(4) Technology for Efficient Packet Access in the Data Link and Physical Layers

Sadayuki Abeta, Motohiro Tanno, Nobuhiko Miki, Atsushi Harada, and Mamoru Sawahashi

This article describes a highly efficient packet access technology for realizing the QoS (packet error rate and delay) required by traffic data in the data link layer and physical layer of broadband wireless access systems.

1. Introduction

Current cellular systems such as the Second-Generation Personal Digital Cellular (PDC) system and the Third-Generation Wideband Code Division Multiple Access (W-CDMA) system are based on circuit switching, in which each user is allocated a dedicated channel. This kind of system is suitable for systems that provide mainly voice, video and interactive services. For the Fourth-Generation (4G) wireless system, however, the future core network will be based on the Internet Protocol (IP) and all traffic, including voice, video conference, video streaming, file transfer and Web browsing, etc., will arrive in bursts of IP packets. Therefore, signal transmission using packet format is also desirable in the Radio Access Network (RAN) as well. Furthermore, traffic data is transmitted to the accessing user via a shared channel, common channel or dedicated channel using the packet format. That allows the base station configuration to involve a smaller scale of wireless circuitry than does circuit switching, which must provide transceiver modules for the respective users. We consider packet based wireless access to also be suitable from the viewpoint of flexibly providing various Quality of Service (QoS) levels (in this paper, we define QoS as the allowable delay and the residual packet error rate), from low-speed data in voice communication to high-speed data at more than 100 Mbit/s over the same air-interface. Therefore, efficient conversion of IP packets to wireless packet is important to achieving shorter transmission delay in the RAN, reduction of the required transmission power and increased link capacity. Also, Automatic Repeat reQuest (ARQ) and channel coding are essential techniques for reducing the impact of transmission errors in the RAN and achieving high reliable packet transmission [1], [2]. Furthermore, it is required the efficient packet allocation to multiple users using a shared channel to improve the system throughput (i.e., increasing the system capacity). Specifically, packet-scheduling technology for allocating the packet signals of multiple users to the time slots of a shared channel is required in the downlink [3]-[5]. For the up-link, the application of efficient packet access by techniques such as random access, in which transmission of the subsequent message data is reserved in the preamble and reservation based packet access, in which a reservation packet is transmitted in advance to reserve transmission slots for the subsequent data packets, [6] are being considered.

In this article, we explain essential technologies for efficient packet transmission such as Hybrid ARQ (H-ARQ), downlink fast packet scheduling and high-efficiency uplink packet access.

2. Signal Conversion in Data Link Layer and Physical Layer

An example of the process of converting IP packets to the radio packet frame in the physical layer is shown in **Figure 1**. We first perform the following signal conversions to obtain the full effect of ARQ control and channel coding in the data link layer and the effect of efficient packet control before mapping the IP packets to the radio packet frame in the physical layer. The data link layer comprises two sub-layers, Radio Link Control (RLC) and Medium Access Control (MAC). In the RLC sub-layer, the IP packets are segmented into processing units that are of a length suitable for retransmission control and reordering processing. Also in the RLC sub-layer, control bits (i.e., the flow control function and protocol error detection and recovery) are attached as header information to RLC-PDUs (Protocol Data Units), which are transferred to the MAC sub-layer. In the MAC sub-layer, on the other hand, multiplexing and separating the logical channel, assignment of correspondence between logical channels and transport channels and processing for priority control and scheduling are performed for efficient packet transmission.

In the broadband wireless access system, the ARQ scheme is applied in the MAC layer to minimize the transmission delay in the RAN. Therefore, The control data (sequence number) needed for retransmission and reordering control are added to the RLC-PDU as the MAC header to form a MAC-PDU. MAC-

