Wireless QoS Control Technologies

Takeshi Yamashita†, Hidetoshi Kayama, Lan Chen, Daisuke Kitazawa, Masayuki Motegi, and Narumi Umeda

Abstract

In the fourth-generation (4G) mobile communication system, all communications including realtime communications is expected to be carried by packets multiplexed on a shared radio channel to improve the radio resource efficiency. Because the delay characteristics and the throughput of a packet radio channel vary greatly with changes in both traffic volume and radio channel quality, NTT DoCoMo proposes novel wireless QoS (quality of service) control technologies to guarantee QoS requirements and achieve highly efficient realtime communication with a wide range of transmission rates.

1. Introduction

The fourth-generation (4G) mobile communication system is expected to provide a wide variety of applications and services through high-data-rate radio channels. All communications including realtime communications such as VoIP, videophones, video streaming, and digital TV broadcasting, will be packet multiplexed on a shared radio channel to improve the radio resource efficiency. These communication services have individual requirements for bandwidth and acceptable transmission delay, i.e., quality of service (QoS). However, packet radio systems suffer from drastic changes in the transmission delay and channel throughput when packet traffic and radio channel conditions fluctuate. To handle these dynamic conditions, a key technology for implementing the system is wireless QoS control for efficient wireless packet multiplexing that takes into account variations in radio propagation and interference while satisfying the QoS requirements of various applications [1], [2]. This article describes four novel wireless QoS control technologies suitable for high-speed packet communication systems: radio-condition-aware admission control, multistage hybrid scheduling, reservation-based reverse channel access protocol, and adaptive battery conservation management. We are implementing these technologies in a prototype 4G system.

† NTT DoCoMo, Inc.
Yokosuka-shi, 239-8536 Japan
E-mail: t_yamasita@nttdocomo.co.jp

2. Features of wireless QoS control

Wireless QoS control is designed to provide the following features.

(1) Guaranteed and best-effort services

Realtime communications has strict requirements regarding the bandwidth and acceptable transmission delay. To cope with these applications, wireless QoS control provides guaranteed services that can provide a flexible guaranteed bandwidth and delay suitable for realtime communications. Best-effort services will also be provided for non-realtime communications and these services will take into account relative packet priorities with/without radio channel conditions.

(2) Smooth working with IP-QoS

By assigning guaranteed-class packets to the expedited forwarding (EF) class for the differentiated services code point (DSCP) and best-effort-class packets to the best effort (BE) class for the DSCP, we ensure that wireless QoS control can work smoothly with Internet protocol (IP) QoS.

(3) Area-free guaranteed services

To improve radio channel efficiency, the adaptive modulation and coding (AMC) scheme is

Table 1. Examples of services provided according to wireless QoS control.

<table>
<tr><th colspan="3">Service/applied technology</th><th>Guaranteed-class services</th><th>Best-effort-class services</th></tr>
<tr><td colspan="3">Services</td><td>• Guaranteed-rate service is provided regardless of variation in interference and user location.
• Guaranteed rate is set flexibly considering a variable rate and the user maximum and minimum requirements.</td><td>Best-effort service</td></tr>
<tr><td colspan="3">Guaranteed QoS</td><td>Transmission rate and delay</td><td></td></tr>
<tr><td colspan="3">Target application</td><td>Realtime application</td><td>Non-realtime application</td></tr>
<tr><td rowspan="3">Technologies</td><td colspan="2">Admission control</td><td>Applied (QoS request and radio conditions are considered.)</td><td>Not applied</td></tr>
<tr><td rowspan="2">Multistage hybrid scheduling</td><td>Scheduling between guaranteed and best-effort classes</td><td colspan="2">Priority control is applied in the order: packets within the guaranteed rate, best-effort packets, and packets that exceed the guaranteed rate.</td></tr>
<tr><td>Scheduling between same-class users</td><td>Guaranteed rate, admitted delay, fairness, etc. are considered.</td><td>Fairness and radio conditions are considered.</td></tr>
</table>

assumed to be used in 4G. This scheme changes the transmission rate according to the signal to interference-plus-noise ratio (SINR). However, from the user's viewpoint, the rate should be constant regardless of his/her location. Our radio-condition-aware admission control enables stable and guaranteed services with a flexible guaranteed rate to be provided to realtime users regardless of the variation in interference and the location of the mobile station. Examples of services that can be provided depending on the type of wireless QoS control used are shown in **Table 1**.

