An Improved Capacity Modeling Approach for 802.11 Protocols

N. Greyling, R. Wolhuter

University of Stellenbosch, Dept. of E & E Engineering

***Abstract*—Despite the widespread use of IEEE 802.11 Wireless Local Area Network (WLAN) protocol, not much analysis has been done on modeling its performance, to determine what amount of users such a network can accommodate. This paper introduces one approach to such a model, assisting with capacity determination of WLAN networks. We introduce a model to estimate the number of users that such a WLAN can accommodate. The model applies to unsaturated traffic conditions and to all traffic arrival patterns. Further, the model tries to emulate such a network as accurately as possible with moderate simplifications. The model provides an improved method of practical WLAN performance estimation.**

***Index Terms*—Unsaturated sources, 802.11 Markov Model, 802.11 throughput analysis.**

I. Introduction

In recent years, the IEEE 802.11 protocol for WLAN's [1] has become popular for wireless and mobile internet users. Access Points (AP's) can be deployed wherever customers need access to fast and mobile information. The IEEE 802.11 standard specifies the Medium Access Control (MAC) layer as well as the physical (PHY) layer. Currently, the standard defines two access coordination functions for the MAC: the contention based Distributed Coordination Function (DCF) and the contention free based Point Coordination Function (PCF). This paper only considers the DCF access method. It should be mentioned; the PCF access method is not mandatory, and therefore, rarely implemented. The DCF has two operating modes: the basic channel access and the RTS/CTS (Request-to-Send/Clear-to-Send) mechanisms.

Existing work on the performance of the 802.11 MAC has focused primarily on the throughput under saturated conditions [3]. Also, work has been conducted by considering unsaturated conditions such as [4], but this method only considers buffers with one packet. Other analytical models have been developed [5] that relies on statistical methods, but fails to model certain events such as pre and post backoff mechanisms. One of the most comprehensive models [6] includes factors such as the buffer occupancy, the number of active transmitting stations, and models the most possible number of states.

A. Our contribution

It is the intention to extend the model in [6] even further, to evaluate more metrics of interest such as network throughput, average number of wireless stations active at any stage, and the delay each station experiences. Also, a wider variety of user classes of slower and faster wireless nodes is to be incorporated. The proposed model will provide a better understanding and estimation of the performance limits of practical wireless data networks.

II. Basic model for saturated sources

This section describes a basic model for the behaviour of saturated sources, following the same approach as [4] with a few minor simplifications. The main assumptions in [6] are as follows:

- A fixed number of competing stations are considered accessing the same wireless channel. Stations always have data to send, and therefore, they operate under saturated conditions.
- There are no hidden terminals.
- Stations are equally likely to access the channel.
- The communication channel is error-free.

A. Markov analysis

In order to develop an analytical model for a system such as this, a typical choice would be to use a discrete Markov chain, and evaluating possible state changes at discrete time instants. A fundamental assumption made is the state of each station is independent of all others, reducing the analysis to monitoring only the behaviour of a single tagged station. Following this approach, we propose the simple Markov chain depicted in Figure 1, which is a modified version of [3] for the saturated case, and the version used in [6].

A typical depiction of the Markov chain under discussion is presented in Figure I. States labeled with b represents the scenario where the backoff counter has decremented to zero and the station attempts to transmits a packet in the current time step. States labeled B represents the scenario where a stations backoff counter is still decrementing.

Each state has an index {0,1,…,m}indicating the backoff stage, where m is the maximum retry limit for a single packet to be transmitted. Using the same notation as in [4], W_i is the contention window size at backoff stage i.

