Proceedings of 2009 APSIPA Summit and Conference, Sapporo, Japan, October 4-7, 2009

Performance Analysis for Total Delay and Total Packet Loss Rate in 802.11 MAC Layer

Nguyen Thi Xuan My*, Chaiyachet Saivichit* , and Yoshikazu Miyanaga†
*Department of Electrical Engineering, Chulalongkorn University, Thailand
E-mail: nguyenxuanmy@gmail.com , chaiyachet.s@chula.ac.th
†Hokkaido University, Sapporo 060-8628 Hokkaido Japan
E-mail: miya@ist.hokudai.ac.jp

Abstract— **In this paper, we introduce the analytical model of 802.11 MAC layer to assess system performance under the effect of external elements such as load and the number of stations. Concurrently, we analyze the effect of these elements in term of total delay time and packet loss rate in network with saturated condition, i.e. when every station always has packet to send. In the scope of this paper, we consider the case of unicast only and assume having no hidden terminal phenomenon.**

I. INTRODUCTION

With demand of communication increasing on various types of information including both time sensitive and insensitive operating in wireless media extensively popular nowadays, the performance of wireless infrastructure needs much more improvement to be able to respond such essential requirement. In the network stack model, media access layer (MAC) is one of the most significant parts contributing in the effect of wireless network performance. The designing of protocols such as routing in upper layer has to depend on the ability of MAC layer. Simulation can only express the phenomenon in limited cases without logical explanation. An analytical model for performance of MAC layer is really necessary to show the dependence of performance parameters that designers have a foundation to rely on to design upper systems.

In network performance parameters, time delay and packet loss are very important parameters. They depend on native parameters such as parameters set in protocol, the size of system buffer, the number of competitive stations in the same area, and load. The objective of this work is to produce the model that present the dependence of performance parameters total delay time and packet loss rate on native parameters, so with the suitable adjustment of native parameters, performance of network can be optimized.

802.11 is standard used for wireless today ([4]). There were many research work done on 802.11 MAC protocol analytical models. The models developed in [1], [3] and [5] were used to evaluate the throughput performance in saturated condition. [3] showed the superiority of media access with RTS/CTS mechanism in most cases compared to the basic access method without this kind of overhead. [6] presents a model to evaluate delay and queue length with condition the buffer size is infinitive. Total delay is also considered in [7] for both unicast and broadcast in non-saturation condition.

Our developed model is based on Bianchi's model ([3]) that demonstrates the transition of transmission intending station states under the form of classical Markov chain, but having some modification to appropriate with the objective of the work. Bianchi's model assumes the number of retransmission for each packet can run to infinite without packet dropping. Our model only considers the finite retransmissions and assesses the packet drop rate in condition native parameters change. Further, we also show the effect of native parameters to the service time for transmission of a packet, the total delay time since packet arrival until it is successfully received by target station, and total packet loss as the cause of full system queue and the number of retransmissions run out. Because what we concern is the performance in boundary condition, the model is set for saturation status of system when every station always having packet to transmit. Media access with RTS/CTS has significant improvement for hidden terminal, which is a common phenomenon in wireless network, so it is chosen to be analyzed. In the scope of this work, we only consider the case of unicast. Multicast case will be mentioned in another.

Before coming to the analytical model, we have a short brief about medium access mechanism of 802.11

II. MEDIUM ACCESS MECHANISM OF 802.11 FOR WIRELESS

The fundamental access method of the IEEE 802.11 MAC is the *distribution coordination function* (DCF) ([4]), and this is the mechanism that is analyzed in our model. The fundamental of DCF is *carrier sense multiple access with collision avoidance* (CSMA/CA). Unlike CSMA/CA mechanism in wireline when successful or collision access is detected by the sense of sending station itself, the wireless sender has to send a *request to send* (RTS) frame to its receiver and only assure its access successfully after receiving the *clear to send* (CTS) frame that is replied by its receiver in given interval of time. This procedure is used to solve the hidden terminal problem in wireless where the receiver announce for other stations around about its current receiving status that every station not in transmission range of current sender but in range of receiver will detain its intended transmission to avoid collision occurring in the receiving node. Data frame will be sent right after the successful medium access without collision by transmissions from other stations. The transmission is completed after the sender receive *acknowledge* (ACK) frame from its receiver.