3. Architecture of packet transmission using wireless QoS control

Wireless QoS control aims to provide a range of guaranteed-rate and area-free realtime communication services regardless of the packet traffic and radio condition fluctuations. Smooth working with IP-QoS and channel efficiency are also taken into account. Based on this concept, we propose a novel wireless QoS control [3]. The architecture for packet transmission is shown in **Fig. 1**. It comprises several control factors for the medium access control (MAC) and radio resource control (RRC) layers and a radio resource control factor over these two layers. These factors can work smoothly with IP-QoS. IP packets are first mapped to the best-effort or guaranteed classes. For guaranteed-class users, radio-condition-aware admission control is carried out to avoid an overload among these users. For best-effort-class users, classification is carried out according to the DSCP implying their relative priorities. In multistage hybrid scheduling, guaranteed scheduling performs traffic shaping to guarantee both the admitted rate and delay requirements of guaranteed-class users, while best-effort scheduling considers radio resource efficiency and fairness among users and classes. In addition, multistage scheduling gives priority to guaranteed-class users. Transmission parameters are determined according to the QoS class and channel conditions based on information sent from the receiver. In the downlink, radio resources can be allocated in a straightforward scheme. In the uplink (reverse link), a reservation-based access protocol is employed to enable priority-based resource allocation required for the wireless QoS control technology (See section 4.).

3.1 Radio-condition-aware admission control (RAC)

These days, many communication applications have functions for responding flexibly to changes in transmission rates by using variable rate coding. Therefore, the guaranteed rate can be set flexibly within a variable rate range, that is the maximum and minimum requirements. This flexible admission control allows more users to be accommodated. Making allowances for user movement and variations in interference, the lowest modulation coding set is used as the channel capacity criterion for admission control decision, as indicated in **Fig. 2(a)**, which shows the transmission rate for three different modulation coding sets (MCSs). To utilize the statistical multiplexing characteristics of packet communications for radio resource efficiency, the resource utilization rate of the guaranteed-class users is also considered. However,

