# Terminal-Assisted Hybrid MAC Protocol for Differentiated QoS Guarantee in TDMA-Based Broadband Access Networks

Seung-Eun Hong, Chung-Gu Kang, and O-Hyung Kwon

**This paper presents a terminal-assisted frame-based packet reservation multiple access (TAF-PRMA) protocol, which optimizes random access control between heterogeneous traffic aiming at more efficient voice/data integrated services in dynamic reservation TDMA-based broadband access networks. In order to achieve a differentiated quality-of-service (QoS) guarantee for individual service plus maximal system resource utilization, TAF-PRMA independently controls the random access parameters such as the lengths of the access regions dedicated to respective service traffic and the corresponding permission probabilities, on a frame-by-frame basis. In addition, we have adopted a terminal-assisted random access mechanism where the voice terminal readjusts a global permission probability from the central controller in order to handle the 'fair access' issue resulting from distributed queuing problems inherent in the access network. Our extensive simulation results indicate that TAF-PRMA achieves significant improvements in terms of voice capacity, delay, and fairness over most of the existing medium access control (MAC) schemes for integrated services.**

**Keywords: Broadband access network, MAC, QoS, TDMA.**

Manuscript received July 12, 2005; revised Mar. 01, 2006.

Seung-Eun Hong (phone: + 82 42 860 1286, email: iptvguru@etri.re.kr) and O-Hyung Kwon (email: ohkwon@etri.re.kr) are with Digital Broadcasting Research Division, ETRI, Daejeon, Korea.

Chung-Gu Kang (email: ccgkang@korea.ac.kr) is with the Department of Radio Communication Engineering, Korea University, Seoul, Korea.

## I. Introduction

Broadband access networks (BANs) such as DOCSIS [1] and IEEE 802.16 [2] have the frame structure where feedback for the uplink transmission is available only at the beginning of the next frame. Both define different types of service flows that should be treated differently by the medium access control (MAC) protocol scheduling process. In both quality-of-service (QoS) architectures, a real-time service flow such as voice uses polling or an unsolicited granting mechanism in a contention-free way, while a best-effort service flow such as data uses a contention-based random access mechanism. However, in order to support packetized voice service, it is generally believed that the packet reservation multiple access (PRMA) [3] scheme with reservation mechanism by random access is efficient.

There have been many different variants of the PRMA protocol to integrate both voice and data services [4]-[8]. Existing PRMA protocols in the literature of QoS-oriented BAN systems with a TDMA frame structure [6]-[9] are classified into two large categories: fixed-request-subframe and dynamic-request-subframe types. In the sequel, MAC protocols can further bifurcate into two structures, one with unified access region [6] and the other with separated access regions [7]-[9] for voice and data. The scheme in [6] introduces the framed pseudo-Bayesian ALOHA with priorities. In [7], throughput efficiency is enhanced by the use of variable request minislots in MAC protocols, and a heuristic approach to compute the number of minislots is presented. In [8], the lengths of separated access regions are fixed.

Although the PRMA is a classical MAC protocol, it is still one of the best alternatives because it is simple and can be efficiently used to transmit variable bit rate video as well as voice and data in a distributed environment such as an access network [10]. Recently, many works have focused on supporting multimedia services over wireless networks, which is challenging due to such constraints as power savings, scarce bandwidth, and a time-varying fading effect [11]-[14]. Cross-layer design engineering holds great promise for addressing these challenges and providing reliable and high-quality end-to-end performance in wireless multimedia communications. To the best of the authors' knowledge, the cross-layer schemes can be adapted into our scheme, although the cross-layer methodologies are beyond the scope of this paper.

In this paper, we propose an access control scheme for adaptively determining the numbers of minislots for separated contention regions and those for information regions of different traffic, and for then determining the corresponding access permission probabilities for contention regions on a frame-by-frame basis.

The rest of the paper is organized as follows. Section II introduces a simple taxonomy of MAC protocols suggested in the literature for integrated voice/data services. The system model and problem statements are described in section III. The terminal-assisted frame-based PRMA protocol is presented in section IV. In section V, simulation results are introduced and discussed. Section VI gives some concluding remarks and future works.

## II. Taxonomy of TDMA MAC Protocols

In the literature of MAC protocols for integrated voice/data services, some points exhibit a large design space. In this section, we introduce a simple taxonomy of the MAC protocols suggested in the literature for integrated voice/data services. Working on a slot-by-slot transmission basis with immediate feedback where terminals can identify the results immediately after transmitting their reservation requests, some variants of classical PRMA are deployed in order to deal with two major problems arising from usage of the fixed parameters, such as voice and data permission probabilities: inflexibility in adjusting to a variable traffic load and undue degradation in data traffic QoS. In an integrated (IPRMA) [4], data terminals that contend successfully also reserve the next min($B$, $K$) free slots to transmit their burst (where $B$ is the length of their burst in packets, and $K$ is a limit enforced by the system). This improves the performance of the data users by introducing a reservation mechanism into data traffic. In the hybrid ALOHA-reservation/R-ALOHA (HAR) scheme [5], permission probabilities are dynamically adjusted according to traffic load.

