Packet Loss Fair Scheduling Scheme for Real-Time Traffic in OFDMA Systems

Seokjoo Shin and Byung-Han Ryu

In this paper, we propose a packet scheduling discipline called packet loss fair scheduling, in which the packet loss of each user from different real-time traffic is fairly distributed according to the quality of service requirements. We consider an orthogonal frequency division multiple access (OFDMA) system. The basic frame structure of the system is for the downlink in a cellular packet network, where the time axis is divided into a finite number of slots within a frame, and the frequency axis is segmented into subchannels that consist of multiple subcarriers. In addition, to compensate for fast and slow channel variation, we employ a link adaptation technique such as adaptive modulation and coding. From the simulation results, our proposed packet scheduling scheme can support QoS differentiations while guaranteeing short-term fairness as well as long-term fairness for various real-time traffic.

Keywords: Packet loss fair scheduling (PLFS), packet scheduling, orthogonal frequency division multiple access (OFDMA), adaptive modulation and coding (AMC).

Manuscript received Dec. 24, 2003; revised June 10, 2004.
This study was supported by research funds from Chosun University, 2003.
Seokjoo Shin (phone: +82 62 230 6030, email: sjshin@chosun.ac.kr) was with ETRI and is currently with the Department of Internet Software Engineering, Chosun University, Gwangju, Korea.
Byung-Han Ryu (email: rubh@etri.re.kr) is with Mobile Telecommunication Research Division, ETRI, Daejeon, Korea.

I. Introduction

The next generation wireless communication systems will provide a wide range of multimedia services while guaranteeing the required quality of service (QoS) of mobile users. Among all of the technical issues that need to be solved in a multimedia network, packet scheduling could be one of the most important issues, since packetized transmission over wireless links makes it possible to achieve a high statistical multiplexing gain.

Scheduling algorithms provide mechanisms for bandwidth allocation and multiplexing at the packet level. Many scheduling algorithms, capable of providing a certain guaranteed QoS, have been studied for wireless networks. In the earliest-due-date first (EDD), each packet from a periodic traffic stream, such as real-time services, is assigned a deadline, and the packets are sent in order of increasing deadlines [1]. The principle of EDD is based on the priority-order-queue-mechanism. Real-time traffic, such as voice and video streaming, is very delay sensitive, but can stand a certain level of packet loss. The service-oriented fair packet loss sharing (FPLS) algorithm is introduced in [2], where a time division code division multiple access (TD-CDMA) system is considered. The basic idea of FPLS is that the packet loss of each user is controlled according to the QoS requirements and the traffic characteristics of all the mobile users sharing the same frequency spectrum in the cell. From the results, FPLS provides a higher spectral efficiency than the GPS algorithm proposed in [3]. In [2], however, the authors did not consider link adaptation techniques such as adaptive modulation and coding (AMC) in a cellular environment. In order to evaluate the performance of a packet scheduling algorithm deployed in a wireless environment more reliably, the wireless channel environment should be considered.

In this paper, we propose a QoS-guaranteed packet scheduling discipline called packet loss fair scheduling (PLFS), in which the packet loss of each user from different real-time traffic is fairly distributed according to the tolerable packet loss requirements of all the user equipment (UE) sharing the same frequency spectrum in a cell. While the basic concept of PLFS is very similar to those shown in [1], [2], we propose a new priority order mechanism with consideration of a wireless cellular environment. Moreover, the proposed scheduling rule is similar with the modified largest weighted delay first (M-LWDF) presented in [4] with the exception of instantaneous QoS distributions for different users. While M-LWDF gives an optimal throughput and may support long-term fairness for real-time users, our proposed scheme supports short-term fairness as well as long-term fairness for diverse real-time traffic.