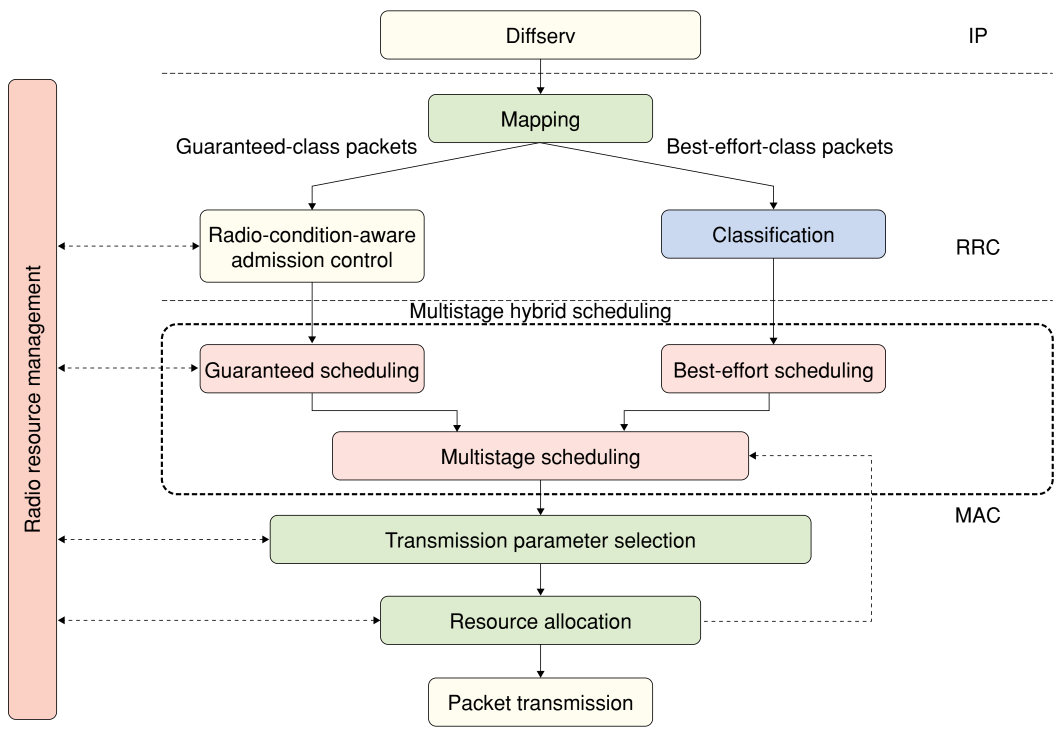


Fig. 1. Architecture of packet transmission.

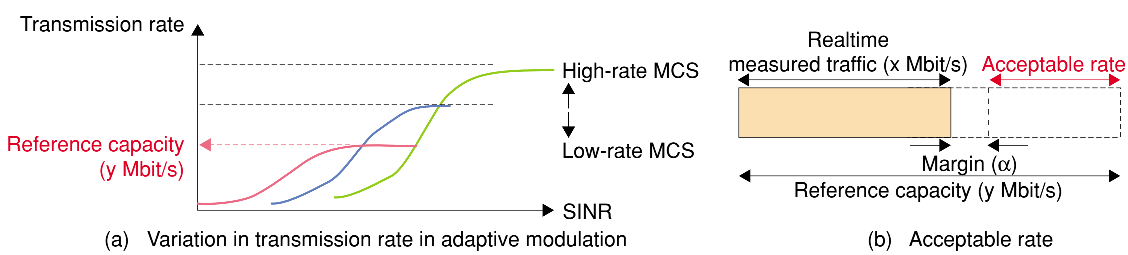


Fig. 2. Principle of radio-condition-aware admission control.

not all the guaranteed-class users encounter poor conditions. To improve the resource utilization efficiency, the measured resource utilization rate for guaranteed-class users is considered, as shown in **Fig. 2(b)**. Here, the total of the bandwidth requests that can be accepted (i.e., the acceptable rate) is given by the reference capacity (y) minus the measured resource utilization rate (i.e., the realtime measured traffic (x)) and a margin (α). That is, the acceptable rate = y - (x + α).

3.2 Multistage hybrid scheduling (MHS)

The scheduling that determines the transmission order when all user packets are multiplexed onto one common radio channel is important for maintaining service quality and supporting various applications and service classes. Furthermore, in wireless communication systems, an important purpose of scheduling is to increase the efficiency of radio resource utilization. There have been many proposals for scheduling that takes into account QoS requests [4] or that aims to increase the radio resource utilization efficiency [5]. However, in the 4G system, there are demands to provide a range of guaranteed rates and maintain the service quality for realtime communications without deteriorating the radio resource utilization efficiency.

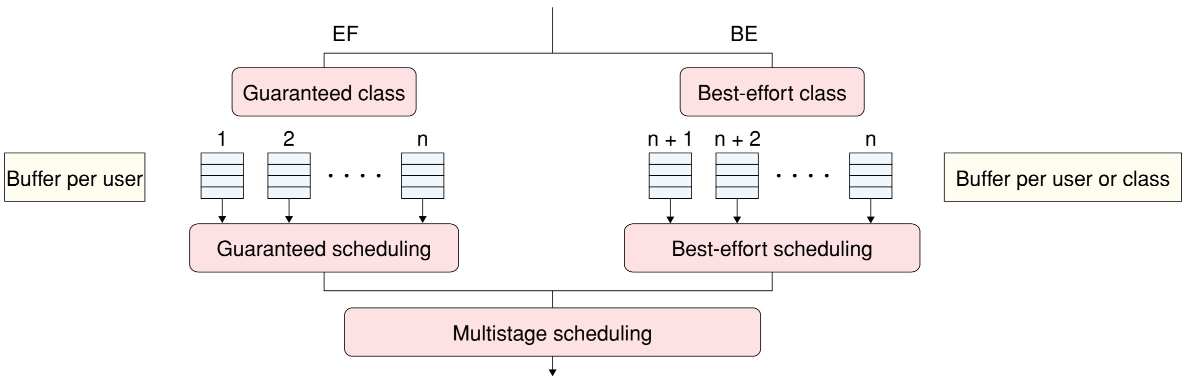


Fig. 3. Principle of multistage hybrid scheduling.

