2005 International Conference on Wireless Networks, Communications and Mobile Computing

Analysis and Evaluation of a New MAC Protocol for Broadband Wireless Access

L. F. M. de Moraes, P. D. Maciel Jr.
High-Speed Networks Laboratory - RAVEL
Systems Engineering and Computer Science Program - COPPE/UFRJ
Tel/Fax: +55 21 2562-7844 - Rio de Janeiro - RJ, Brazil
{moraes,pdmjr}@cos.ufrj.br

Abstract— **Broadband Wireless Access (BWA) systems offer a solution for broadband access and high data rate transmission of multimedia services with distinct Quality-of-Service (QoS) requirements through a wireless medium. IEEE 802.16 Standard specifies the PHY and MAC layers for BWA systems. However, in spite of including the possibility for QoS support, the standard does not define how to schedule different types of traffic. This article propose a new MAC protocol for BWA systems that incorporates a traffic scheduling mechanism based on messages and/or stations priorities. An analytical model to evaluate the performance of the proposed protocol is also developed and results obtained for the messages waiting times of different traffic classes are presented.**

I. Introduction

The IEEE 802.16 standard [1] for Broadband Wireless Access (BWA) proposes a wireless high-speed and high performance access system, with service differentiation for distinct classes of traffic and Quality-of-Service (QoS) requirements. However, in spite of mentioning and providing the means to support different types of traffic, the IEEE 802.16 standard leaves it open in terms of an implementation which is able to guarantee QoS to different applications. In other words, it does not specify how to efficiently schedule the traffic related to different applications in order to meet their specific requirements. On the other hand, proposals suggested to manipulate different types of traffic in 802.16, which are available in the literature [2]- [6], are mainly related to signaling issues and also do not specify any scheduling algorithm or admission control mechanism to handle QoS for different types of traffic.

This article presents the proposal of a MAC reservation protocol with a traffic scheduling mechanism based upon priority rules, for use in BWA systems and which can be easily implemented as part of the IEEE 802.16 standard to support aggregate level and relative QoS. The access control scheme utilizes TDMA periods in order to identify users' demands for transmissions across the channel. Those reservation periods are then followed by transmission intervals, during identified and reserved terminals are able to send their data in accordance with priorities rules devised to satisfy the relative QoS for different types of traffic applications.

The remainder of this text is organized as follows. Section II gives a short description of IEEE 802.16 standard. The proposed media access protocol is described in Section III. An analytical model, which allow to obtain the average message waiting time for different priority classes, is provided in Section IV. Section V presents some numerical results and the paper is concluded with a few discussions concerning the proposal in Section VI.

II. IEEE 802.16 Standard

We give next a brief description of the PHY and MAC layers of IEEE 802.16, as well as the QoS architecture specified by the standard. We noted that 802.16 only describes support for the implementation of the described architecture, without specifying explicitly how a particular solution should be exactly implemented.

A. PHY and MAC Layers

The basic architecture consists of one Base Station (BS), and one or more Subscriber Stations (SSs). The BS is the main node, which is responsible for coordinating the whole communication process among the SSs. Therefore, it is assumed that there is no direct communication taking place between SSs, which must rely upon the BS to communicate with each other. Transmissions are assumed to take place through two independent channels: a Downlink Channel (DL), where the data flow is directed from the BS to the SSs, and an Uplink Channel (UL), which is used to send data from the SSs to the BS. Hence, there is no contention associated with the the DL channel, while the UL channel must be shared by the SSs through the use of some multiple access control protocol.

The standard provides the flexibility of two channel allocation schemes: Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD). Basically, in the FDD the DL and UL use different frequencies, while in the TDD, both channels share the same frequency and the data in each channel (DL and UL) are transmitted in different time slots. The channel is assumed to be time slotted and composed of fixed-length frames. Each frame is divided into DL and UL sub-frames. The duration of each of these sub-frames is dynamically controlled by the BS.

Transmission in the DL is relatively simple because only the BS transmits during the associated sub-frame. Data are transmitted in broadcast to all the SSs; each SS captures only those packets which are destined to itself. For the UL, the BS determines the number of slots to be allocated for each SS in the correspondent sub-frame. This information is then broadcast by the BS in an UL-MAP message at the beginning of each frame. The UL-MAP contains specific data (Information Element – IE) that include the transmission opportunities, that is, the time slots during which the SS can transmit during the UL sub-frame. After receiving the UL-MAP message, the stations transmit their data in pre-defined time slots as indicated in the IE. A scheduling module for the UL is necessary to be kept in the BS in order to determine the transmission opportunities (IEs) using the bandwidth requests (BW-Request) sent by the SSs. The Figure 1 illustrates the structure of the MAC frame in the TDD allocation scheme.

