IEICE TRANS. COMMUN., VOL.E106–B, NO.8 AUGUST 2023

PAPER

HARQ Using Hierarchical Tree-Structured Random Access Identifiers in NOMA-Based Random Access*

Megumi ASADA†, Nobuhide NONAKA††, *Members*, *and* Kenichi HIGUCHI†a), *Senior Member*

SUMMARY We propose an efficient hybrid automatic repeat request (HARQ) method that simultaneously achieves packet combining and resolution of the collisions of random access identifiers (RAIDs) during retransmission in a non-orthogonal multiple access (NOMA)-based random access system. Here, the RAID functions as a separator for simultaneously received packets that use the same channel in NOMA. An example of this is a scrambling code used in 4G and 5G systems. Since users independently select a RAID from the candidate set prepared by the system, the decoding of received packets fails when multiple users select the same RAID. Random RAID reselection by each user when attempting retransmission can resolve a RAID collision; however, packet combining between the previous and retransmitted packets is not possible in this case because the base station receiver does not know the relationship between the RAID of the previously transmitted packet and that of the retransmitted packet. To address this problem, we propose a HARQ method that employs novel hierarchical tree-structured RAID groups in which the RAID for the previous packet transmission has a one-to-one relationship with the set of RAIDs for retransmission. The proposed method resolves RAID collisions at retransmission by randomly reselecting for each user a RAID from the dedicated RAID set from the previous transmission. Since the relationship between the RAIDs at the previous transmission and retransmission is known at the base station, packet combining is achieved simultaneously. Computer simulation results show the effectiveness of the proposed method.

key words: *random access, NOMA, grant-free NOMA, HARQ, packet combining, packet collision*

1. Introduction

Random access [1]–[5] that actualizes multi-packet reception [6] based on non-orthogonal multiple access (NOMA) with an advanced transceiver [7] reduces the control-signaling overhead and transmission delay compared to scheduling-based access. Therefore, an advanced random access protocol needs to be investigated to accommodate uplink transmission for massive machine-type communications (mMTC) and ultra-reliable low latency communications (URLLC) in the fifth-generation mobile communication system (5G NR) [8], [9] and beyond [10].

In this paper, we consider random access-based grant-free data transmission [4]. Thus, users transmit a preamble and data jointly within a random-access packet, which is referred to as two-step random access. This is under investigation by the 3rd Generation Partnership Project (3GPP) [11], [12] with the aim of reducing the control-signaling overhead and achieving a shorter transmission delay. In such a random-access scheme with multi-packet reception, the system prepares multiple random access identifiers (RAIDs) in advance to decode simultaneously transmitted packets using multiuser detection/decoding in NOMA. Here, the RAID functions as a separator for simultaneously received packets that use the same channel in NOMA. The transmitting user sends uplink random access packets generated using an individually selected RAID [4], [11], [12]. The RAID can be a scrambling code in the 4th and 5th generation mobile communication systems, Long Term Evolution (LTE)/LTE-Advanced and NR [9], [13], [14]; a channel interleaver in interleave division multiple access (IDMA) [3], [15]; or a generation matrix of sparse transmission signals in the time and frequency domains in low density signature (LDS)- and sparse code multiple access (SCMA)-based approaches [16], [17].

Inherently, random access is integrated with retransmission control. This is because random access causes decoding errors when multiple users transmit packets in the same time/frequency channel. In conventional random access such as ALOHA [18] and carrier sense multiple access (CSMA) [19] used in wireless LANs, multi-packet reception is not considered with priority, and using a random retransmission timing backoff is a basic strategy that is employed to resolve packet collisions when attempting retransmission. Although such retransmissions in random access have also been studied for NOMA-based random access that actualizes multi-packet reception [20], [21], in this paper, we consider the improvement in the spectrum efficiency when retransmissions are also performed in the same channel. However, the proposed method described in this paper can be combined with the above-mentioned approach that prepares multiple time/frequency channels for retransmission.

In NOMA-based random access, which enables multi-packet reception, decoding errors of received packets occur when the received signal power decreases due to channel fading or when the number of simultaneously received packets exceeds the decoding capability of an advanced receiver such as a successive interference canceller (SIC). Packet combining [22] between the initial and retransmitted packets based on hybrid automatic repeat request (HARQ) is very powerful in reducing such packet decoding errors.

Manuscript received January 4, 2023.
Manuscript publicized February 21, 2023.