To satisfy these demands, multistage hybrid scheduling (MHS) has been proposed [6], [7]. Its principle is shown in **Fig. 3**. Here, we explain the downlink transmission case. As described above, each packet is classified as either a guaranteed- or best-effort-class packet; that is, IP packets with a header of EF and BE are mapped into the guaranteed and best-effort classes, respectively. Either each user has his/her own individual buffer or best-effort-class packets with the same DSCP can share a buffer because, from the viewpoint of the QoS requirement, flow control based on the guaranteed rate and acceptable delay are necessary for individual guaranteed-class users, while class-based flow control is sufficient for best-effort-class users. However, for scheduling that takes into account the radio resource utilization efficiency, it is necessary to consider the radio propagation conditions of each user, and individual buffers are required not only for the guaranteed-class users, but also for best-effort-class users.

In MHS, a different scheduling scheme is used between the guaranteed- and best-effort-class users. Guaranteed scheduling considers guaranteeing both the admitted rate and delay requirements of guaranteed-class users, while best-effort scheduling considers the radio resource efficiency and fairness among all best-effort-class users. Examples of scheduling applicable to both types of users are described below.

(1) Guaranteed class
- Weighted round robin (WRR): Each user in turn sends a weighted number of packets.
- Header early first (HEF) [8]: A time stamp is attached to each packet. The time stamps of the packets at the tops of the buffers are compared and the packet with the earliest time stamp is scheduled first.

In both of these methods, the number of packets transmitted within one turn depends on the guaranteed rate declared by the users.

(2) Best-effort class
- Proportional fairness (PF) [9]: The ratio of instantaneous SINR to averaged SINR is measured for each user. The packets are scheduled based on this ratio in order from highest to lowest. This method takes into account the radio resource efficiency and fairness among users.
- Class-based queuing (CBQ): When users of the same DSCP class share a buffer and DSCP class-based QoS control is used, weights are assigned to each class according to its priority and packets are taken from the buffers according to their weights.

(3) Scheduling between guaranteed and best-effort classes (multistage scheduling)

In multistage scheduling, guaranteed-class packets within the guaranteed rate are selected first. Then, packets of best-effort-class users are selected for transmission. Finally, packets that exceed the guaranteed rate are selected. Resources are assigned on a block basis, and in the case of a guaranteed-class user, the guaranteed rate is also considered during resource allocation. Here, the transmission block size of a user is determined according to his/her radio channel conditions using a channel quality indicator during transmission.

3.3 Performance of RAC and MHS

Computer simulation results for the satisfaction rate for guaranteed-class users and the transmission rate for best-effort-class users are shown in **Fig. 4**. In these figures, round robin (RR) indicates the performance of a method without RAC and MHS, while priority scheduling (PS) indicates the performance of a method that gives unconditional priority to the guaranteed-class users. We can see that wireless QoS can increase the satisfaction rate of guaranteed-class users without greatly degrading the transmission rate of best-effort-class users by setting the guaranteed rate dynamically within the range between the maximum and minimum rates requested by users and by performing scheduling that takes into consideration the guaranteed rates. RAC and MHS make it possible to aim for efficient utilization of the radio resources while guaranteeing the delay and transmission rate for guaranteed-class users. These methods can simultaneously maintain some degree of fairness among best-effort-class users by taking into account the radio resource efficiency and fairness [3], [6]-[8].

