Performance Analysis of IEEE 802.11 DCF under Limited Load

Yong Shyang Liaw, Arek Dadej, Aruna Jayasuriya
Institute for Telecommunications Research, University of South Australia
Yong.Liaw@postgrads.unisa.edu.au, Arek.Dadej@unisa.edu.au, Aruna.Jayasuriya@unisa.edu.au

The goal of effective distributed network control is that each wireless node operates near its maximum capacity, with bounded queue delay. This "optimum" operating point, where maximum throughput is achieved, is usually located well below the saturation point. The choice of optimum operating point requires knowledge of MAC performance below saturation, hence a limited load model of MAC operation needs to be studied for useful insights. In this paper, we extend the analytical model of IEEE 802.11 DCF from [2] to study throughput performance under limited load, where the node's transmit queue may be at times empty.

I. INTRODUCTION

The performance of IEEE 802.11 DCF MAC has been well studied under the saturation conditions, where each node always has packets in its transmission queue [2, 3, 4, 7]. In [2], an analytical Markov model is used to study the IEEE 802.11 DCF under saturation. Alternatively, [3, 4] use a model based on average values to provide a closed-form solution for saturation throughput. These works provide useful insights into the IEEE 802.11 performance and in performance tuning of the MAC protocol [7, 8]. However, all are based on the assumption of saturation, and fail to capture the throughput behaviour below saturation.

A robust analytical model is required to study the DCF throughput under a full range of traffic loads. The benefits of such studies include traffic admission policies suitable for use at individual nodes, under fully distributed network control. It is with this in mind that our analytical model is built; we are particularly interested in expressing the throughput performance as a function of queue utilisation measurable locally at every node. Furthermore, we extend our analysis to account for the presence of hidden terminals [5, 10].

There are some works on finite load models for IEEE 802.11 [6, 9, 11]. However, these models are either based on simplifying assumptions that limit their use near the optimum throughput point, or use variables not measurable locally at a node, hence are not suitable for distributed network control.

The main contribution of our work is the limited load Markov model of IEEE 802.11, and its analysis based on variables observable locally at each node. We offer insight into the throughput behaviour of 802.11 over a wide range of traffic loads. The paper is organized as follows. Section II presents the model and its analysis. Section III presents numerical and simulation results. We conclude in Section IV.

II. LIMITED LOAD ANALYTICAL MODEL & ANALYSIS

A. Analytical Model

We extend the discrete-time Markov model from [2] by adding another state (i.e. E) to represent the node with empty queue, as shown in Fig. 1. The parameters used in the model are summarised in Table 1. Time is divided into slots of duration σ. A packet is discarded after m failed retransmissions. We assume that unsuccessful transmissions are solely due to collisions (perfect channel). The model does not account for packet capture effect. Similarly as in [2], a key assumption in this model is that collisions happen with the same probability p regardless of the number of retransmission attempts.

When a packet arrives in state $b_{E,}$, the system enters with transition probability p_{e0}/W_0 one of the back-off states $b_{0,i}$. The p_{e0} depends on the packet arrival distribution function. At i^{th} back-off stage, the back-off timer assumes a random value k chosen from {0, 1, 2, …, W_i-1} with uniform distribution, where W_i is given by (1). Each time slot, the back-off timer is decremented by 1, and when zero is reached, the packet is transmitted. When the transmission collides with another transmission (with probability p), the next $(i+1)^{th}$ back-off stage is entered with probability p/W_{i+1}. There is a maximum of m back-offs. At m^{th} failed retransmission, the packet is discarded. After a successful transmission (with probability 1-p) or after m^{th} failed retransmission, the transmission queue is empty with probability q and non-empty with probability 1-q. If not empty, the system enters state $b_{i,k}$ to serve the next packet in the queue. When empty, the system enters state b_E, and waits for a packet arrival. As soon as a packet arrives in state b_E, the system enters $b_{0,k}$ to contend for the channel.

$$W_i = 2^i W_0 \tag{1}$$

B. Transmission probability in a slot

To solve for p, we balance the equations from the Markov chain in Fig. 1 and obtain the following equations for k ϵ {0, …, W_i -1}. It shall be noted that y in Fig.1 denotes not a state, but the sum of the probabilities of successful transmission and discarding a packet, as given in (4).