The orthogonal frequency division multiplexing frequency division multiple access (OFDM-FDMA) system is considered in this paper, since the most suitable modulation choice for beyond third-generation mobile communication systems seems to be OFDM due to its high data rate transmission capability with sufficient robustness to radio channel impairments. In addition, the link adaptation techniques have been widely studied to overcome low wireless channel efficiency. AMC is one of the compromising techniques providing flexibility to choose an appropriate modulation and coding scheme (MCS) for the channel conditions based on either UE measurement reports or network determination. We adapt the AMC technique to the proposed packet scheduling algorithm according to the received signal-to-interference ratio of each UE.

The rest of this paper is organized as follows. In sections II and III we describe the system model and introduce our proposed PLFS algorithm, respectively. Our simulation environment and simulation results are shown in sections IV and V. We conclude the paper in section VI.

II. System Model

We consider an OFDMA cellular system with packetized transmission. One central base station (BS) and multiple distributed users are set to be a cell. The basic frame structure of the considered OFDMA system is shown in Fig. 1 in the context of a downlink in a cellular packet network, where the time axis is divided into a finite number of slots within a frame and the frequency axis is segmented into subchannels that consist of multiple subcarriers. In Fig. 1, the basic resource space of the packet transmission is defined as a basic unit (BU), which corresponds to a slot in time and a subchannel within a frequency. Therefore, there are $N_{slot} \cdot N_{sub}$ BUs in a frame, where N_{slot} is the number of slots in a frame and N_{sub} is the number of subchannels within a frequency. In this system, a maximum of N_{sub} users can transmit their packets simultaneously in each slot without intra cell interference.

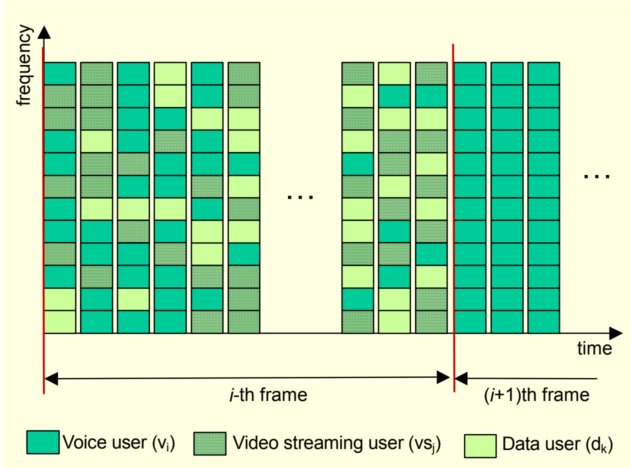

Fig. 1. The proposed frame structure of OFDMA.

BU_{ij}, where $0 \le i \le N_{slot} - 1$ and $0 \le j \le N_{sub} - 1$, delivers a certain amount of information bits defined as $C(BU_{ij})$. $C(BU_{ij})$ is highly dependent on the channel condition of its assigned UE. The instantaneous system capacity in bits per second, $R_c(t)$, is represented as follows:

$$R_c(t) = F_s \sum_{i=0}^{N_{slot}-1} \sum_{j=0}^{N_{sub}-1} C(BU_{ij}), \qquad (1)$$

where F_s is the number of frames in a second. The system capacity, $R_c(t)$, changes in time randomly.

In addition, for each connection request, c_i, where $1 \le i \le N_{user}$, the QoS parameter $Q(c_i)$ is carried in the request. N_{user} represents the number of active users in a cell. In a trivial case, $Q(c_i)$ specifies the requested average transmission rate (*bit/s*) for the connection. For each new connection request, c_{new}, the network determines the admission of a new call if, and only if, the following condition is satisfied:

$$Q(c_{new}) \le \overline{R}_c - \sum_{i=1}^{N_{user}} Q(c_i), \qquad (2)$$

where $\overline{R}_c$ is the estimated average transmission rate of the previous seconds. Then, (2) becomes a call admission control strategy.

The packet scheduler schedules up to N_{sub} users among all active users in every slot instant. Each UE has its own buffer to queue the packets for transmission. The packet size of all buffers is assumed to be identified. At the *i*-th time slot, the packet scheduler selects the N_{sub} highest priority buffers after calculating the priority of each buffer based on the uplink feedback information and buffer management information. The scheduling discipline is described in more detail in the next section.

