ESTIMATION AND APPLICATION OF END-TO-END DELAY UNDER UNSATURATION TRAFFIC IN WIRELESS AD HOC NETWORKS

Dong Linfang[1], Shu Yantai[1], Chen Haiming[1], Ma Maode[2]
[1]Tianjin University, Tianjin, China, [2]Nanyang Technological University, Singapore

Abstract - In this paper, we analysis the average packet delay on IEEE 802.11 DCF under unsaturation traffic in multi-hop ad hoc networks. We employ a Markov chain model to analysis the probability of transmission at each node in an arbitrary slot, and derive the channel access delay. We model each node using an M/G/1 queue and derive the queueing delay. The model is extended from analyzing the single-hop average packet delay to evaluating the end-to-end packet delay in multi-hop ad hoc networks without assuming the traffic to be in a saturation state. We have done extensive simulation to validate our analysis. The analytic and the simulation results match very well. The analysis results can be used to evaluate the fairness level of wireless channel.

Keywords - IEEE 802.11 DCF, Packet delay, Multi-hop, Unsaturation traffic

I. INTRODUCTION

The IEEE 802.11 is the standard [1] to specify Wireless Local Area Networks (WLANs). Although it specifies two fundamental access mechanisms, Distributed Coordination Function (DCF) and Point Coordination Function (PCF), DCF is more widely accepted by the researchers as well as by the telecommunication industry. Therefore, in this paper, we focus our analysis on the multi-hop end-to-end delay performance of the DCF mechanism.

There have been considerable amounts of work on the performance evaluation of IEEE 802.11 DCF through either simulation or mathematical modeling. The performance of the DCF protocol was studied by simulation in [2-3]. Some analytic models to study the performance of DCF scheme under the saturated or unsaturated traffic in single-hop or multi-hop ad hoc networks have been proposed. The proposals in [4-6] studied the performance in terms of throughput and the models in [7-9] studied the performance in terms of access delay and average packet delay in single-hop ad hoc networks under saturated load. Both of the analytic model in [10] on throughput and the delay analysis in [11-12] were set up under unsaturated traffic in single-hop scenarios. Only the models in [13-14] investigated the throughput with unsaturated traffic in multi-hop ad hoc networks. However, there has been no solution to present the end-to-end delay under unsaturated traffic in a multi-hop scenario to the best of our knowledge.

In this paper, we evaluate the average packet delay in an IEEE 802.11 MAC under unsaturation traffic in multi-hop ad hoc networks. We employ a Markov chain model to analyze the probability of transmission at each node in an arbitrary slot, and derive the wireless channel access delay. We model each node as an M/G/1 queue and derive the queueing delay at a node. Thus the average delay from the time a packet arrived at current node to the time it can be successfully received by the next hop node can be calculated. The model is extended from analyzing the single-hop average packet delay to evaluating the end-to-end packet delay in multi-hop ad hoc networks without assuming the traffic being in a saturation state. The major contribution of this paper is to analyze the average packet delay for IEEE 802.11 DCF by using an analytic model under unsaturated traffic in multi-hop ad hoc networks. To the best of our knowledge, it has never been presented before, especially for multi-hop ad hoc networks.

The paper is organized as follows. In Section II, we introduce and derive the Markov chain model for unsaturated sources. Then we provide a detail analysis of packet service time and queuing time in single-hop ad hoc networks in Sections III. Section IV extends the analysis to the multi-hop ad hoc networks. The analytic results were compared with data derived from simulation in Section V. Section VI demonstrates how to estimate the fairness level of wireless channel based on the delay. We conclude the paper in Section VII.

II. MARKOV MODEL FOR IEEE 802.11 DCF UNDER UNSATURATION TRAFFIC

We model the DCF scheme at MAC layer using an M/G/1 queue. We proposed a Markov chain model to describe the scheme in Fig.1, which evolves from the model presented in [4]. We have included the NoPK state in which the node does not have any packet to transmit. Packets arrive at nodes according to a Poisson process with average rate λ packets/sec. We assume that the wireless channel condition is ideal (error-free and no capture).