The IEEE 802.16 [1] utilizes random access and *piggybacking* in the UL sub-frame to send transmission opportunity requests by the SSs. This one is responsible for establishing a reservation period at the beginning of each UL, so that the SSs can place reservations to transmit in the next UL sub-frame (or later, depending on the occurrence or not of collisions). The standard defines the *binary*

0-7803-9305-8/05/$20.00 ©2005 IEEE

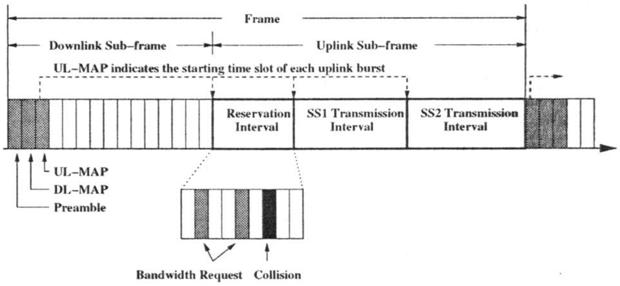


Fig. 1. MAC frame structure in TDD scheme.

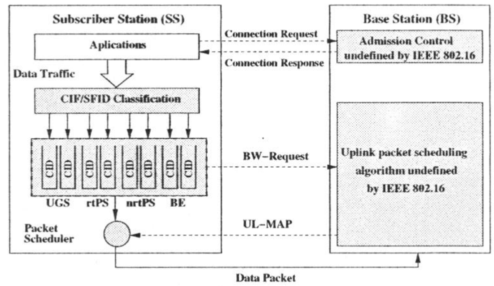


Fig. 2. QoS architecture of IEEE 802.16.

truncated exponential backoff algorithm for collision resolution in this interval. An SS detects the occurrence of collision if the UL-MAP of the next frame does not contain any transmission opportunities addressed to it. However, the 802.16 only defines the signaling mechanisms for QoS, such as BW-Request and UL-MAP; it does not define the UL scheduler, the mechanism that determines the IEs in UL-MAP. Another feature of the standard is the support for transmission opportunity requests based on connection (*Grants per Connection* - GPC) or by station (*Grants per Subscriber Station* - GPSS). In the GPSS, each station requests transmission opportunities as a set for all the services it maintains, and is responsible to allocate the opportunities received among its different types of flows.

B. QoS Architecture

The IEEE 802.16 supports many traffic types (data, voice, video) with different QoS requirements. In this context, the MAC layer defines QoS signaling mechanisms and functions for data control transmissions between the BS and the SSs. In addition, the standard defines four types of data flows, each one with distinct QoS requirements, and an appropriate policy for the UL scheduler [1]:

1) *Unsolicited Grant Service* (UGS): for applications that need constant bandwidth allocation.
2) *Real-Time Polling Service* (rtPS): for applications that have specific bandwith requirements and maximum acceptable delay.
3) *Non-Real-Time Polling Service* (nrtPS): for applications that are intolerant to delay and require a minimum bandwidth allocation.
4) *Best Effort Service* (BE): for applications that receive the remaining bandwidth after the allocation to the three previous types of services.

For UGS traffic, BW-Request is not necessary. For other types, the actual queue length is included in the BW-Request message to represent the current demand for transmission. In summary, the IEEE 802.16 specifies: the signaling mechanism for information exchange between the BS and the SSs, as the connection configuration, BW-Request and UL-MAP; and the scheduling of UL for UGS traffic. The standard does not define: the scheduling of UL for rtPS, nrtPS and BE services; admission control and traffic policing.

Figure 2 shows the QoS architecture present in 802.16. The UL packet scheduling (UPS) module is found in the BS and controls all the packet transmissions in the UL. As the protocol is connection-oriented, the application should establish a connection between the BS and the associated service flow (UGS, rtPS, nrtPS or BE). The BS identifies the connections by assigning a unique Connection ID (CID) to each one. The 802.16 defines the signaling process for the establishment of a connection (Connection-Request and Connection-Response) between SS and BS, but does not specify the rules for admission control.

All the packets in the application layer of an SS are classified in accordance to the CID and forwarded to an appropriate queue. The SS recovers the packet in the queue and sends it into the network at the time slot determined in the UL-MAP sent by BS. The UL-MAP is defined by the UPS module based on the BW-Request messages that report the actual size of the queue for each connection in the SS.