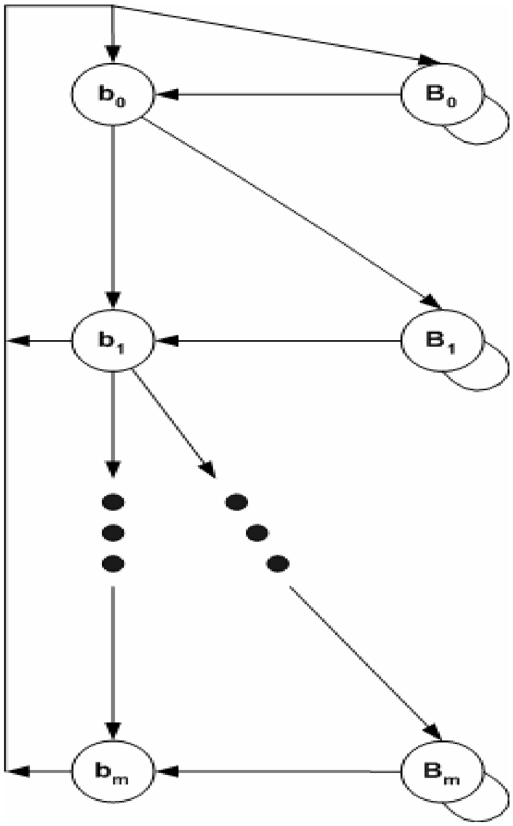


Fig. 1. Markov chain for saturated sources.

We have W_i = min {2^i CW_{min}, CW_{max}}, where CW_{min} and CW_{max} are constant parameters representing the minimum and maximum backoff values.

The remainder of this section is to give a brief overview of [6]. The transition probabilities from state s_i to s_f is $P(s_i, s_f)$, which are presented in Table I, where *i* denotes the initial and *f* the final state. To improve the readability of the transition probabilities we have put α_i =2/W_i and β_i =2/(W_i-1). Note including the factor of 2 in the numerator in α_i and β_i compensates for the picking a backoff value of 0 or 1 has the same effect as starting a transmission at the beginning of the next interval [6].

The only unknown value to be determined from the transition probabilities is *p*, which is the collision probability seen by a station transmitting on the channel.

The stationary distribution of the Markov chain is denoted by $\pi = \{\pi_s\}$ where s is the generic state of the model. This is similar to Bianchi's model, where we have $\pi_{bi} = p\pi_{bi-1}$, for all *i*>0.

As described in [6], all states belonging to backoff stage *i* with a backoff counter value greater than one, have been collapsed into a single state B_i. This has been done to reduce the number of states, which decreases the models complexity as it is extended to compensate for more variables. Using this approach sacrifices little accuracy.

TABLE I
TRANSITION PROPERTIES FOR BASIC MODEL

$$\alpha_i = 2/W_i \quad \beta_i = 2/(W_i - 1)$$

s_i	s_f	$P(s_i, s_f)$	*Condition*
b_i	b_0	(1-*p*) α_0	$0 \le i < m$
	B_0	(1-*p*)(1-α_0)	
	b_{i+1}	$p\alpha_{i+1}$	
	B_{i+1}	$p(1-\alpha_{i+1})$	
b_i	b_0	α_0	i=m
	B_0	1-α_0	
B_i	b_i	β_i	$0 \le i < m$
	B_i	1-β_i	

The result is the backoff is modeled as geometrically distributed random variable instead of a uniformly distributed random variable. To obtain the same performance metrics to that of the collapsed version of the backoff values, with that of the more exact model, having the uniformly distributed random variable, the following condition has to be satisfied,

$$\pi_{Bi} = \pi_{bi} \frac{(W_i - 1)(W_i - 2)}{2W_i}$$

for all values of *i*. Thanks to the particular structure of the Markov chain, all probabilities can easily be computed. All probabilities can be related to π_{b0}, and thus can be calculated by normalizing the overall sum of the probabilities to one.

After which, one can calculate the probability τ that a station transmits in a given time-slot as $\tau = \sum_{i=0}^{m} \pi_{bi}$. To compute the conditional collision probability *p*, the fundamental assumption used is the individual stations are independent. With this assumption it is possible to monitor only the behaviour of a single station. This results in the following expression,

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

The two unknowns to be determined are τ and *p*, and can be calculated by a simple iterative process as in [4].

The probabilities that the channel is occupied with a successful transmission (Π_s), a collision (Π_σ), or an idle slot (Π_c), is respectively computed as follows:

$$\Pi_s = n\tau(1 - \tau)^{n-1}$$
$$\Pi_\sigma = (1 - \tau)^n$$

$$\Pi_c = 1 - \Pi_\sigma - \Pi_c$$

Finally, the aggregate packet throughput *Tp* is determined (expressed in packets/s). This is given by,

$$T_p = \frac{\Pi_s}{\Pi_s T_s + \Pi_c T_c + \Pi_\sigma \sigma}$$

where the denominator, calculates the average duration of a time step.

