Admission Control Policy
for WLAN Systems based on the "Capacity Region"

A. Barbaresi, S. Barberis, P. Goria
Telecom Italia LAB - Via G. Reiss Romoli, 274 – I-10148 Torino (TO), Italy
e-mail: {andrea.barbaresi, sergio.barberis, paolo.goria}@telecomitalia.it

ABSTRACT

This paper deals with the definition of an Admission Control (AC) policy for Wireless Local Area Network (WLAN) systems applicable to Real-Time (RT) services. The proposed AC policy is based on an analytical model, able to predict the performance of the Distributed Coordination Function (DCF) of IEEE 802.11 Medium Access Control (MAC). Results provided by the model for the considered real-time services are exploited in order to identify the maximum number of users for each class (i.e. services mix) that a 802.11b,a,g Basic Service Set (BSS) can support, taking into account the Quality of Service (QoS) constraints. The work has been carried out within the framework of EVEREST project, in which specific radio resource management strategies and algorithms for supporting QoS in heterogeneous beyond 3G networks are investigated.

I. INTRODUCTION

The main objective of the EVEREST project [1] is to investigate and propose mechanisms and algorithms that can handle the expected traffic growth of high demanding QoS services in a heterogeneous network which comprises 2G and 3G cellular systems as well as WLANs.

At present, any scenario considering also WLAN hotspots within a 2G and 3G cellular heterogeneous network, have to face the fact that WLANs lack of any QoS management mechanisms. In particular, WLANs are "no-blocking" systems: new users entering into the system try to access the shared medium for transmitting and receiving data. Consequently, the QoS of all the users in the BSS degrades in terms of throughput, delays, jitter and transmission errors, with an increasing number of users. This is particularly true for the IEEE 802.11 MAC based on the DCF access method [2], that makes the system performance very sensitive to the number of stations and their traffic profiles. On the basis of these considerations, the need of an AC scheme should be evident: only best effort services can be supported without any policy for blocking the admission of new users when specific load conditions occur.

It is worth to note that one of the recent major interests concerning the WLANs within the standardization fora is focused on the QoS improvements as a whole, not only as far as the AC mechanism is concerned. Several enhancements for QoS management are going to be specified by the upcoming IEEE 802.11e standard [3]. In this paper we propose an AC policy applicable to IEEE 802.11a,b,g systems, without taking into account any specific QoS mechanisms that will be available in near future WLANs.

The paper is organized as follows: Section II introduces the reference architecture for interworking between WLAN and cellular networks. Next, in Section III, the analytical model used to estimate the throughput perceived by each user in an IEEE 802.11 WLAN system is described. Sections IV and V present how to exploit analytical results in order to derive the AC policy. Finally, a summary of the overall work is addressed in Section VI.

II. COUPLING WLANS WITH CELLULAR NETWORKS: THE EVEREST APPROACH

Up to now, efforts for integrating WLANs within 2G and 3G cellular networks are based on a loose coupling approach, trying to minimize the impacts on the network equipments. Beside this fact, it seems reasonable to think that the level of integration could be in the future tighter: a loose coupling does not seems to be adequate to meet very high demanding requirements specified by 3GPP in [4], in which six interworking scenarios have been described and characterized.

For this reason, in EVEREST, tight/very-tight coupling architecture between a mobile core network and WLANs is considered. Basically, the tight coupling approach has the general target to make possible to consider a WLAN BSS as a UMTS Radio Access Network (i.e. URAN in 3GPP standard, [5]), connected to the 3G mobile core network. According to this guide principle, when a tight coupling approach is used, WLAN can be considered as a specific concretization of a broadband radio access network within the UMTS system.

A possible architecture where WLAN are tightly coupled with cellular network is depicted in Figure 1.