The model consists of an aggregation of states that a node can reside in. The points C_0, C_1 and C_2 in Fig.1 represent connection points, which are not states. The queue will be checked at point C_0 after each successful transmission or after having reached the maximum number of the retransmissions *m*. If there is a packet, the node enters into backoff state directly. Otherwise, the node enters into the post-backoff state (on the left part of the Markov chain in Fig.1). This is because that after each successful transmission, the node must enter into a backoff even if there is no packet in the queue. If no packet arrives during the post-backoff, the node will enter the empty queue state (the probability is $p_{\rm nk}$) and stay at this state until a new packet arrives. If a packet arrives at the node with an empty queue, the node senses the medium If the medium is idle (no transmission), the node transits to state FirstPK and sends the packet immediately. On the other hand, if the medium is sensed busy, the node goes to point C_1 and then enters into backoff and the packet will be transmitted after the backoff timer reaches 0. When a packet arrives, the medium is sensed idle with probability $p_{\rm idle}$ and it is sensed busy with probability $p_{\rm busy}$. For a node, we use the tuple (i, k) to represent the different states in the backoff stages, with i being the backoff stage number i = 0,1, … m', …, m, and k being the value of the backoff timer in the range [0, $W_{\rm i}$-1]. $W_{\rm i}$ is the size of the CW at stage i and is computed by $W_{\rm i}=2^i W_0$, if $i \le m'$. Otherwise, if $i \in [m', m]$, $W_{\rm i}$ is kept at its maximum value $W_{\max}=2^{m'} W_0$. With m we denote the maximum number of packet retransmissions before the packet is dropped. According to [1], the default value for m' is 5 and it is 7 for m. $b_{\rm i,k}$ denotes the probability to be on the state (i, k). For the state Empty, we denote its state probability with $b$empty. q denotes the probability of having an empty queue. The probability of failure is denoted by p. We use $b_{\rm C0}$ and $b_{\rm C1}$ to denote the probability that the node arrives at point C_0 and C_1, respectively.

In order to get the expression of the state steady probabilities of a node, in the first step, we need to express all state probabilities in terms of $b_{0,0}$. Then, we use the normalization condition to obtain $b_{0,0}$ itself. From the balance equation in the steady state, we can obtain the following relations:

$$b_{0',k} = \frac{W_0 - k}{W_0} q b_{C0} \tag{1}$$

$$b_{i,0} = p^i b_{0,0} \tag{2}$$

$$b_{i,k} = \frac{W_i - k}{W_i} p^i b_{0,0} \tag{3}$$

$$\begin{aligned}\tau &= \sum_{i=0}^{m} b_{i,0} + b_{empty} p_{idle} (1 - e^{-\lambda\sigma}) \\ &= b_{0,0}\left(\frac{1-p^{m+1}}{1-p}\right) + b_{empty}(1-\tau)^n (1 - e^{-\lambda\sigma})\end{aligned} \tag{4}$$

τ denotes the probability that a node transmits in a randomly selected time slot, and p denotes the collision probability of a transmitted node. σ denotes the duration of one system time slot. It is known that a node can send a packet only at state (i,0) or state FirstPK. The probability of state FirstPK equals the probability that during a slot σ , at least one packet arrives the empty queue and the medium is sensed idle. Therefore, the node transmits from state Empty to state FirstPK and sends the packet immediately. The probability $b_{\rm empty}$ to be in state Empty is equal to:

$$b_{\rm empty} = b_{\rm C0}\, q\, p_{nk} \tag{5}$$

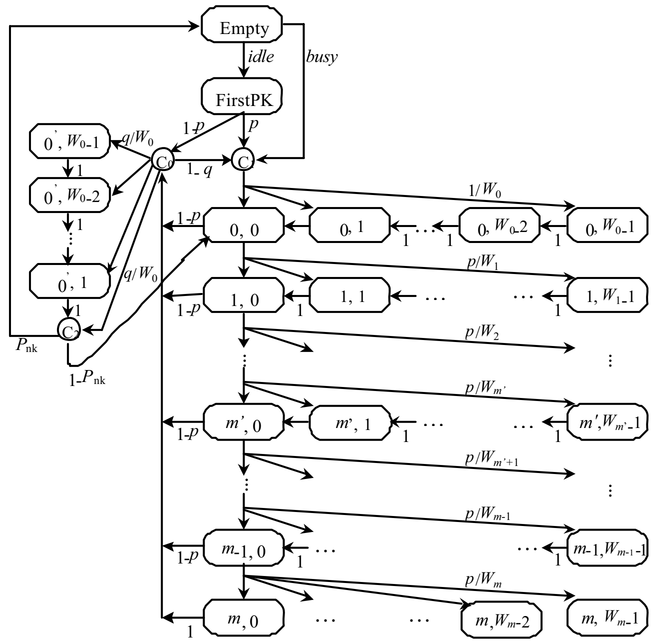

Fig. 1. Finite Load Markov Chain Model

p_{nk} means that the probability of not receiving any packet during the time spent in the post-backoff stage:

$$p_{nk} = e^{-\lambda (W_0+1)\overline{\sigma}/2} \tag{6}$$

$b_{\rm C0}$ can be obtained from:

$$\begin{aligned} b_{C0} &= \sum_{i=0}^{m-1} b_{i,0}(1-p) + b_{m,0} + (1-p) b_{empty} p_{idle} \\ &= (1-p) b_{0,0} (p^0 + \cdots + p^{m-1}) + p^m b_{0,0} + (1-p) b_{empty} p_{idle} \\ &= b_{0,0} + (1\text{-}p)\, b_{empty} p_{idle} \end{aligned} \tag{7}$$

$b_{0,0}$ is obtained by using the normalization condition:

$$1 = \sum_{i=0}^{m} \sum_{k=0}^{W_i - 1} b_{i,k} + \sum_{k=1}^{W_0+1} b_{0',k} + b_{empty}(1 + p_{idle}) \tag{8}$$

III. DELAY ANALYSIS OF SINGLE-HOP AD HOC NETWORKS

In this section, we analyze the average packet delay in single-hop ad hoc networks. In a single-hop ad hoc network, all the nodes can hear each other. In our model, we use an M/G/1 queue to analyze the queueing delay of the DCF scheme. Let q denote the probability of having no packet in the buffer. In an M/G/1 queue, q is simply equal to q=MAX (0,1- λ D), where D is the average service time that a packet spends in the MAC layer (from the packet leaves the MAC buffer until it is successfully transmitted or reaches the retry limit). We consider two cases: one with an empty queue, another with a nonempty queue. Therefore, D has to be conditioned on q and is equal to:

$$D = (1\text{-}\, q)\, E(S_{\rm b}) + q T_{no} \tag{9}$$

where $E(S_b)$ and T_{no} denotes the average service time of a packet that at its arrival, it finds the queue being non-empty and empty, respectively.

$$E(S_b) = \overline{\sigma}\sum_{s=0}^{m} p^s \frac{W_s - 1}{2} + T_s \sum_{s=0}^{m} p^s (1-p) + T_c\left(\sum_{s=1}^{m} p^s (1-p)s + (m+1)p^{m+1}\right) \quad (10)$$

where T_s is the time for a successful transmission. And T_c is the average time the channel is sensed busy by each station during a collision. Let H = PHY_{hdr} + MAC_{hdr} be the packet header, and δ be the propagation delay. T_s and T_c equal to:

$$T_S = DIFS + RTS + CTS + H + E[P] + ACK + 3SIFS + 4d$$
$$T_c = DIFS + RTS + CTS + SIFS + \delta \quad (11)$$

where $E[P]$ is the average packet length.

The first term in (10) accounts for the total time needed to attain a transmission state, which is called (i, 0) in Fig.1. The second term is the expected value of the time needed to actually accomplish the physical transmission and the receipt of the ACK. The third term accounts for the expected time of collisions.