Thus, the contention among intended sending station will occur in medium access time. A packet when arrives at the time of channel idle, the station will send RTS frame to access medium immediately. Otherwise, it will detain until the channel comeback idle state, and then experience a back-off period before accessing the channel again. Back-off is a mechanism that the station chooses a value randomly from an interval *[0, W-1]*, in which *W* is known as the size of back-off window, to be the initial value for its counter. If the channel is idle, the counter will be reduced *one* after each certain period time. Reversely, if the channel is sensed busy, the counter will be detained until the channel back to idle. When the value of the counter reaches *zero*, RTS transmission will be taken immediately without sensing the channel is busy or not. If the access fails because of collision with other station, the procedure is reset from back-off period. The retransmission is allowed in limit number of times. When reaching the maximum number of transmission, but packet is still not delivered, it will be dropped. After each time of transmission but not success, the size of back-off window will be doubly increased until it reaches maximum size (Fig. 1). Beyond that value, the size of window will be kept constant. After finishing transmission of a frame, the station experiences a back-off period again before comeback the idle state or start another new transmission with the minimum back-off window size.

In 802.11 standard, the station suppose the channel idle when it senses no transmission from other stations on the channel at its location after an interval time of DCF interframe space (DIFS). Frames in the same transmission session are spaced by short interframe space (SIFS). Because the model is used for saturation condition, so in order to guarantee the fairness among stations, our model concern the case one DATA frame per transmission only. That means if a transmission is successful, it will have a sequence of frames RTS, CTS, DATA, ACK with minimum space SIFS between them. Fig. 2 illustrates this procedure.

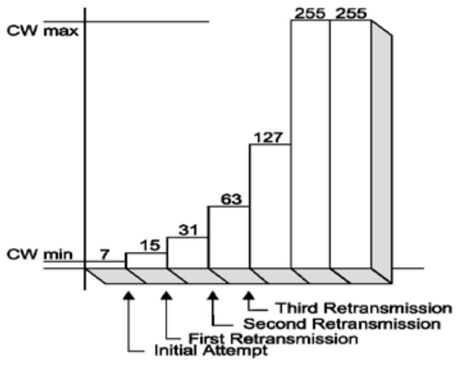

Fig. 1 Example of exponential increase of back-off window

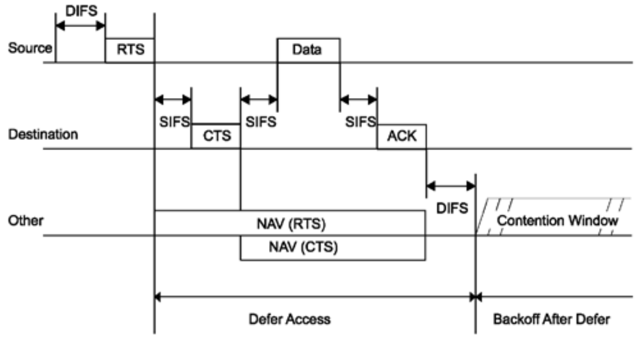

Fig. 2 Media access mechanism with RTS/CTS

III. ANALYZED MODEL

As mentioned above, this model is built based on Bianchi's model ([3]), using 2-D Markov chain to describe the state transformation of a station in condition of saturation that the station always have packet to send (Fig. 3). Each state in the model *B(i, j)* is representative by combination of two indexes, in which *i* is the order number of rows that denotes the current number of retransmission stage, *j* is order number of column that denotes the current value of the back-off counter. The state transformation is not based on normal time system. It uses the unit for each transformation as the back-off slot. In each back-off slot, the channel can be idle or busy, so the time size of each slot can be different. After each transmission regardless of success or collision, because all stations always detains their transmission in at least the interval time of DIFS to assure the channel idle, the chance for the back-off counter to reduce its value is 1. As diagram presentation: *m* is the maximum number of retransmission for each packet; W_i is the size of back-off window in stage of retransmission *i*; W_0 is the minimum window size and W_k is the maximum window size. That mean with *i* not greater than *k*, $W_i = 2^i W_0$; otherwise, $W_i = 2^k W_0$.