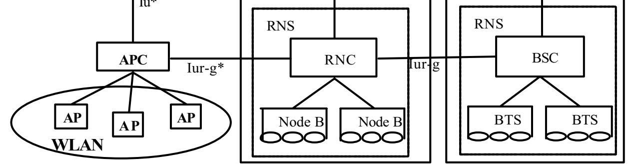

Figure 1: Tight coupling WLAN with cellular networks.

According to the figure, the Access Points (APs) in charge of the wireless communication with the users are linked to an Access Point Controller (APC). Within the WLAN system, the APC has the role of the Radio Network Controller (RNC) of UTRAN and it is linked to the SGSN of the mobile core network via an interface similar to the Iu-PS (Iu* in the figure). This interface should support at least the subset of functionalities of the Iu that are compatible with the specific WLAN technology considered. In the case of "very-tight" coupling, also the hypothesis of a direct interface between RNC/BSC and the APC can be assumed, in the same way of the Iur-g interface between GERAN and UTRAN [6]. This interface allows some decisions or procedures to be executed without relying on SGSN functions.

In this work, the AC algorithm presented in Section V, is supposed to run inside the APC, receiving services requests via the Iu* interface.

III. ANALYTICAL MODEL FOR MAC 802.11 DCF

The analytical model to evaluate the saturation throughput of MAC 802.11 DCF access method in ideal channel condition was first proposed in [7],[8] and further developed in this work to fulfil the requirements of EVEREST project. The generalization made for the project is two-fold: i) it eliminates the assumption of saturation, and (ii) this in turn allows to consider stations with different traffic patterns, characterized by mean offered load and mean packet length. Another extension made allows a full distinction of traffic patterns between uplink and downlink. The final model is very general and allows the analysis of the behaviour of a WLAN BSS in realistic situations (e.g., in case of asymmetric traffic and differentiated service classes).

III.1 PACKET TRANSMISSION PROBABILITY

We consider a set of stations grouped in R classes, and fully described by the vectors:

- Number of stations per class: $\vec{n} = [n_1 \quad n_2 \quad \ldots \quad n_R]$.
- Initial value of the contention window in slots: $\vec{W} = [W_1 \quad W_2 \quad \ldots \quad W_R]$.
- Maximum number of backoff stages: $\vec{m} = [m_1 \quad m_2 \quad \ldots \quad m_R]$,.
- Capacity of the MAC buffer in packets: $\vec{L} = [L_1 \quad L_2 \quad \ldots \quad L_R]$.
- Mean payload length in bit: $\overrightarrow{Ep} = [Ep_1 \quad Ep_2 \quad \ldots \quad Ep_R]$.
- Probability of packet generation per slot: $\overrightarrow{Pin} = [Pin_1 \quad Pin_2 \quad \ldots \quad Pin_R]$

(1)

As first proposed by [6], an integer time scale is adopted, in which t represents a slot number. To describe the behaviour of a given network node by means of a Markov chain, three stochastic processes are defined:

- The node's backoff stage: $i(t)$.
- The node's backoff time counter: $b(t)$.
- The number of packets in the node's buffer: $n(t)$.

These three processes give rise to a tri-dimensional chain with states labelled by the 3-vector $\{i,b,n\}$. According to this definition, the packet transmission probability corresponds to the probability of being in any of the states $\{1,\mathbf{0},n\}$. The chain, as well as the full procedure to calculate the transmission probability is well-known and described in [6]: it is based on the key assumption of a constant collision probability for each class, denoted by p(r).

The development comprises the generation of an equivalent two-dimensional chain by means of state aggregation, the calculation of the transition probabilities of the new chain, and the final solution based on these. All the quantities involved are vectors; therefore each of the next equations actually is a set of R equations, one per each class.