The collision probability p in single-hop ad hoc networks is equal to the probability that at least one of the other n-1 nodes transmits a packet. Therefore, p can be written as:

$$p = 1-(1-\tau)^{n-1} \quad (12)$$

The probability that at least one node transmits is:

$$p_{tr(n)} = 1-(1-\tau)^n \quad (13)$$

The probability that the transmitted packet is successful is:

$$p_{s(n)} = \frac{n\tau(1-\tau)^{n-1}}{p_{tr(n)}} \quad (14)$$

$\overline{\sigma}$ is the average time between successive counter decrements,

$$\begin{aligned}\overline{\sigma} &= (1-p_{tr(n-1)})\sigma + p_{tr(n-1)}p_{s(n-1)}(T_s+\sigma) \\ &\quad + p_{tr(n-1)}(1-p_{s(n-1)})(T_c+\sigma) \\ &= (1-\tau)^{n-1}\sigma + (n-1)\tau(1-\tau)^{n-1}(T_s+\sigma) \\ &\quad + \left[1-(1-\tau)^{n-1}-(n-1)\tau(1-\tau)^{n-1}\right](T_c+\sigma)\end{aligned} \quad (15)$$

Now we can get:

$$E(S_b) = \frac{\overline{\sigma}W(1-(2p)^{m+1})}{2(1-2p)} - \frac{\overline{\sigma}(1-p^{m+1})}{2(1-p)} + T_s(1-p^{m+1}) + T_c p\frac{1-p^{m+1}}{1-p} \quad (16)$$

The average service time of a packet when it arrives an empty queue can be found:

$$T_{no} = (1-t)^n\, t_{idle} + [1-(1-t)^n]\, t_{busy} \quad (17)$$

When a packet arrives in an empty queue, two cases exist. One is that the packet arrives when the channel is idle; while another is that the channel is busy.

$$t_{idle} = p[T_c + E(S_b)] + (1-p)T_s . \quad (18)$$
$$t_{busy} = E(S_b) \quad (19)$$

Finally, combining (16) and (17) the expression of average service time of packet can be obtained.

Under M/G/1 assumption, an accurate analysis of the average packet delay requires the knowledge of the second moment of the service time of packets at the MAC layer. Note that this quantity is not easy to obtain, we use a simple analysis to get the expected waiting time E[W] in the M/G/1 buffer. We assume that the average service time of packets equals to the sum of an exponential random variable with expected value $1/\mu$ and T_s. Let X denotes the average service time of packets. We can obtain μ from the following relation, where D is given in formula (9):

$E[X] = T_s + 1/\mu = D$.
$D[X] = 1/\mu^2$
$E[X^2] = D[X] + (E[X])^2 = 1/\mu^2 + D^2 = (D - T_s)^2 + D^2$

where E[X] and D[X] are the mean and the variance of X, respectively. We now have the second moment of X and use it in the Pollaczek-Khinchin formula [15] to get the expected waiting time E[W] in the M/G/1 buffer:

$$E[W] = \frac{\lambda\overline{X^2}}{2(1-\rho)} = \frac{\lambda[(D-T_s)^2 + D^2]}{2(1-\rho)} \quad (20)$$

where $\rho = \lambda D$.

The average packet delay for our system is the sum of average service time E[X] and the obtained waiting time E[W], *i.e.* E[$Delay$] = E[X] + E[W].

IV. DELAY ANALYSIS OF MULTI-HOP AD HOC NETWORKS

A single-hop ad hoc network is a fully connected network, in which there is no hidden terminal problem. While in multi-hop ad hoc networks, the presence of hidden terminals substantially degrades the performance of the protocol. The IEEE 802.11 standard provides a four-way handshaking technique to solve this problem. We generalize the analysis one-hop by one-hop in the same method as used in single-hop scenario to analyze the performance of multi-hop ad hoc networks with considering hidden terminal problem.