III. Proposed Protocol

The proposed priority schedule protocol is based on RPAC (Reservation-Priority Access Control) described in [7] and [8], where reservation periods use TDMA with one time slot allocated per station in the network. After reservation periods, the stations are allowed to transmit their messages in accordance to the established priority rules. Therefore, a broadcast communication channel is assumed, as is the case for the BWA systems considered in this paper.

The main advantage of using a TDMA (fixed assignment) scheme for bandwidth requests is to provide a simple and efficient reservation. On the other hand, the use of polling or probing discipline in the reservation period of the UL sub-frame requires either multiple switching of the wireless devices from the transmission to the reception mode and vice-versa, and that the stations are in the same range [9]. A disadvantage of TDMA (as with any fixed assignment scheme) is the waste of channel bandwidth due to pre-allocated slots to stations that might be idle (i.e., with no messages to be transmitted). However, we note that this waste is relatively small (specially for medium and high traffic load [10]) if compared to the time that a message stay buffered due to the inefficiency of the collision resolution algorithm [11]. In addition, as will be seen, the length of reservation periods in the proposed protocol are fixed and relatively small as compared with the (random) periods where messages are transmitted across the channel. Therefore, the assignment rule utilized to make the reservations will have a minor impact in the overall performance of the proposed protocol.

The proposed protocol uses TDMA by the SSs to send bandwidth requests, where one slot is allocated for each station of the network. These requests are centralized by the BS, that coordinates the access to the transmition channel. Therefore, it is assumed that the each of the SSs are within the range of the BS; however, different stations are not limited to be in the line of sight of each other. Another relevant characteristic of this protocol is the incorporation of a traffic scheduler that uses priority rules, supporting an aggregate level and relative QoS based on messages or stations, as will be exposed ahead.

The communication channel is assumed to be completely synchronized by the BS. Therefore, the time axis is subdivided into fixed intervals called slots, each of which has duration equal to τ seconds. Transmissions are synchronized and can only be started at the beginning of each time slot. In each reservation interval, time slots are assigned to network stations using TDMA (one slot per station in each reservation interval). During this period and using their pre-allocated slots, stations must send all the necessary information concerning the messages residing in their respective transmission buffers to the BS. In the paper, we only use the TDD allocation scheme; however, the following results can be easily extended to the FDD scheme.

The MAC frame structure for the proposed protocol is illustrated in Figure 3. Differently from the 802.16 standard, the frame lengths are not fixed in the proposed scheme. Basically, the length of each frame will depend upon the number of packets arriving to the stations in the previous frame. Reservation periods are located at the end of the UL sub-frame (not at the beginning as in 802.16). Those periods are used by each of the stations to inform the BS about the services for which a bandwidth reservation is being requested, as well as the number of packets to be transmitted for each of those services. After processing all the requests, the BS sends, in the DL sub-frame of the next frame, a UL-MAP with transmission opportunities to all reserved stations.

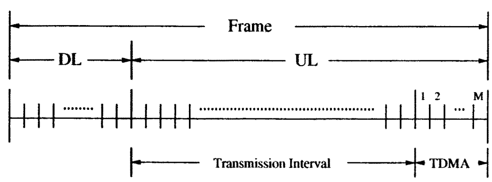


Fig. 3. MAC frame structure for the proposed protocol.

For analysis of the protocol, the activity in the channel can be seen as a sequence of reservation, downlink, and transmission intervals, where each pair of DL and UL constitutes a transmission cycle, as illustrated in Figure 4. It is important to note that there is a difference between the MAC frame and the transmission cycle, despite both having an equal size, because the reservation period is fixed. In fact, the n-th cycle is formed by the reservation period of $(j-1)$-st frame, plus the downlink and transmission intervals of the j-th frame. With this definition of the transmission cycles, the analytical approach described in [7] can be used for the analysis of the average waiting-time of the messages, as will be seen in Section IV. Referring to the picture in Figure 4, we define L_n^R, L_n^{DL} and L_n^T, respectively, as the lengths (or duration, given by the number of slots) of the reservation, the downlink, and the transmission intervals in the n-th cycle. Thus, $L_n = L_n^R + L_n^{DL} + L_n^T$ represents the total length of the n-th cycle. Following the TDMA protocol, each reservation period is composed by M slots ($L_n^R = M; n = 1, 2, 3, ...$), where M is the number of stations in the network. During this period, of duration of $M\tau$ seconds, each station is associated to a single slot in a fixed manner.

