So we can compute p_c under different values of λ' and n with the help of (1) and (2) using some recursive algorithm. Thus we can obtain the distribution of MAC service time at different offered load verified by simulation results in Fig. 3.

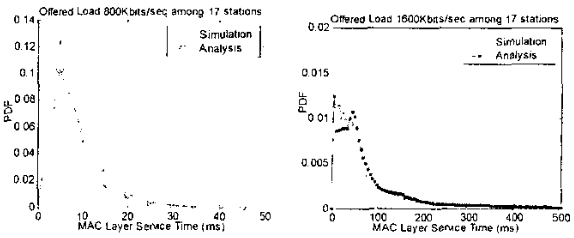

Figure 3. Simulation Results of MAC Layer Service Time

D. Performance Metrics of the Queueing System

The average queue length, blocking probability, and average waiting time including MAC service time are given by

$$L = \sum_{i=0}^{K} i \times p_i, \; P_B = p_k = 1-\frac{1}{\pi_0+\rho}, \; W = \frac{L}{\lambda(1-P_B)} \tag{13}$$

E. Throughput

If we know the blocking probability p_B, then the throughput S can be computed easily by

$$S = (1 - p_B)(1 - p_c^{\alpha+1}), \tag{14}$$

where $p_c^{\alpha+1}$ is the packet discard probability due to transmission failure.

F. Results

Fig.4 shows the results for the major performance metrics. All of them have a dramatic change around the traffic load of 1.1-1.5 Mbits/sec. This is because that the collisions increase significantly around this traffic load, resulting in much longer MAC service time for each packet. Simulation results are shown in Fig.5. We use the same network topology as in section III, and the packet arrivals are Poisson.

From the results, we observe that all the metrics are dependent on the collision probability p_c. Fig.4 shows that p_c mainly depends on the total traffic in the non-saturated scenario. On the other hand, p_c is affected by the total number of packets attempting to transmit by all neighboring stations. In non-saturated case, when all arriving traffic is immediately served by the MAC layer, the queue length may reach zero and the corresponding station will not compete for the medium. However, in the saturated scenario, i.e. the stations always have packets to transmit, the total number of packets attempting to transmit equals to the total number of neighboring stations, hence p_c is mainly dependent on the total number of neighboring stations.

The MAC layer service time shows similar change at different offered load, because it is dependent on the p_c. All other performance metrics are dependent on the distribution of the MAC layer service time, so they show the similar change in the figures. Average queue length is almost zero at the non-saturated status and reaches almost maximum length at the saturated status. Average waiting time for each packet in the queue almost equals zero at non-saturated status and reaches several seconds at the saturated status. Queue blocking probability is zero at the non-saturated status when traffic load is low, and linearly increases with the offered load at the saturated status. Throughput linearly increases with the offered load at the non-saturated status and maintains a constant value with different total number of transmitting stations at the saturated status. The packet discarding probability at MAC layer is much smaller than the queue blocking probability.

In summary, all these results indicate that IEEE 802.11 MAC works well at the non-saturated status at low traffic load while its performance dramatically degrades at the saturated status, especially for the delay metric. Besides, at the non-saturated status, the performance is dependent on the total traffic and indifferent to the number of transmitting stations. At the saturated status, the number of transmitting stations is much more important to the whole performance. The similar phenomena have been observed for the distribution of MAC service time as in section III.