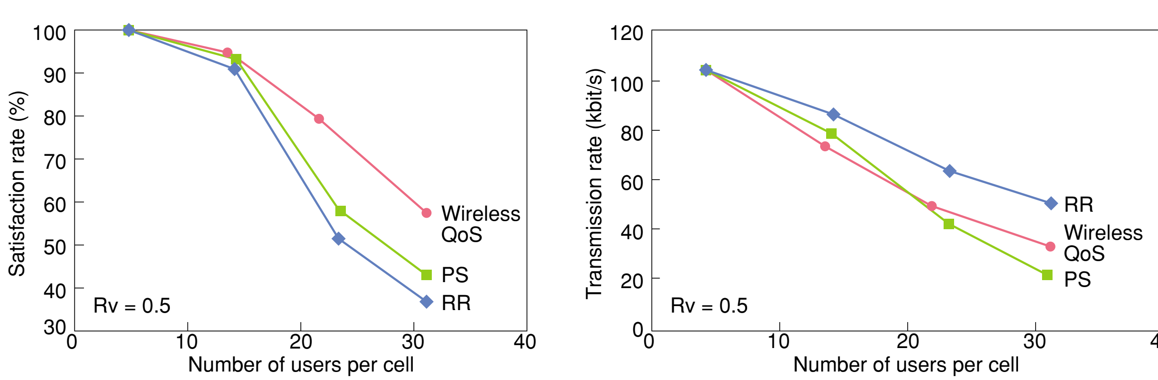
4. Reservation-based reverse channel access protocol [10]

For downlink packet transmission, resource allocation on a packet-by-packet basis is relatively easy because the base station can manage all the transmissions on the downlink packet channel. On the other hand, packets in the reverse channel are generally transmitted in a random access manner. In this manner, a massive transmission with the same timing causes packet loss due to multiple access interference. As a result, each realtime packet may suffer an unacceptable transmission delay caused by a random access retry under relatively heavy traffic conditions. According to the wireless QoS architecture (See section 3.), all packets are categorized into two classes: guaranteed and best-effort. Realtime packets are categorized into a guaranteed class, in which the transmission rate and delay are guaranteed by RAC and MHS. Incidentally, in MHS, the guaranteed-class packets that are below the guaranteed rate are always transmitted before best-effort-class packets. In the downlink channel, the transmission order of the packets can be easily arranged because all packets are assembled at a base station and arranged by centralized control. On the other hand, packets in the reverse channel are not; they are accumulated in distributed queues from individual mobile stations. To satisfy the QoS requirement for realtime communications and to increase the radio resource utilization efficiency in such a distributed queue environment, two schemes have been proposed.

(1) Prioritized resource allocation (PRA) [11]

This is a reservation-based scheme. Two different resource allocation areas in the time domain are established for the guaranteed and best-effort classes. The guaranteed-class packets have a wider allocation area so that they can obtain radio resources before the best-effort-class packets can. With this scheme, almost all the bandwidth can be occupied by guaranteed-class packets. **Figure 5** shows an example of the time slot allocation in the PRA scheme.

(2) Reservation-based random access protocol with a temporarily assigned dedicated control channel



(a) Satisfaction rate for guaranteed-class users

(b) Transmission rate for best-effort-class users

Rv = 0.5 means 50% of all users are guaranteed-class users.

Fig. 4. Performance of RAC and MHS.

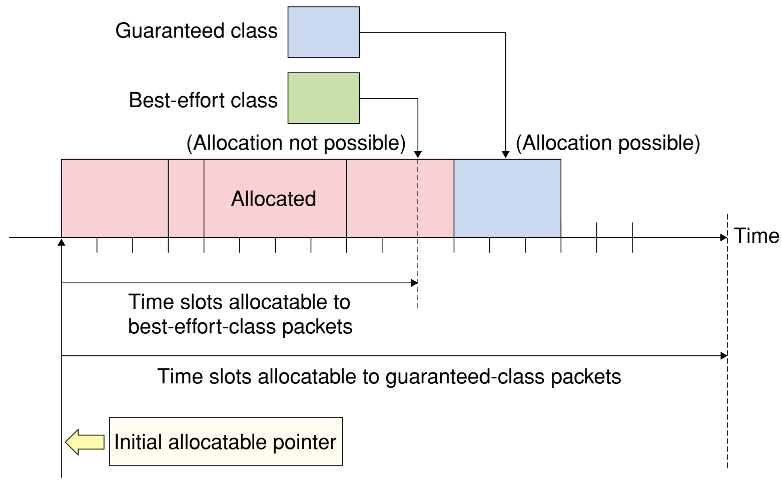


Fig. 5. Example of time slot allocation of PRA.