In addition, under the wireless environment, we adapt the AMC technique to the proposed packet scheduling algorithm by assigning K MCS levels depending on UE measurement reports. In the frequency selective fading environments, UE measurement reports of different subchannels have different values. User k receives SNR_k values, $SNR_k^{\,0}, \cdots, SNR_k^{N_{sub}-1}$, from predetermined pilot signals corresponding to each subchannel. The SNR is measured as the ratio of pilot signal power to background noise when we assume that there is no other cell interference at all. More specifically, the SNR of the n-th subchannel allocated for the k-th user can be represented as

$$SNR_k^n = \frac{P_p h_{k,n}^2}{N_0 \frac{B}{N_{sub}}}, \tag{3}$$

where $h_{k,n}$ is a random variable representing the fading of the k-th user and n-th subchannel. P_p is the transmitted power of the pilot signal. N_0 is the noise power spectral density and B is the total bandwidth of the system. The channel gain, $h_{k,n}^2$, of subchannel n of user k is given by

$$h_{k,n}^2 = \left|\alpha_{k,n}\right|^2 \cdot PL_k. \tag{4}$$

Here, PL_k is the path loss for user k and is defined by

$$PL_k = PL(d_0) + 10\beta \log\left(\frac{d_k}{d_0}\right) + X_\sigma, \tag{5}$$

and $a_{k,n}$ is the short scale fading for user k and subchannel n. The reference distance and distance from BS to user k are d_0 and d_k, respectively. The path loss component is β, while X_σ represents a Gaussian random variable for shadowing with standard deviation σ.

To reduce the signaling overhead, the UE selects $N(< N_{sub})$ SNR_k values for the feedback of channel quality information. After receiving the UE's channel quality information index, the BS allocates the appropriate BUs to the UE if the user is selected for the scheduling in the current slot.

The MCS level is classified by the required SNR strength, SNR_{req}, and maps to the number of packets in a BU. The mapping between the MCS level and the number of packets is shown in Table 1, where we assume that all subchannels are allocated with equal power, i.e., 1 W. From the channel condition of user k, $\overline{A_k}$ is defined as the moving average of the number of transmittable packets in a BU from the previous trials with window size WS.

Table 1. Transmission mode with convolutionally-coded modulation.

Index	SNR_{req} (dB)	Packet/BU	Modulation	Coding rate
AMC_1	1.5	1	BPSK	1/2
AMC_2	4.0	2	QPSK	1/2
AMC_3	7.0	3	QPSK	3/4
AMC_4	10.5	4	16QAM	1/2
AMC_5	13.5	6	16QAM	3/4
AMC_6	18.5	9	64QAM	3/4

III. Packet Loss Fair Scheduling Algorithm

Among the diverse objectives for fairness to incorporate multimedia services, we focus on the fair QoS-guarantee scheduling rules for real-time traffic. The proposed PLFS algorithm is concerned with satisfying the different required QoS evenly for real-time traffic. A fairness guarantee for real-time traffic is assumed to be achieved when the current packet loss rate (PLR) is distributed proportionally equal for all users at any instant in time. In other words, our proposed scheme ensures short-term fairness as well as long-term fairness, simultaneously. In our context, short term corresponds to the slot time duration, since our proposed scheduling scheme decides corresponding users according to the fairness of all users at each time slot. As noted in [6], short-term fairness implies long-term fairness but not vice versa.

A PLR occurring in real-time traffic is defined as the sum of the packet error rate (PER) resulting from the channel impairments and packet dropping rate (PDR) calculated from packets exceeding the required maximum delay, D_{max}, in each traffic. The PLR should be less than a certain determined threshold, $PLR_{req,i}$, for user i; that is,

$$PLR_i(t) = PER_i(t) + PDR_i(t) \le PLR_{req,i}. \tag{6}$$

The scheduler easily calculates the number of dropped packets of each user, or $PDR_i(t)$, from the transmission buffer management information. On the other hand, $PER_i(t)$, can be calculated by an individual user's ARQ mechanism when we assume that a transmitted packet is fed back its acknowledgement information through the dedicated uplink acknowledgement channel, regardless of the traffic type.

