A Simple Model of the IEEE 802.11 MAC Protocol with Heterogeneous Traffic Flows

B. Bellalta, M. Oliver, M. Meo and M. Guerrero

Abstract — We present a model of the IEEE 802.11 MAC protocol considering heterogeneous non-saturated traffic flows in a single-hop scenario. We validate the model comparing several performance measures with simulation results, showing the model accuracy and ability to catch the effect of flow multiplexing over the shared channel.

Keywords — WLAN, IEEE 802.11 model, heterogeneous traffic, streaming and elastic flows.

INTRODUCTION

INTERNET traffic is mainly composed by data flows carrying data generated by the file transfer (documents, e-mails, web pages, songs, images, etc.) between two end points. These data flows use TCP as transport protocol and have the characteristic to adapt their bandwidth to the state of the network. For this reason, they are sometimes called elastic flows. Due to the increment of the access network capacity, the utilization of novel services like the transmission of voice and video (streaming), possibly in real time, is increasing their popularity and their use begins to change the traditional Internet traffic profiles. The main difference of these streaming flows with respect to the elastic flows is the strict bandwidth requirement they need to perform correctly.

The fast deployment of the WLAN coverage in business, home and public open zones will increase the use of this technology by consumers to access the Internet; therefore, the radio interface technology and protocols will have to manage efficiently these heterogeneous Internet traffic profiles. The current technology used for WLAN is based on the IEEE 802.11b standard which defines the MAC and PHY layers. The behavior of the MAC protocol has been extensively studied during last years, remarking the papers of Bianchi and Tay and Chua [2,3], referred in most of posterior works. In [2,3,4], the authors propose an analytical model to evaluate several performance metrics of a single-hop wireless LAN (collision probability, throughput, delay, ... etc.), presenting elegant solutions for the MAC behavior under saturation conditions (all nodes always have a packet ready to be transmitted). These results can be applied to understand how the MAC protocol performs but have a limited applicability to real scenarios since real networks in steady state will never work well in saturation conditions. Therefore, in order to design, analyze and plan WLAN networks we need a model capable to catch the traffic characteristics of each node and evaluate the network performance in non-saturation conditions [5,6,7,8].

II. MODEL FOR HETEROGENEOUS TRAFFIC PROFILES

In this section we introduce our analytical model. For a general description of the IEEE 802.11 MAC layer, refer to [1].

A. Mobile Station

The MAC layer of a mobile node is modeled as an M/M/1/Q queue where Q is the queue length (in packets). The offered traffic (from network layer) to MAC layer is

\[ u_k(n) = \lambda_k X_k(n) \]  

where \( X_k(n) \) and \( \lambda_k \) are the service time and the packet arrival rate for the node \( k \) when there are \( N \) nodes in the network (the size of the set of nodes \( \psi \) is \( N \)). We define \( n \)
as a vector of $N$ positions, where $g_k$ refers to mobile node $k$, with $k \in \Psi$. The traffic offered to the network by node $k$ is

$$p_k(n) = \lambda_k \left(1 - P_{\text{ssl}}(n)\right) X_k(n)$$

(2)

where the blocking probability $P_{\text{ssl}}(n)$ for an M/M/1/Q queue is computed from

$$P_{\text{ssl}}(n) = \frac{(1 - v_\Psi(n)) v_\Psi(n)^n}{1 - v_\Psi(n)^{n+1}}$$

(3)

with the special case where $v_\Psi(n) = 1$ which gives a blocking probability of $P_{\text{ssl}}(n) = 1/(Q+1)$. Note that the probability that a node has at least one packet buffered is equal to the queue utilization parameter $p_k(n)$.

B. Traffic Flows and Bandwidth Sharing

The simplest way to classify the Internet traffic flows is to consider only two types of flows: elastic (TCP-like) and streaming (UDP-like). A streaming flow is characterized by the bandwidth required during the flow sojourn time. If the network can satisfy all the requested bandwidth, the data will be delivered timely. To characterize this type of flows we need to know the required bandwidth ($B_e$), the packet arrival process ($\lambda_e$) and packet length statistics ($L_e$). In opposition, an elastic ideal flow tries to use all the available bandwidth, sharing it fairly with the other active flows in the system. Thus, the arrival rate of packets for an elastic flow adapts to the system state, and it is equal to the inverse of the service time observed by the flow. This situation can be modeled maintaining the queue utilization of the node with an elastic flow equal to $p_e = 1$.