†The authors are with the Graduate School of Science and Technology, Tokyo University of Science, Noda-shi, 278-8510 Japan.
††The author is with NTT DOCOMO, INC., Yokosuka-shi, 239-8536 Japan.

*The material in this paper was presented in part at the 94th IEEE Vehicular Technology Conference (VTC2021-Fall), Online, 2021.

a) E-mail: higuchik@rs.tus.ac.jp

DOI: 10.1587/transcom.2023EBP3001

Packet combining in HARQ improves the signal-to-noise plus interference power ratio (SINR) or error correction coding gain at retransmission. When multiple users select the same RAID, simultaneously transmitted packets that are generated using the same RAID cannot be separately decoded on the base station receiver side [23]. Therefore, RAID collisions must also be resolved during retransmission in HARQ.

In NOMA-based random access, the base station can recognize the RAID of the received packet through active user detection and channel estimation [24], [25] using a RAID-specific preamble added to the received packet. However, the user index (user ID) is written in the data part of the packet, so it will be recognized only when the received packet is correctly decoded. In other words, retransmission control is directed by and responds to the RAID.

For this reason, the base station must recognize the relationship between the RAIDs of the initial packets and the retransmitted packets to be combined. However, if the same RAID is used for the initial and retransmitted packets, packet collision resulting from duplicate RAID selection by multiple users cannot be resolved even with retransmission. On the other hand, if the RAID of the retransmitted packet is randomly reselected by the user, the RAID collision can be resolved at retransmission, but packet combining becomes impossible. Furthermore, when channel-dependent RAID selection is assumed such as in the RAID-linked receiver beamforming (BF) method [26], [27], reselection of a RAID that does not match a given user channel at retransmission degrades the transmission performance due to the reduced received signal power after BF.

As related work, in the discussion of two-step random access by the 3GPP, when decoding of the random access packet fails, the process is assumed to fall back to the conventional four-step random access scheme [11], [12]. This approach is simple, but fails to gain the packet combining effect. In [28]–[30], packet combining at retransmission is investigated for random access based on NOMA including code division multiple access (CDMA), but RAID collision is not taken into account. In practice each user independently selects the RAID used for transmission (in the case of CDMA, the RAID is a spreading code including a scrambling code). When the selected RAID is the same among multiple simultaneous transmitting users, the receiver cannot separate the packets of different users, which results in a high decoding-error probability for the received packets. This reduces the transmission efficiency. In [31], [32], advanced retransmission control to reduce the number of packet collisions is investigated, but packet combining during retransmission is not achieved. In [33], a method for blind detection of the RAID equivalent from the received signal is proposed assuming binary phase-shift keying (BPSK) data modulation. However, there is no consideration of error correction coding, which is essential to achieve low error rates, and insufficient consideration given to the case where RAIDs collide. Therefore, the above-mentioned conventional methods can only perform packet combining or resolve RAID collisions at retransmission.

This paper proposes an efficient HARQ method that simultaneously achieves packet combining and resolution of the RAID collisions during retransmission in NOMA-based random access systems for multi-packet reception. The proposed method employs novel hierarchical tree-structured RAID groups in which the RAID for the previous (including initial) packet transmission has a one-to-one relationship with the set of RAIDs for retransmission. The proposed method resolves RAID collisions at retransmission through individual reselection of RAIDs at each user from the RAID set for retransmission, which is dedicated to the RAID selected at the previous transmission. Since the relationships between the RAIDs at the previous transmission and retransmission are known at the base station, packet combining is achieved simultaneously. Computer simulation results show the effectiveness of the proposed method. We note that the contents of this paper are based on [34], but include enhanced evaluation and discussions.

The remainder of the paper is organized as follows. First, Sect. 2 describes NOMA-based random access with multi-packet reception and the challenges it faces. Section 3 describes the proposed method. Then, Sect. 4 presents numerical results based on computer simulations. Finally, Sect. 5 concludes the paper.

2. NOMA-Based Random Access for Multi-Packet Reception and Its Challenges

Figure 1 shows the basic operational flow of random access with multi-packet reception. The base station periodically broadcasts a set of available RAIDs, $\mathcal{M}$, via the downlink. Each user terminal k that has uplink information individually selects RAID $i \in \mathcal{M}$ from the informed set. At the user k transmitter, based on the selected RAID i, channel coding, data modulation, and physical channel mapping are applied to the transmission information bit sequence. For example, when we assume scrambling code-based NOMA, the channel coded bit sequence is scrambled using the i-th scrambling code. The transmission information bit sequence includes user ID k. One packet comprises data and preamble parts. The preamble sequence is dedicated to RAID i. When multiple users attempt random access simultaneously, multiple packets are transmitted using the same time/frequency channel.