TABLE I. PARAMETERS FOR THE MODEL

p	Probability a transmission is unsuccessful/collided.
q	Probability an empty queue is found after a packet is served.
$b_{i,k}$	State being in i[th] back-off stage, with back-off timer of value k.
b_E	State being in queue empty state.
$P_{i,k}$	Stationary probability being in state $b_{i,k}$.
P_E	Stationary probability being in state b_E.
p_{e0}	Transition probability from state b_E to a non-empty state.
W_0	The initial maximum contention window.
W_i	Maximum contention window at i[th] back-off stage.
m	The maximum number of retransmissions allowed.

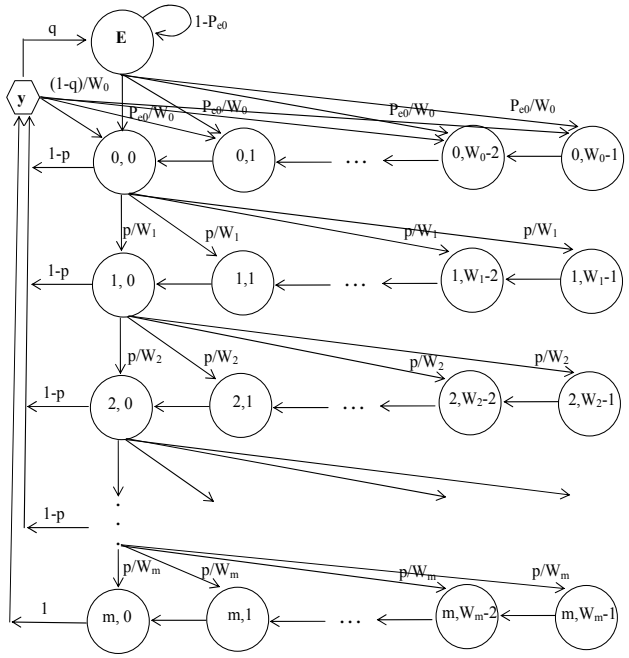


Figure 1. IEEE 802.11 Markov model for limited load

$$P_{0,k} = \frac{W_0 - k}{W_0} \cdot \left[p_{e0} . P_E + (1 - q) y \right] \quad (2)$$

$$P_{i,k} = \frac{W_i - k}{Wi} . p^i . P_{0,0} \qquad \text{for } 0 < i \le m \quad (3)$$

$$y = \sum_{i=0}^{m} (1 - p) . P_{i,0} + p . P_{m,0} = P_{0,0} \quad (4)$$

$$P_E \cdot p_{e0} = q \cdot y \quad (5)$$

With (4) and (5), (2) and (3) can be rewritten as (6).

$$P_{i,k} = \frac{W_i - k}{Wi} . p^i . P_{0,0} \quad (6)$$

for i ϵ {0, …,m}, k ϵ {0,…, W_i -1}

From the normalising condition and (6), we obtain an expression for $P_{0,0}$ in terms of p and P_E.

$$\sum_{i=0}^{m} \sum_{k=0}^{W_i - 1} P_{i,k} + P_E = 1 \Rightarrow P_{0,0} = \frac{2(1 - P_E)}{\beta} \quad (7)$$

$$\beta = \frac{1}{2} \sum_{i=0}^{m} \sum_{k=0}^{W_i - 1} p^i \frac{W_i - k}{W_i} = W_0 \frac{1 - (2p)^{m+1}}{1 - 2p} + \frac{1 - p^{m+1}}{1 - p}$$

for $0 \le p < 1$.

The node will only transmit in state $b_{i,0}$, hence the probability p_{tx} that a node will transmit in an arbitrary time slot is the probability of being in any of the states $b_{i,0}$.

$$p_{tx} = \sum_{i=0}^{m} P_{i,0} = \varepsilon P_{0,0} \text{ where } \varepsilon = \sum_{i=0}^{m} p^i = \frac{1 - p^{m+1}}{1 - p} \quad (8)$$

We note that (6), (7) and (8) are independent of q and p_{e0}, i.e. independent of packet arrival distribution. So far, (6) and (8) are equivalent to expressions obtained (in terms of $P_{0,0}$) in [2] for a saturated system.