Fig.2 gives a hidden area (H_A) of node S. We assume that the transmission range is R and n nodes are randomly placed within the range. We will calculate the number of hidden terminals of node S, which is denoted by H_n. If each neighbor of node S has the same probability to be selected as a receiver of S, the average distance d between node S and the reciver is given by $d = \frac{2}{3}R$. According to [16], H_A can be computed by:

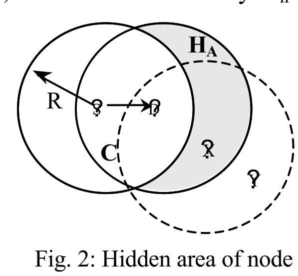

Fig. 2: Hidden area of node S

$$H_A = \pi R^2 - 2R^2\left(\arccos(\tfrac{d}{2R}) - \tfrac{d}{2R}\sqrt{1-(\tfrac{d}{2R})^2}\right) \quad (21)$$

Then we can get the number of hidden terminals of node S:

$$H_n = \frac{H_A}{\pi R^2} n \approx 0.4164\, n \quad (22)$$

Because of the hidden nodes, the collision probability p in multi-hop ad hoc networks is much higher than that in single-hop ad hoc networks. When node S initiates communication by transmitting an RTS in a given time slot, the transmission will be successful if and only if all of the following events happen in the same slot. Firstly, no RTS is transmitted by any node in the transmission areas of node S and D. Secondly, no CTS is transmitted by any node which located in the transmission area of node D. For the first event, if a hidden node X transmits an RTS, it will collide with the RTS from node S at the node D. For the same reason, a CTS transmitted by a node in the transmission area of D will definitely effect node D's reception of the RTS from node S. The probabilities of the both events can be analyzed as following.

Let p_1 denotes the probability of the first event. There are average H_n nodes in the hidden area, thus

$$p_1=(1-\tau)^{n-1+Hn} \quad (23)$$

Let p_2 denotes the probability of the second event. The probability of node X transmitting a CTS equals to the probability of successfully receiving an RTS and the destination of the RTS is just node X. The probability of successfully receiving an RTS is $\tau(1-p)$. Since the destination node is selected from the neighbors with equal probability, the probability of node X is selected as the destination is $1/n$. Because of the first condition, an RTS received by X must be transmitted by a D's hidden node, Y, for example. There are H_n nodes in this area. So, p_2 can be written as follows:

$$p_2 = \left(1 - \frac{H_n}{n}\tau(1-p)\right)^{n-1} \quad (24)$$

Using (22) and (23), the probability of failure, p, can be expressed as:

$$p = 1 - p_1 p_2 \quad (25)$$

After p was founded, we obtain average packet delay for 1-hop in multi-hop ad hoc networks with the same method as used in single-hop analysis. If the average number of hops counted per packet is n_p, and the average delay per hop is $E(T)$, the average end-to-end delay will be $n_p*E(T)$. We should note that if the number of packet transmissions, on average, at each hop is 2, the actual load of every node will be doubled of the input load from the outside. Therefore, we can calculate the average delay per hop $E(T)$ using $n_p\lambda$ as equivalent arrival rate.

V. PERFORMANCE EVALUATION

To verify our analytic results, we compared our analytic results with simulation results obtained from the Qualnet simulator [17]. We validate our delay analysis with two scenarios: single-hop and multi-hop ad hoc networks.

All the parameters used in our analytic model and simulation follow the parameters in paper [4] for DSSS. RTS/CTS mode of 802.11 DCF is used as the MAC layer protocol. The minimum backoff window size is 32 and maximum window size is 32×2^5. The transmission rates are equal for all transmitters. The buffer size of each node is set to 50 packets in the simulations. The number of the queueing packets is not more than 50 in the simulation. So, the buffer space can be regarded as infinite. There is no mobility considered.

In the single-hop scenario we studied the case of 10 and 20 active nodes. The packet arrival rate was increased so than the system load reached saturation gradually. Fig.3 compares the medium access delay for the analytic results and the simulation results under different system load in the case of 10 and 20 active nodes. We can see that the medium access delay increases when the system load increases. When the system load is large enough, however, the medium access delay will not increase. This is mainly due to the reason that packets arrive so fast that the system cannot consume them. The increased number of active nodes causes the increase of collision probability, and thus the medium access delay become longer. The analysis result is about 5% higher than the simulation result. Fig.4 compares the simulation and analytic results on the average delays for the 10 and 20 node case. It can be seen when the channel utilization factor gradually increases to 1, the average delay approaches infinity. This is because that when the system load is high, the queue becomes very long. The analysis result matches the simulation very well.