At the base station, first the received preamble signal is used to detect the set of active users and estimate its channel [24]. In random access, the user index (user ID) of the user who transmits a packet is unknown in advance, and active user detection means detecting the set of RAIDs, $\mathcal{M}_{\text{det}}$, each of which is used in a received packet. It should be noted that due to the failure of active user detection, it is possible that $\mathcal{M}_{\text{det}}$ may not include the RAID of the packet that was actually sent (this corresponds to RAID miss detection), and/or $\mathcal{M}_{\text{det}}$ may contain a RAID that is not actually sent (this corresponds to a RAID false alarm). Subsequently, based on detected RAID set $\mathcal{M}_{\text{det}}$ and its channel estimate, the base

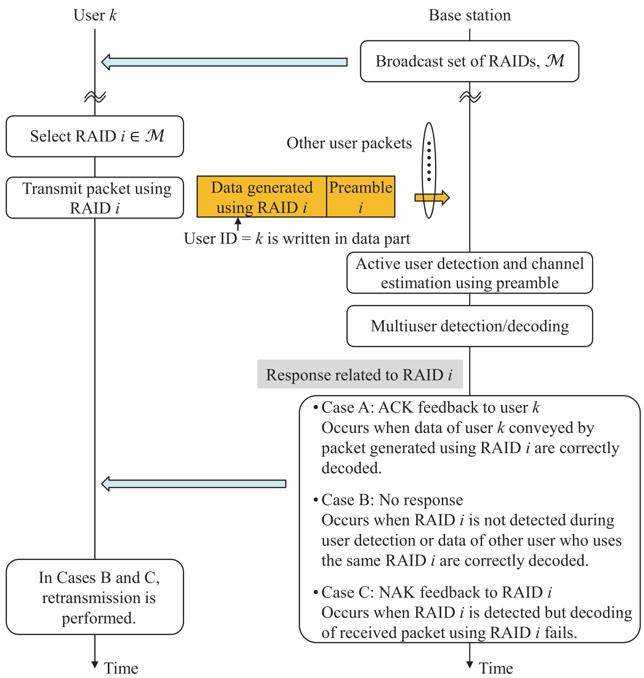


Fig. 1 Basic operational flow of random access with multi-packet reception.

station decodes the packet using a multi-user receiver such as a SIC. When multiple users select the same RAID i at the same time, the packet decoding results of RAID i become whether all users fail to decode or only one user succeeds in decoding.

The feedback from the base station regarding RAID i, which is used by user k, conforms to the following three cases from the viewpoint of user k.

A. Case A: ACK to User k

Case A occurs when the information bit sequence of user k carried by the packet using RAID i is correctly decoded. Since the base station recognizes the user ID of user k from the decoded data, it sends an acknowledgement message (ACK) to user k instead of RAID i to resolve the potential RAID collision.

B. Case B: No Response

Case B occurs when RAID i is miss detected during the active user detection at a base station. This case also occurs from the viewpoint of user k when RAID i is simultaneously used by another user l and the information bit sequence of user l is correctly decoded from the packet of RAID i. In this case, an ACK is sent to user l from the base station (this corresponds to Case A for user l).

C. Case C: NAK to RAID i

Case C occurs when the information bit sequence sent by the packet using RAID i is not correctly decoded by the advanced receiver such as the SIC after RAID i is correctly detected in the active user detection process at the base station. To inform this to the user who used RAID i, the base station sends a negative acknowledgement message (NAK) for RAID i. This case includes the situation where RAID i is also used by another user l and neither the information of user k nor that of user l are correctly decoded.

In Case A, user k completes data transmission. However, when Case B and Case C occur, user k will retransmit. Here, in Case B, since the base station does not recognize the previous packet transmission of user k, there is no room for improving the performance of the retransmission by using packet combining etc.

In the retransmission in Case C, since the base station recognizes the previous packet transmission using RAID i, there is a possibility to improve the random access performance during retransmission. More specifically, for the decoding error of the received packet when the channel is faded or when the number of simultaneously received packets exceeds the decoding capability of the advanced receiver such as the SIC, the packet combining between the previous (including initial) and retransmitted packets based on HARQ is very powerful in reducing the error rate after retransmission. For the decoding error of the received packet due to RAID collision, resolving the RAID collision at the retransmission attempt should be considered.