III. MODEL FOR UNSATURATED SOURCES

In this section we describe the approach to deal with stations that have unsaturated sources. This implies that a stations queue might become empty, and thus the analysis in section II is invalid. The remainder of this section gives a brief overview of [6].

We assume packets arrive from a higher layer at the MAC buffer according to some external process with a rate of λ packets/s. The model is general enough to compensate for any traffic arrival process, but here we used a Poisson arrival process. We also assumed that if a packet arrives at an otherwise empty queue, a new backoff value is chosen. The result is a simpler model, sacrificing little accuracy, especially at higher traffic loads [6].

The resulting model, comprises of the states belonging to the set $\{b_{i,j,k}, B_{i,j,k}\}$, where *b* and *B* have the same meanings as in Section II-B. The three indices $0 \le i \le m$, $0 \le j \le K$ and $0 \le k \le n$, respectively indicate the backoff stage (i), the number of packets currently stored in the buffer (j), and the active number of stations (k). Figures 2 and 3 depict the Markov model. The transition probabilities from state s_i to s_f are $P(s_i, s_f)$, where *i* denotes the initial and *f* the final state, of which some are presented in Table II. The transition probabilities for different channel probabilities (q_{good}, q_{bad}, q_{idle}) have been omitted but are discussed later in this section. The MAC buffer (j) at each station is assumed to be of finite size, with a maximum value of K. If a new packet arrives and the MAC buffer has reached maximum capacity, the packet cannot be stored and will immediately be discarded. Further, the extra index *k* indicates the number of active stations, keeping track of the number of stations having at least one packet in their queue.

The number *k* of active stations may fluctuate during a time step because of the following events: i) one or more of the stations having an empty buffer may receive new packets to be transmitted, increasing *k*; ii) a station that has only one packet successfully transmits, leaving the buffer empty, and decreases *k*. The possibility exists that these two events can occur simultaneously.

The number of stations joining the competing set depends on the duration of the time step Δ and the current number of competing stations *k*. During the interval Δ, the number of stations joining the competing set depends on the probability of q, where the probability of at least on packet arriving at the stations queue is $q = 1 - e^{-\lambda\Delta}$. The number of stations joining the competing set is modeled according to a binomial distribution with parameters (k, q, N), with k indicating the number of active stations, q the packet arrival probability and N the total number of stations present.

The possible time durations Δ directly relates to the probabilities that the channel is occupied by a successful (q_{good}), idle (q_{idle}) or collision (q_{bad}) time-slot. This implies, to compute such probabilities, the probability $\tau(C)$ that a tagged station transmits in a given time step should be determined. Where, C indicates the number of competing stations, and $0 \le C \le n$ (n being the fixed maximum number of wireless stations present).

The probability $\tau(C)$, $C \ge 1$, is given by:

$$\tau(C) = \frac{\sum_{i=0}^{m}\sum_{j=1}^{K}\pi_{bi,j,C-1}}{\pi_{b0,0,C} + \pi_{B0,0,C} + \sum_{i=0}^{m}\sum_{j=1}^{K}(\pi_{bi,j,C-1} + \pi_{Bi,j,C-1})}$$

where, we have C=k if j=0, and C=k+1 if j>0. If C=0, we have $\tau(0)$ =0. Note that, the probability $\tau(C)$ may vary from one state to another, depending on the value of C.