C. *Solving for probability of collision, p*

1) *No Hidden Terminals*

In a neighborhood with N identical nodes where every node can sense other nodes' transmissions (no hidden terminals), a transmission results in a collision when any of the (N-1) neighbors also transmits in the same time slot. Eq. (8) must satisfy the constraint given by (9), and the system may be solved numerically for a given value of P_E.

$$p = 1 - (1 - p_{tx})^{N-1} \Rightarrow p_{tx} = 1 - (1 - p)^{1/(N-1)} \quad (9)$$

2) *Hidden Terminals*

Let us assume that in addition to the (N-1) direct neighbors, there are, on average, N_{ht} hidden terminals in the region, as in Fig. 2. A collision with a neighboring node can occur only when the nodes transmit in the same slot, otherwise the nodes would be able to sense busy channel. However, a hidden terminal cannot sense our node, hence can begin the colliding transmission in any slot. Therefore, for a successful transmission, the N-1 neighbors must not transmit in the same slot, and the N_{ht} hidden terminals must not transmit during the entire transmission period.

Let L_{tx} be the average number of slots in a transmission period (t_{tx}), and ξ be the average slot duration (i.e. average time spent in a state). We categorise states in the Markov model in Fig. 1 into 3 types, empty-queue state (EQ), transmission state (TX) and backoff state (BO). The average time spent in each state type before a transition occurs is different. Let t_{eq}, t_{tx} and t_{bo} be the average times spent in states of type EQ, TX and BO respectively (they are functions of p (and P_E), given in the later section). Then L_{tx} and ξ are given by (10), and p_{tx} by (11). Eq. (8) must satisfy constraints given by (10) and (11).

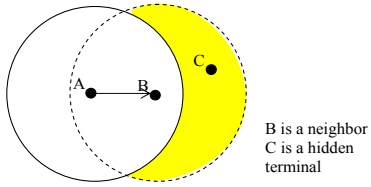


Figure 2. Transmission in the presence of hidden terminal

$$L_{tx} = \frac{t_{tx}}{\xi}$$

$$\xi = P_E t_{eq} + p_{tx} t_{tx} + t_{bo} \sum_{i=0}^{m} \sum_{k=1}^{W_i - 1} P_{i,k} \qquad (10)$$

$$= P_E t_{eq} + (\varepsilon t_{tx} + \frac{\alpha}{2} t_{bo}) p_{0,0}$$

$$\alpha = \frac{1}{2} \sum_{i=0}^{m} \sum_{k=1}^{W_i - 1} p^i \frac{W_i - k}{W_i} = W_0 \frac{1 - (2p)^{m+1}}{1 - 2p} - \frac{1 - p^{m+1}}{1 - p}$$

$$p = 1 - (1 - p_{tx})^{N - 1 + L_{tx} N_{ht}} \qquad (11)$$

$$\Rightarrow p_{tx} = 1 - (1 - p)^{1/(N - 1 + L_{tx} N_{ht})}$$

For uniform node distribution over a large area with node density D, it can be shown that N_{ht} is given by (12). Eq. (8), (10), (11) and (12) may be solved for p iteratively. We note that $N_{ht} = 0$ is a special case without hidden terminals.

$$N_{ht} = \frac{3\sqrt{3}}{4} R^2 D = \frac{3\sqrt{3}}{4\pi} N \qquad (12)$$

D. Packet Service Rate

Next, we find an expression for μ, the packet service rate (rate at which packets are removed from the queue), in terms of P_E and p. Let t_{access} be the average time spent backing off before a packet is transmitted (i.e. MAC access time), and t_{tx} be the packet transmission time (i.e. the slot time for TX state). Since at i^{th} back-off stage the average number of back-off slots is $W_i/2$, and the system enters the next back-off stage with probability p, the average access time t_{access} is given by (13).