In contrast with the variants of PRMA schemes in the above, QoS-oriented BAN systems with a TDMA frame structure work on a delayed feedback basis, where terminals are informed of the results at least one TDMA frame after transmitting their reservation requests over the request subframe. There is thus a need to design an efficient random access control since each terminal has at best one opportunity for a reservation request per TDMA frame. Existing MAC protocols in the literature [6]-[9] are classified into two large categories: fixed-request-subframe and dynamic-request-subframe types. In the sequel, MAC protocols can further bifurcate into two structures: one with unified access region and the other with separated access regions for voice and data. In the former structure, collisions not only among homogeneous request packets but also between heterogeneous (that is, voice and data) request packets can take place, where Frigon's system [6] belongs. The major shortcoming of this approach is that real-time QoS such as voice packet-delay and voice packet-drop probability may be swayed by non-real-time traffic or non-real-time traffic QoS such as data delay, which should be over-sacrificed. In order to overcome this problem, Frigon's system introduces the framed pseudo-Bayesian ALOHA with priorities. From now, we will abbreviate the system to FPBP. MAC protocols in the latter structure divide the request subframe into two access regions for supporting the two types of traffic. Thus, only collisions between homogeneous request packets occur, in which the systems of Sriram and Magill [7], Li [8], and Ren and others [9] are put. In [7], throughput efficiency is enhanced by the use of variable request minislots in MAC protocols (shortly, VR-MAC) and a heuristic approach to compute the number of minislots is presented. It is however, not flexible in that the fixed number of minislots for real-time traffic is used and the multiplication factor in the heuristic formula for non-real-time traffic is difficult to determine. In [8], the lengths of separated access regions are fixed. Both, however, are not flexible in that the fixed numbers of minislots for voice or data is used. In [9], an enhanced PRMA protocol with an intelligent transmission controller (E-PRMA/ITC) using fuzzy control and a neural network is introduced, which provides a key idea in our work. Employing a similar framework, we note that E-PRMA/ITC has two major drawbacks: estimating the number of contending users through the received signal power under the assumption of perfect power control, and introducing a fixed reservation subframe size.

## III. System Model and Problem Statements

In infrastructure-based BANs, the access point (AP) and the terminal equipment (TE) communicate with each other using

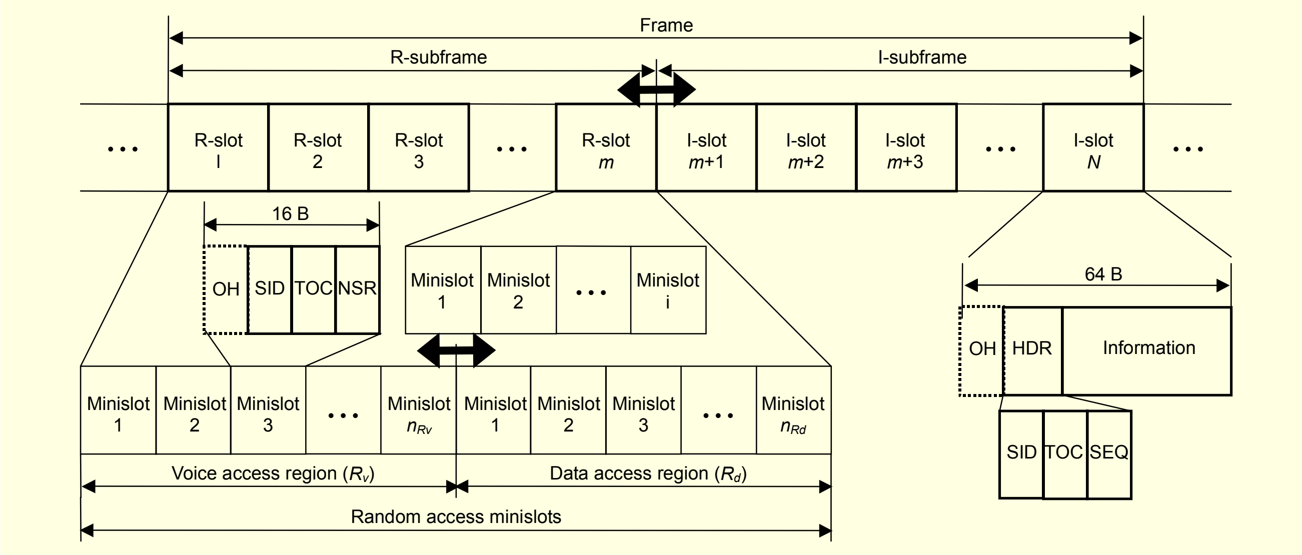



Fig. 1. Frame and slot format for framed random access protocol (OH: physical overhead, SID: source identifier, TOC: type of traffic class, NSR: number of slots requested, HDR: header, and SEQ: sequence number).

two logically separated links: An AP uses the downlink to broadcast control traffic and information traffic to the TE, while the TE uses the uplink to transmit related traffic to the AP.

As shown in Fig. 1, an uplink frame is composed of $N$ time slots (with a duration of $T$) within which each frame is further divided into a request subframe (R-subframe) and an information subframe (I-subframe). The R-subframe is composed of $m$ reservation slots (R-slots), each further subdivided into $i$ minislots, with each minislot holding a reservation request packet [15]. The $m \times i$ minislots in the R-subframe are partitioned into two access regions, one each for voice and data. The boundary between these two access regions is specified by $n_{Rv}$, which is the number of minislots assigned for the voice access region ($1 \leq n_{Rv} \leq m \times i$ ). In the voice access region, only voice users can send their reservation request packets using a slotted ALOHA protocol with an access permission probability of $p_v$. For the data access region, there will be the associated probability of permission probability $p_d$. Contending voice or data terminals have permission to transmit a request at the start of an R-subframe if they select a smaller random number than the corresponding access permission probability value, which can vary on a frame-by-frame basis. Voice (data) terminals with permission transmit their requests using one minislot selected randomly out of $n_{Rv}$ ($n_{Rd}$). If the request is unsuccessful because of collisions with other requests in the same class terminals, then the request re-attempts the next frame, (with a new permission probability). Once such a request is successful, the process of allotting information slots (I-slots) in the I-subframe over the next frames is initiated. In the I-subframe, there are $N$-$m$ I-slots, each holding one information packet.