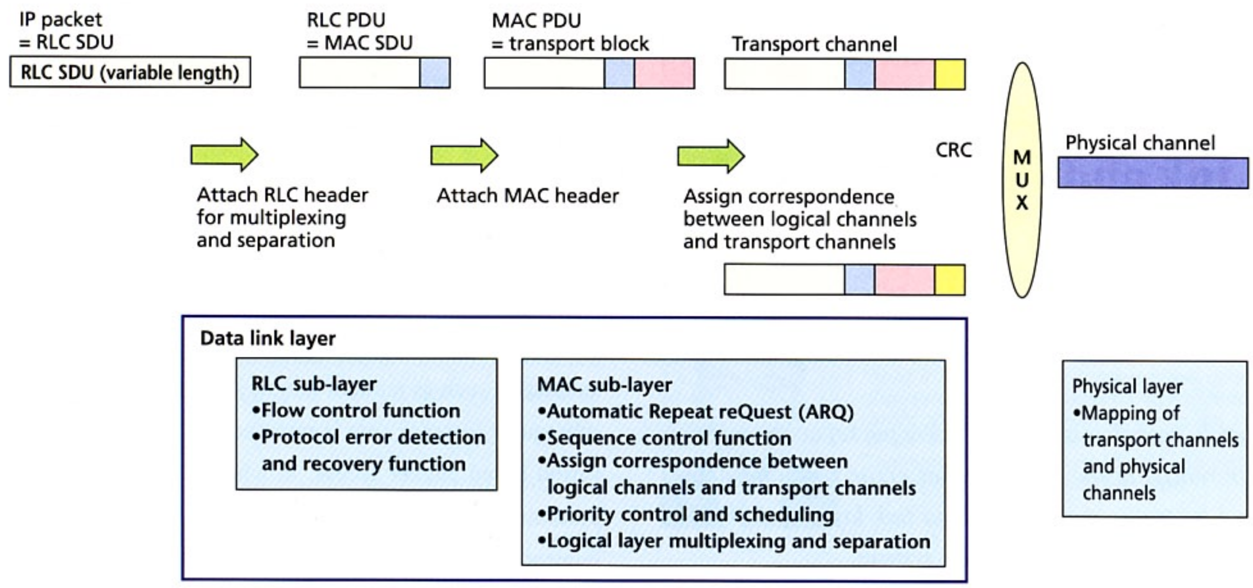

Figure 1 Conversion of IP packets to physical layer signal format

PDUs are mapped into a transport channel and transferred to the physical layer. In the physical layer, various processing such as adding an error detection code, coding with error correction code, interleaving, spreading, and code multiplexing are performed to generate the radio packet frame.

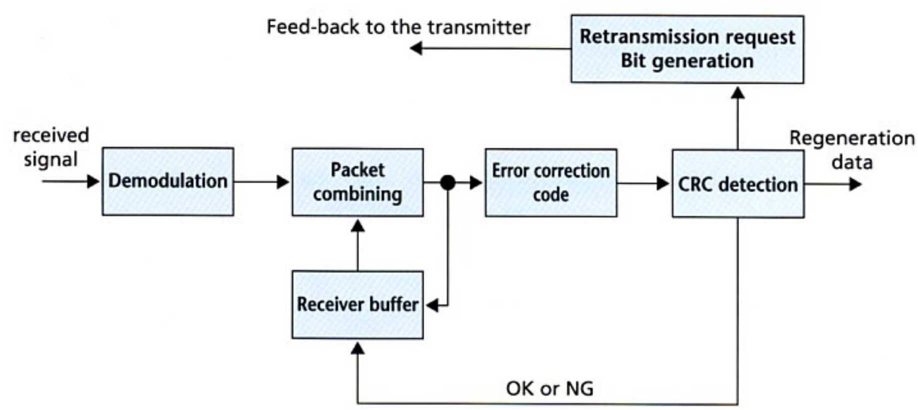

Figure 2 Receiver configuration for packet combining type H-ARQ

3. ARQ Scheme in MAC Layer

ARQ in the MAC layer is an effective technique because packet errors that occur in the RAN are efficiently compensated with shorter control delay. This technique is also employed in High Speed Downlink Packet Access (HSDPA) based on the W-CDMA air-interface. In broadband wireless access systems, ARQ in the MAC layer is also essential for attaining short delay. The configuration of the receiver for H-ARQ with packet combining in the MAC layer is shown in **Figure 2**. After decoding the received packet, packet error is detected using the Cyclic Redundancy Check (CRC) code. If a packet error is detected, the packet is not discarded but is stored in the receiver. In the packet combiner block, the newly received packet and the corresponding formerly received packet stored in the receiver buffer are combined according to the packet number reported by control channel. There are two main methods being studied for retransmitting and combining the packets in practical systems (**Figure 3**).

- Chase Combining Method (Type-I H-ARQ with Packet Combining) [1]

After receiving the retransmission packet, the packet is combined with the packet stored in the receiver buffer using Maximal Ratio Combining (MRC) and error correction code processing is performed. The Chase combining method can improve the both Signal to Interference and Noise power Ratio (SINR) of the received signal and decrease packet error.

- Incremental Redundancy Method (Type-II Method) [2]

After encoding the transmit data sequences with a coding rate *R*', punctured coding is performed according to elimination rules that differ with the number of transmission times (The signal is encoded at rate *R* (> *R*') and transmitted). Concretely, after turbo encoding at rate *R*', the data sequence is punctured with a parity bit at fixed intervals, thus generating a data sequence that is encoded at a rate *R* that is greater than *R*'.