A new transmission starts at the back-off stage $i = 0$. The value *j* in every back-off stage is chosen randomly in the respective window size interval $[0, W_i - 1]$. The station always transmit *RTS* frame to access the medium when it reaches state *B(i, 0)*. If the transmission is success, it comes back to stage *0* to start the procedure of new transmission for another packet; otherwise, it continues current transmission with higher order stage. For the simplicity of analysis, on each attempt to access the medium at state *B(i, 0)*, the probability of collision is assumed to be independent and the same in every transmission stage of every station. Because the value for initializing the back-off counter is chosen uniformly randomly, with collision probability *p*, state transition probability from state *B(i-1, 0)* to next higher stage states is identical of p/W_i. In the other hand, if transmission is successful, it comes to states at the first transmission stage 0 with successful probability $(1-p)/W_0$. Stage *m* is the last stage for the retransmission of a packet. After this stage if packet is still not delivered, it will be dropped and the state is transformed to initial stage state with probability $1/W_0$. In equilibrium, applying global equation system, we found the probability of state B(i, j) as following:

$$B(0,0) = \frac{2}{\frac{(W_k+1)(p^{k+1}-p^{m+1})}{1-p}+\frac{(W_0-2W_k p^{k+1})}{1-2p}+k+1} \quad (1)$$

$$B(i,j) = \frac{W_i-j}{W_i} p^i B(0,0) \quad (2)$$

Because the station attempts to access medium at state *B(i, 0)*, the transmission probability τ of the station is the summation of the probability of those states:

$$\tau = \sum_{i=0}^{m} B(i,0) = \frac{1-p^{m+1}}{1-p} B(0,0) \quad (3)$$

From [3], collision probability can be calculated from transmission probability as following:

$$p = 1-(1-\tau)^n \quad (4)$$

With *n* is the number of other competitive stations around in the same area. The equation system of (3) and (4) allows us to find the value of τ and *p* base on parameters W_0, *m*, *k* and *n*.

Thus, with the known value of W_0, *m*, *k* and *n*, the information of τ and *p* are available. Based on this, we can find the other parameters. In this work, we will use this model to find the service time, total delay time, and packet loss caused by full system queue and dropped when beyond the number of retransmission.

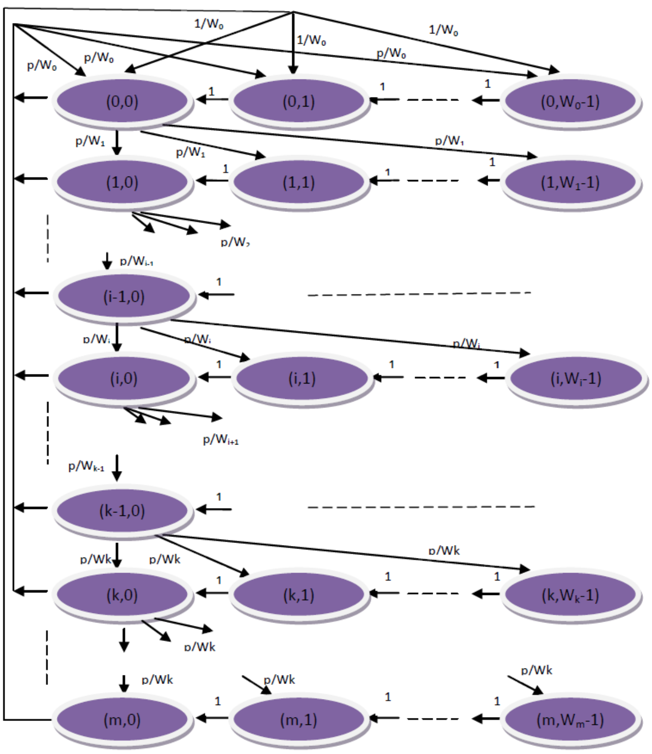

Fig. 3 Back-off state transition model

IV. System Performance Parameters

A. Service time and packet drop rate.

We know that a successful transmission will include a sequence of exchanging consecutive frames *RTS*, *CTS*, *DATA*, *ACK* spaced at least of SIFS time. Therefore, the size *ts* of back-off slot having successful transmission is:

$$t_s = T_{RTS} + \sigma + SIFS + T_{CTS} + \sigma + SIFS + T_{DATA} + \sigma + SIFS + T_{ACK} + \sigma + DIFS \quad (5)$$