A request for an I-slots reservation, transmitted on a minislot, contains the source terminal identifier (SID), type of traffic class (TOC), that is, voice or data, and the number of I-slots requested (NSR). Each request packet is assumed to amount in total to 16 bytes, including physical overhead (OH). Each I-slot consists of a physical overhead plus a header and information field conveying a packet. The header contains an SID, TOC, and packet sequence number (SEQ). Although these header fields are generally used in reservation-based TDMA, the lengths may be different between the systems. However, in this paper, we assume the same fields with the same lengths in order to compare the schemes in a fair way.

In the integrated voice/data services literature, the goal of a design for service integration is to guarantee a differentiated QoS for individual service while sharing network resources efficiently and maximizing system resource utilization. We note that, in the reservation-based MAC protocol literature, the most pivotal factor affecting the delay-related QoS is the contention phase, where request messages are sent with a random access protocol, for example, a slotted ALOHA protocol.

In particular, when we consider the multimedia information sources that produce a traffic mix of real-time messages with delay constraint and random bursty messages without any delay constraint, what is the prioritized but 'fair (without undue sacrifice of low priority traffic, and in the same priority traffic) access' scheme is an open issue. In fact, the problem of providing a QoS guarantee in a distributed environment such as an access network has become an important research problem. The uplink queues are located at the terminals, and thus they are regarded as remote queues in terms of the AP. The

AP has no idea about the queue status unless the terminals exchange such information.

The specific contributions of this work are as follows: 1) to consider the most general and flexible frame structure in which the lengths of random access regions assigned to respective service traffic and the corresponding permission probabilities are adaptively determined frame-by-frame, 2) to investigate the performance of the central controller, which attempts to optimize the most general form of the frame structure in 1), and 3) to introduce a hybrid type of access control mechanism enforced by so-called terminal-assisted random access control, which readjusts the access permission probability in the voice terminal so that an unfair access control due to the remote distributed queuing problem in the BAN can be overcome.

## IV. TAF-PRMA Protocol

The proposed TAF-PRMA protocol is composed of two major components: centralized controller in the AP, which determines system state variables, and a distributed controller in the voice terminal, which readjusts the access permission probability, informed by the AP via downlink signaling according to its local states.

### 1. Centralized Controller in AP

The multi-QoS frame structure under consideration consists of a series of time slots, which are partitioned into information regions and two contention minislot regions, each for voice and data traffic. Each time slot consists of $M$ minislots of fixed length, and furthermore, the frame length is given by $L_f$ minislots. Let $L^{(i)}$ $(1 \le i \le 4)$ denote the number of minislots assigned for an individual region as illustrated in Fig. 2, where the index $i$ represents the order of determining its length. Denoting the access permission probabilities associated with the access minislots for voice and data terminals by $p^{(v)}$ and $p^{(d)}$, respectively, a system state of the AP in each frame can be defined by ($L^{(1)}$, $L^{(2)}$, $L^{(3)}$, $L^{(4)}$, $p^{(v)}$, $p^{(d)}$). The objective of this letter is to present a specific approach to dynamically determine the system state on a frame-by-frame basis so as to minimize the average packet delay for data traffic while maintaining the maximum allowable packet drop rate ($P_{drop}^{\max}$)

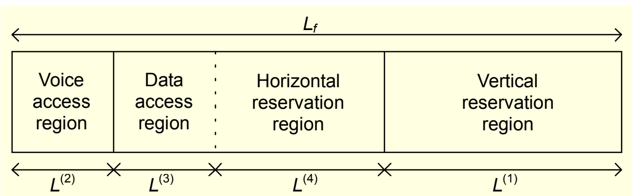



Fig. 2. Access and reservation regions in the TDMA frame.

and the fairness for voice traffic.

At the end of the k-th frame, the AP updates two sets of reserved terminals, $\mathbf{V}_k$ and $\mathbf{H}_k$. Terminal set $\mathbf{V}_k$ is a set of vertically reserved terminals that are periodically granted only one slot per frame over consecutive frames. Along with the reserved voice terminals, $\mathbf{V}_k$ also includes the reserved data terminals, which successfully request transmitting more than or equal to $\eta$ data packets. As in [3], [4], only one voice packet is generated from an active terminal in each frame. Meanwhile, $\mathbf{H}_k$ is a set of horizontally reserved terminals, which includes the data terminals other than the ones in $\mathbf{V}_k$. Those in $\mathbf{H}_k$ share the remaining slots after being granted to those in $\mathbf{V}_k$. Let $\tilde{s}_k^i$ and $s_k^i$ be the number of minislots requested by a newly reserved data terminal $i$ and the number of minislots not yet granted to the reserved data terminal $i$ during the k-th frame, respectively. By additionally defining the set $\tilde{\mathbf{H}}_k = \{i \mid 0 < \tilde{s}_k^i < \eta \cdot M\}$, $\tilde{\mathbf{V}}_k$ as the set of newly reserved voice terminals, and $\breve{\mathbf{V}}_k$ as the set of voice terminals transited into an inactive state, we can describe the updates of both sets $\mathbf{V}_k$ and $\mathbf{H}_k$ as follows:

$$\mathbf{H}_k = \mathbf{H}_{k-1} - \{i \mid s_{k+1}^i = 0,\ i \in \mathbf{H}_{k-1}\} + \tilde{\mathbf{H}}_k, \tag{1}$$

$$\begin{aligned}\mathbf{V}_k &= \mathbf{V}_{k-1} - \breve{\mathbf{V}}_k + \tilde{\mathbf{V}}_k \\ &\quad - \{i \mid s_{k+1}^i = 0, i \in \mathbf{V}_{k-1}\} + \{j \mid \tilde{s}_k^j \ge \eta \cdot M\}.\end{aligned} \tag{2}$$