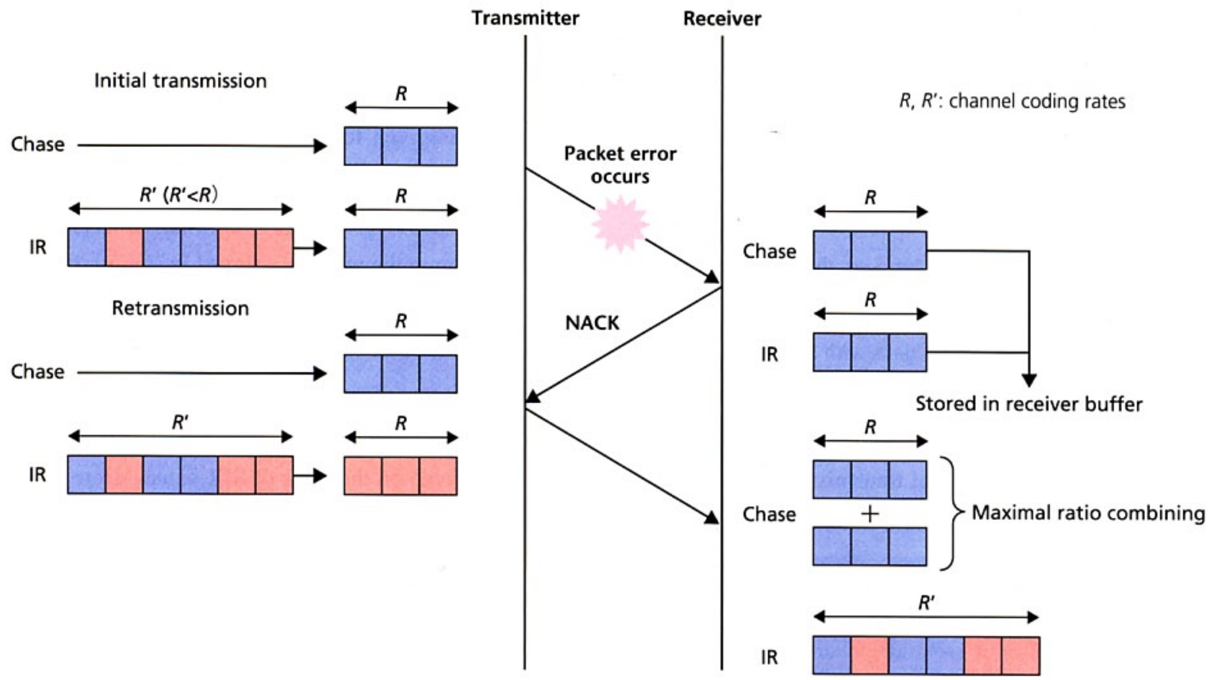

Figure 3 Principle of the packet combining type H-ARQ

Accordingly, an error correction decoding in a lower coding rate than R at the time of transmission can be achieved by combining the packet received the first time and stored in the receiver buffer with the retransmission packet at the receiver, because a coded sequence that is different from the first packet is transmitted as the retransmission packet. That is to say, the incremental redundancy method increases the coding gain by combining packets, and can thus achieve a greater packet error reduction effect than that in the Chase combining method.

The throughput performance for both H-ARQ and Adaptive Modulation and channel Coding (AMC) is plotted as a function of the average received E_s/N_0 (signal energy to noise power density ratio per data symbol) in **Figure 4** [7], with the packet combining scheme as a parameter. Here, the modulation scheme and coding rate (Modulation and channel Coding Scheme (MCS)) used in AMC are both optimized for each combining scheme. Since the effect of AMC and the packet errors caused mainly by errors in MCS selection depend greatly on the maximum Doppler frequency, f_D, the performance is shown for f_D = 20 Hz, at which the MCS selection errors are few, and for f_D = 400 Hz, at which MCS selection errors are frequent. For f_D = 20 Hz, the optimum MCS is selected for both methods, and there is no striking difference in throughput between the Chase combining method and the incremental redundancy method. For f_D = 400 Hz, on the other hand, retransmission caused by MCS selection error increases, and we can see that the incremental

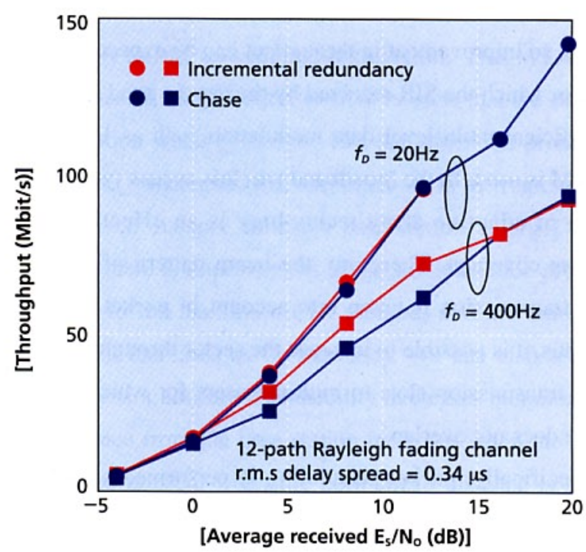

Figure 4 Comparison of the Chase combining method and the incremental redundancy method

redundancy method outperforms the Chase combining method. This is because a large coding gain is obtained with the retransmission in incremental redundancy method.