(RAP-TDC) [12], [13]

In this protocol, only a low-rate temporarily dedicated control channel (TDCCH) is "temporarily" assigned between a base station and a mobile station during a packet burst. In this channel, reservation signals for access and transmission control parameters are transmitted simultaneously. After a reservation has been accepted, user packets are transmitted on assigned common packet channel (CPCH) slots. This scheme prevents the packet-by-packet setup delay by maintaining transmission parameters using the TDCCH. In addition, the mobile station can transmit its reservation signal without any conflict with other reservation signals, so this scheme enables the short delay packet transmission that is required for realtime communications. The interleave reservation mechanism that enables us to remove the channel inefficiency caused by the round trip delay of the reservation protocols is also employed. In addition, because the reservation signal is transmitted without collisions, when the reservation is denied because of a temporary lack of resources, the terminal can continuously retry the request. Transmission control signals, such as the transmission power control (TPC) command, channel quality indicator, and acknowledgment signals, are continuously exchanged in the TDCCH to enable immediate packet (re)transmissions in both directions. If no packets are transmitted in the CPCH during a predetermined time, which is referred to as the release delay time below, then the TDCCH is released. It takes a little time to reestablish the TDCCH, so the release delay time should be adaptively set according to the QoS requirement of the transmitted packets. In addition, it is effective to coordinate the release delay time with a dormant control timer such as the timer described in section 5.

5. Adaptive battery conservation management

In existing mobile communication systems, a mobile station intermittently receives paging signals that are periodically transmitted by a base station during its stand-by time (which is defined as the idle mode in this paper). This scheme is generally used for battery conservation. Moreover, since the termination of a session can be clearly defined in a connection oriented manner, the mobile station can immediately transit to the idle mode after a session ends. On the other hand, in connectionless communications, session initiation and termination are generally indistinct, so it is effective to employ a timer that is initiated after each packet has been received or transmitted. When this timer expires, it initiates the transition to idle mode. When a packet addressed to a mobile station that is in idle mode reaches a base station, the base station informs the mobile station by sending packet arrival notifications (i.e., by paging it) using a periodically transmitted signal. Therefore, this scheme involves a wake-up delay. In the conventional battery conservation method, because there is only one timer for entering idle mode and one paging interval, the requirement for the acceptable transmission delay of realtime communications cannot be satisfied when applying this method to the 4G system. To overcome these problems, adaptive battery con-

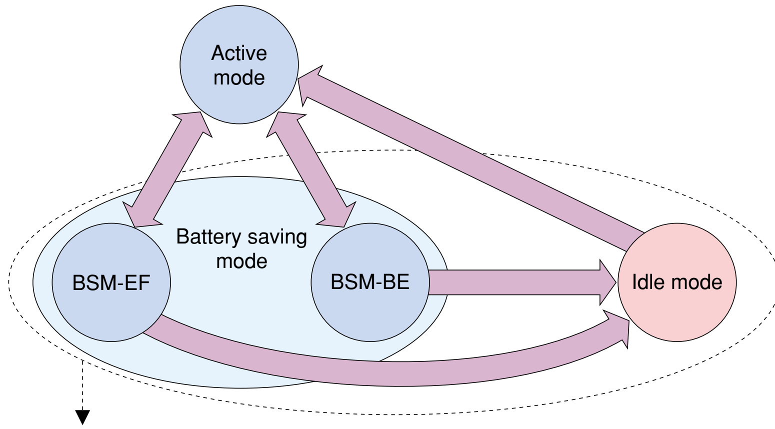


Fig. 6. State transitions in adaptive battery conservation management.

servation management (ABCM), which takes into consideration the QoS requirements, has been proposed for connectionless communication systems such as the 4G system [14].

In packet communications, the communication state of the mobile station can be classified into three states where i) packets are transmitted and received continuously, ii) no packets are transmitted or received for a long time, and iii) no packets are transmitted or received for a short time. ABCM pays attention to this feature, and a new state called battery saving mode is defined to improve battery conservation.