Given the arrival rates of voice and data traffic, denoted by $\hat{\lambda}^{(v)}$ and $\hat{\lambda}^{(d)}$, respectively, the number of voice and data terminals to transmit the reservation request packet in the ($k$+1)th frame, denoted by $\hat{N}_{k+1}^{(v)}$ and $\hat{N}_{k+1}^{(d)}$, respectively, can be estimated recursively using the contention traffic load in the previous frame. In this paper, we consider a frame-based version of the pseudo-Bayesian algorithm in [16]:

$$\hat{N}_{k+1}^{(S)} = \hat{\lambda}^{(S)} + n_k^{(S)} \cdot \max\left(0,\ \frac{\hat{N}_k^{(S)}}{L_k^{(\cdot)}} - 1\right) + c_k^{(S)} \cdot \left(\frac{\hat{N}_k^{(S)}}{L_k^{(\cdot)}} + \frac{1}{e-2}\right), \tag{3}$$

where $c_k^{(S)}$ and $n_k^{(S)}$ denote the numbers of minislots that have and have not been collided, respectively, over the (voice or data) access regions of the $k$-th frame, and $L_k^{(\cdot)}$ is given by $L_k^{(2)}$ or $L_k^{(3)}$ if $S = v$ or $d$, respectively.

A specific procedure of determining a system state of the AP for the $k$-th frame is described so as to optimize the overall performance.

#### *A. Granting I-slots to the Set of Vertically Reserved Terminals*

First of all, the highest priority is given to the vertically reserved terminals, and thus the number of reserved minislots is set by

$$L_k^{(1)} = |\mathbf{V}_{k-1}| \cdot M , \tag{4}$$

where $|\cdot|$ denotes the number of members in the set.

*B. Determining the Length and Permission Probability of Random Access Region for Voice Terminals*

For the immediate reservation request of voice traffic, the number of random access minislots for such traffic is determined by the estimated number of voice terminals, that is,

$$L_k^{(2)} = \min(\hat{N}_k^{(v)},\ L_f - L_k^{(1)}) , \tag{5}$$

and thus the permission probability is set to

$$p_k^{(v)} = \min(1,\ L_k^{(2)} / \hat{N}_k^{(v)}) . \tag{6}$$

*C. Determining the Length and Permission Probability of Random Access Region for Data Terminals*

The remaining $L_f - L_k^{(1)} - L_k^{(2)}$ minislots are separated for the random access and horizontally reserved regions for data traffic. Since $L_k^{(4)}$ is underutilized if $L_k^{(3)}$ is set too small or vice versa, $L_k^{(3)}$ and $L_k^{(4)}$ must be properly balanced out to maximize the slot efficiency. Using a flow balance equation in a fully loaded system [17], that is,

$$L_k^{(3)} \cdot \frac{1}{e} \cdot \overline{m}_k = L_k^{(4)} , \tag{7}$$

the number of random access minislots for data traffic is computed as follows:

$$L_k^{(3)} = \min\left( \frac{e}{e + \overline{m}_k} (L_f - L_k^{(1)} - L_k^{(2)}),\ \hat{N}_k^{(d)} \right), \tag{8}$$

where

$$\overline{m}_k = \frac{1}{W} \sum_{l=1}^{W} \frac{1}{|\widetilde{\mathbf{H}}_{k-l}|} \sum_{\forall i \in \widetilde{\mathbf{H}}_{k-l}} \widetilde{s}_{k-l}^{\,i} , \tag{9}$$

with $W$ being a moving average window size whose unit is a TDMA frame. According to (8), the permission probability for data terminals is set to

$$p_k^{(d)} = \min(1,\ L_k^{(3)} / \hat{N}_k^{(d)}) . \tag{10}$$

*D. Granting I-slots to the Set of Horizontally Reserved Terminals*

Finally, the number of horizontally reserved minislots is set to

$$L_k^{(4)} = \min\left( \sum_{\forall i, i \in \mathbf{H}_{k-1}} s_k^i,\ L_f - \sum_{i=1}^{3} L_k^{(i)} \right) . \tag{11}$$

If $\Delta = L_f - \sum_{i=1}^{4} L_k^{(i)} > 0$, then the remaining minislots are additionally used for the data access region. Thus, taking into consideration the two values in the procedure in IV.1.C, $L_k^{(3)}$ and $p_k^{(d)}$ are set as follows:

$$L_k^{(3)} \leftarrow L_k^{(3)} + \Delta , \quad \text{and} \quad p_k^{(d)} = \min(1,\ L_k^{(3)} / \hat{N}_k^{(d)}) , \tag{12}$$

which allow for accepting more data traffic.