In order to validate our delay analysis for a multi-hop network with different offered load, a 120-node network was considered, in which the nodes were placed randomly in a square service area of $1500\times1500\text{m}^2$. All nodes transmission radius was 250 meters. Each node acted both as transmitter and receiver. Thus the average number of active node in the transmission range was about 10. We also assume that the network is relatively stable during the transmission of data packet or control message. Routes had been determined.

Fig.5 shows the 1-hop and 2-hop access delay in multi-hop scenario respectively. We can see that the access delay reached the maximum delay rapidly in 2-hop case. This is mainly due to the reason that in 2-hop case the load of the medium was increased. Therefore the collision probability and the access delay were increased. The difference between the analysis result and the simulation is within 5%. Fig.6 gives the average end-to-end packet delay for 1-hop and 2-hop cases in the multi-hop scenario. The packet delay increased more rapidly with the increase of packet arrival rate in 2-hop case than in 1-hop case. The simulation result is about 5% higher than the analysis result.

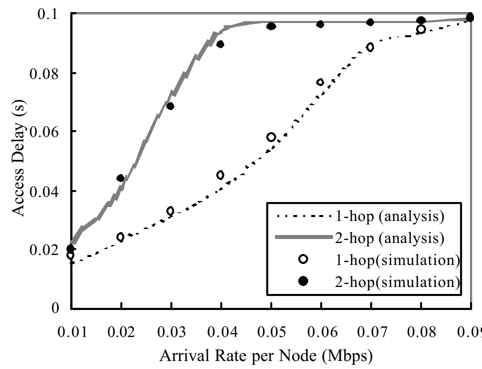

Fig. 5 Access Delay in Multi-Hop Scenario

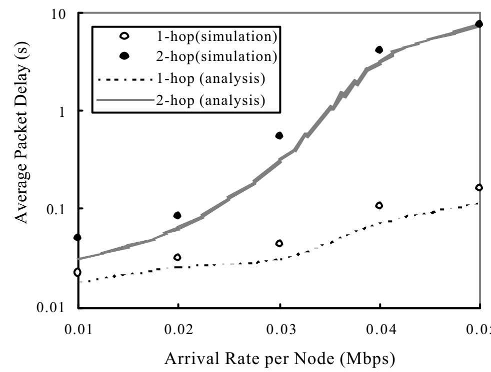

Fig. 6 Average Packet Delay in Multi-Hop Scenario

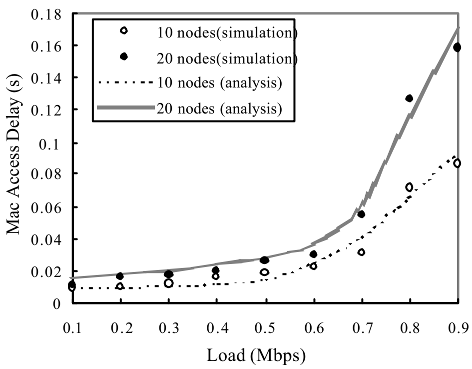

Fig. 3 Access Delay in Single-Hop Scenario

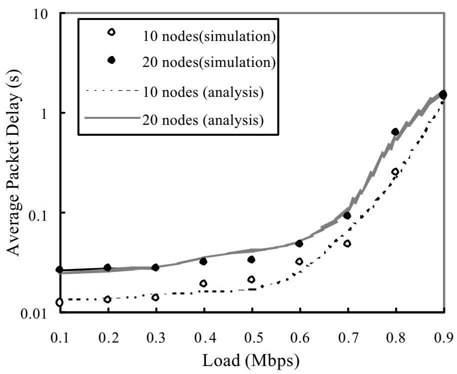

Fig. 4 Average Packet Delay in Single-Hop Scenario

VI. APPLICATION OF END-TO-END DELAY

In order to detect the fairness level, the throughput of data transmission is an often-used metric. In fact, we can also use end-to-end delay to evaluate the fairness level of the wireless channel. It is believed that the delay of data can reflect the fairness of shared link more promptly. For each packet's transmission, the more delay, the more severe congestion and competition. Nodes that share the same wireless channel have different average packet delay due to the fact that they have different number of competitors. Thus some nodes may lie in an unfair situation.