A priority order queue mechanism is applied for supporting the PLR of each active user fairly. The PLFS rule schedules the highest priority user among all the users. After that, the scheduler selects the next highest priority users continuously until all the subchannels in a slot are occupied. The scheduler updates the priority of each active user in every time slot before scheduling. For a predetermined set of parameters related to the required QoS of each user, the rule is given by,

$$j = \max\left[\left(\frac{A_k(t)}{\bar{A}_k}\right)\cdot\left(\frac{PLR_i(t)}{PLR_{req,i}\cdot D_{\max,i}}\right)\right], \qquad (7)$$

where $A_k(t)$ is the state of the channel in terms of the MCS level of user k at time t. The key feature of this algorithm is that a scheduling decision depends on both current channel conditions and the current packet loss of different users. The term, $\frac{A_k(t)}{\bar{A}_k}$, becomes the proportionally fair queuing presented in [7], while $\frac{PLR_i(t)}{PLR_{req,i}\cdot D_{\max,i}}$ renders the scheduling rule to be the packet loss fair queuing and can provide QoS differentiations between different users.

Similar approaches are shown in [4], where the decision rule is based on the weighted delays of packets in addition to the wireless channel conditions. This rule balances different users' probabilities of a deadline violation, approximately. In the context of diverse real-time traffic with different QoS requirements, however, the rule is not sufficient to guarantee the short-term fairness in terms of all users' packet loss at any instant in time.

IV. Simulation Environments

For the system level simulation, we consider two types of real-time traffic such as voice and video streaming. A voice source creates a pattern of talkspurts and gaps that are assumed to have an exponentially distributed duration [8]. These are assumed to be statistically independent of each other. The mean duration of the talkspurts and gaps are 1 s and 1.35 s, respectively, which leads to a voice activity factor of 0.43 [9].

On the other hand, a streaming video is modeled as a variable bit rate characterized by a Pareto distribution. The modeling is composed of continuous video-frames, where each video-frame is divided into a fixed number of video-packets [10]. Moreover, the size of a video-packet is determined by Pareto distribution. The parameters of a video streaming model where the generation rate is 32 kbps is listed in Table 2, and an example of a video traffic pattern is shown in Fig. 2. Note that a variable length video packet is segmented into fixed-length packets before being stored in the scheduler buffer.

In this paper, we consider a radio cell with N_{user} active UEs and a centralized BS. The BS has a packet scheduler in its MAC layer to obtain a high statistical multiplexing gain, where N_{user} buffers are directly connected to the packet scheduler. We assume that the fixed-length packets (or MAC PDUs) are stored into the scheduler buffers from the upper layer.

In a wireless cellular network, the channel state varies randomly in time on both slow and fast scales. As slow channel

Table 2. Parameters of 32 kbps video streaming model.

Information types	Distribution	Distribution parameters
Inter-arrival time between the beginning of each video-frame	Deterministic	100 ms
Number of video-packet in a video-frame	Deterministic	8
Video-packet size	Truncated Pareto (Max.=125 bytes)	k=20, α=1.1
Inter-arrival time between video-packets in a video-frame	Truncated Pareto (Mean=6 ms, Max.=12.5 ms)	k=2.5, α=1.2