First, some quantities are defined to shorten the final expression. In what follows, the convention is that parenthesized super indexes, as well as sub indexes are positions in a vector.

$$\tau_B^{(r)} = \frac{2\left(1-2p^{(r)}\right)}{\left(1-2p^{(r)}\right)\left(W_0^{(r)}+1\right)+p^{(r)}W_0^{(r)}\left(1-\left(2p^{(r)}\right)^{m^{(r)}}\right)} \quad (2)$$

This is the transmission probability of a class-r station given that its buffer is not empty. This is equivalent to the transmission probability calculated in [6], where the nodes are assumed to always have a packet to transmit.

Now we define a conditioned success probability, that is, the probability that the previous conditioned transmission occurs and that a collision does not.

$$P_{succB}^{(r)} = \tau_B^{(r)}\left(1-p^{(r)}\right) \quad (3)$$

Finally we define the next quantity, which will appear in the final expression.

$$P_v^{(r)} = \frac{p_{in}^{(r)} - p_{succB}^{(r)}}{\left(1-p_{in}^{(r)}\right)\left(p_{in}^{(r)}\left[\frac{p_{in}^{(r)}\left(1-p_{succB}^{(r)}\right)}{p_{succB}^{(r)}\left(1-p_{in}^{(r)}\right)}\right]^{L^{(r)}} - p_{succB}^{(r)}\right)} \quad (4)$$

Finally, (5) is obtained, which expresses the transmission probability of a generic terminal in a slot.

$$\tau^{(r)} = \frac{1/\left(1-p^{(r)}\right)}{D} \quad (5a)$$

$$D = 1 + \frac{\left(1-p_{in}^{(r)}\right)p_v^{(r)}}{p_{in}^{(r)}} + \sum_{k=1}^{W_0^{(r)}-1}\left[1-\frac{k}{W_0^{(r)}}\right] +$$

$$+\sum_{i=1}^{\left(m^{(r)}-1\right)}\sum_{k=0}^{\left(W_i^{(r)}-1\right)}\left[\left(1-\frac{k}{W_i^{(r)}}\right)\left(p^{(r)}\right)^i\right]+$$

$$+\frac{\left(p^{(r)}\right)^m}{1-p^{(r)}}\sum_{k=0}^{\left(W_m^{(r)}-1\right)}\left[1-\frac{k}{W_m^{(r)}}\right] \qquad (5b)$$

At this point, we recall that the probability of generating a packet in a slot was defined in (1) as a parameter specifying the traffic pattern of a class. Since for RT services requesting a fixed amount of bandwidth it is desirable to specify the traffic pattern in terms of the offered data rate, let us consider the link between those values (E_{TS} is the mean slot duration):

$$DataRate^{(k)}=\frac{P_{in}^{(k)}E_P^{(k)}}{E_{TS}} \qquad (6)$$

This expression shows that it is not possible to know beforehand the load a station will offer to the network as a result of the choice of *Pin*, or, conversely, knowing the data rate a station should offer does not allow the calculation of the parameter *Pin*, since they are related through an output of the problem. Therefore, the final solution is obtained through the solution of the next nonlinear system of 3R equations.

$$\begin{aligned}\tau &= \tau(p,P_{in})\\ P_{in} &= P_{in}(\tau,p)\\ p &= p(\tau)\end{aligned} \qquad (7)$$

The input to the model is thus:

- Number of stations per class: $\vec{n}=[n_1 \quad n_2 \quad \dots \quad n_R]$.
- Initial value of the contention window in slots: $\vec{W}=[W_1 \quad W_2 \quad \dots \quad W_R]$.
- Maximum number of backoff stages: $\vec{m}=[m_1 \quad m_2 \quad \dots \quad m_R]$.
- Capacity of the MAC buffer in packets: $\vec{L}=[L_1 \quad L_2 \quad \dots \quad L_R]$.
- Mean payload length in bit: $\vec{Ep}=[Ep_1 \quad Ep_2 \quad \dots \quad Ep_R]$.
- Offered data rate in bit/s: $\vec{DR}=[DR_1 \quad DR_2 \quad \dots \quad DR_R]$

(8)