$$t_{access} = \frac{\sum_{i=0}^{m} p^i (\frac{W_i}{2} t_{bo})}{\sum_{i=0}^{m} p^i} = \frac{\phi W_0 t_{bo}}{2\varepsilon} \qquad (13)$$

$$\text{where } \phi = \sum_{i=0}^{m} p^i W_i = \frac{1 - (2p)^{m+1}}{1 - 2p}$$

The transmission time t_{tx} depends on the payload size and the access mechanism (i.e. basic access or RTS/CTS). Note that the RTS/CTS mechanism combats the hidden-terminal problem, but amplifies the exposed-terminal problem, with a potentially greater performance degradation effect than hidden terminal [10]. For this reason, we limit our analysis to the basic access mechanism. For basic access mechanism, t_{tx} is given by (14), where H is the packet header, PL the payload size, τ propagation delay, and t_{succ} and t_{coll} the times taken by successful and unsuccessful transmissions respectively.

$$t_{tx} = (1 - p) t_{succ} + p t_{coll}$$
$$t_{succ} = H + PL + SIFS + DIFS + ACK + 2\tau \qquad (14)$$
$$t_{coll} = H + PL + DIFS + \tau$$

At the beginning of a BO-slot, the node senses either an idle channel with probability P_{n-1}, or a busy channel followed by an idle slot with probability (1- P_{n-1}). The average busy period is the transmission period t_{tx}, hence the average time spent in a BO state, t_{bo}, is given by (15).

$$t_{bo} = 1 - P_{n-1} + P_{n-1} t_{tx}$$
$$P_{n-1} = 1 - (1 - p_{tx})^{N-1} \qquad (15)$$

The average time taken to transmit a packet, t_s, is the total of the transmission time and MAC access time, as given in (16), and the average transmission rate (μ) is $1/t_s$. However, only (1-p) proportion of packets transmitted are successful, p_{disc} proportion are discarded after maximum number of retransmissions (m), and the rest (i.e. p-p_{disc}) result in a collision after which the packet is re-inserted into the queue. Hence the effective packet service rate (μ_{eff}) is a sum of the rate of successful transmissions (μ_{succ}) and the rate at which packets are discarded (μ_{disc}), as in (17). For small p and $m > 1$, p_{disc} is negligible, hence we may approximate μ_{eff} with μ_{succ}.

$$t_s = t_{access} + t_{tx} \qquad (16)$$

$$\mu_{eff} = \mu_{succ} + \mu_{disc} \qquad (17)$$

$$\text{where } \mu_{succ} = \frac{1 - p}{t_s},\ \mu_{disc} = \frac{p_{disc}}{t_s}$$

$$p_{disc} = \frac{p . P_{m,0}}{\sum_{i=0}^{m} P_{i,0}} = \frac{(1 - p) p^{m+1}}{1 - p^{m+1}} P_{0,0}$$

E. Throughput

For limited load where the arrival rate λ is less than μ_{eff}, the node's throughput (T_{node}) is the portion of traffic that arrived minus the portion that is discarded i.e. λ*(1-P_{disc}). To derive the throughput of a node, it is necessary to know the distribution of packet service time. Here, we assume an idealized M/M/1/k_q model for the transmission queue, hence the packet arrival rate (λ), throughput at a node (T_{node}), and the total normalized throughput for a neighborhood of N identical nodes, T_{total}, are given by (18), where ρ is the traffic intensity.

$$\lambda = \rho . \mu_{eff}$$
$$T_{node} = \begin{cases} \rho \cdot \mu_{succ} & for\ \rho < 1 \\ \mu_{succ} & for\ \rho \geq 1 \end{cases} \qquad (18)$$
$$T_{total} = \frac{N \cdot T_{node} \cdot PL}{ChannelRat\ \ e}$$

For M/M/1/k_q queue, ρ is related to the queue utilization as in (19), where k_q is the size of the queue buffer. It is noted that the model in Fig. 1 uses time slot of varying duration, and the probabilities P_E and $P_{i,k}$ are relative frequencies of visiting the corresponding states. Let S_E and $S_{i,k}$ be the equivalent probabilities with respect to arbitrary time. They are given by (20), and specify the distribution of time the node spends in the corresponding states; in particular S_E is the proportion of time the queue is empty, hence (1- S_E) is the queue utilization.