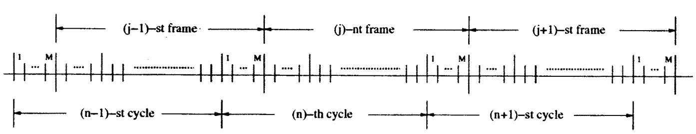


Fig. 4. Consecutive transmission cycles.

We note that the length of a current (say, the n-th) cycle will depend on the number of messages that arrived during the previous (say, the $(n-1)$-st) cycle. This happens because request for the messages that arrived during the $(n-1)$-st cycle will be transmitted in the reservation period of the n-th cycle. After this period, the BS performs the centralized processing of the transmission opportunities and sends a UL-MAP in the DL sub-channel, still in the n-th cycle. Afterwards, the stations transmit their messages in the transmission interval of the same cycle, following the priorities established in the UL-MAP. Therefore, messages arriving during an ongoing cycle get transmitted only in the subsequent cycle.

We propose a medium access protocol with priorities based on messages and/or stations, in accordance to the 802.16 protocol, that uses GPC or GPSS admissions. We assume that, following the reservation phase and the DL sub-frame, the channel is allocated to the stations following the sequence $1, 2, 3, ..., M$. Thus, according to the priority rules used to determine the order in which the messages should be transmitted during the transmission period, the following versions of the protocol are considered:

- **Version I**, in which, for any $p, q \in \{1, ..., P\}$ such that $p < q$, all the class-p messages are transmitted before any messages of class q, independent of which station it belongs to. For messages belonging to the same class but in distinct stations, the order of the transmissions is according to the order in which the stations access the channel (first station 1 and last station M). For messages in the same station with the same class of priorities, the transmissions occur by order of arrival.
- **Version II**, in which, for any $i, j \in \{1, ..., M\}$ such that $i < j$, all the messages in station i are transmitted before any message in station j, independently of it's priority class. In any terminal, the messages are transmitted in accordance with their priority classes and in order of arrival, in the case of belonging to the same class; that is, at each station, the priority discipline HOL (Head-Of-the-Line) is applied with the highest priority assigned to class 1 and the lowest assigned to class P.

The behavior of the channel, according to Versions I and II of the proposed protocol, is illustrated in Figures 5 and 6, respectively. Note that, in Version II, a higher priority station transmits all it's messages before one of lower priority. So, unlike Version I, it is possible that messages of lower priority are transmitted before messages with higher priorities. We observe that a rigid admission control carried by the BS and also by the SSs is necessary so that the heavy traffic of a specific class (or station) does not overload (and hog) the channel, affecting the response time of the others.

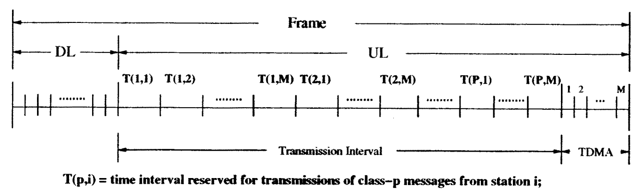


Fig. 5. Version I of proposed protocol.

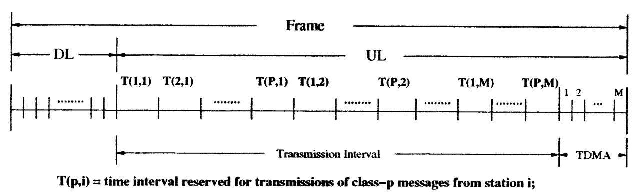


Fig. 6. Version II of proposed protocol.

IV. Analytic Model

The technique used to obtain the average waiting-time for the messages is similar to the method used in [7] and [12]. Due to space limitations, the reader should look at the above-mentioned references for additional and more specific details. We outline first the analysis for Version I of the proposed protocol. In the sequence, we present an adaptation of the result obtained to get the analysis for Version II.

A. Version I

The system considered has one BS and M ($M \geq 1$) client stations (SSs), each of which has an infinite buffer-size and is already associated with the base station. The transmission channel is assumed to be error-free, with a transmission rate equal to C bit/s. In general, the messages generated at each station are composed of a random number of fixed units of data, called packets, each of which contains μ^{-1} bits. The transmission time of each packet is made equal to a time slot (τ). Therefore, $\tau = (\mu C)^{-1}$.