In which, T_{RTS}, T_{CTS}, T_{CTS}, , T_{DATA}, T_{ACK} are the time to transmit frames RTS, CTS, DATA, ACK , respectively. T_{DATA} is taken as the average time of DATA frames; it depends on the length of packet. σ is the propagation delay time for each frame transmission. *DIFS* is the time that stations have to wait after the transmission finish for the transition to their next state. If the transmission is unsuccessful by collision, the *CTS* transmission will not happen, and the time for back-off slot having collision is:

$$t_c = T_{RTS} + \text{DIFS} \quad (6)$$

Seen from a station when in back off state *B(i , j)*, with $j \neq 0$, the back-off slot is idle when there is no transmission from the other stations, having a successful transmission when there is transmission from only one of the others, and there is collision when there is transmission from at least two from the others. Therefore, the probabilities corresponding to each case are:

$$P_{idle} = (1-\tau)^n \quad (7)$$

$$P_s = n(1-\tau)^{n-1} \quad (8)$$

$$P_c = 1-(1-\tau)^n - n(1-\tau)^{n-1} \quad (9)$$

As a result, the average size of these back-off slots is:

$$T_b = P_{idle}\,\delta + P_s t_s + P_c t_c \quad (10)$$

In which, δ is the size of back-off slot when it is idle.

For the time spent on each transmission stage, station have to experience *k* back-off slots (*B(i, j)* with *j≠0*) and one transmission slot (*B(i, 0)*). Because *k* is chosen uniformly randomly in the interval *[0, W_i -1]*, average time in each transmission stage T_i is:

$$T_i = \sum_{k=0}^{W_i-1} \frac{1}{W_i} k T_b + T_t = \frac{W_i-1}{2} T_b + T_t \quad (11)$$

T_t is time size of the station's transmission slot. If its transmission success, T_t = ts; otherwise, T_t = tc. A packet is delivered unsuccessfully when it experience maximum allowed number of retransmission *m*, but still get collision, results in being dropped, so probability of packet drop P_d is:

$$P_d = p^{m+1} \quad (12)$$

And the time these packets have to experience is:

$$T_d = \sum_{i=0}^{m} T_i = \sum_{i=0}^{m} \left(\frac{W_i-1}{2} T_b + t_c\right) \quad (13)$$

For a packet being delivered successful at retransmission stage *i*, it has to experience the first *i* transmission stages collision and be successful in the final stage, so the time T_{si} it have to experience to be successful at stage *i* is:

$$T_{si} = \sum_{j=0}^{i-1} \left(\frac{(W_{i-1}-1)}{2} T_b + t_c\right) + \frac{W_i-1}{2} + t_s \quad (14)$$

As a result, mean service time for each successful packet deliver is:

$$T_s = \sum_{i=0}^{m} \frac{p^i(1-p)}{1-P_d} T_{si} \quad (15)$$

And general service time T_{sv} is the result of following equation:

$$T_{sv} = (1 - P_d)T_s + P_d T_d \quad (16)$$

However, total delay time and total packet lost are just important elements need to be concerned. We will present the computation of these parameters in the next section.

B. Total delay time and total packet loss rate

Using *M/M/1/K* queue model for the transmission of a given station, from queuing theory, we can represent the relationship of probability of station state having *i* packets in the system π_i, $i \geq 1$, with probability of state having only one packet in the system that is being in served π_1 as following:

$$\pi_i = \rho^{i-1}\pi_1 \quad (17)$$

In which, $\rho = \lambda/\mu$, with λ is arrival rate of packet to system transmission queue, $\mu = 1/T_{sv}$ is service rate

From (17), we found the probability of packet loss P_{loss} when queue is full in condition station always has packet to send by following equation:

$$P_{loss} = \frac{\pi_K}{\sum_1^K \pi_i} = \frac{\rho^B(1-\rho)}{1-\rho^{B+1}} \quad (18)$$

In which, $B = K-1$ is the length of queue buffer.

Thus, packet loss because of two reasons, that is when queue is full and when the retransmission beyond the maximum allowable number. As a result, total packet loss rate $P_{totalloss}$ is computed as:

$$P_{totalloss} = P_{loss} + (1 - P_{loss})P_d \quad (19)$$

In condition of saturated condition, stations always have packet to send, delay time in queue is only meaningful when the number of packet in the queue greater than 0 (no delay time), and less than *B* (no loss). In such condition, also from queuing theory, average queue length *E[Q]* is found as:

$$E[Q] = \frac{\rho}{1-\rho} - \frac{B\rho^B}{1-\rho^B} \quad (20)$$

Applying Little's theorem, we can find delay time T_{delay} in queue by the equation following:

$$T_{delay} = \frac{E[Q]}{\lambda(1-P_{loss})} \quad (21)$$

The total delay time since packet arrival at queue until it is delivered successfully is the summation of delay time it experiences in queue and service time of transmission:

$$T_{totaldelay} = T_{delay} + T_{sv} \quad (22)$$

So far, we have presented computation of some performance parameters. The computation is suitable for general case, except some special case with specific value such as *p=1* or *p=0.5*, but it is straightforward to replace such value for computation, so we do not introduce it here. However, the results that we show in the next section also account these special cases.