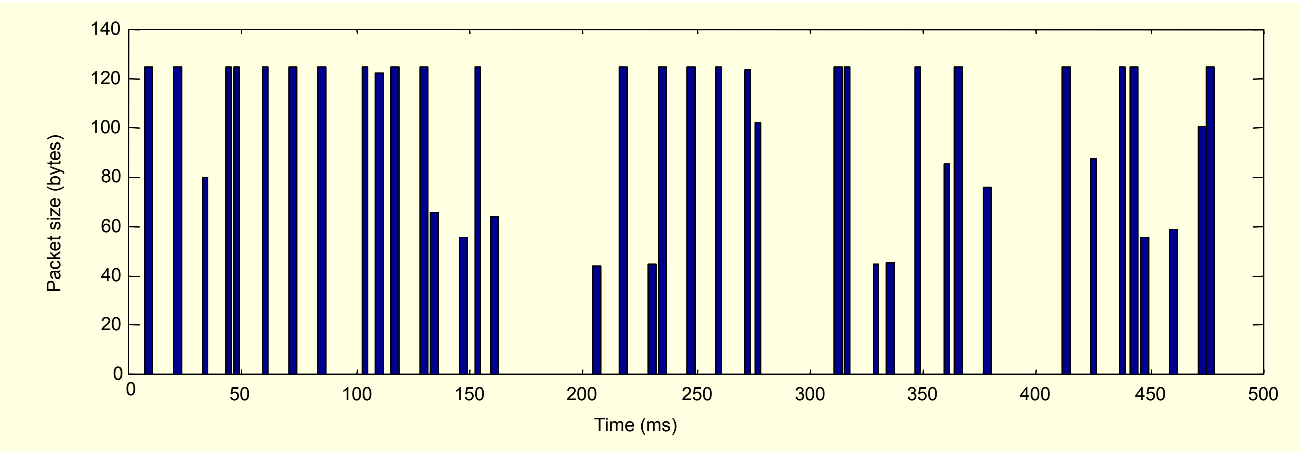

Fig. 2. Example of the video streaming model.

variation depends mostly on user location and the interference level, the normalized power from the BS is adapted to the diverse MCS level in the AMC technique according to the received SNR_k for a user k.

The set of the simulation parameters is listed in Table 3.

Table 3. Simulation parameters.

Parameters		Value
Number of subcarriers in OFDM		1536
Number of subchannels		12
Frame length (ms)		20
Slots per frame		20
Maximum packet loss rate (PLR_{req})	Voice	10^{-2}
	Video streaming	10^{-3}
Packet size (byte)		44
Source data tare (kbps)	Voice	16
	Video streaming	32
Maximum packet delay (ms)		Variable
Cell radius (km)		1
User distribution		Uniform
BS transmission power (W)		12
Path loss model	α	4
	σ (dB)	8

V. Simulation Results

The packet loss rate (PLR) and channel utilization (μ) are considered as the performance measures for the proposed PLFS algorithm. The current system load over the average transmissible system load, $\overline{L}$, is defined as μ. $\overline{L}$ is calculated by

$$\overline{L} = \frac{1}{6}\sum_{i=1}^{6} N_{sub} \cdot N_{slot} \cdot F_s \cdot AMC_i \,, \tag{8}$$

where AMC_i is as defined in Table 1.

Through a computer simulation, the PER is ignored based on the assumption that the channel condition is well estimated and predicted. Figures 3 through 5 show the performances of the proposed scheme when the system supports an equal number of users in both real-time traffic, $N_{voice}=N_{video}$, simultaneously.

The simulation results show that the proposed PLFS algorithm gives fair resource sharing in terms of the PLR requirements of real-time traffic, as shown in Fig. 3. A fair distribution of the PLR experienced by each user at a time is maintained for all of the different real-time traffic. For each given PLR requirement, when D_{max}=100 ms for example, the number of concurrent voice and video users are both 115. Note that the higher D_{max} is, the higher the number of users who can be supported.

As an example, Fig. 4 shows the short-term behavior of $PLR_i(t)$for all users when the numbers of simultaneous users in both real-time traffic are 100 and D_{max}=100 ms. It is noteworthy that the instant PLR values of all users are distributed fairly at a certain point.