Messages arriving at each station, belong to one of the different P classes. We assume that class-1 messages have the highest priority and class-P messages lowest. At each station, the arrival of messages is characterized by a Poisson point process, such that λ_i^p (messages per slot) is the average arrival rate of class-p messages to station i. The number of packets that compose the m-th class-p message at the i-h station is denoted by $B_{i,m}^p$ ($i = 1, 2, ..., M;\ p = 1, 2, ..., P$). For each $i = 1, 2, \ldots M$, the random variables (r.v.s) $\{B_{i,m}^p; m \geq 1\}$ are assumed to be independent and identically distributed (i.i.d.), with distribution $\beta_{i,j}^p = P(B_{i,m}^p = j)$ ($j = 1, 2, ...;$), average b_i^p, and second moment $b_{2,i}^p$. As defined in Section III, $L_n = L_n^R + L_n^{DL} + L_n^T$ is the length of the n-th cycle, where L_n^R is the length of the n-th reservation interval, L_n^{DL} is the size of the DL sub-frame belonging to the n-th cycle, and L_n^T is the size of the n-th transmission interval. We assume that $L_n^R = M$ (where M is the number of stations in the network) with duration of $M\tau$ seconds.

The r.v.s representing the number of class-p messages arriving at the terminal i during the $k-th$ slot, $\{N_{i,k}^p; k \geq 1\}$ ($i = 1, 2, ..., M; p = 1, 2, ..., P$) are Poisson with mean λ_i^p (messages/slot), independently of the arrival process for other classes. Hence, the r.v.s $N_{i,k}$, representing the total number of messages arriving at the terminal i during the k-th slot, are also Poisson, with average given by $\lambda_i = \sum_{p=1}^{P} \lambda_i^p$; for each $i \in \{1, 2, ..., M\}$.

Let $W_{i,n}^p$ denote the waiting time (measured in slots) for the n-th class-p message arriving at station i. The latter is defined as the number of slots from the arrival to station i of the n-th class-p message, until the slot where its transmission across the channel is initiated. In [7], a reservation scheme with the same structure defined for the proposed access protocol was analyzed, and the steady-state results for $W_{i,n}^p$ were obtained by using a Markov Ratio Limit Theorem (see [7] for details). By noting the similarities between the reservation scheme in [7] and the one defined here, we use the results from that paper to obtain the following expression for the steady-state average of class-p messages at station i, $W_i^p = \lim_{n\to\infty} W_{i,n}^p$, for Version I of the proposed protocol:

$$\overline{W_i^p} = M + E[DL] + [\frac{(1+\rho_i^p)}{2} + \sum_{j=1}^{p-1}\sum_{g=1}^{M} \rho_g^j + \sum_{j=1}^{i-1} \rho_j^p] \frac{E[L^2]}{E[L]} - \frac{1}{2}, \quad (1)$$

where $\rho_i^p = \lambda_i^p b_i^p$ is the traffic in the terminal i due to the messages of class p, and $E[DL] = \lim_{n\to\infty} E[L_n^{DL}]$. The expression above for $\overline{W_i^p}$ is still given as a function of $E[L]$ and $E[L^2]$, the first and second steady-state moments of L_n. Following up with the analysis to find the first and second moments of the cycle length, we have (see [7] for details):

$$E[L_n] = E[L_n^R + L_n^{DL} + L_n^T] = M + E[L_n^{DL}] + \sum_{p=1}^{P}\sum_{i=1}^{M} \rho_i^p E[L_{n-1}]. \quad (2)$$

Then, making $\rho = \sum_{p=1}^{P}\sum_{i=1}^{M} \rho_i^p < 1$ and taking the limits (with $n \to \infty$) in both sides of the equation (2) we obtain $E[L]$:

$$E[L] = \frac{M + E[DL]}{1-\rho};\ \rho = \sum_{p=1}^{P}\sum_{i=1}^{M} \rho_i^p < 1. \quad (3)$$

In a similar way, through a recursive equation for $E[L_n^2]$ and assuming $\rho < 1$, we obtain the equation (4) for $E[L^2]$.

$$\begin{aligned} E[L^2] &= \frac{1}{1-\sum_{p=1}^{P}\sum_{i=1}^{M}(\rho_i^p)^2}\Big\{M^2 + E[DL]^2 + 2ME[DL] \\ &+ \Big[2\rho(M + E[DL]) + \sum_{p=1}^{P}\sum_{i=1}^{M}\lambda_i^p b_{2,i}^p\Big]E[L] \\ &+ \Big[\sum_{p=1}^{P}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}\rho_i^p\rho_j^p + \sum_{p=1}^{P}\sum_{\substack{q=1\\ q\neq p}}^{P}\sum_{i=1}^{M}\sum_{k=1}^{M}\rho_i^p\rho_k^q\Big]E^2[L]\Big\} \end{aligned} \quad (4)$$

Finally, substituting equations (3) and (4) in (1), we have a closed expression for the average waiting-time of the class-p messages in the station i with the Version I of the proposed protocol.