2. Distributed Controller in Voice Terminal

In this paper, we introduce a hybrid access control mechanism through which the voice terminal readjusts the access permission probability ($p^{(v)}$), informed by the AP via downlink signaling, according to its local states. The probability may not be accurate since it was determined by various estimates in the AP. We note that each voice terminal has additional and more accurate information on dropped packets and backlogged packets, which can therefore be used for reconfiguring access permission probability in a timely and accurate manner. As the terminal is directly involved with the random access control, it is referred to as terminal-assisted random access hereafter. In order to fairly maintain the pre-specified level of QoS degradation for voice traffic, our method presents a higher $p^{(v)}$ with regard to a terminal that has failed many times (due to the collision or suppression) and to a terminal that suffers higher packet dropping probability. The access permission readjustment procedure can be described as follows:

$$\widetilde{p}_j^{(v)} = \begin{cases} \alpha_j \cdot (p^{(v)} + \beta_j) & \text{if } P_{drop} < P_{drop}^{\max} , \\ \alpha_j \cdot p^{(v)} + 0.5 & \text{if } P_{drop} \geq P_{drop}^{\max} , \end{cases} \tag{13}$$

where $\alpha_j$ and $\beta_j$ are the factors that govern the short-term and long-term fairness, respectively. In a case in which the reservation requests consecutively fail, $\widetilde{p}_j^{(v)}$ must be increased appropriately, for example, by setting $\alpha_j = c \cdot (1/2)^{K-i}$ for the $i$-th trial, $i = 1, 2, \cdots, K$ . Here the maximum of $\alpha_j$ should be 0.5 so that $\widetilde{p}_j^{(v)}$ may be bounded to 1, and thus $c$ is set to 0.5. On the other hand, the packet drop rate requirement can be maintained by setting $\beta_j = P_{drop} / P_{drop}^{\max}$. In the case where $P_{drop} / P_{drop}^{\max} \geq 1$, we set $\alpha_j \cdot \beta_j = 0.5$ such that a state of urgency is accordingly reflected onto $\widetilde{p}_j^{(v)}$ .

## V. Simulation Results and Discussions

In this section, we present the simulation results for the performance of the proposed TAF-PRMA in terms of system capacity, voice packet dropping probability, data packet delay, and fairness. Five other systems [3], [5]-[8] are selected for our quantitative comparison, with the objective of highlighting the merits and demerits of different systems providing the integrated voice/data services. Three [6]-[8] of these schemes under comparison have the same frame structure, but are different in how the access regions are partitioned. For example, VR-MAC in [7] and Li in [8] consider the separate access regions for voice and data traffic while FPBP in [6] does not. In these systems, the respective values of the parameters, that is, the numbers of minislots for R-subframe ($n_R$), voice access region ($n_{Rv}$), and data access region ($n_{Rd}$), are determined to satisfy the follow equation:

$$n_R = n_{Rv} + n_{Rd} = i \cdot m, \quad m = 1,2,\cdots$$

where $i$ is the number of minislots per slot.

More specifically, in VR-MAC, $n_{Rv}$ is fixed to 2; in FPBP, $n_R$ is fixed to 8; and in Li, $n_{Rv}$ is fixed to 2 and $n_{Rd}$ is fixed to 6. Li's system is particularly modified into two variants: 1) Li-FPB determines permission probabilities with a framed pseudo-Bayesian algorithm, and 2) Li-Perf determines permission probabilities with perfect terminal information. In addition, two other systems [3], [5] without frame structure are considered. Other parameters used for the integrated voice and data systems are described in Table 1.

Concurrently served with 20 data users, each with an average arrival rate of 10,248 bps, the packet dropping probability ($P_{drop}$) of the voice user and packet delay performance of the data user are compared against other frame-based schemes in Figs. 3(a) and 3(b), respectively. It is shown

Table 1. Simulation parameters.

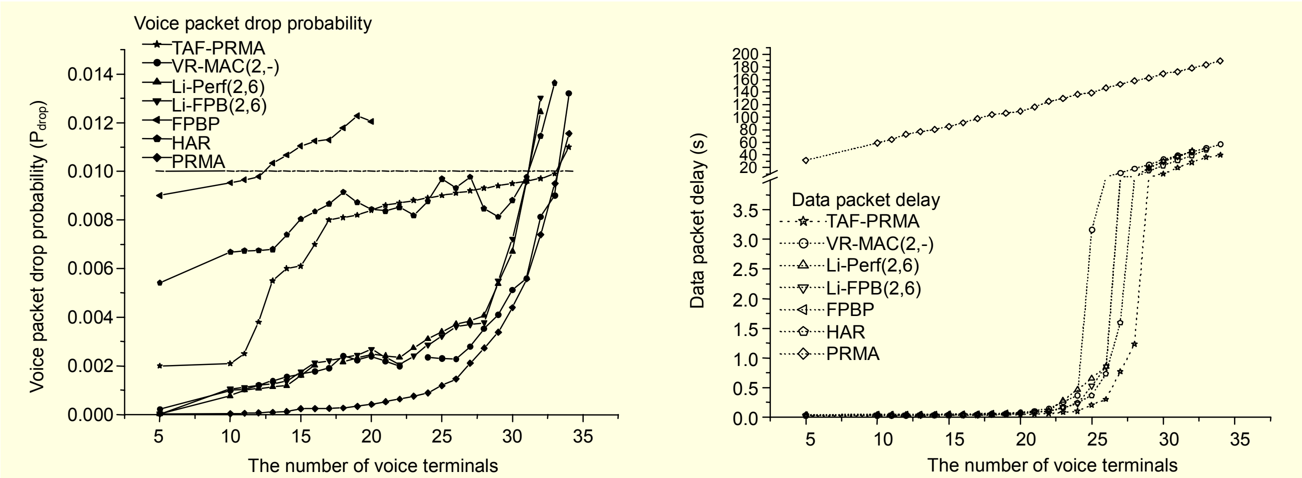
| | |
|---|---|
| Frame duration | 16 ms |
| Number of slots per frame, $N$ | 20 |
| Slot duration | 800 µs |
| Number of minislots per slot, $i$ | 4 |
| Channel rate | 720 kbps |
| Average talkspurt duration | 1.00 s |
| Average silence duration | 1.35 s |
| Voice source rate | 32 kbps |
| Maximum tolerable transfer delay | 32 ms |
| Average data message size | 4 packets |
| Data rate of data user | Variable |
| Air interface overhead | 64 bits |
| Number of active voice (data) users | Variable |
| Simulation time | 500 s |