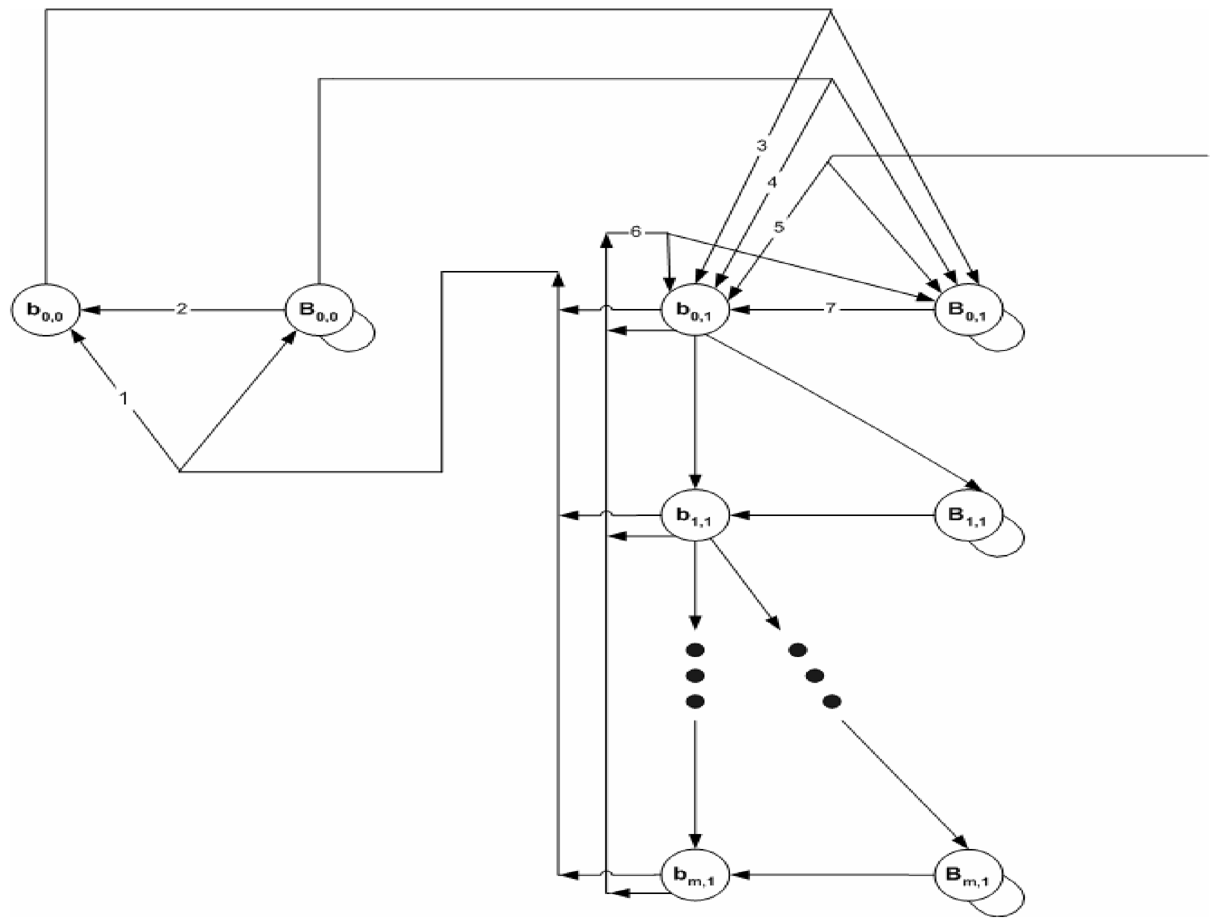
A tagged station transmits in all states, given that j>1. The conditional collision probability is given by

$$p(C) = 1 - [1 - \tau(C)]^{C-1}$$

Thus, with probability 1-p(C), the transmission is successful. The probabilities $\Pi_S(C,k)$, $\Pi_C(C,k)$, and $\Pi_\sigma(C,k)$ respectively indicates the channel is occupied with a successful transmission, a collision, or an idle time-slot. Further, these probabilities are dependant on k and the total number of competing stations C:

$$\Pi_\sigma(C,k) = [1-\tau(C)]^k$$
$$\Pi_S(C,k) = k\tau(C)[1-\tau(C)]^{k-1}$$
$$\Pi_C(C,k) = 1 - \Pi_\sigma(C,k) - \Pi_C(C,k)$$

The last parameter of interest here is P_E, which is the probability that after a successful transmission of a packet, a station other than the tagged one finds its queue empty. This is a critical quantity to our model estimate, as we do not keep track of the buffer occupancy at other stations.



0,0

and $B_{0,0}$ and the transitions from them are clearly illustrated.