B. Version II

The analysis for Version II of the proposed protocol follows in a direct manner noting that, according to this scheme, the messages are transmitted in the same order as the Version I, with the classes of the messages exchanged for the numbers of the stations and vice-versa (see Figures 5 and 6). Therefore, the expression for $\overline{W_i^p}$ in the Version II is analogous to that of Version I, changing only the i for the p and the M by the P and vice-versa. Thus, we obtain the follow expression for the steady-state average waiting-time of the class-p messages, at station i with the Version II of the proposed protocol.

$$\overline{W_i^p} = M + E[DL] + [\frac{(1+\rho_i^p)}{2} + \sum_{j=1}^{i-1}\sum_{k=1}^{P} \rho_j^k + \sum_{j=1}^{p-1} \rho_j^p] \frac{E[L^2]}{E[L]} - \frac{1}{2}, \quad (5)$$

with $E[L]$ and $E[L^2]$ given by the equations (3) and (4), respectively. In the next section, some numerical results will be presented to illustrate the performance of the proposed protocols with message-based priorities.

V. Numeric Results

To evaluate the level of differentiation obtained with the described protocols, we considered two distinct scenarios, where in each scenario there is a differentiated probability between four types of traffic classes ($P = 4$), as shown in Table I.

Traffic Classes	Scenario I	Scenario II
Class 1	40%	10%
Class 2	30%	20%
Class 3	20%	30%
Class 4	10%	40%

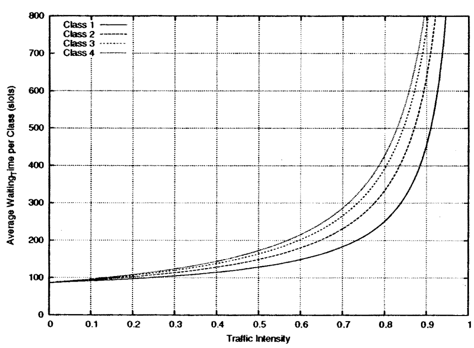
TABLE I
TRAFFIC SCENARIOS USED.

The differences between the scenarios is that, in Scenario I, there is a greater probability for the classes of high priority, while in the Scenario II the classes of low priority prevail over those of high priority. Thus, it is possible to compare which is the influence of a greater load of flows of lower priority over those of higher priority and vice-versa.

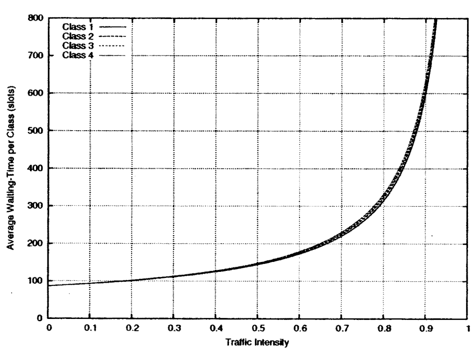
In each scenario there are 10 stations ($M = 10$) with a balanced traffic between these, that is $\lambda_i^p = \lambda^p/10$, where λ^p represents the rates of messages from class p and λ_i^p represents the rate of messages of class p in the station i. We assume that the number of packets in each message of class p in the station i is constant with average $b_i^p = 5$ and $b_{2,i}^p = 25$, for each $p = 1, 2, 3, 4$; and $i = 1, ..., 10$. So, the average waiting-time for the class p is given by:

$$\overline{W^p} = \sum_{i=1}^{10} \frac{\lambda_i^p}{\lambda^p} \overline{W_i^p}.$$

The Figures 7 and 8 illustrates the average waiting-time in the queue for each class of priorities related to the offered traffic in the channel. The behavior of Versions I and II of the proposed protocol in Scenario I is shown in Figures 7(a) and (b) respectively. In the same manner, Figures 8(a) and (b) shown the behavior of both versions in Scenario II.