V. PERFORMANCE ANALYSIS

In this section, we take the parameters used in [3] as example in the model to discuss about system performance analysis (Table 1).

Table 1 Example parameter used in model ([3])

Packet payload	8184 bits
MAC Header	272 bits
PHY Header	128 bits
ACK	112 bits + PHY header
RTS	160bits + PHY header
CTS	112 bits + PHY header
Channel bit rate	1Mbit/s
Propagation Delay	1µs
Slot Time	50µs
SIFS	28µs
DIFS	128 µs

We can easily see the relationship among parameters that have been presented in the last sections by showing them intuitively in diagrams. With $W_0 = 8$, $k=3$, $m=5$, $B=5$ as an example, Fig. 4 represents the change of total delay time and packet loss rate when load and number of the stations change. It show that packet loss is only small at the corner where number of the stations and arrival rate small, while total delay increases with number of stations at the first stage, then decreases after the number of stations greater than a certain value. Load doesn't affect to total delay much, except its value is small. In order to explain that phenomenon, we look at elements constitutes total delay time and total packet loss as shown in Fig. 5. As in (22), total delay time is the summation of delay in queue and service time. Service time only depends on W_0, *k, m, n*, it is not depends on λ. When λ is small, the queue is not full, so the loss rate is small. When λ increase, the queue length increase, so the loss rate also increase. Because λ increase when service time does not increases, when ρ is very large, in denominator of (21), we have:

$$\lambda(1 - P_{loss}) = \rho\mu\left(1 - \frac{\rho^B(1-\rho)}{1-\rho^{B+1}}\right) \approx \mu \quad (23)$$

That mean $T_{delay} = (B\text{-}1)T_{sv}$. As the result, when λ is large, the delay in queue only depends on service time, and total delay depends on service time too. Similarly, packet dropped in period of transmission does not also depend on λ, so from (19), total packet loss depends on packet loss in queue linearly, and change with λ in the same form. This result also shows that in order to decrease total delay time, we can decrease delay time in queue by decrease buffer length. However, decreased buffer length is main cause contributing to total packet loss, especially when total packet lost initially affected by packet loss in queue.