4. Fast Packet Scheduling in Downlink

The concept of fast packet scheduling in the downlink is shown in **Figure 5**. The purpose of the fast packet scheduler is to improve throughput at the base station while satisfying various QoS requirements and maintaining fairness in packet allocation among users . Accordingly, the scheduling method allocates transmission slots on the basis of the criteria described below.

First, data packet QoS (especially for allowable delay) is given highest priority. For example, the allowable delay for Real-Time (RT) traffic such as voice and video is short, but the allowable delay for Non-Real-Time (NRT) traffic such as data downloading is long. Accordingly, services that involve traffic of different QoS can be provided efficiently by giving higher priority to the user having the traffic data with shorter allowable delay.

The second priority factor is the type of packet, i.e., whether it is a retransmission packet or initial transmission packet. That makes it possible to reduce the increase in delay that accompanies retransmission by giving higher priority to the user having the retransmission packets.

The third priority factor is the radio channel condition of each user. Generally, the Signal to Interference power Ratio (SIR) is used as an indicator of radio channel conditions. In mobile communication, the propagation circumstances vary greatly, so improvement in throughput can be expected if conditions for which the SIR received by the user is good are selected and efficient multi-level data modulation such as 16QAM and 64QAM is used. In the broadband wireless access system, application of adaptive array technology is an effective way to increase coverage. Therefore, the beam pattern of directional beam transmission is taken into account in packet scheduling [8]. Thus, it is possible to increase the sector throughput by allocating transmission slots to multiple users for which the beam pattern does not overlap.

Specifically, packet scheduling is performed as described below.

i) The packets for each user that arrive at the base station are stored in a buffer for that user.

ii) Priority is given to packets on the basis of the allowable delay for the type of traffic and according to the time of storage in the buffer.

iii) In addition to the priority assigned in step ii), higher priority is given to retransmission packets. Priority is controlled so that packets that have particularly severe allowable delay requirements are given highest priority.

iv) In addition to the priorities assigned in steps ii) and iii), priority is given on the basis of SIR values. Here, For RT traffic data with a rigid delay requirement, precedence is placed on faster transmission slot assignment based on the instantaneous received SIR over fairness of slot assignment for all access users.

This is because data packets of RT traffic arrive at the BS through an IP-based core network almost periodically, so the transmission slots must be assigned to a larger number of access users in descending order of the received SIR from the highest within the required time delay. On the other hand, for NRT traffic data with a lax delay requirement, the packet scheduler assigns transmission slots taking into account a sufficient degree of fairness for the slot assignment for all access users. Two methods have been proposed for packet scheduling based on received SIR, the maximum CIR method and the proportional fairness method [3]. The Minimum Throughput Assured Instantaneous-SIR (MTA-ISIR) method, which achieves a sector throughput close to that of the maximum CIR method while maintaining the fairness among the access users, has also been

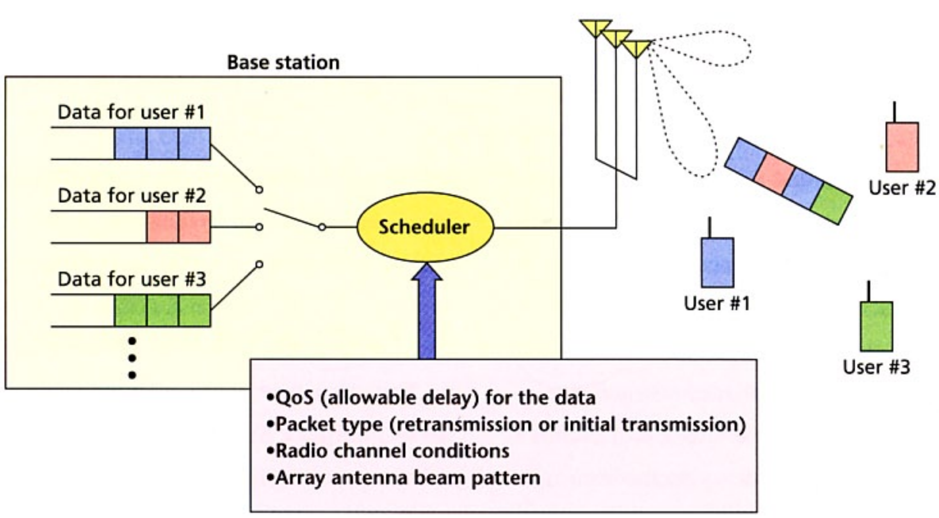

Figure 5 Conceptual diagram for the fast packet scheduler

proposed [5].