that the TAF-PRMA achieves the highest voice user capacity among all the schemes considered in the simulation, supporting a maximum of 33 voice users while guaranteeing the packet dropping probability bound ($P_{drop}^{\max} = 0.01$). Even if the same level of capacity is achieved by the VR-MAC and PRMA schemes, their data packet delay performances are significantly sacrificed, and resultantly these two schemes are not flexible enough to optimally trade off the voice and data performance. The corresponding performance gain of TAF-PRMA is further manifested by the data packet delay presented in Fig. 3(b). In TAF-PRMA, the data packet delay increases rapidly from above 25 voice users, which is the largest number among competitive schemes, while the voice QoS is still satisfied. In other words, with that point as the beginning, the system is saturated and data packet delay performance is sacrificed due to the prioritized resource allocation to voice users.

This phenomenon of becoming flat is due to the change of the right axis scale from 0.5 to 20. Without the axis scale change, the data packet delay will be shown as going to infinity.

To illustrate a specific example regarding the QoS guarantee capabilities of some access schemes, $P_{drop}$ is measured as varying the traffic load offered by data users.

In this simulation, the numbers of voice and data users are fixed to 20, while the offered data traffic load is determined by the rate of each data user, which varies from 2,048 bps to 20,480 bps. As shown in Fig. 4, those schemes with dedicated access regions, including the TAF-PRMA, guarantee the voice QoS irrespective of the data traffic load. We note that some schemes, for example, PRMA [3] and HAR [5], also meet the voice QoS, while setting the access permission probability of data users as low as necessary; this however, sacrifices the delay performance due to a limited data access capability. Since a QoS performance tends to depend on the number of data users in HAR, we also consider a special case where the maximum number of requests is limited to 5, allowing for simulating the effect of increasing the data users. This particular case is referred to as HAR* in Fig. 4. While $P_{drop}$ in HAR increases rapidly up to 14000 bps and thereafter decreases, $P_{drop}$ in HAR* goes beyond the voice QoS limit from 11000 bps and up. This phenomenon in HAR is explained by the fact that the access trials by data terminals increase as their data rates increase up to 14000 bps, and thereafter the trials decrease due to concatenation requests, that is, requests for more multiple packets. The decreased access trials reduce the collisions between data and voice access trials so that $P_{drop}$ declines, which is applicable to FPBP, as shown in Fig. 4. The dependence of voice QoS in both systems (HAR and FPBP) upon data terminals raises difficulties in guaranteeing the voice packet dropping probability bound, which is demonstrated by the simulation of HAR*.




(a) Voice packet dropping probability versus number of voice terminals

(b) Data packet delay versus number of voice terminals

Fig. 3. Voice packet dropping probability and data packet delay versus number of voice terminals. The number of data terminals is fixed at 20.

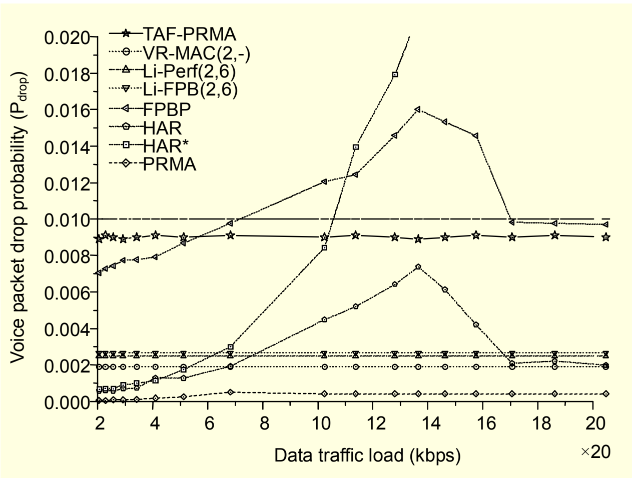



Fig. 4. Voice packet dropping probability versus data traffic load. The numbers of voice and data users are each fixed at 20.

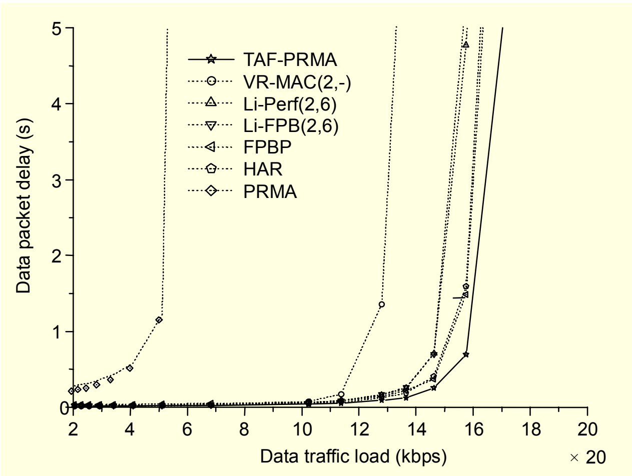



Fig. 5. Data packet delay versus data traffic load. The numbers of voice and data users are each fixed at 20.