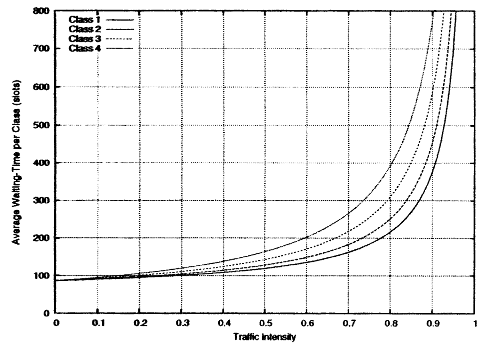
(a) Version I



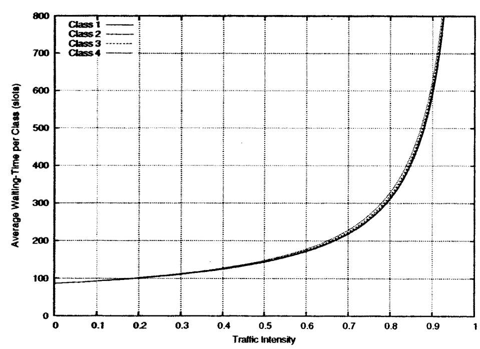
(b) Version II

Fig. 7. $\overline{W^p}$ in Scenario I: Version I (a) and Version II (b).

By the figures, it is possible to observe a more evident differentiation for a heavy traffic in the channel and that, with the raise of traffic intensity, the waiting time in the queue grows for all classes. However, this differentiation is smaller for Version II of the proposed protocol, as shown by Figures 7(b) and 8(b). This happens because, in Version I, the priorities between classes prevail over the priorities between stations, occurring the opposite in the Version II where the



(a) Version I



(b) Version II

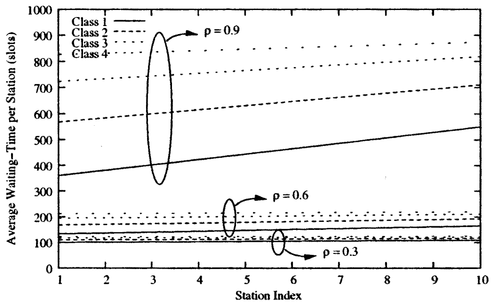
Fig. 8. $\overline{W^p}$ in Scenario II: Version I (a) and Version II (b).

priorities between the stations superpose. Through Figures 7(a) e 8(a) we observe that the waiting-time for the traffic of higher priority (class 1) is smaller in respect to other classes, even in the Scenario II where there is a greater probability of the low-priority traffic. With this, it can be perceived that the proposed protocols manage to differentiate efficiently the traffic classes, guaranteeing a smaller waiting-time in the queue for the messages of greater priority.

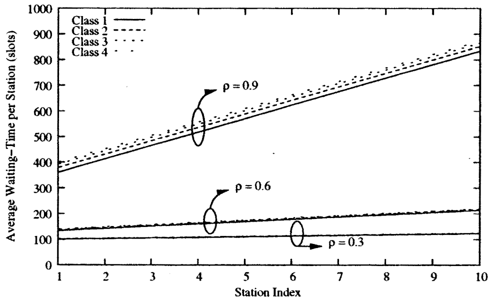
Figures 9 and 10 presents the average waiting-time in the queue in function of terminals for three values of traffic intensity ($\rho = 0,3; 0,6$ e $0,9$). In this way, the average waiting-time in the queue for the station i is given by the equation below, where λ_i represents the message rate in the i station.

$$\overline{W_i} = \sum_{p=1}^{4} \frac{\lambda_i^p}{\lambda_i} \overline{W_i^p}$$

The behavior of both versions of the proposed protocol in Scenario I is illustrated in Figures 9(a) and (b). Next, Figures 10(a) and (b) introduce, respectively, the Version I and II with the utilization of Scenario II. From Figures 9(a) and (b), it can be observed that for fixed p values, the difference in the waiting-time in the queue between the station in Version I is smaller than that of Version II for, as has been said before, in Version I, the priorities are defined at first by classes and not by stations. Therefore, the variation of the waiting time between classes is smaller for Version II that promotes a differentiation in the media access between the stations in such a way that, increases the average waiting-time in the queue for the messages of the station with a low priority. In respect to the change between the traffic scenarios, the second version behaves in a similar way for the Scenario I and II, as illustrate Figures 9(b) and 10(b). On the other hand, in Version I the waiting-time for the classes are slightly smaller