The operating principle of the MTA-ISIR method is shown in **Figure 6**. Similar to the Maximum CIR method, the MTA-ISIR method basically assigns transmission slots with priority according to the instantaneous received SIR of each user. However, in order to achieve fair transmission slot assignment, the instantaneous received SIR value, $Q^i_{inst.}$, of users whose average received SIR, Q^i_{avg}, is lower than the prescribed threshold value, Q_w, is emphasized by multiplying by a weighting factor so that the average value of the received SIR after weighting is equal to Q_w (e.g., user 2 in the example shown in Fig. 6). The scheduler assigns slots to users on the basis of the weighted instantaneous received SIR, Q^i_{inst}', with priority given users with higher values. Here, the value of Q^i_{inst}', is expressed by Eq. (1).

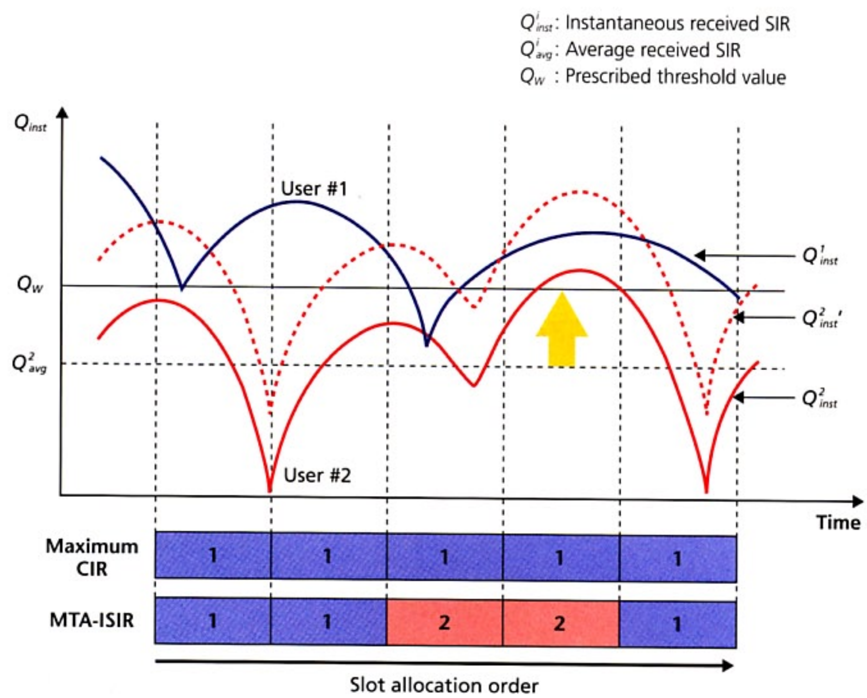

Figure 6 Principle of the Minimum Throughput Assured Instantaneous-SIR method

$$Q^i_{inst}{}' = \begin{cases} Q^i_{inst} & \text{for the users such as } Q^i_{avg} > Q_w \\ Q^i_{inst} \cdot \dfrac{Q_w}{Q^i_{avg}} & \text{for the users such as } Q^i_{avg} \leq Q_w \end{cases} \quad (1)$$

In this algorithm, users for which $Q^i_{avg} < Q_w$ are regarded as having a average received SIR of Q_w, so the opportunities for slot allocation among such users are more or less equal, so slots are also allocated to users whose average received SIR values are small. Accordingly, the proportion of small average received SIR users whose throughput is below the required throughput can be reduced. On the other hand, preserving the priority in slot allocation for users that have high instantaneous received SIR makes it possible to improve peak throughput and sector throughput compared to the proportional fairness method, in which slot allocation is based on the ratio of each user's instantaneous received SIR to the average received SIR.

The slot allocation probability of the MTA-ISIR method is shown in **Figure 7** as a function of the normalized cell distance from the cell site. The performances for the proportional fairness method and the maximum CIR method are also shown as references. The slot allocation probability is defined here as the probability of a slot being allocated to a user that has data in the base station buffer awaiting transmission. The MTA-ISIR method allocates a higher ratio of slots to users that are closer to the base station when Q_w is low, and the allocation probability decreases as Q_w increases. For users that are in the region where the normalized distance from the base station is large, on the other hand, there is improvement in the slot allocation probability relative to the maximum CIR method, and the amount of improvement increases as Q_w increases. We can also see that, although the slot allocation probability decreases monotonically as the distance from the base station increases when the MTA-ISIR method is used, when the normalized distance is greater than 0.6, the allocation probability is nearly constant. Thus, the MTA-ISIR method maintain the fairness by implementing the same slot allocation for users that are distant from the base station as does the proportional fairness method, and throughput that is near that of the maximum CIR method is achieved for users in the cell neighborhood.