In Fig. 5, we plot the average data packet delay versus the data traffic load with 20 voice users and 20 data users. As expected, the delay of VR-MAC starts to increase rapidly from a comparatively low traffic load, simply because even in an overbooking state, where the value of $q$ is very large, it allocates minislots in the same manner as in an otherwise state, aside from an incorrect multiplication factor as we mentioned previously. Improvements in the maximum data rate admissible in the TAF-PRMA scheme ranges from 7% to 200% over the other schemes, assuming that the delay QoS is imposed by the maximum data packet delay of 1 second. Such an improvement is simply attributed to the design objective of TAF-PRMA, which adjusts the R-subframe to minimize the voice access region under QoS constraint, and balances the data access region so that the data packet delay can be minimized.

Figure 6 depicts the system throughput, defined as the ratio of total number of the received packets to total number of offered slots, versus the data rate again with 20 voice users and 20 data users. The maximum achievable throughput of the TAF-PRMA is close to 0.95, which turns out to be a 16% to 150% improvement over the other schemes. From Figs 3 through Fig. 6, we note that the TAF-PRMA outperforms other existing schemes in bandwidth utilization/efficiency because the superiorities in the voice capacity, delay, and throughput were caused by the efficiency of bandwidth utilization.

The fairness index, defined as a ratio of the average voice packet dropping probability for the worst 5 users to that for the best 5 users, is presented in Table 2. This particular index will indicate how the resource is fairly shared among the voice users from a QoS perspective. In terms of our fairness index, the TAF-PRMA achieves the best fairness performance among all the access schemes. In particular, the corresponding

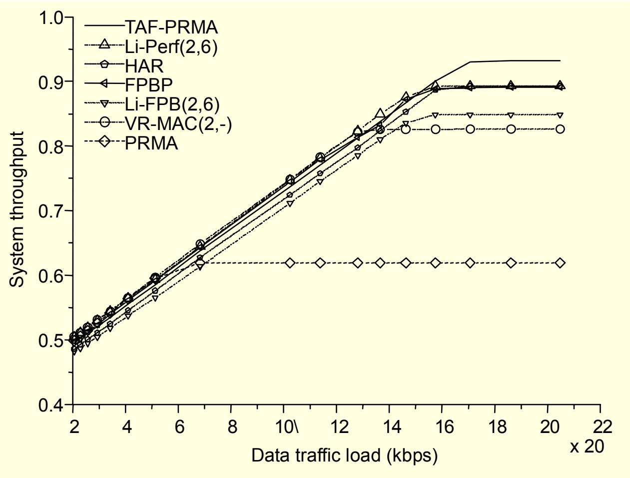



Fig. 6. System throughput versus data traffic load. The numbers of voice and data users are each fixed at 20.

Table 2. Fairness index versus the number of voice terminals.

| No. of voice terminals | Fairness index | | | |
|---|---|---|---|---|
| | TAF-PRMA | VR-MAC(2, -) | Li-Perf(2,6) | Li-FPB(2,6) |
| 10 | 0.801831 | 0.288618 | 0.369746 | 0.373401 |
| 20 | 0.661650 | 0.374390 | 0.283680 | 0.316166 |
| 30 | 0.547478 | 0.195291 | 0.249466 | 0.293913 |

improvement becomes more significant as the number of voice users increases. The fair QoS guarantee is mainly attributed to the terminal-assisted adaptation mechanism in the TAF-PRMA. In contrast to the TAF-PRMA, all the other systems do not introduce any fairness mechanisms, resulting in a situation where a comparative predominance among their systems is non existent.

## VI. Conclusions and Future Works

In this paper, a hybrid medium access controller is designed to provide the bandwidth-efficient and QoS guaranteed integrated voice/data services over dynamic reservation TDMA-based broadband access networks. It partitions an R-subframe, which consists of the random access minislots for reservation requests, into two separate minislot regions for voice and data users, respectively, which provides a useful means of guaranteeing QoS for the real-time service. A boundary between these two access regions and the length of R-subframe are the critical parameters that govern the overall resource utilization efficiency, along with the access permission probabilities associated with each access region. Our extensive comparative studies indicate that it outperforms other existing schemes in almost every performance aspect.

However, many technical issues still need to be addressed. How to adapt a scheduling algorithm into the proposed optimal frame structure is the most relevant open issue. In addition, we will expand our developed scheme to the orthgonal frequency division multiplexing (OFDM) or orthgonal frequency division multiple access (OFDMA) systems such as IEEE 802.16 based fixed broadband wireless access system and WiBro, and devise how to integrate the adaptive modulation and coding scheme with the optimal frame structure.

## References


[1] CM-SP-RFIv2.0-I08-050408, *Data-Over-Cable Service Interface Specification - Radio frequency Interface Specification*, Apr. 2005.

[2] IEEE Std. 802.16-2004, *Air Interface for Fixed Broadband Wireless Access Systems*, IEEE, Oct. 2004.

[3] D. J. Goodman et al. “Packet Reservation Multiple Access for Local Wireless Communications,” *IEEE Trans. Commun.*, vol. 37, no. 8, Aug. 1989, pp. 885-890.

[4] W. C. Wong and D. J. Goodman, “A Packet Reservation Multiple Access Protocol for Integrated Speech and Data Transmission,” *IEE Proc.-1*, vol. 139, Dec. 1992, pp. 607-612.

[5] W.-S. Jeon, D.-G. Jeong, and C.-H. Choi, “An Integrated Services MAC Protocol for Local Wireless Communication,” *IEEE Trans. Commun.*, vol. 47, no. 1, Feb. 1998, pp. 352-364.