The ABCM state transition diagram for the mobile station is shown in **Fig. 6**. Three mobile station states are defined: the active, idle, and battery saving modes. For the battery saving mode, two submodes (BSM-EF and BSM-BE) are defined that correspond to the two QoS classes. The guaranteed-class mobile station transits to BSM-EF and the best-effort-class mobile station transits to BSM-BE. The transition to the corresponding state is triggered by the transmission and reception of packets or by the expiration of a timer. The battery saving mode is an intermediate state between the active and idle modes, and the control signal from the base station is received intermittently, as in the idle mode. The difference between the battery saving and idle modes is that the intermittent receiving cycle is shorter in the former, and BSM-EF is assigned a shorter period than BSM-BE because it has stricter delay requirements. Furthermore, the timer for the transition from active mode to battery saving mode is also set corresponding to the guaranteed and best-effort classes, so the timer for the guaranteed class is longer than that for the best-effort class. This method makes it possible to guarantee the packet transmission delay for the guaranteed class, although battery conservation is equivalent to that of the conventional methods without an intermediate state like the battery saving mode. For the best-effort class, on the other hand, effective battery conservation is possible (although an increase in delay equivalent to at most the intermittent receiving cycle is unavoidable). Simulation results indicate that ABCM achieves better battery conservation than conventional methods.

6. Conclusion

This article described some wireless QoS control technologies that have been proposed for the 4G system and are currently under evaluation. In the future, we plan to evaluate them in more detail by simulation and test their compatibility and effectiveness for various applications in a radio environment using a prototype system.

References

[1] Y. Yamao, N. Umeda, T. Otsu, and N. Nakajima, “Fourth Generation Mobile Communications System—Issues Regarding Radio System Technologies—,” IEICE Trans., Vol. J83-B, No. 10, Oct. 2000.

[2] S. Ohmori, Y. Yamao, and N. Nakajima, “The future generations of mobile communications based on broadband access technologies,” IEEE Communications Magazine, pp. 134-142, Dec. 2000.

[3] L. Chen, H. Kayama, Y. Kawabe, and N. Umeda, “Performance of a Novel Wireless QoS Proposed for Broadband CDMA Packet Communication Systems,” IEEE Proc. VTC-2004 Spring, 2004.

[4] P. Ferguson and C. Huston, “Quality of Service,” John Wiley & Sons, 1998.

[5] Y. Cao and V. O. K. Li, “Scheduling Algorithms in Broad-Band Wireless Networks,” Proceedings of IEEE, Vol. 89, No. 1, pp. 76-87, Jan. 2001.

[6] L. Chen, H. Kayama, and N. Umeda, “Wireless QoS Architecture for High Speed CDMA Packet Cellular system,” IEICE 2003 Annual

Conference, B-5-138, 2003.

[7] L. Chen, H. Kayama, and N. Umeda, "Radio Condition Aware Admission Control and Hybrid Scheduling for High Speed CDMA Packet Cellular System," IEICE 2003 Annual Conference, B-5-139, 2003.

[8] D. Kitazawa, L. Chen, H. Kayama, and N. Umeda, "A Study on Packet Scheduling Considering Buffering Delay for High Speed CDMA Cellular System," IEICE 2003 Annual Conference, B-5-140, 2003.

[9] A. Jalali, R. Padovani, and R. Pankaj, "Data Throughput of CDMA-HDR a High Efficiency, High Data Rate Personal Communication Wireless System," IEEE Proc. VTC-2000 Spring, 2000.

[10] H. Kayama, H. Qiu, and N. Umeda, "Prioritized Reservation Type MAC Protocol for Broadband CDMA Packet Communication Systems," IEEE Proc. VTC-2004 Spring, 2004.

[11] H. Kayama, H. Qiu, and N. Umeda, "Prioritized Resource Allocation (PRA) for high-speed uplink packet channel," IEICE 2003 Annual Conference, B-5-145, 2003.

[12] M. Nagatsuka, Y. Ishikawa, J. Hagiwara, T. Nakamura, and S. Onoe, "Capacity evaluation of a Media Access Control method in W-CDMA Packet Mobile Communication," IEICE 1998 Annual Conference, B-5-164, 1998.

[13] H. Qiu, H. Kayama, and N. Umeda, "A Study on Reservation Access Protocol with Dedicated Control Channel (RAP-DC) in High-speed Packet Communication CDMA System," IEICE 2003 Annual Conference, B-5-146, 2003.