5. Efficient Uplink Packet Access Method

A configuration for packet access in the up-link is shown in **Figure 8**.

When data that is to be transmitted is generated, the mobile terminal first transmits a reservation packet, which contains QoS information such as the allowable delay and residual packet error rate in addition to the user ID, the data size of the data packet, the user/control data identifier, etc. It also includes data on the state of the channel condition measured at the mobile terminal. Based on the reservation packet from each user, the base station performs reservation control (scheduling) and transmits a reply packet to the mobile terminal. The reply packet contains an access permission/denial identifier, scheduling information, radio link parameters of the data packet, transmission timing control information, etc. The mobile terminal transmits data packets according to the contents of the reply packet. This kind of control allows efficient packet access according to radio link and traffic conditions.

Two methods for configuring the reservation packet are being considered. One is to use layered signatures. The signatures comprise two layers. The lower layer signatures are utilized for user identification and path search at the base station, whereas the higher layer signatures convey reservation control information [9]. This approach has the merit of not requiring decoding processing because the signature detection and reservation control data detection are performed at the same time, and it is suitable for transmitting relatively short reservation control information. The other method is to use the pilot and data symbols. This method can transmit relatively long reservation control information and high reliable transmission can be attained by application of channel coding.

For the uplink, an efficient packet access method in which radio link parameter control is performed according to the allowable delay has been proposed, as shown in **Figure 9** [10],[11]. The mobile terminal informs the base station of the allowable delay by means of the reservation packet, and the base station controls the transmission power offset of the data packets that correspond to the reservation packet. When the allowable delay requirement such as for RT traffic is severe, a large transmission power offset is applied to achieve small delay. When the allowable delay such as for NRT traffic is relatively lax, a small transmission power offset is applied, thus reducing the total received E_b/N_0 (the signal energy per bit to background noise power spectrum density ratio) by making use of the time diversity effect of packet retransmission and combining and so increasing the system capacity.

The computed total required average received E_b/N_0 for error-free transmission is shown in **Figure 10** as a function of the transmission power offset of the data packet from the reservation packet. (Note that both the signal energy of the reservation and data packets are included in E_b.) As the transmission power offset becomes small, the total received E_b/N_0 decreases because of the time diversity effect. As the offset becomes even smaller, the received power per packet becomes too small and the received E_b/N_0 becomes