III.2 CALCULATION OF SOME RELATED PROBABILITIES

Once the transmission probability is obtained, other quantities of interest can be calculated. The probability that at least one transmission occurs in a slot is calculated in (9). It is the complementary to the probability that no transmission occurs.

$$P_{tr}=1-\prod_{r=1}^{R}\left(1-\tau^{(r)}\right)^{n^{(r)}} \qquad (9)$$

The next expression, equation (10), gives the probability that a packet transmitted by a class-r terminal suffers a collision.

$$p^{(r)}=1-\left[\left(1-\tau^{(r)}\right)^{n^{(r)}-1}\prod_{j\neq r}^{R}\left(1-\tau^{(j)}\right)^{n^{(r)}}\right] \qquad (10)$$

This expression assumes independent behaviour among stations (quite realistic). The first factor inside parenthesis stands for the probability that no other station of the class transmits, while the subsequent product stands for the probability that no station of the other classes does so.

Define the success probability as the probability that in any time slot, a station performs a successful transmission. This is the probability that a station transmits multiplied by the probability that its packet does not collide (the complement of (10)). This probability is given by.

$$P_{succ}^{(r)}=\tau^{(r)}\left(1-p^{(r)}\right) \qquad (11)$$

From this, the total success probability, defined as the probability that in any slot time there is a successful transmission, regardless of the class of the station, is calculated as the next expression shows.

$$P_{succ}^{tot}=\sum_{r=1}^{R}n_rP_{succ}^{(r)} \qquad (12)$$

III.3 DETERMINATION OF THE THROUGHPUT

The throughput (rate of the payload bit) obtained by each station of a given class can be calculated as the mean number of bit successfully transmitted in a slot divided by the mean slot duration:

$$T_{eff}^{(k)}=\frac{P_{succ}^{(k)}E_P^{(k)}}{E_{TS}} \qquad (13)$$

Equation (13) is very useful, because it provides the throughput each single station obtains as a result of all the service requests in the network. This in turn gives a beforehand assessment on the network conditions upon arrival of a new service request and hence has great use in eventual policies of admission control.The aggregate throughput (total effective throughput accounting all the stations of all classes) can be calculated from (13). It gives a performance metric of the network as a whole.

$$T_{eff}^{ag}=\sum_{k=1}^{R}T_{eff}^{(k)}n^{(k)} \qquad (14)$$

III.4 INCLUSION OF AN ACCESS POINT

An AP can be easily included in the system to simulate an infrastructure network. Recall that the IEEE 802.11 standard does not give any access priority to the AP in DCF, that is, during the period in which the stations contend to access the medium, the AP behaves simply as another station. Therefore, having established the network configuration in terms of total number of stations, grouped in classes, and traffic patterns for each classes, it is enough to add a new class with only one member (the access point), whose traffic pattern is calculated according to the traffic patterns of the stations.

In the previous analysis, only uplink traffic has been considered, deriving the throughput that stations obtain when they try to transmit packets with a given pattern. With the introduction of an AP also the downlink traffic, and even asymmetric traffic patterns, can be considered.

With respect to the inputs specified in (8), the additional parameters required are the offered data rate and the packet length in downlink. So, considering also the AP, the input to the problem is now completely specified by eight vectors:

$$\vec{n}, \vec{W}, \vec{m}, \vec{L}, \overrightarrow{DR}_{UL}, \overrightarrow{EP}_{UL}, \overrightarrow{DR}_{DL}, \overrightarrow{EP}_{DL} \qquad (15)$$

The input to the model is the set of vectors given in (15) where one of the classes is the AP alone, and the others are the original nodes in the network. The data rate and packet length in (8) are replaced by the corresponding uplink and downlink parameters.

At the output, the model provides the throughput obtained by each stations. This is directly the uplink throughput obtained by the stations. The traffic obtained by the AP, on the other hand, is actually the aggregate downlink obtained by the stations. In the case this aggregate throughput equals the aggregate request of all the nodes, the request in downlink of each node is satisfied.