However, the conventional approach cannot simultaneously achieve packet combining and resolution of the RAID collisions. At retransmission, each user selects the RAID again, and the effect of retransmission changes depending on this selection method in the conventional method as follows.

(1) *Reselection of the Same RAID for Retransmission*

In this case, since the base station can recognize the relationship of RAIDs between the initial transmission packet and the retransmitted packet, these two packets can be combined. Through packet combining, the received signal power or error correction code capability after retransmission increases. However, if multiple users select the same RAID in the initial transmission, there is a problem that RAID collision will again occur even after retransmission.

(2) *Random Reselection of RAID for Retransmission*

In this case, even if there are RAID conflicts in the initial transmission, a RAID collision can be resolved at retransmission since the RAID is randomly reselected at each user independently. However, since the base station cannot recognize the relationship between the RAIDs in the initial transmission and retransmission, the base station cannot perform packet combining. Furthermore, when channel-dependent RAID selection is assumed such as in the RAID-linked receiver BF method [26], [27], reselection of the RAID, which does not match a given user channel at retransmission, reduces the received signal power after receiver BF, which represents the combined received signal power from among all receiver antennas. This degrades the transmission performance. References [26] and [27] include some quantitative evaluations on the impact of the unmatched receiver BF applied to the user channel on the achievable packet error rate (PER).

3. Proposed Method

The proposed method simultaneously actualizes packet combining and resolution of RAID collision at retransmission in HARQ. To achieve this, we employ novel hierarchical tree-structured RAID groups. Figure 2 shows the proposed hierarchical tree-structured RAID group configuration. Unlike the conventional method, in the hierarchical tree-structured RAID group configuration, a RAID set for initial transmission and that for retransmission are prepared separately. Let $\mathcal{M}$ be the set of RAIDs for initial transmission. The number of RAIDs prepared for initial transmission is denoted as $M = |\mathcal{M}|$. Term T denotes the maximum number of retransmissions allowed by the HARQ protocol. The RAID set for retransmission is defined separately for the respective number of retransmission attempts. Furthermore, each RAID set for retransmission is linked to the RAID used in the previous transmission. The set of RAIDs that can be used for the t-th retransmission ($t = 1, \dots, T$) of the user who used RAID i in the $t-1$-th transmission ($t-1 = 0$ corresponds to initial transmission) is denoted as $\mathcal{R}_i^{(t)}$. The number of RAIDs in $\mathcal{R}_i^{(t)}$ for all i is denoted as $R^{(t)}$, and $R^{(t)}$ is set equal to or greater than two.

The proposed method is applied to retransmission in Case C described in Sect. 2. In Case B, since the base station does not recognize the reception of the corresponding packet, the user repeats the process equivalent to that for the initial transmission, which is the same as in the conventional retransmission process. On the other hand, in Case C, the base station recognizes the RAID of the received packet in decoding error. By utilizing this, when the base station feeds back a NAK for RAID i used at the $t-1$-th transmission in Case C of the proposed method, the base station also notifies the index of the RAID set for the next retransmission, $\mathcal{R}_i^{(t)}$, to users who used RAID i at the $t-1$-th transmission via the downlink. This is a kind of on-demand-type RAID allocations for retransmission. The user who replies with a NAK to the previously transmitted packet using RAID i recognizes RAID set $\mathcal{R}_i^{(t)}$ and randomly selects a RAID from $\mathcal{R}_i^{(t)}$ for the next t-th retransmission. Here, if the decoding error of the previously transmitted packet using RAID i is caused by a RAID collision with other users, the user-dependent random RAID reselection from $\mathcal{R}_i^{(t)}$ yields an opportunity to resolve the RAID collision at the t-th retransmission. For example, when two users incur a RAID collision at the $t-1$-th transmission, the probability that the RAID collision is resolved in the next t-th retransmission is $1-(1/R^{(t)})$. When the proposed method is applied to the RAID-linked receiver BF method in [26], [27], the received BF vector associated with RAID i in the initial transmission is also applied to all the RAIDs in all the retransmission RAID groups that are linked to the initial RAID i.