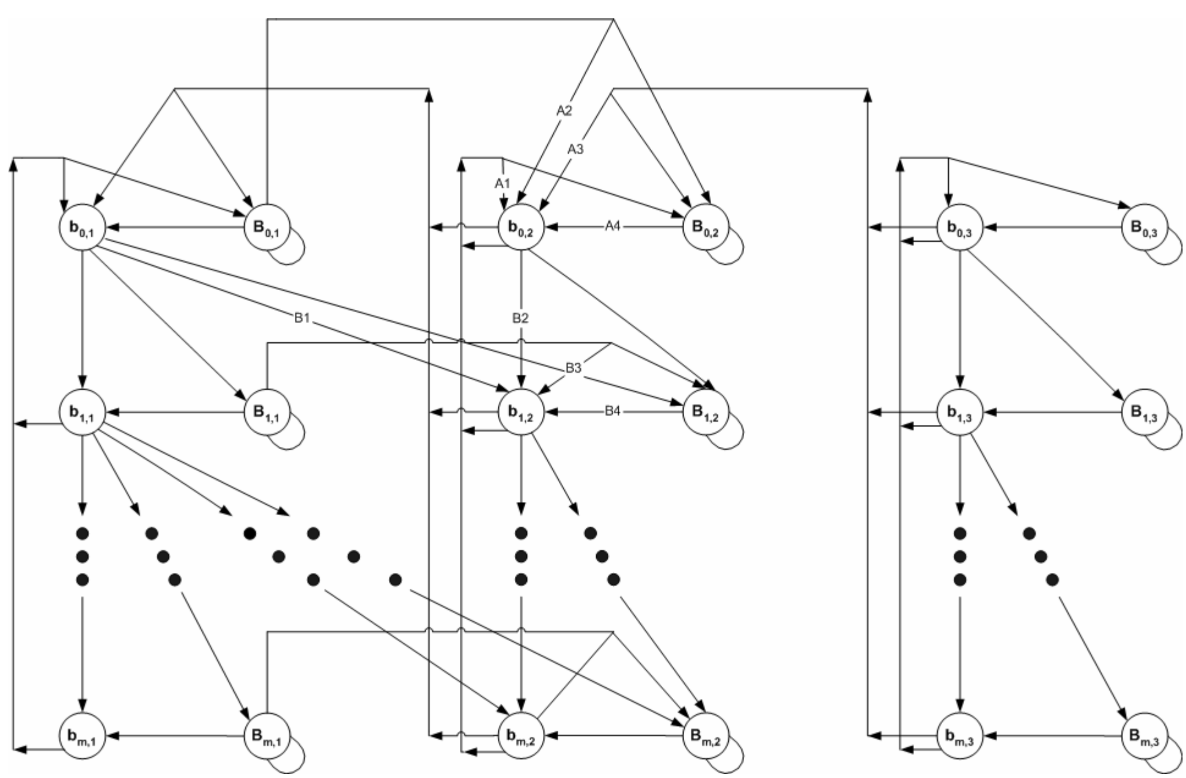


TABLE II
TRANSITION PROPERTIES FOR IMPROVED MODEL

$$\alpha_i = 2/W_i \quad \beta_i = 2/(W_i - 1)$$

s_i	s_f	$P(s_i, s_f)$	Condition
$b_{i,j,k}$	$B_{i,j+1,k+1}$	$(1-\alpha_0)q$	$i=0, j=0$
$B_{i,j,k}$	$B_{i,j+1,k+1}$	$(1-\alpha_0)q$	$i=0, j=0$
$b_{i,j,k}$	$B_{0,j-1,k}$	$(1-\alpha_0)(1-p)(1-q)$	$0 \le i<m$, $0<j<K$, $0<k \le N$
	$b_{0,j-1,k}$	$(1-p)\alpha_0(1-q)$	
	$B_{i+1,j,k}$	$(1-\alpha_0)(1-q)p$	
	$b_{i+1,j,k}$	$\alpha_0(1-q)p$	
	$b_{0,j,k}$	$(1-p)\alpha_0 q$	
	$B_{0,j,k}$	$(1-p)(1-\alpha_0)q$	
	$b_{i+1,j+1,k}$	$p\alpha_{i+1}q$	
$B_{i,j,k}$	$B_{i,j+1,k}$	$(1-\beta_i)q$	$0 \le i<m$, $0<j<K$, $0<k \le N$
	$b_{i,j,k}$	$\beta_i(1-q)$	
	$B_{i,j,k}$	$(1-\beta_i)(1-q)$	
	$b_{i,j+1,k}$	$\beta_i q$	
$b_{i,j,k}$	$b_{0,j-1,k}$	$\alpha_0(1-q)$	$i=m$, $0 \le j \le K$, $0<k \le N$
	$B_{0,j-1,k}$	$(1-\alpha_0)(1-q)$	
	$B_{0,j,k}$	$(1-\alpha_0)q$	
	$b_{0,j,k}$	$\alpha_0 q$	

In this table q represents the probability of a new packet arrival, the use of q_{good}, q_{bad} and q_{idle} was omitted due space limitations. The variable p indicates the conditional collision probability.