[6] J.-F Frigon and V. C. M Leung, “A Pseudo-Bayesian Aloha Algorithm with Mixed Priorities,” *Wireless Networks*, vol.7, no.1, 2001, pp. 55-63.

[7] K. Sriram and P. D. Magill, “Enhanced throughput Efficiency by Use of Dynamically Variable Request Minislots in MAC Protocols for HFC and Wireless Access Networks,” *Telecommunication Systems 9*, vol. 9, no. 3, 1998, pp. 315-333.

[8] Z. Li, C.-M. Zhao, and X.-H. You, “Protocol for Integration of Voice and Data over Frame-Based PRMA,” *IEE Electron. Lett.*, vol. 34, no. 10, 1998, pp. 959-960.

[9] F.-C. Ren, C.-J. Chang, and R.-G. Cheng, “A QoS-guaranteed Fuzzy Transmission Controller for Dynamic TDMA Protocol in Multimedia Communication Systems,” *IEE Proceedings-Communications*, vol. 149, no. 6, Dec. 2002, pp. 292-298.

[10] S. Elnoubi and A.M. Alsayh, “A Packet Reservation Multiple Access (PRMA)-Based Algorithm for Multimedia Wireless System,” *IEEE Trans. Vehicular Tech.*, vol. 53, no. 1, Jan. 2004, pp. 215-222.

[11] S. Khan et al., “Application-Driven Cross-Layer Optimization for Video Streaming over Wireless Networks,” *IEEE Comm. Magazine*, vol. 44, no. 1, Jan. 2006, pp. 122-130.

[12] C. Y. Sin, P.C. Cosman, and L.B. Milstein, “Multimedia Communication over OFDM Mobile Wireless Networks: A Cross-Layer Diversity Approach,” *Proc. ICC,* 2005, pp. 1276-1281.

[13] A. Ksentini and A. Gueroui, “Toward an Improvement of H.264 Video Transmission over IEEE 802.11e through a Cross-Layer Architecture,” *IEEE Comm. Magazine*, Jan. 2006, pp. 107-114.
[14] S. Ryu et al., “Wireless Packet Scheduling Algorithm for OFDMA System Based on Time-Utility and Channel State,” *ETRI Journal*, vol. 27, no. 6, Dec. 2005, pp. 777-787.
[15] K. C. Chua, “Minislotted Packet Reservation Multiple Access,” *IEE Electron. Lett.*, Oct. 1993, pp. 1920-1921.
[16] R. L. Rivest, “Network Control by Bayesian Broadcast,” *IEEE Trans. Info. Theory*, vol. IT-33, no. 3, May 1987, pp. 323-328.
[17] P. D. Roorda, V.C.M. Leung, “Dynamic Control of Time Slot Assignment in Multi-Access Reservation Protocols,” *IEE Proc.-Commun.*, no. 3, Jun. 1996, pp. 167-175.

**Seung-Eun Hong** was born in Gangneung, Korea, in 1971. He received the BS, MS, and PhD degrees in electrical engineering from Korea University in February 1997, February 1999, and August 2003. While working on his PhD dissertation from September 2002, he also was with Electronics and Telecommunication Research Institute (ETRI), where he worked on designing and implementing a wireless MAC protocol. From September 2003 to March 2005, he was a Senior Engineer with Samsung Electronics, where he worked on developing an IEEE 802.16e based WiBro system. Since March 2005, he is once again with ETRI, as a Senior Member of Research Staff. His research interests include broadcasting systems (digital cable TV, DMB, WiBro MBS) and wireless multimedia communication systems development, including MAC protocol design and implementation, fair scheduling, mobile ad-hoc networking, and quadruple play service.

**Chung-Gu Kang** received the BS degree in electrical engineering from the University of California, San Diego in 1987 and the MS and PhD degrees both in electrical and computer engineering from University of California, Irvine, in 1989 and 1993. While working on his PhD dissertation from June 1991 to May 1992, he also was with Aerospace Corporation in El Segundo, California, as a part-time member of technical staff (MTS). In 1993, he joined Rockwell International, Inc. in Anaheim, California, where he worked on the signaling system no. 7 and other telecommunication systems development. Since March 1994, he has been with the Department of Radio Communication Engineering at Korea University, Seoul, Korea, where he is currently a Professor. In the academic year of 2000 to 2001, he was a Visiting Associate Professor at the University of California, San Diego, where he was also affiliated with the Center for Wireless Communications. His research interests include next generation mobile radio communication system and broadband wireless networks, with special emphasis on physical layer/medium access control layer design and performance analysis. His recent research is focused on MIMO/multiple access schemes for mobile broadband wireless access systems and MAC/Routing protocols for mobile ad hoc networks. He has over 100 refereed publications in international journals and conferences in the areas of communications networks, CDMA cellular systems, OFDM systems, and wireless local area/personal area networks. He served as an associate publication editor of JCN (Journal of Communication and Network) and TPC member of various international conferences. He has also been involved with various standardization activities, including ITU-R JRG 8A-9B. He is currently a vice chair of 2.3GHz Portable Internet Project Group (PG05) and a chairman of service/network working group of PG05 in the Telecommunications Technology Association of Korea. He is a member of IEEE COMSOC, IT, and VT, and a member of KICS and KITE.

**O-Hyung Kwon** received the BS, MS, and PhD degrees in electrical engineering from Sogang University, Seoul, Korea, in 1981, 1983, and 2004. Since 1983, he has been with ETRI. From 1999, he has been a Team Leader working on developing digital cable TV systems. His research interests are in digital watermarking, conditional access system (CAS), and downloadable CAS.