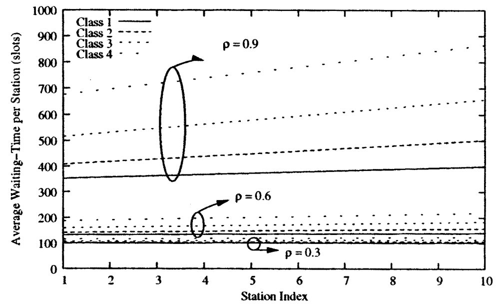


(a) Version I

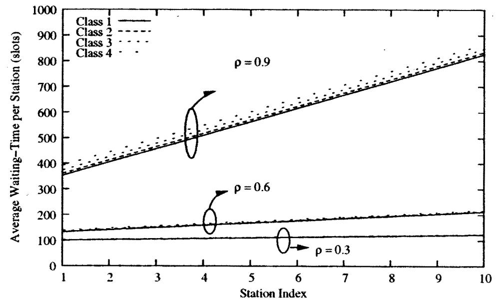


(b) Version II

Fig. 9. $\overline{W}_i$ in Scenario I: Version I (a) and Version II (b).



(a) Version I



(b) Version II

Fig. 10. $\overline{W}_i$ in Scenario I: Version I (a) and Version II (b).

in Scenario II, preserving, however, the differentiation between traffic classes, as can be seen in Figures 9(a) and 10(a).

VI. Conclusions

In this paper we proposed a new MAC protocol for BWA systems that incorporating traffic scheduling functions with message- or station-based priorities. Moreover, we presented an analytic model for the mean message waiting-time to two versions of the proposed protocol, under the assumptions of Poisson arrivals and general distributions of message lengths. From the results, we conclude that the proposed protocol can provide service differentiation between distinct traffic types, even under heavy load, decreasing the mean waiting-time for the highest priority class. Furthermore, Version I presented a greater degree of fairness in the access to the media between the stations of the network than Version II, taking into account a greater static precedence between the stations.

As future work, we intend to evaluate the behavior of the proposed protocols with the inclusion of variable priorities between stations, where an improvement of the fairness degree in the access to the media in the Version II is expected. Moreover, we intend to include to the protocols an admission control mechanism so that the overload of a specific type of traffic doesn't affect the response time of the others.

Acknowledgment

This work was supported by FINEP, through project grant number 1956/04.

References

[1] IEEE 802.16, *IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Std. 802.16, Oct. 2004.

[2] M. Hawa and D. W. Petr, "M/G/FQ: Stochastic Analysis of Fair Queueing Systems," in *IEEE 2nd Inter. Conf. on Networking (ICN 2002)*, Atlanta, Georgia, Aug. 2002.

[3] ——, "Quality of Service Scheduling in Cable and Broadband Wireless Access Systems," in *Tenth Inter. Workshop on Quality of Service (IWQoS 2002)*, Miami, Florida, May 2002.

[4] M. Hawa, "Stochastic Evaluation of Fair Scheduling with Applications to Quality-of-Service in Broadband Wireless Access Networks," Ph.D. dissertation, Univ. of Kansas, Aug. 2003.

[5] K. Wongthavarawat and A. Ganz, "IEEE 802.16 Based Last Mile Broadband Wireless Military Networks with Quality of Service Support," *IEEE Military Comm. Conf. (MILCOM'03)*, vol. 2, Oct. 2003.

[6] X. Xu, D. Liang, H. Jiang, and X. Lin, "Dynamic Bandwidth Allocation in Fixed BWA Systems," *Proc. of Inter. Conf. on Comm. Technology (ICCT'03)*, vol. 2, Apr. 2003.

[7] L. F. M. de Moraes and A. N. L. Valverde, "Waiting-Time Analysis of a Reservation Access-Control Scheme with message-based priorities," *Third IFIP TC 6/wg 7.3 Inter. Conf. on Data Communication Systems and Their Performance*, June 1987.

[8] L. F. M. de Moraes, "Priority Scheduling in Multiaccess Communication," *In: Stochastic Analysis of Computer and Communication Systems*, vol. edited by H. Takagi; North-Holland, 1990.

[9] C. Papachristou and F. Pavlidou, "Collision-Free Operation in Ad Hoc Carrier Sense Multiple Access Wireless Networks," *IEEE Comm. Letters*, vol. 6, no. 8, 2002.

[10] J. Capetanakis, "Generalized TDMA: The Multi-Accessing Tree Protocol," *IEEE Trans. on Comm.*, vol. COM-27, no. 10, Oct. 1979.

[11] K. Sriram, "Performance of MAC Protocols for Broadband HFC and Wireless Access Networks," *Advances in Performance Analysis*, vol. 1, no. 1, 1998.

[12] I. Rubin and L. F. M. de Moraes, "Message Delay Analysis for Polling and Token Multiple-Access Schemes for Local Communication Networks," *IEEE JSAC*, vol. SAC-1, no. 5, Nov. 1983.