$$S_E = \frac{1 - \rho}{1 - \rho^{k_q + 1}} = 1 - U_q \qquad (19)$$

$$S_E = \frac{P_E t_{eq}}{P_E t_{eq} + P_{tx} t_{tx} + \sum_{i=0}^{m}\sum_{k=1}^{W_i} P_{i,k} t_{bo}} = \frac{P_E t_{eq}}{P_E t_{eq} + (\varepsilon \cdot t_{tx} + \frac{\alpha \cdot t_{bo}}{2}) P_{0,0}}$$

$$S_{i,0} = \frac{P_{i,0} t_{tx}}{P_E t_{eq} + P_{tx} t_{tx} + \sum_{i=0}^{m}\sum_{k=1}^{W_i} P_{i,k} t_{bo}} = \frac{P_{i,0} t_{tx}}{P_E t_{eq} + (\varepsilon \cdot t_{tx} + \frac{\alpha \cdot t_{bo}}{2}) P_{0,0}} \quad (20)$$

$$S_{i,k} = \frac{P_{i,k} t_{bo}}{P_E t_{eq} + P_{tx} t_{tx} + \sum_{i=0}^{m}\sum_{k=1}^{W_i} P_{i,k} t_{bo}} = \frac{P_{i,k} t_{bo}}{P_E t_{eq} + (\varepsilon \cdot t_{tx} + \frac{\alpha \cdot t_{bo}}{2}) P_{0,0}} \quad \text{for } i \geq 1$$

$P_E.t_{eq}$ is the proportion of time the queue is empty in an average time slot. Eq. (20) gives a unique solution for any given value of λ, for $0 < t_{eq} < 1/\lambda$. Then, for simplicity, a solution is obtained when $P_E = S_E$, and t_{eq} is given by (21).

$$S_E = P_E \Rightarrow t_{eq} = \frac{(\varepsilon \cdot t_{tx} + \frac{\alpha \cdot t_{bo}}{2}) P_{0,0}}{1 - P_E} \quad (21)$$

F. Queue Delay

For M/M/1/k_q queue, the queue delay is given by

$$t_{queue} = \left(\frac{\rho}{1-\rho} - \frac{(k_q+1)\rho^{k_q+1}}{1-\rho^{k_q+1}} \right) \frac{1}{\lambda} \quad (22)$$

III. Simulations & Results

Table II shows the default parameters used for both simulations and numerical evaluation. U_q is approximated by ρ, since a buffer size (k_q) of 60 packets is assumed, and T_{node} is taken as equivalent to λ since the p_{disc} is negligible in (18).

A. Without Hidden Terminals

Fig. 3 shows the total system throughput (T_{total}) in respect to queue utilization (U_q) at the node. The point of maximum throughput is at the saturation point (i.e. $U_q = 1$) only for less than 10 nodes in the neighborhood. For $N \geq 10$, the maximum throughput is achieved below the saturation point (i.e. $U_q < 1$). The maximum throughput point shifts towards lower U_q with increasing N. This result agrees with the results in [11].

Fig. 4 shows that maximum throughput degrades slower than the saturation throughput, and approaches an asymptotic value with large N. Significant benefits can be obtained through distributed control of queue utilization at each node.

TABLE II. VALUES FOR SIMULATION & NUMERICAL EVALUATION

Payload size (PL)	8184 bits
MAC header	272 bits
PHY header	128 bits
ACK	112 bits + PHY header
Channel Bit Rate	1 Mbit/s
Propagation delay (τ)	1 μs
Slot time (σ)	20 μs
SIFS	10 μs
DIFS	50 μs
Initial contention window (W_0-1)	31
Maximum number of retransmissions (m)	5