Figure 5 shows the channel utilization according to various D_{max}. In the figure, μ does not exceed 1 even though the traffic load is very high. The reason for this is because the traffic generation rate of the traffic sources is lower than the average

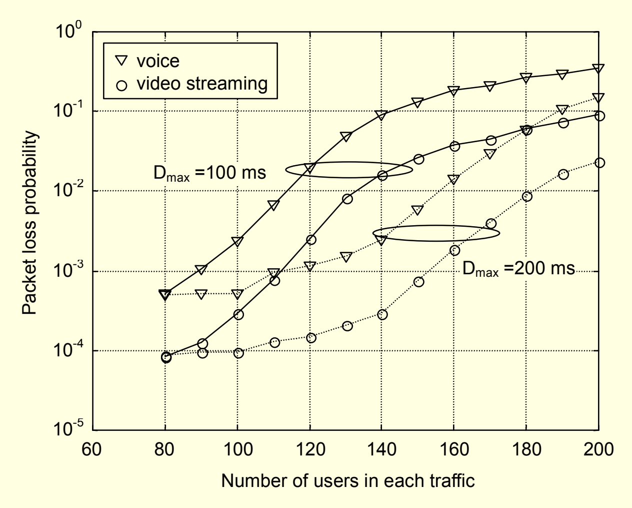

Fig. 3. Packet loss rate vs. the number of concurrent voice and video users when D_{max} = 100 and 200 ms, respectively.

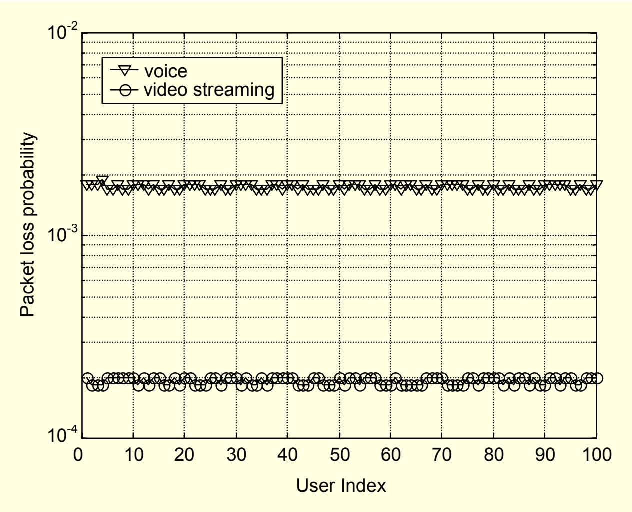

Fig. 4. PLR of each user at a certain simulation time instant when the system is occupied with 100 voice and 100 video streaming users under the condition D_{max} = 100 ms.

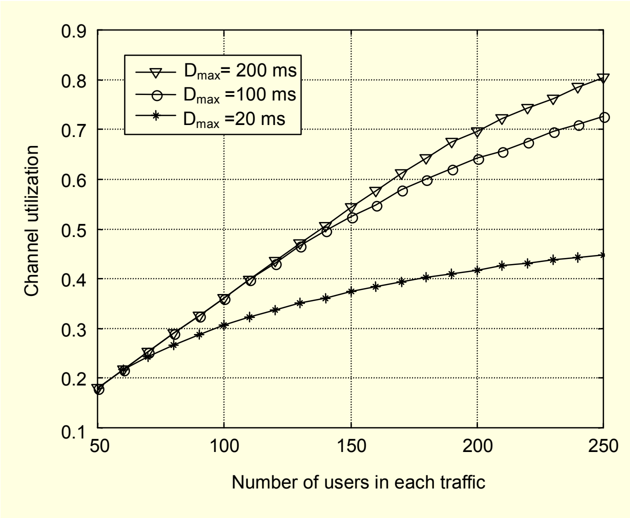

Fig. 5. Channel utilization vs. the number of concurrent voice and video users when $D_{max}=20$, 100, and 200 ms, respectively.

transmission rate determined by the AMC scheme. Note that μ is getting higher when D_{max} increases.