[14] M. Motegi, H. Kayama, and N. Umeda, "Adaptive Battery Conservation Management for Multimedia Mobile Packet Communications," IEICE Technical Reports, NS2001-242 IN2001-198, Mar. 2002.

Takeshi Yamashita

Research Engineer, Communication Systems Laboratory, Wireless Laboratories, NTT DoCoMo, Inc.

He received the B.E. and M.E. degrees from Nagaoka University of Technology, Nagaoka, Niigata in 1996 and 1998, respectively. During 1998-2002, he was with NTT DoCoMo, Inc., Radio Network Development Department, where he was involved in the development and study of mobile satellite communication systems. Since 2002, he has been with the Communication Systems Laboratory of the Wireless Laboratories. His research interests include handoff, wireless QoS, and moving networks. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Hidetoshi Kayama

Executive senior researcher, Director of Innovative Radio Transmission Laboratories, DoCoMo Beijing Communications Laboratories.

He received the B.S. and M.S. degrees in electrical engineering and the Ph.D. degree in informatics from Kyoto University, Kyoto in 1987, 1989, and 2004, respectively. In 1989, he joined NTT Radio Communication System Laboratories, where he engaged in R&D of PHS packet data systems, 19-GHz WLANs, and 802.11a. From 1998 to 2004, he was with NTT DoCoMo, Inc. where he engaged in R&D of PHS and 4G mobile communication systems. His research interests include access protocols, QoS control, system architecture, and transmission technologies for packet radio systems. He is a member of IEICE and IEEE. He received the scholarship encouragement award from IEICE in 1995 and the best paper award of the International Conference on Telecommunications at ICT2002.

Lan Chen

Director, Advanced Radio System Lab, Executive senior researcher, DoCoMo Beijing Communications Laboratories.

She received the M.S. and Ph.D. degrees in communications and computer engineering of Kyoto University, Kyoto. She joined the Wireless Laboratories of NTT DoCoMo in 1999, where she researched wireless QoS control and radio resource management for 4G and beyond. Her research interests include wireless QoS, radio resource management, packet scheduling, and admission control. She received the best paper award at ICT2002. She is a member of IEEE and IEICE.

Daisuke Kitazawa

Corporate Strategy & Planning Department, NTT DoCoMo, Inc.

He received the B.S. and M.S. degrees in communication engineering from Osaka University, Suita, Osaka in 1998 and 2000, respectively. At university, he studied mobile communication systems, multi-layer (hierarchical) cellular systems, and software-based multimedia communication systems. He was especially interested in Software Radio and he devoted two years toward his masters degree to improve the system capacity of a software-based mobile communication system using two types of access schemes, TDMA and CDMA. He made two presentations about his research to IEICE. In 2000, he joined the Wireless Laboratories, NTT DoCoMo, Inc. and researched beyond-IMT-2000 (4G) mobile communication systems, especially packet scheduling, admission control, and radio resource management. In 2003, he moved to the corporate strategy & planning department.

Masayuki Motegi

Radio Network System Group, IP Radio Network Development Department, NTT DoCoMo, Inc.

He received the M.E. degree in electrical and computer engineering from Yokohama National University, Yokohama, Kanagawa in 2001 and the same year he joined the Wireless Laboratories, NTT DoCoMo, Inc., Yokosuka, Kanagawa. In 2004, he moved to the IP Radio Network Development Department. He is a member of IEICE.

Narumi Umeda

Director, Communication Systems Laboratory, Wireless Laboratories, NTT DoCoMo, Inc.

He received the B.E. and M.E. degrees in electronic engineering from Hokkaido University, Sapporo, Hokkaido in 1985 and 1987, respectively. He joined NTT Laboratories in 1987. In 1992, he transferred to NTT Mobile Communications Network, Inc. (now NTT DoCoMo, Inc.). He has been engaged in research on the radio link control for the personal digital cellular (PDC) system and IMT-2000. His current research interests are in the radio control for 4G mobile communications systems (beyond 3G). He was a co-recipient of the Japan Institute of Invention and Innovation (JIII) Imperial Invention Prize in 1998 and the best paper award at ICT2002. He is a member of IEICE and IEEE.