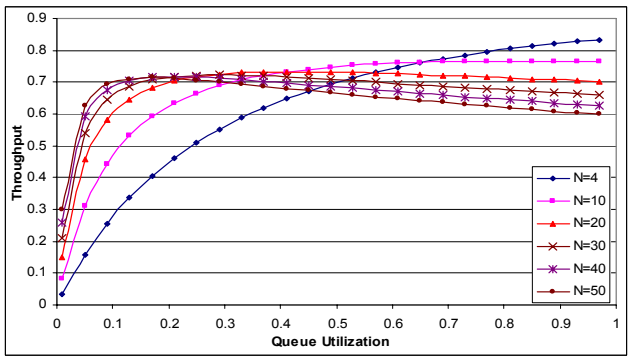


Figure 3. System throughput versus queue utilization

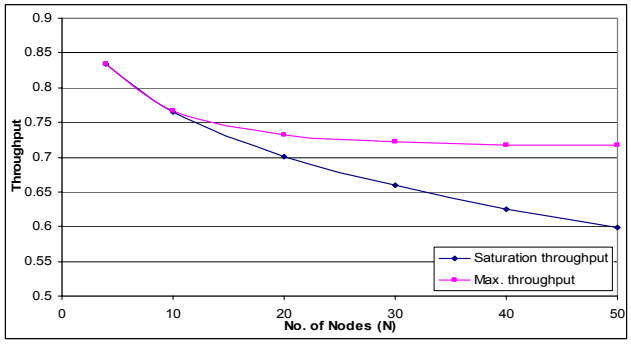


Figure 4. Maximum and saturation throughput

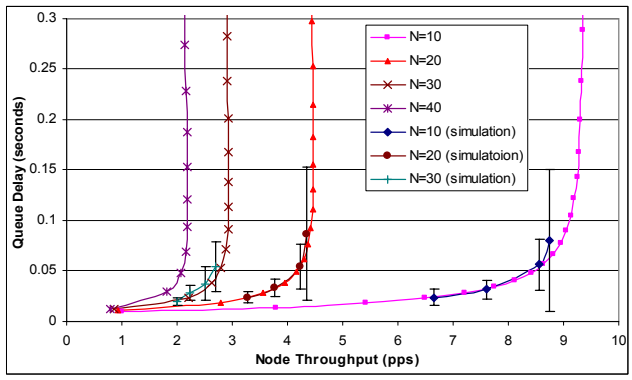


Figure 5. Queue delay

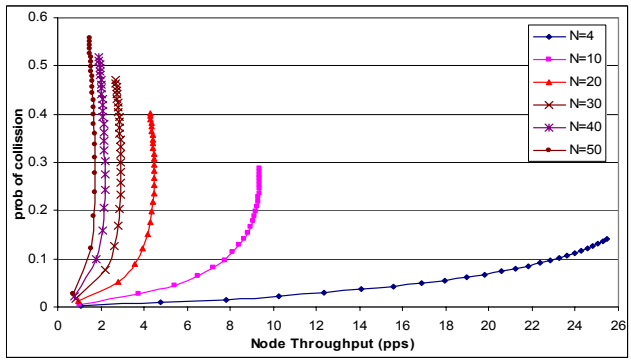


Figure 6. Transmission collission probability

U_q is very sensitive to the traffic arrival rate near the maximum throughput point. A small increase in λ will cause a significant increase in U_q, and consequently a dramatic increase in queue delay as shown in Fig. 5. This can be explained by sharp increase in collision probability when the throughput exceeds certain threshold, as illustrated in Fig. 6. As a result, operating at saturation point is not desirable as it dramatically increases queue delay and reduces throughput for larger numbers of nodes. For maximum throughput with low delay, the "optimum" operating point for IEEE 802.11 MAC is at the "knees" of the delay curves shown in Fig. 5.

B. *Hidden Terminals*

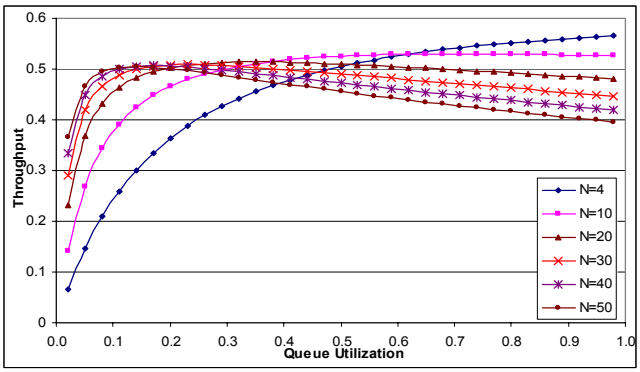


Figure 7. Throughput performance in the presence of hidden terminals

Fig. 7 depicts throughput performance of the IEEE 802.11 in the presence of hidden terminals. As expected, the maximum and saturation throughputs are lower than those without hidden terminals. Maximum and saturation throughput start to differ significantly at N=10. For a given N, the maximum throughput occurs at only slightly lower U_q than without hidden terminals. We conclude that the main effect of hidden terminals is throughput reduction and shifting of the maximum throughput point towards lower values of U_q.