IV. CAPACITY REGION FOR THE WLAN BSS

In the previous section, a mathematical model to estimate the throughput perceived by users in an IEEE 802.11 WLAN system, when their traffic profiles are specified, has been described. Users (stations) belonging to different classes of real-time services can be considered by the model and each service class is characterized by means of the mean data length [bit] of user data packets and the mean offered bit rate [bit/s]. Ideal channel condition is assumed by the model, so that no transmission errors occur and the throughput perceived by the users is only affected by packet collisions.

The model has been applied considering the real-time services envisaged by the EVEREST scenario: the WLAN system is supposed to be used to offer video telephony and video streaming services (conversational and streaming class packet services). The bit rates and packet lengths, as well as the requested minimum bandwidth for the above mentioned real-time services are summarized by Table1.

The following input parameters for MAC DCF access method were considered in the analytical model (corresponding to IEEE 802.11b technology):

- B: overall bandwidth = 11 Mbit/s
- W: minimum contention window = 32 slots
- m: number of backoff stages = 5
- L: buffer length = 10 SDU
- HMAC: MAC header size = 272 bit
- SIFS: Short interframe spacing = 10 µs.
- DIFS: DCF Interframe spacing = 50 µs.
- ACKl: ACK packet size = 112 bit

Table 1: Considered real-time services.

Service	**Bit rate / Packet length**	**Requested QoS (guaranteed bit rate)**
Video Telephony	UL:64, DL:64 kbit/s UL:1024, DL:1024 byte	UL:58, DL:58 kbit/s
Video Streaming Business	UL:16, DL:128 kbit/s UL:128, DL:2048 byte	UL:8, DL:112 kbit/s
Video Streaming Consumer	UL:16, DL:64 kbit/s UL:128, DL:1024 byte	UL:8, DL:58 kbit/s

According to the considered services and the used parameters, the analytical model provides the results depicted in the below Figure 2 (reported as an example).

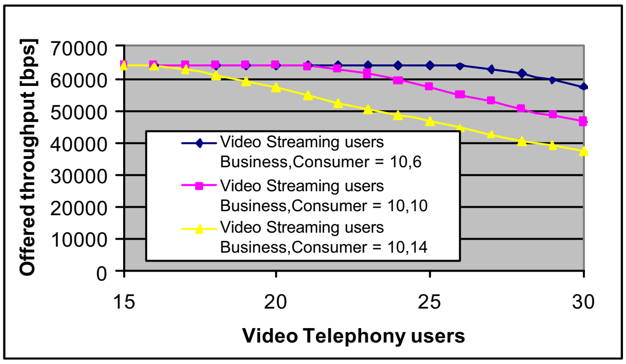

Figure 2: Offered throughput per user (IEEE 802.11b).

Taking into account also the QoS constrains, it is possible to exploit results coming from the model to identify the service mixes compliant with a minimum level of offered throughput (uplink and downlink) per user, i.e. the "Capacity Region". For the three types of services of Table1, the following capacity region can be obtained:

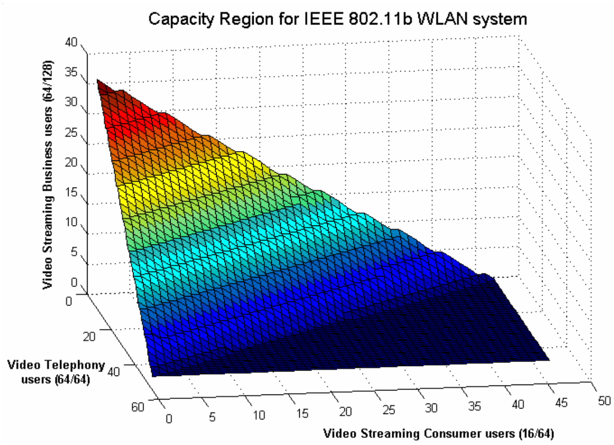

Figure 3: Example of capacity region.