From ideality, we expect that a link of bandwidth $B$ will be perfectly shared between $n_e$ elastic flows and $n_i$ streaming flows if each elastic flow uses $B_e = (B - n_i B_i)/n_e$ bandwidth units (assuming that $B - n_i B_i > 0$ and $B_i = 0$ if $B - n_i B_i < 0$ [9]. In a WLAN, the Distributed Coordination Function (DCF) can be viewed as a long term fair scheduling function (processor sharing). However, the main drawback is that it is a node based scheduling that shares the bandwidth without considering the flow dynamics.

C. IEEE 802.11 MAC protocol

As the attempt rate of each node is a regenerative process, the renewal-reward theorem can be used to compute the node average attempt rate [7]. Thus, the probability that node $k$ transmits in a random slot is

$$\pi_k(n) = \frac{N_{\text{ssl}}(n)}{N_{\text{ssl}}(n) + \text{Idle}_k(n)}$$

(4)

where $N_{\text{ssl}}(n)$ is the number of transmissions needed to receive a packet correctly, $E\{W_k(n)\}$ is the average number of slots (uniformly distributed) of the backoff procedure and

$$\text{Idle}_k(n) = \frac{1}{\alpha_k(n)} \left(1 - \frac{1}{\lambda_k(1 - P_{\text{ssl}}(n))} - X_k(n)\right)$$

(5)

is the average number of slots during which a node is idle (i.e., the node has no packets to transmit buffered in the MAC transmission queue). When a node is saturated (there are no idle slots) equation (4) is the same presented in [2,3]. The collision probability can be expressed as a function of the transmission probability conditioned to the probability that a packet is ready in another node. In this case, the probability that a node transmits a packet is

$$\tau_k(n) = \frac{p_k(n)}{E\{W_k(n)\} + 1}$$

(6)

The assumption that the node state is independent from that of other nodes has been demonstrated imprecise for a non-saturated situation [6] but, as we show in our results, this assumption is accurate enough without introducing the extra complexity needed to model the dependencies among nodes. Therefore, as the probability that a station has a packet ready to be transmitted is $p_k(n)$, the conditional collision probability can be expressed as the probability that at least two nodes transmit at the same temporal slot

$$\pi_k(n) = 1 - \prod_{\Psi \neq k} (1 - \tau_k(n))$$

(7)

It can be extrapolated from (7) that flows which require low bandwidth have a higher collision probability when they share the channel with flows consuming a high amount of bandwidth.

Using (7) and considering that the channel is error free (retransmissions are only caused by collisions), the average number of transmissions needed to correctly send a packet can be approximated by

$$N_{\text{ssl}}(n) = \frac{1}{1 - p_k(n)}$$

(8)

The service time $X_k(n)$ depends on the number of necessary transmissions and the frame structure defined in the standard [1]. If $T(L_e)$ is the average time needed to transmit an IP packet and $\overline{T}_e(L_e)$ is the average time spent in collision when the node $k$ collides, the service time can be computed as

$$X_k(n) = N_{\text{ssl}}(n) (E\{W_k(n)\} \alpha_k(n) + \overline{T}_e(L_e) + T(L_e))$$

(9)

Note that previous equations are valid for both the BA (Basic Access) and the RTS/CTS mechanism and we only need to consider the different frame durations $T(L_e)$ and $\overline{T}_e(L_e)$.

In previous equations, the $\alpha_k(n)$ parameter corresponds to the average slot duration observed by a node. This parameter relates the different states of the channel: empty, in a collision and in a successful transmission. When the channel is sensed busy, the back-off counter is frozen and restarts its count-down when the channel is detected idle again. Thus, the average slot duration can be computed as (note that we add an extra slot time due the fact that the backoff counter is only decreased by one after the channel is sensed idle for a full slot duration)

$$\pi_k(n) = \frac{p_k(n)}{E\{W_k(n)\} + 1}$$

(6)

The assumption that the node state is independent from that of other nodes has been demonstrated imprecise for a non-saturated situation [6] but, as we show in our results, this assumption is accurate enough without introducing the extra complexity needed to model the dependencies among nodes. Therefore, as the probability that a station has a packet ready to be transmitted is $p_k(n)$, the conditional collision probability can be expressed as the probability that at least two nodes transmit at the same temporal slot

$$\pi_k(n) = 1 - \prod_{\Psi \neq k} (1 - \tau_k(n))$$

(7)

It can be extrapolated from (7) that flows which require low bandwidth have a higher collision probability when they share the channel with flows consuming a high amount of bandwidth.