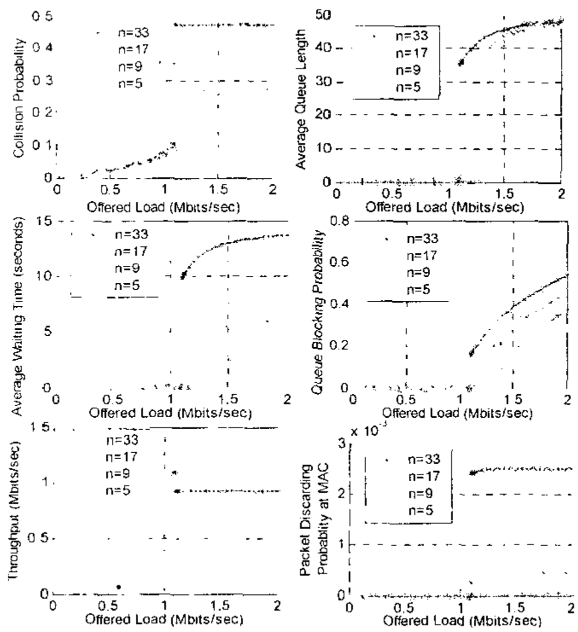

Figure 4. Queue Characteristics

V. CONCLUSION

In this paper, we have derived the probability distribution of the MAC layer service time. To obtain this distribution, we expand the Markov chain model to the more general case for the exponential backoff procedure in IEEE 802.11 MAC protocols. Accurate discrete probability distribution and approximate continuous probability distributions are obtained in this paper. Based upon the distribution of the MAC service time, we come up with a queueing model and evaluate the performance of the IEEE 802.11 MAC protocol in Wireless LANs in terms of throughput, delay, and other queue characteristics. Our results show that at the non-saturated status, the performance is dependent on the total traffic and indifferent to the number of transmitting stations. And at saturated status, the number of transmitting stations affects the performance more significantly.

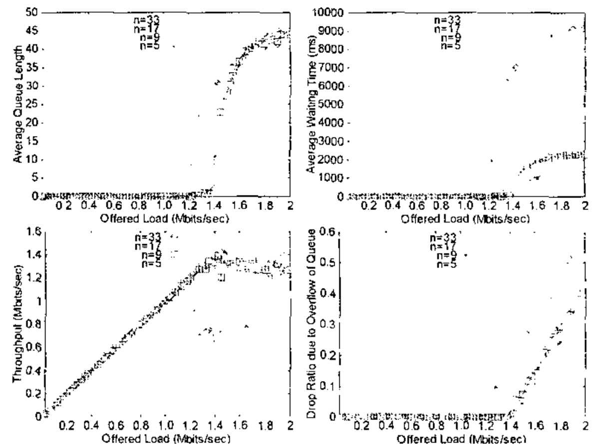

Figure 5. Simulation Results of Queue Characteristics

REFERENCES

[1] J. Broch et al., "A performance Comparison of Multihop Wireless Ad Hoc Network Routing Protocols," Proc. IEEE/ACM MOBICOM'98, Oct. 1998, pp. 85-97

[2] C. Perkins et al., "Performance Comparison of Two On-demand Routing Protocols for Ad Hoc Networks", IEEE Personal Communications, Feb. 2001, pp. 16-28

[3] G. Holland and N. Vaidya, "Analysis of TCP Performance over Mobile Ad Hoc Networks," ACM MOBICOM 1999

[4] S. Xu and T. Safadawi, "Does the IEEE 802.11 MAC Protocol Work Well in Multihop Wireless Ad Hoc Networks?" IEEE Communications Magazine, June 2001, pp. 130-137

[5] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 Protocol: Design and Performance Evaluation of an Adaptive Backoff Mechanism," IEEE journal on Selected Area in communications, V18, N9, September 2000

[6] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," IEEE Journal on Selected Area in Communications, V18, N3, March 2000

[7] C. H. Foh and M. Zukerman, "Performance Analysis of the IEEE 802.11 MAC Protocol," European Wireless 2002, Feb. 2002, Florence, Italy

[8] D. W. Choi, "Frame alignment in a digital carrier system-a tutorial", IEEE Communications Magazine, Volume: 28 2, Feb. 1990, Page(s): 47-54

[9] H. Wu, Y. Peng, K. Long, et al. "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement," Proc. IEEE INFOCOM'2002, June, 2002.

[10] IEEE standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ISO/IEC 8802-11: 1999(E), Aug. 1999

[11] K. Fall and K. Varadhan, editors, "NS Notes and Documentation," The VINT Project, UC Berkeley, LBL, USC/ISI, and Xerox PARC, April, 2002

[12] D. Gross and C. M. Harris, "Fundamentals of Queueing Theory," 3rd ed., John Wiley & Sons, Inc. 1998