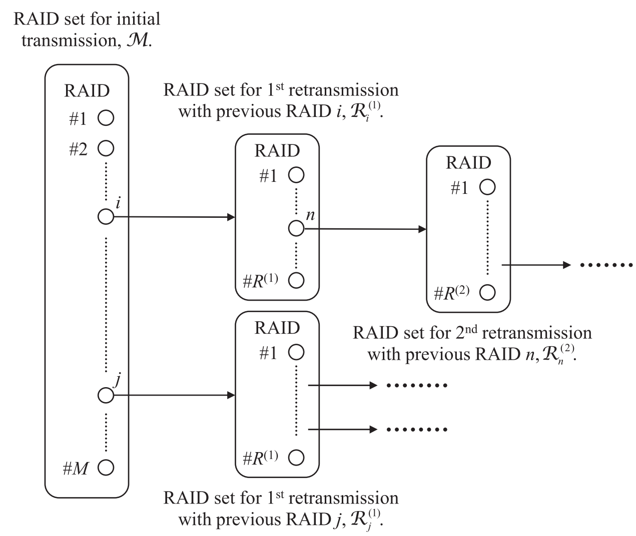


Fig. 2 Proposed hierarchical tree-structured RAIDs.

Table 1 Operation of proposed method.

Operation of proposed method for user k using RAID i at the t−1-th transmission

Feedback from Base Station to t−1-th Packet Transmission	User k Action at Subsequent Time t	Base Station Receiver Process at Time t	
Case A: ACK to user k	Complete packet transmission	None	
Case B: No response	Recognize miss detection of previous packet and reattempt initial packet transmission	Decode packet as initial transmission	
Case C: NAK to RAID i with $\mathcal{R}_i^{(t)}$	Randomly select RAID n from $\mathcal{R}_i^{(t)}$ and retransmit packet using RAID n	Receive single RAID from $\mathcal{R}_i^{(t)}$	Decode after combining packet i in t−1 and n in t
		Receive multiple RAIDs from $\mathcal{R}_i^{(t)}$	Decode packet n solely

The receiver operation at the base station in the proposed method is explained hereafter. When the base station identifies the reception of the packet using the RAID contained in $\mathcal{R}_i^{(t)}$ from the active user detection process, it recognizes that this packet is a retransmission of the previous packet using RAID i. Therefore, the packet combining of these two packets is activated and this brings about improvement in the SINR or error correction code capability. However, when multiple RAIDs in $\mathcal{R}_i^{(t)}$ are detected in the active user detection process, the base station recognizes the RAID collision of the previous packet using RAID i. In this case, packet combining is not activated to avoid degradation due to the combining of the conflicted RAID packets. Table 1 summarizes the operation of the proposed method.

4. Numerical Results

4.1 Simulation Parameters

The performance of the proposed method is evaluated based on computer simulations. Table 2 gives the major simulation parameters. Random access with a 4.6-MHz transmis-

Table 2 Simulation parameters.

Transmission bandwidth		4.6 MHz
Packet length		0.5 ms (7 DFT blocks)
RAID		Scramble code
Channel coding	Turbo code	Code rate = 1/3
	Repetition code	Code rate = 1/10 (10 repetitions)
Data modulation		QPSK
Number of users that transmit initial packets simultaneously		K = 20 (except for Fig. 7)
Number of receiver antennas		100
Receiver BF		RAID-linked receiver BF method [26, 27]
Channel model		Six-path block Rayleigh, rms delay spread = 1 µs
Maximum number of retransmissions		T
Packet combining		LLR-based combining
Number of RAIDs in proposed method	Initial transmission	M = 200 (except for Fig. 6)
	Retransmission	R for all $R^{(1)}, \ldots, R^{(T)}$

sion bandwidth and 0.5-ms time slot, which is equal to the packet length, using the RAID-linked receiver BF [26], [27] is assumed as the baseline random access method. A combination of the turbo and repetition codes is used for channel coding. The coding rate for the turbo code is 1/3, which is used in LTE [13]. The number of repetitions in the repetition coding is set to 10. The scrambling code is used as the RAID. We assume bit-level scrambling. The scrambling code is randomly generated where its length is equal to that of the coded bit sequence and each element of the scrambling code takes '1' or '0'. The bit-level scrambling is achieved by calculating the exclusive-OR operation between each of the encoded bits and elements of the scrambling code. Quadrature phase-shift keying (QPSK) data modulation is assumed. We assume discrete Fourier transform (DFT)-spread orthogonal frequency division multiplexing (OFDM)-based single-carrier transmission [13], [14]. One packet comprises 7 DFT blocks with the packet length of 0.5 ms including a cyclic prefix (CP). One DFT block is used for transmission of the preamble comprising a Zadoff-Chu sequence. The base station performs received-RAID detection (active user detection) and channel estimation using the received preamble signal [25]. The number of users that transmit the initial packet simultaneously at each random access time slot, K, is set to 20, except for the evaluation in Fig. 7. For example, K of 20 corresponds to the situation where 40,000 user terminals are associated with a base station and each user terminal transmits one random-access packet per second on average. In Fig. 7, K is parameterized. Six-path block Rayleigh fading with the rms delay spread of 1 µs is simulated using the tapped delay line model [35] as the channel model. Assuming that the transmission power control compensates for the distance-dependent path loss and random shadowing between each user and the base station, the received signal-to-noise ratio (SNR) of all users is assumed to be the same.