Using (7) and considering that the channel is error free (retransmissions are only caused by collisions), the average number of transmissions needed to correctly send a packet can be approximated by

$$N_{\text{ssl}}(n) = \frac{1}{1 - p_k(n)}$$

(8)

The service time $X_k(n)$ depends on the number of necessary transmissions and the frame structure defined in the standard [1]. If $T(L_e)$ is the average time needed to transmit an IP packet and $\overline{T}_e(L_e)$ is the average time spent in collision when the node $k$ collides, the service time can be computed as

$$X_k(n) = N_{\text{ssl}}(n) (E\{W_k(n)\} \alpha_k(n) + \overline{T}_e(L_e) + T(L_e))$$

(9)

Note that previous equations are valid for both the BA (Basic Access) and the RTS/CTS mechanism and we only need to consider the different frame durations $T(L_e)$ and $\overline{T}_e(L_e)$.

In previous equations, the $\alpha_k(n)$ parameter corresponds to the average slot duration observed by a node. This parameter relates the different states of the channel: empty, in a collision and in a successful transmission. When the channel is sensed busy, the back-off counter is frozen and restarts its count-down when the channel is detected idle again. Thus, the average slot duration can be computed as (note that we add an extra slot time due the fact that the backoff counter is only decreased by one after the channel is sensed idle for a full slot duration)
\[ \alpha_i(n) = p_{s,i}(n) + p_{s,i}(n)(\bar{T}^*(L_s) + \sigma) + p_{s,i}(n)\left(\bar{T}(L_s) + \sigma\right) \]

where \( p_{s,i}(n) \), \( p_{e,i}(n) \) and \( p_{s,i}(n) \) are the probabilities related to the channel events. Note that \( \bar{T}^*(L_s) \) and \( \bar{T}(L_s) \) are the average duration of a successful transmission of another node or a collision between two other nodes. The probability to sense the channel empty is the probability that any of the other mobile stations transmits in a randomly chosen slot

\[ p_{s,i}(n) = \prod_{j \neq i} \left(1 - \tau_j(n)\right) \]

(12)

and the computation of the probability to observe a collision in the channel is simply

\[ p_{c,i}(n) = 1 - p_{s,i}(n) - p_{e,i}(n) \]

(13)

Once computed the service time, the expected throughput for a node \( k \) can be computed as

\[ S_i(n) = \frac{P_i(n)}{X_i(n)} L_s \]

(14)

To calculate the \( E\{W_i(n)\} \) parameter we use the same expression obtained by Bianchi and Tey and Chua [2,3], assuming that the number of retransmissions before dropping a packet is infinite

\[ E\{W_i(n)\} = \frac{1 - p_i(n) - p_i(n)2p_i(n)^n}{1 - 2p_i(n)} \frac{CW_{\text{min}}}{2} - \frac{1}{2} \]

(15)

III. MODEL VALIDATION

Due to the existing dependence of previous expressions with the queue utilization \( P_i(n) \) and the fact than (6) and (7) form a set of non-linear equations, we use iterative numerical techniques to solve the model. To validate it, we have built a simulator of the IEEE 802.11 MAC protocol using the COST (Component Oriented Simulation Toolkit) simulation package [10]. The value of the parameters used both in simulation and in the model are shown in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIFS</td>
<td>10 µs</td>
</tr>
<tr>
<td>DIFS</td>
<td>50 µs</td>
</tr>
<tr>
<td>EIFS</td>
<td>364 µs</td>
</tr>
<tr>
<td>σ (SLOT)</td>
<td>20 µs</td>
</tr>
<tr>
<td>CW_{\text{min}},CW_{\text{max}} m</td>
<td>32, 1024, 5</td>
</tr>
<tr>
<td>MAC header</td>
<td>240 bits + 32 bits (FCS)</td>
</tr>
<tr>
<td>PHY header</td>
<td>192 bits @ Rbasic</td>
</tr>
<tr>
<td>L_{data,k}</td>
<td>PHY h + MAC h + L_k</td>
</tr>
<tr>
<td>L_{ack}</td>
<td>112 + PHY h @ Rbasic</td>
</tr>
<tr>
<td>R_{data}</td>
<td>2 Mbps</td>
</tr>
<tr>
<td>R_{basic}</td>
<td>1 Mbps</td>
</tr>
</tbody>
</table>

The modeled system has two traffic profiles: \( n_s \) streaming flows with a bandwidth requirement of \( B_s=64 \text{ Kbps} \) and a packet length of \( L_s=1024 \text{ bits} \) and \( n_e \) elastic traffic flows with a variable bandwidth (the maximum available) and packet length \( L_e=4096 \text{ bits} \). Each flow is associated to a single node which sends the data to a randomly selected neighbor in a single-hop wireless network. Each node has a MAC queue of length \( Q=50 \text{ packets} \) and the MAC layer uses the BA mechanism.