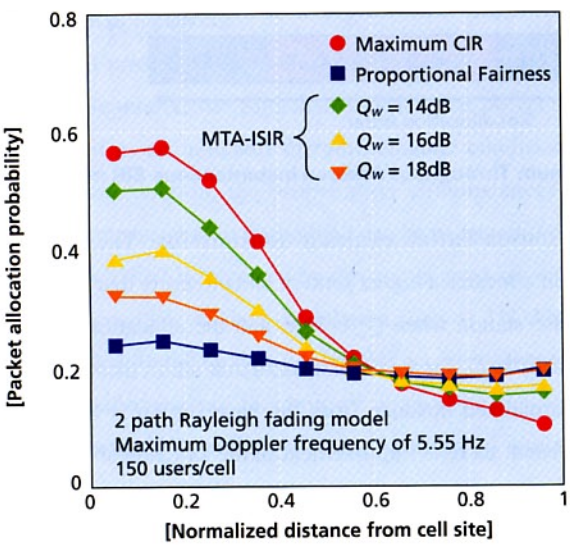

Figure 7 Packet allocation rate

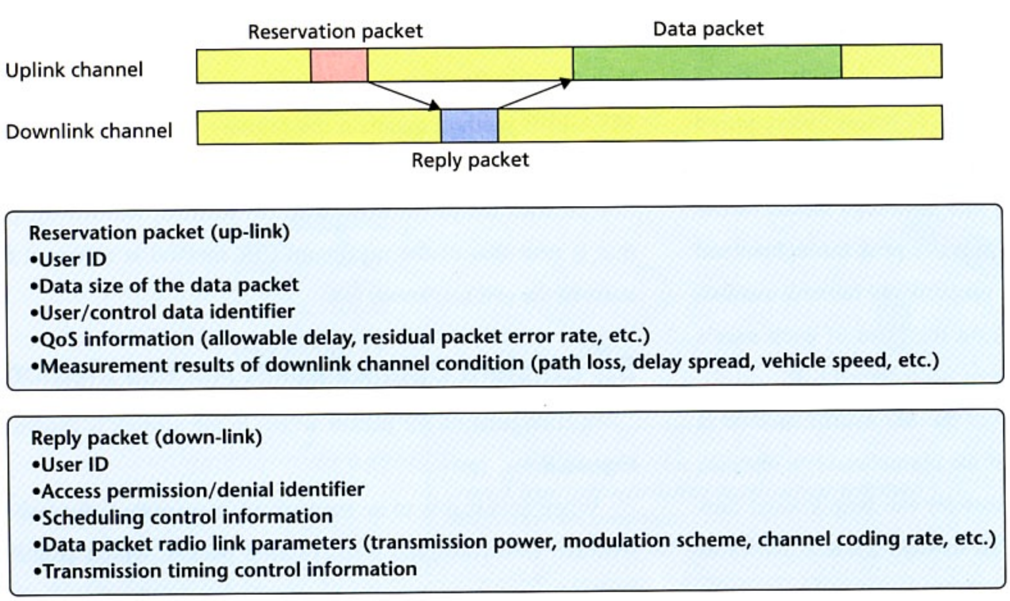

Figure 8 Simplified configuration for packet access in the up-link (reservation packets and data packets)