We measure one-hop delay in multi-hop ad hoc network, using that to detect congestion. We can control drop rate of the queue to guarantee fairness. In another scheme, every node monitors its packet delay separately. If the delay exceeds the threshold we set, the traffic of this node should be paused so that other nodes will have chance to send packets.

VII. SUMMARY AND CONCLUSION

In this paper, we have employed a Markov chain model to analysis the channel access delay. We have modeled each node by using an M/G/1 queue and have derived the queueing delay. The model has also been extended from analyzing the single-hop average packet delay to evaluating the end-to-end packet delay in multi-hop ad hoc networks under different traffic loads. To our knowledge, our solution is the first analysis on the end-to-end packet delay of the multi-hop ad hoc networks with finite load. Simulations have been conducted to verify the analytical results. And the simulation results have been matched quite well by the analytical results.

References

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

[2] H. S. Chhaya and S. Gupta, "Performance of Asynchronous Data Transfer Methods of IEEE 802.11 MAC Protocol," IEEE Personal Communications, vol. 3 No. 5, 1996, 8-15.

[3] G. Bianchi, L. Fratta, and M. Oliveri, "Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for

802.11 Wireless LANs, Proceeding of PIMRC96, vol. 2, 1996, 392-396.
[4] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Communications*, vol. 18, 2000, 535-547. Number 3.
[5] Y. C. Tay and K. C. Chua, "A Capacity Analysis for the IEEE 802.11 MAC Protocol," *ACM/Baltzer Wireless Networks*, vol. 7, 2001, 159-171.
[6] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of Reliable Transport Protocol over IEEE 802.11 Wireless LAN: Analysis and Enhancement. Proceedings of INFOCOM 2002, 599-607.
[7] G. H. Wang, Y. T. Shu, O. Yang, and L. Zhang, "Delay Analysis of the IEEE 802.11 DCF," in Proc. IEEE PIMRC'03, 2003, 7-11.
[8] P. Chatzimisios, A. C. Boucouvalas and V.Vitsas, "IEEE 802.11 Packet Delay – A Finite Retry Limit Analysis", IEEE Globecom 2003
[9] Marcelo M. Carvalho J. J. Garcia-Luna-Aceves, "Delay Analysis of IEEE 802.11 in Single-Hop Networks", IEEE ICNP 2003.
[10] F. A.-Shabdiz, S. Subramaniam, "A Finite Load Analytical Model for the IEEE 802.11 Distributed Coordination Function MAC", *WiOpt'03 workshop*, INRIA Sophia Antipolis, 2003.
[11] G. R. Cantieni, Q. Ni, C. Barakat, T. Turletti "Performance Analysis under Finite Load and Improvements for Multirate 802.11," To appear in Elsevier Computer Communications Journal, special issue on Performance Issues of Wireless LANs, PANs, and Ad Hoc Networks, 2004.
[12] O. Tickoo, B. Sikdar, "Queueing Analysis and Delay Mitigation in IEEE 802.11 Random Access MAC based Wireless Networks," *IEEE Infocom* 2004.
[13] Y. Chen, Q-A. Zeng, and D. P. Agrawal, "Performance of MAC Protocol in Ad Hoc Networks," Proceeding of the *Communication Networks and Distributed Systems Modeling and Simulation Conference* (CNDS'03), Orlando, Florida, 2003.
[14] F. Alizadeh-Shabdiz, S. Subramaniam, "MAC Layer Performance Analysis of Multi-Hop Ad Hoc Networks," *The IEEE Global Telecommunications Conference* (*Globecom*) 2004.
[15] D. Bertsekas, R. Gallager, "Data Networks, second edition," *Prentice-Hall International Editions*, 1987.
[16] H. Takagi, L. Kleinrock, "Optimal Transmission Range of Randomly Distributed Packet Radio terminals," *IEEE Transactions on Communications*, vol. COM-32, No. 3 1984, 246-57.
[17] Scalable Network Technologies, "Qualnet simulator version 3.6," ww.scalable-networks.com.