The approach used to calculate P_E is dependent on the number of competing stations C and the backoff stage i. A good approximation by [6] is whenever the tagged station finds itself at backoff stage i, and other stations are at backoff stage h, h differs from i by at most one, that is $|h-i| \le 1$. The assumption behind this is the backoff stage of the tagged station will not significantly differ from that of the other competing stations. The estimate of P_E is defined as follows,

$$P_E(C,i) = \frac{\sum_{h:|h-i|\le 1} \pi_{bh,1,C-1}}{\sum_{h:|h-i|\le 1} \sum_{j=1}^{K} \pi_{bh,j,C-1}} e^{-\lambda T_S}$$

This model is not only capable of computing averages, but also the entire distribution of the queue.

IV. RESULTS

TABLE III
MODEL PARAMETER SETTING FOR MAC AND PHYICAL LAYER

SIFS	10 μs
DIFS	50 μs
EIFS	634 μs
σ	20 μs
Basic rate	2 Mbps
Data rate	11 Mbps
PLCP length	1 Mbps
MAC header (RTS, CTS, ACK, DATA)	(20, 14, 14, 28) bytes at Basic rate
Packet payload	500 bytes
CW_{min}, CW_{max}	(32, 1023)
Short retry limit	6
Long retry limit	3

To validate our analytic model, a simulation was done in C++ of which the results are printed here. The model is to undergo further simulation to compare the results to ns-2 [7] for different network topologies, number of nodes as well as the load on the network. Although, at the time this paper was done the ns-2 simulations were not finished, the results matched those in [6].

The parameters for the MAC and physical layers that were used are specified in table III.

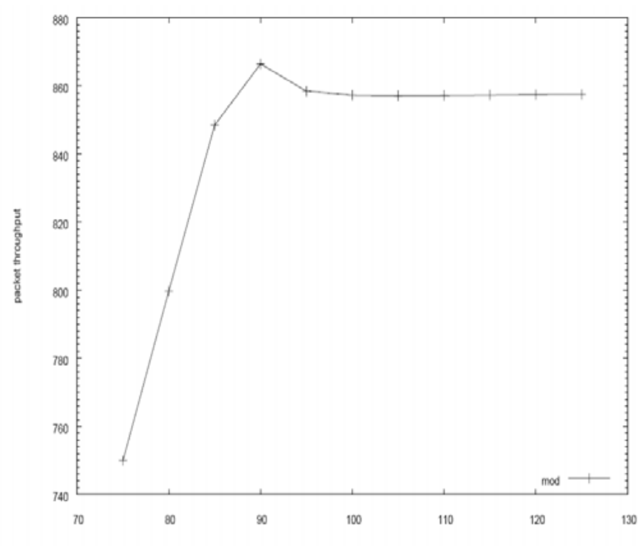


Fig. 4. Packet throughput as a function of the packet arrival rate (Pkts/s).

Figure 4 illustrates a typical metric were the packet arrival rate is varied and the packet throughput is measured. Note, in this figure the packet arrival rate is per station and 10 stations were present, thus the total packet arrival rate will be 10 times this. The model captures important features such as,

- the linear relationship between the offered load and aggregate throughput at low traffic arrival rates (smaller than 80 packets/s for specific example) which corresponds to the system being unsaturated.

- the limiting behaviour at higher traffic (larger than 90 packets/s for specific example) loads which correspond to saturation.
- the complex transition between under loaded condition and saturated conditions (between 80 and 90 packet/s for specific example).