The number of receiver antennas at the base station is 100. We use the SIC in [4] as a multiuser receiver and the maximum number of iterations in the SIC process is set to eight. Packet combining is achieved by adding the log-likelihood ratio of each coded bit. Except for the evaluation in Fig. 6, the M of the proposed method is fixed to 200. The maximum number of retransmissions in HARQ, T, is parameterized. For simplicity, all of $R^{(1)}, \ldots, R^{(T)}$ in the proposed method are set to the same value, R, where R is parameterized. For comparison, we also evaluate the conventional method in which a single RAID group $\mathcal{M}$ is used for initial transmission and retransmissions. The average total number of available RAIDs per random access time slot in the proposed method, $\overline{N}_{\mathrm{RAID}}^{(\mathrm{Prop.})}$, for the given K, M, and $R^{(t)}$ can be roughly estimated as

$$\text{Rough estimate of } \overline{N}_{\mathrm{RAID}}^{(\mathrm{Prop.})} = M + \sum_{t=1}^{T} K Q_{t-1} R^{(t)}, \qquad (1)$$

where Q_{t-1} is the temporal PER at the $t-1$-th retransmission. We note that Eq. (1) is shown just for a purpose to provide an insight to the on-demand usage of RAIDs for retransmissions in the proposed method. The actual $\overline{N}_{\mathrm{RAID}}^{(\mathrm{Prop.})}$ can be less than the above estimate since the same set of RAIDs for retransmission is used for multiple users when the packet decoding error is caused by RAID collision. In the following performance evaluations, Eq. (1) is not used for calculating $\overline{N}_{\mathrm{RAID}}^{(\mathrm{Prop.})}$. Instead, actual $\overline{N}_{\mathrm{RAID}}^{(\mathrm{Prop.})}$ is experimentally measured using computer simulations for given system conditions including K, M, T, and R. The number of RAIDs in the conventional method is set to the measured $\overline{N}_{\mathrm{RAID}}^{(\mathrm{Prop.})}$ for fair comparison in all figures except for Fig. 6. In Fig. 6, the average PER after HARQ process is measured as a function of the average total number of available RAIDs per random access time slot. The HARQ ACK/NAK feedback is assumed to be transmitted to the user without error. In the following evaluations, the average PER represents the average residual PER after HARQ process.

4.2 Simulation Results

Figure 3 shows the average PER as a function of the SNR per receiver antenna. In addition to the proposed method with $T = 1$ and $R = 4$, the conventional method with and without packet combining with $T = 1$ and the case without HARQ are tested. The RAID collision probability with regard to the number of RAIDs, M, and the number of simultaneously transmitting users, K, is represented as $P_{\mathrm{collision}}(M, K) = 1 - ((M-1)/M)^{K-1}$. As a reference, $P_{\mathrm{collision}}$ with $M = 200$ and $K = 20$ is shown in Fig. 3. The observed PER floor in a high SNR region without HARQ is dominated by RAID collisions. The reason why the PER floor is slightly lower than the RAID collision probability is that even if the RAIDs collide, when there is a clear difference in received signal power between collided users, the information sequence of the user with a higher received signal power may be correctly decoded, which is referred to as the capture effect.