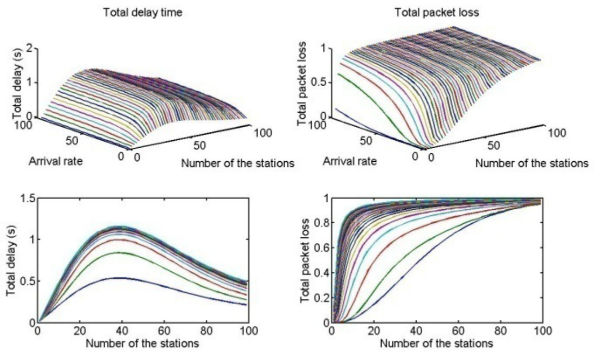

Fig. 4 Total delay time and packet loss rate

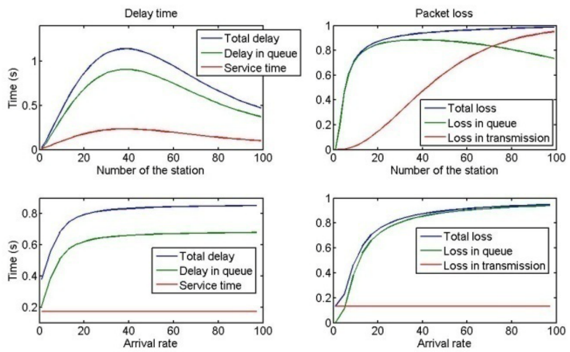

Fig. 5 Delay time and packet loss component analysis

It is not difficult to understand why the total delay time increases with the number of stations in the first stage, and then decreases later. When the number of station is small, there are many idle slots and the time size of idle slots is very small compared to slots having transmission. More station, more slot having transmission result in service time increase. Because there are still few stations, the contention is small, and the number of successful transmission is large. When the number of station increases, the collision increases also, making the number of successful transmission decrease. As we know, the slots that collision occurs have time size much smaller than the slot with transmission successful, so when collision increases, the service time obviously decreases. This reduction of service time helps packet lost rate in queue decrease, but raising the packet drop in transmission. As consequence, total packet loss is still high. Access medium by back-off mechanism is a solution to try to decrease the collision among transmission in condition of high density of stations; however it will make increased delay time, affecting to time sensitive applications. Therefore, choosing suitable parameter such as $W_{0,}$ *m*, *k* is the problem to be solved.

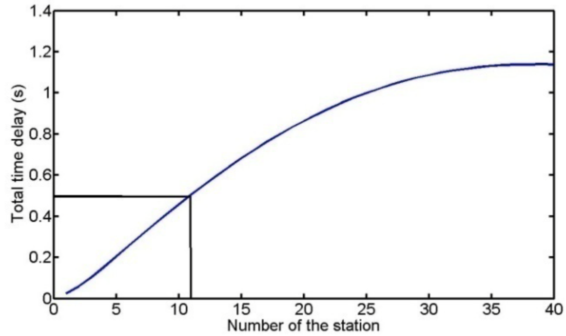

Fig. 6. Dependence of total delay on number station, λ=37

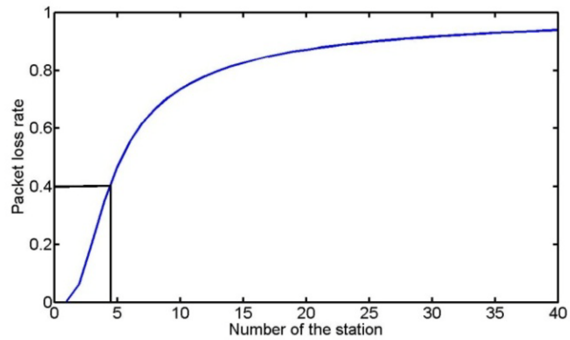

Fig. 7 Dependence of total packet lost on number of station, λ=37

Each application has a certain constraint of packet loss and delay time, so with the given condition, based on this model, we can find the potential of the system. For example, the constraint for an application is packet loss rate ≤ *0.4* and delay time ≤ *0.5s*. In condition of load *λ=20*, in Fig. 6, we can see system can accept maximum *11* stations for delay time. However, for satisfying packet loss rate constraint, as shown in Fig. 7, only 4 stations can be accepted. Consequently, with maximum of 4 stations operates in the same area, the requirement is satisfied.

VI. Conclusion

In this paper, we have discussed the model that can be used to analyze the performance of system at MAC layer of 802.11 in saturation condition. We have also analyzed the effect of load and the number of stations to the performance of system, particularly total delay time and packet loss rate. Through this analysis, we obtained an important judgment that is: in order to improve one of performance parameters that are total delay time or packet loss base on native parameters adjustment, there is always inverse effect to the other. By this model, we can find the system capability in boundary condition given. In this model, however, we have not still considered the effect of hidden terminal and mobility. We will proceed on these problems in future.

REFERENCES

[1] A. Kumar, E. Altman, D. Miorandi, M. Goyall, "New insights from a fixed point analysis of single cell IEEE 802.11 WLANs", *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE,* March 2005

[2] C. Siva Ram Murthy and B.S Manoj, "Ad hoc wireless networks, architectures and protocols", *Prentice Hall,* 2004

[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, March 2000.

[4] IEEE Computer Society, "IEEE Std 802.11™-2007", chap. 11, *IEEE*, June 2007.

[5] I. Koukoutsidis, V. A. Siris, "Modeling approximations for an IEEE 802.11 WLAN under Poisson MAC-level arrivals", *Proc. of IFIP Networking 2007, Atlanta, USA*, May 2007

[6] O. Tickoo, B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks", *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, 2004.

[7] R. Oliveira, L. Bernado, P. Pinto, "Modeling delay on IEEE 802.11 MAC Protocol for unicast and broadcast nonsaturated traffic", *Proc. of IEEE Wireless Communications and Networking Conference (WCNC'07), Hong Kong*, Mar. 2007