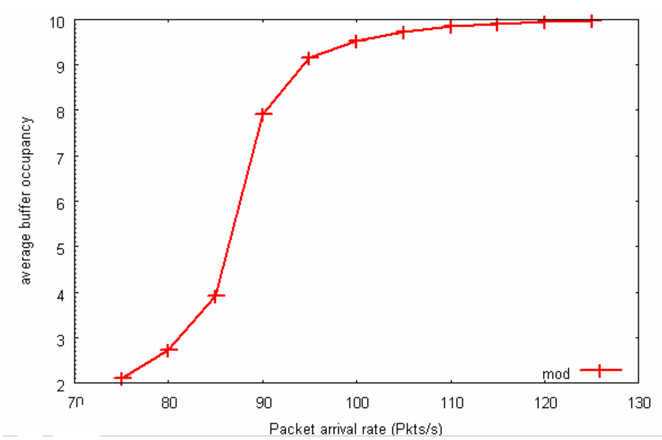


Fig. 5. Average buffer occupancy as a function of the packet arrival rate (Pkts/s).

Figure 5 illustrates the average buffer occupancy when the packet arrival rate varies. Clearly it agrees with the results obtained in Figure 4, because this is also metric used to derive the system change over from unsaturated to saturated conditions.

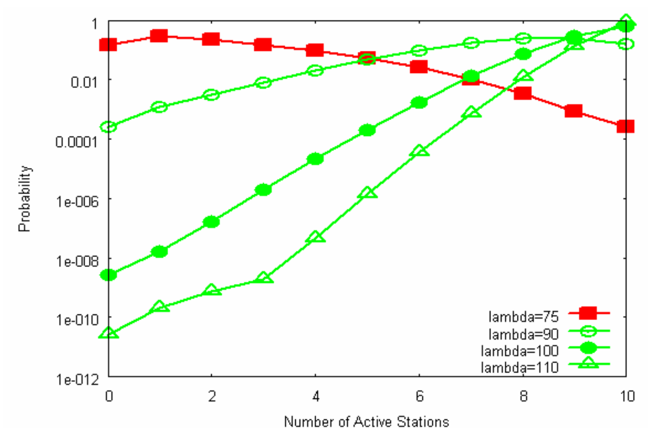


Fig. 6. The probability to find an active number of stations at any given time, as a function of the packet arrival rate.

Figure 6 compliments Figures 4 and 5 further, because the probability of finding more active stations increases as the packet arrival rate (lambda) increases.

V. Conclusion

A model is presented here of IEEE 802.11 MAC layer that, according to a first round of simulations, performs accurately in terms of most network conditions. It predicts the throughput, buffer occupancy, the average number of connections, channel probabilities and packet losses. To our knowledge, this is one of the most accurate and flexible models currently available. This model should be valuable in future network capacity planning, as well as future research.

References

[1] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications ISO/IEC 8802-11:1999(E); ANSI/IEEE Std 802.11, 1999 Edition, USA.

[2] IEEE 802.11 WG, Draft Supplement to Part 11: Wireless Medium Access Control and Physical layer (PHY) specifications: Medium Access Control Enhancements for Quality of Service, IEEE 802.11e/ Draft 3.0, Aug. 2002.

[3] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function", IEEE Journal on Selected Areas in Communications, Vol. 18, No. 3, pp. 535-547.

[4] D. Malone, K. Duffy, and D.J. Leith, "Modelling the 802.11 distributed coordination function in non-saturated heterogeneous conditions", IEEE-ACM Trans. On Networking, vol. 15, No. 1, pp. 159-172, Feb. 2007.

[5] O. Tickoo and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks", in Proceedings of IEEE INFOCOM, pp. 1404-1413, Hong Kong, China, March 2004.

[6] Michele Garetto, Carla-Fabiana Chiasserini, "Performance Analysis of the 802.11 Distributed Coordination Function under Sporadic Traffic", Networking 2005, Waterloo, Canada.

[7] K. Fall and K. Varadhan, editors, "ns notes and documentation", The VINT Project, UC Berkley, LBL, USC/ISI, and Xerox PARC, November 1997.

Neville Greyling received his degree in Electronics Engineering from the University of Pretoria (South Africa) in 2006. Currently he is doing his Masters degree at the Electrical and Electronics department at the University of Stellenbosch (South Africa). His research interests are wireless MAC layer protocol design, wireless communications, and network capacity planning.

Riaan Wolhuter holds a PhD degree in Electronics Engineering from Stellenbosch University (South Africa). He is a senior researcher in Telecommunications Systems and Communications strategies in the Dept of Electrical and Electronic Engineering.