In Fig. 1, we show the aggregate throughput for the streaming flows when there are 0, 2 or 5 elastic flows in the system. The throughput is plotted versus the number of streaming flows active in the network. When no elastic flows are active, we observe how the network reaches the saturation throughput at \( n_s=11 \) streaming flows (in this point the queue of all streaming flows is saturated, see Fig 3). For \( n_s=2 \), the saturation throughput is reached at \( n_s=6 \) and for \( n_s=5 \) it is reached at \( n_s=2 \). We can conclude that, in contradiction with the assumption, elastic flows interfere with streaming flows. The reason is that the MAC protocol tries to give the same access probability to all nodes with a packet ready to be transmitted and therefore the elastic flows get more transmission opportunities than streaming flows. Thus, a simple solution is to change the “fair” idea implicit on the DCF scheduling function from “node” to “flow” point of view. This can be done by properly adjusting the MAC parameters.

In Fig. 2, we present the aggregated throughput for the elastic flows. When there are no streaming flows, the aggregated elastic throughput is higher with 2 elastic flows than with 5. This is motivated by the extra contention delay of the multiple access mechanism which introduces and the increment of the collision probability. As the number of streaming flows increases, the throughput of elastic flows progressively reduces as it would be expected. This throughput reduction is inversely proportional to the number of elastic flows. As we can see, with 2 elastic flows, the throughput reduction is more rapid than with 5 elastic flows. Another interesting result is the existence of an inflexion point and a change on the throughput reduction curve when the network enters in saturation. These results can be justified by the long term fair sharing provided by the MAC protocol which is proportional to the number of active nodes and the queue utilization of each node and, therefore, the elastic flows have a higher share of the bandwidth.

Those results are confirmed by Fig. 3, in which queue utilization for streaming flows is plotted versus the number of elastic flows. This is the key parameter of the model, since it links the reciprocal impact between the traffic characteristics and the network state. For example, with 2 elastic flows, the queue utilization for a single streaming flow is 0.5, while, without elastic flows, for a single streaming flows it is 0.1. Obviously, an increment of the traffic load offered to the network, increases the collision probability, which it is plotted in Fig. 4, where the dependencies with the queue utilization parameter are clearly shown.
Finally, in Fig. 5 we plot the throughput for elastic flows for $n_s=0$, 2 and 5 when the number of elastic flows increases. As it can be observed, part of the bandwidth share of the streaming flows is captured by the elastic flows, increasing its throughput, which obviously causes a proportional reduction of the throughput of streaming flows.

Fig. 1. Aggregate throughput for streaming flows.

Fig. 2. Aggregate throughput for elastic flows.

Fig. 3. Queue utilization for streaming flows.

Fig. 4. Collision prob. observed by streaming flows.

Fig. 5. Aggregate throughput for elastic flows.

IV. CONCLUSIONS

In this paper we present an analytical model of the IEEE 802.11 MAC considering the presence of heterogeneous traffic profiles. We validate the model considering two type of flows: elastic and streaming flows. The results show the accuracy and applicability of the model to reflect the main dynamics of WLAN networks based on the random access MAC protocol (DCF).

We observe how, due to the node long term share of the MAC protocol, the elastic flows cause a clear degradation in the performance of streaming flows, which leads to a clear reduction of the capacity of real-time flows. In future work, we will consider TCP persistent connections (including the ACK feedback traffic) in a similar way than [11] to find a general model to evaluate WLAN in infrastructure mode or ad-hoc networks with both TCP and UDP traffic simultaneously; however, our preliminary results show similar tendencies as the ones observed in this paper. Another point which remains open is the computation of the probability density function of the number of simultaneous active nodes in heterogeneous conditions and its application to the computation of the conditional collision probability and other parameters of the model.

Finally, it is important to remark that the model can be easily applied to evaluate traffic differentiation
mechanisms, admission control schemes, MAC multi-rate capabilities, etc, by simply setting different MAC parameters for each flow.

REFERENCES