C. *Validation of the Model*

To verify the analytical model, simulations were performed using OPNET to observe the maximum throughput point. To avoid capture effect, all N wireless nodes were placed at the same location. Each node generated packets at Poisson arrival rate λ, and transmitted to a node over 1 hop. The total system throughput was measured over a period of 60 seconds for a range of arrival rates. The same parameters as in Table II were used. The transmission buffer was set to 60 packets.

Fig. 8 shows the results. For N=4, the maximum throughput occurs at saturation, and for $N \geq 10$ the throughput peaks before the saturation point. The maximum throughput closely matches that from the analytical model. At 90% of maximum arrival rate, queue utilization is relatively low, close to the analytical results, as shown in Table III.

IV. CONCLUSIONS

We have presented an analytical model to study the performance of IEEE 802.11 DCF under a full range of load conditions, in respect to queue utilization before saturation. We have shown that for large numbers of nodes the maximum throughput is achieved well before the saturation, and have validated our model via simulations. Our model offers a useful method of estimating the optimum operating point for the purpose of distributed network control, based on U_q.

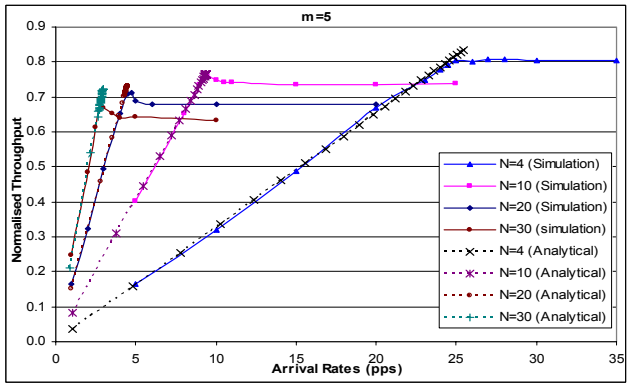


Figure 8. Throughput performance of IEEE 802.11

TABLE III. ESTIMATED QUEUE ULTILIZATION

N	U_q (Analytical)	U_q (Simulation)
10	0.33±0.04	0.31±0.09
20	0.17±0.04	0.18±0.05
30	0.09±0.04	0.082±0.029

REFERENCES

[1] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specificatoions", IEEE standard, 1999 Edition.

[2] Guiseppe Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordinated Function", IEEE J-SAC, Vol. 18, No 3, March 2000.

[3] Y. C. Tay, K. C. Chua, "A Capacity Analysis for the IEEE 802.11 MAC Protocol", Wireless Networks, 7, Jan 2001.

[4] J. Yeo, A. Agrawala, "Packet Error Model for the IEEE 802.11 Mac Protocol", IEEE International Symposium on personal, Indoor and Mobile Radio Communications (PIMRC), Sep 2003.

[5] F. A. Tobagi, L. Kleinrock, "Packet Switching in Radio Channels: Part II – The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution", IEEE Transactions on Communications, Vol. COM-23, No. 12, 1975.

[6] F. Alizadeh-Shabdiz, Suresh Subramaniam, "A Finite Load Analytical Model for the IEEE 802.11 Distributed Coordination Function MAC", WiOpt'03, March 2003.

[7] Federico Cali, Marco Conti, Enrico Gregori, "IEEE 8-2.11 Protocol: Design and Performance Evalution of an Adaptive Backoff Mechansim", IEEE J-SAC, Vol. 18, Sept 2000.

[8] Haitao Wu, et al, "Performance of Reliable Transport Protocol over IEEE 802.11 Wireless LAN: Analysis and Enhancement", Proc. of INFOCOM'02, June 2002.

[9] O. Tickoo, B. Sikdar, "A Queue Model for Finite Load IEEE 802.11 Random Access MAC", Proc. of ICC 2004, June 2004.

[10] Aruna Jayasuriya, et al, "Hidden vs. Exposed Terminal Problem in Ad hoc Networks", Proc. of Australian Telecommunication Networks and Applications Conference , Dec 2004.

[11] C. L. Fullmer, et al, "Solutions to Hidden Terminal Problems in Wireless Networks", Proc. of ACM SIGCOMM 97, Sept 1997.