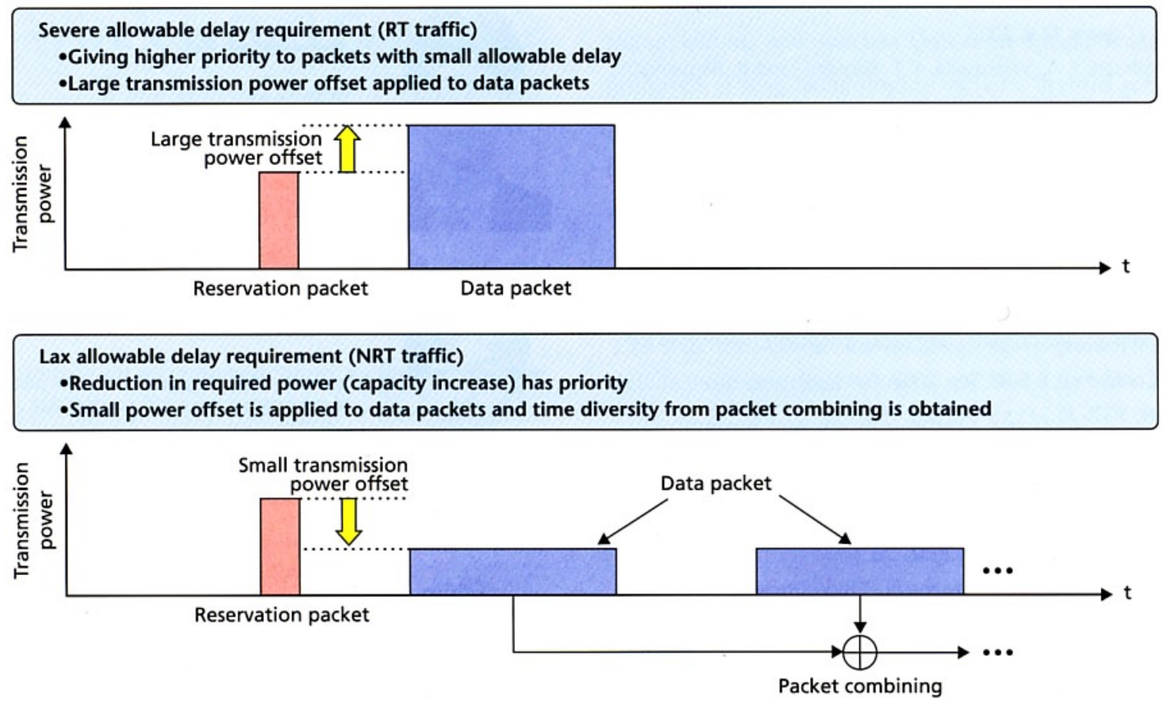

Figure 9 Time diversity proposal

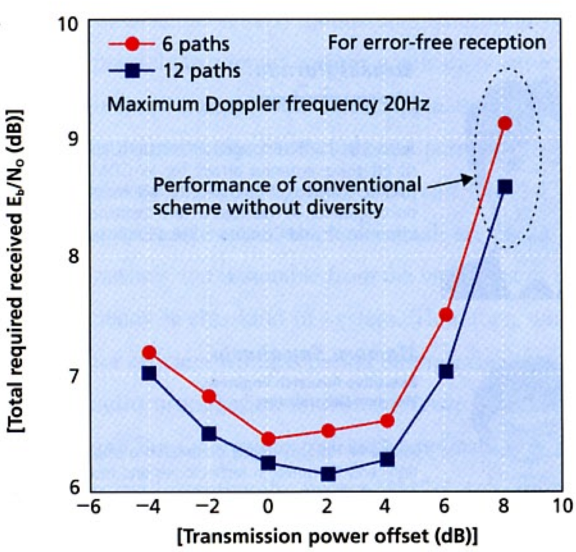

Figure 10 Time diversity characteristics

large due to retransmission. Since the conventional scheme does not control the power offset based on the allowable delay, a large transmission power offset is used to satisfy the requirement of rigid allowable delay. In contrast to that, the proposed method reduces the total received E_b/N_0 by 2 dB or more compared to conventional methods by applying a 2-dB transmission power offset to traffic that has relatively lax allowable delay requirements. In that way, the proposed method can increase system capacity while satisfying the respective allowable delay requirements of the different types of traffic.

6. Conclusion

We described efficient packet access technology for achieving the QoS (packet error rate and allowable delay) required by traffic data in the data link layer and the physical layer in a broadband wireless access system.

In the future, we plan to test the effectiveness of this efficient packet access technology in broadband channels through laboratory and field experiments, verify the effectiveness of hand-over, and evaluate the IP packet transmission performance.

REFERENCES

[1] D. Chase: "Code combining—a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," IEEE Trans. Commun., Vol. 33, No. 5, pp. 385-393, May. 1985.

[2] J. Hagenauer: "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," IEEE Trans. Commun., Vol. 36, No. 4, pp. 389-400, Apr. 1988.

[3] A. Jalali, R. Padovani, and R. Pankaj: "Data Throughput of CDMA-HDR a High Efficiency-High Data Rate Personal Communication Wireless System," IEEE VTC2000-Spring, pp. 1854-1858, Tokyo, May. 2000.

[4] Y. Ofuji, S. Abeta, and M. Sawahashi: "Comparison of Packet Scheduling Algorithms Focusing on User Throughput in High Speed Downlink Packet Access", IEICE Trans. Commun., Vol. E86-B No. 1 pp. 132-141, Jan. 2003.

[5] Y. Ofuji, S. Abeta, and M. Sawahashi: "Fast Packet Scheduling Algorithm Based on Instantaneous SIR with Constraint Condition

Assuring Minimum Throughput in Forward Link," IEEE WCNC pp. 860-865, New Orleans; Mar. 2003.

[6] D.J. Goodman, R.A. Valenzuela, K.T. Gayliard, and B. Ramamuithi, "Packet Reservation Multiple Access for Local Wireless Communications," IEEE Trans. Commun., Vol. 37, No. 8, pp. 885-890, Aug. 1989.

[7] N. Miki, S. Abeta, and M. Sawahashi: "Evaluation of Throughput Employing Hybrid ARQ Packet Combining in Forward Link OFCDM Broadband Wireless Access," IEEE PIMRC2002, pp. 394-398, Sep. 2002.

[8] Y. Ofuji, S. Abeta, and M. Sawahashi : "Packet Scheduling Algorithm with Beam Forming in High-Speed Downlink Packet Access" 2002 IEICE Annual Conference, B-5-88, Sep. 2002.

[9] 3GPPRAN, 3GTS 25.211 V3-8-0, Sep. 2001.

[10] Y. Iizuka, M. Tanno and M. Sawahashi: "Random Access Channel Transmission Utilizing Soft-Combining of Retransmitted Message Packets for Broadband Multicarrier/DS-CDMA," IEICE Technical Report, RCS2002-195, pp. 19-24, Nov. 2002.

[11] Y. Iizuka, M. Tanno, and M. Sawahashi, "Efficient random access channel transmission method utilizing soft-combining of retransmitted message packets according to QoS," IEEE ICCS2002, pp. 441-445, Nov. 2002.

ABBREVIATIONS

AMC: Adaptive Modulation and channel Coding
ARQ: Automatic Repeat reQuest
CIR: Carrier to Interference power Ratio
CRC: Cyclic Redundancy Check
H-ARQ: Hybrid ARQ
HSDPA: High-Speed Downlink Packet Access
IP: Internet Protocol
IR: Incremental Redundancy
MAC: Medium Access Control
MCS: Modulation and channel Coding Scheme
MRC: Maximal Ratio Combining
MTA- ISIR: Minimum Throughput Assured Instantaneous-SIR
MUX: MUltipleXer
NACK: Negative ACKnowledgement
NRT: Non Real Time
PDC: Personal Digital Cellular
PDU: Protocol Data Unit
QoS: Quality of Service
RLC: Radio Link Control
RT: Real Time
SINR: Signal to Interference and Noise power Ratio
SIR: Signal to Interference power Ratio
W-CDMA: Wideband Code Division Multiple Access