VI. Conclusion

In this paper, a PLFS packet scheduling scheme has been proposed for real-time traffic based on the OFDMA system, in which the packet loss of each user is fairly distributed according to the QoS requirements. From the computer simulation, we can see that the results verify the PLFS to give fair resource sharing between real-time users. Furthermore, the results show that D_{max} could be an important parameter to determine the total channel utilization.

References

[1] C. Lin and J. Layland, "Scheduling Algorithms for Multiprogramming in a Hard Real-Time Environment," *J. ACM*, vol. 20, Jan. 1973, pp. 46-61.

[2] V. Huang and W. Zhuang, "Fair Packet Loss Sharing (FPLS) Bandwidth Allocation in Wireless Multimedia CDMA Communications," *Proc. Int'l Conf. 3G Wireless and Beyond*, May 2001, pp. 198-203.

[3] A. Elwalid and D. Mitra, "Design of Generalized Processor Sharing Schedulers Which Statistically Multiplex Heterogeneous QoS Classes," *Proc. IEEE INFOCOM*, vol. 3, Mar. 1999, pp. 1220-1230.

[4] M. Andrews, K. kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing Quality of Service over a Shared Wireless Link," *IEEE Commun. Mag.*, Feb. 2001, pp. 150-154.

[5] D. Kim, B. Ryu, and C. Kang, "Packet Scheduling Algorithm Considering a Minimum Bit Rate for Non-Realtime Traffic in an OFDMA/FDD-Based Mobile Internet Access System," *ETRI J.*, vol. 26, no. 1, Feb. 2004, pp. 48-52.

[6] C. Koksal, H. Kassab, and H. Balakrishnan, "An Analysis of Short Term Fairness in Wireless Media Access Protocol," *Proc. ACM SIGMETRICS*, June 2000.

[7] A. Jalali, R. Padovani, and R. Pankaj, "Data Throughput of CDMA-HDR a High Efficiency-High Data Rate Personal Communication Wireless System," *VTC2000*, vol. 3, Apr. 2000, pp. 1854-1858.

[8] S. Shin, J.A. Lee, and K. Kim, "A Modified Joint CDMA/PRMA Protocol with an Access Channel for Voice/Data Services," *IEICE Trans. Funda.*, vol. E82-A, no. 6, June 1999, pp. 1029-1031.

[9] D.J. Goodman, R.A. Valenzuela, K.T. Gayliard, and B. Ramamurthi, "Packet Reservation Multiple Access for Local Wireless Communications," *IEEE Trans. Commun.*, vol. 37, no. 8, Aug. 1989, pp. 885-890.

[10] WG5 Evaluation AHG, *1xEV-DV Evaluation Methodology-Addendum (V6),* July 2001.

Seokjoo Shin received the BE degree in avionics engineering from Hankook Aviation University, Korea in 1997, and the MS and PhD degrees in information and communications from Gwangju Institute of Science and Technology (GIST), Korea, in 1999 and 2002. He joined the Mobile Telecommunication Research Laboratory, Electronics and Telecommunications Research Institute (ETRI), Korea, in 2002. In 2003, he joined the Faculty of Engineering, Chosun University, where he is currently a Full-time Instructor in the Department of Internet Software Engineering. His research interests include the resource management and packet transmission protocols for wireless multimedia applications and networking. Dr. Shin is a member of IEEE and IEICE.

Byung-Han Ryu received the BS and MS degrees in industrial engineering from Hanyang University and Seoul National University, Seoul, Korea, in 1985 and 1988, respectively, and the PhD degree in information and computer sciences from Osaka University, Osaka, Japan, in 1997. He joined the Switching Technology Division, Electronics and Telecommunications Research Institute (ETRI), Korea, in 1988. He is currently with the Broadband Mobile MAC Research Team, Mobile Telecommunication Research Division, as a team leader. From 2000, he has been a principal member of engineering staff in ETRI. His research interests include communication network design, mobile communication systems, and performance modeling and evaluation. Dr. Ryu is a member of KIIE, KICS, IEEK, and IEICE.