Table 2 specifies some points of the capacity region depicted in Figure3.

Table 2: Capacity region.

Video Telephony (64/64 kbit/s)	Video Streaming Consumer (16/64 kbit/s)	Video Streaming Businness (16/128 kbit/s)
0	0	38
1	0	35
…	…	…
46	0	0
…	…	…
28	10	6
…	…	…
1	45	0
0	47	0

V. ADMISSION CONTROL POLICY

The capacity region can be used by an AC algorithm running inside the APC, to keep the number of active users for each class of services within the range of values able to respect the QoS constrains.

In more detail, the AC algorithm should perform the following steps:

- For each station requesting a new association to the WLAN BSS, retrieve the type of service requested by the user;
    - If the acceptance of the new user brings the load operating point of the WLAN BSS outside the QoS region (according to the a-priori evaluated capacity region), reject the request of association;
    - If the new user can be admitted without compromising the minimum level of offered throughput per each class of users, accept the request of association and update the load operating point of the system (i.e. number of active users within the systems).
- For each station requesting a de-association from the WLAN BSS, update the load work point of the system (i.e. number of active users within the systems);

From an operational point of view, it is useful to remark that the IEEE 802.11 standard explicitly foresees that a station should be authorized by the AP before exchanging data, by means of the "association procedure". According to this, a station that wishes to enter into the WLAN system sends a "request of association" frame to the AP and then waits for a response. Taking advantage of this fact, the response of the request of association can depend on the decision of the above mentioned AC algorithm. As an alternative, the proposed AC algorithm can be performed on the AddTS phase defined in 802.11e [3].

It is also worth to note that the proposed AC makes use of a capacity region which can be evaluated a-priori on the basis of the expected classes of users. With this assumption, the computational complexity of the analytical model does not affect the algorithm performances. In this way, the WLAN BSS can be used keeping the control of how many users the BSS can support, respecting a minimum QoS level in terms of guaranteed bandwidth.

VI. CONCLUSIONS

In this work a mathematical model to estimate the throughput perceived by different classes of real-time users in an IEEE 802.11 WLAN BSS was described. Results provided by the model can be used to derive the capacity region of the WLAN BSS for the considered real-time services.

The capacity region can be exploited by an AC algorithm so that the maximum number of users the hot-spot can support will be never exceeded. In this way the throughput offered by the BSS to each user remains above the specified threshold.

The proposed AC policy has been investigated within the context of EVEREST project, in which WLANs are supposed to be used as an alternative radio access system within a heterogeneous network.

REFERENCES

[1] F. Casadevall et al., "Overview of the EVEREST project", IST Mobile Summit 2004.

[2] IEEE. "Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications". ANSI/IEEE Std 802.11, 1999 Edition.

[3] Draft Supplement to Part 11: Wireless Medium Access Control (MAC) and Physical Layer Specification: MAC Enhancements for Quality of Service (QoS), IEEE 802.11 WG Std. 802.11e, 2002.

[4] 3GPP TR 22.934, "Feasibility study on 3GPP system to Wireless Local Area Network (WLAN) interworking"

[5] 3GPP TS 23.101, "General UMTS Architecture".

[6] 3GPP TR 43.930, "Iur-g interfaces".

[7] G. Bianchi. Throughput evaluation of the IEEE 802.11 distributed coordination function. In Proc. 5th Int. Wshp. on Mobile Multimedia Communications, pages 307-318, Berlin, 1998

[8] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed coordination Function", IEEE Journal on selected areas in communications, vol. 18, No. 3, March 2000.

ACKNOWLEDGEMENTS

This work has been performed in the framework of the project IST-EVEREST (www.everest-ist.upc.es), which is partly funded by the European Community. The Authors would like also to acknowledge Eng. Javier Araque Quijano for his contribution.