The floor of the average PER of the conventional method is reduced by retransmission. This is because even

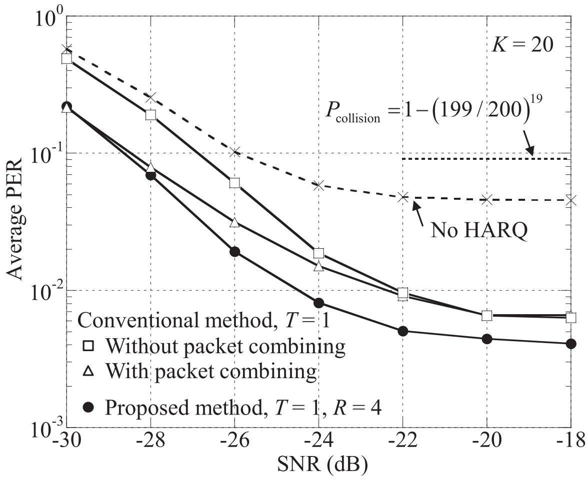


Fig. 3 Average PER as a function of SNR.

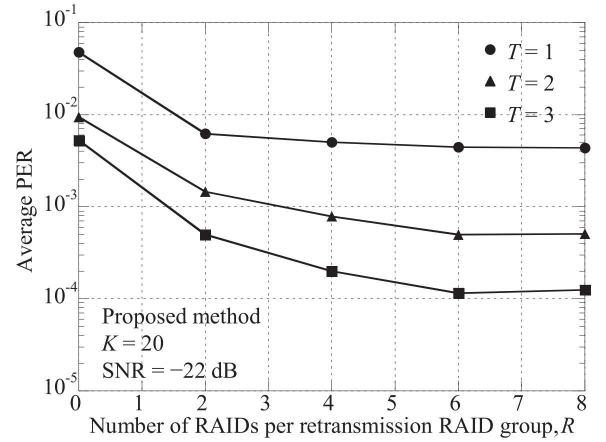


Fig. 4 Average PER as a function of the number of RAIDs per retransmission RAID group, R.

if the RAIDs collide at the initial transmission, when one of the users succeeds in decoding, the RAID conflicts will be resolved at retransmission. However, in the conventional method with packet combining, the improvement in the PER floor by retransmission is relatively small because packet combining degrades the decoding performance when a RAID collision occurs. The proposed method further reduces the PER floor compared to the conventional method. This is because, according to the exclusive definition of the RAIDs for retransmission, even if the decoding of the packet with RAID collision fails, the RAID collision can be resolved by retransmission. Furthermore, since the proposed method can conduct packet combining at retransmission, the received signal power after packet combining increases. Even in the conventional method, the effect of packet combining is observed in a relatively low SNR region. However, in the conventional method, the effect of packet combining is limited since RAID collisions between users with different numbers of retransmission attempts occur. Therefore, the proposed method is able to reduce the required SNR to achieve the same average PER compared to that for the conventional method.

Figure 4 shows the average PER as a function of the number of RAIDs per retransmission RAID group, R. The maximum number of retransmissions, T, is parameterized. The SNR is set to −22 dB. The performance level at $R = 0$ shows the conventional method without using packet combining. The average PER decreases as R increases. This is because the resolution probability of RAID collision increases since the users incurring RAID collision individually and randomly reselect the RAID for retransmission from the R candidates. This effect is especially significant when T is set high. This is because there is an opportunity to resolve the RAID collision at each retransmission attempt up to T times.

Figure 5 shows the average PER as a function of the SNR with the maximum number of retransmissions, T, as a parameter. The R of the proposed method is set to four. For comparison, the results of the conventional method with-

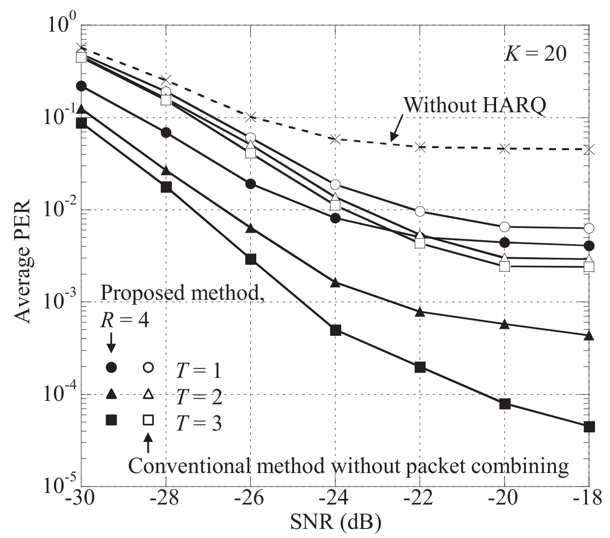


Fig. 5 Average PER as a function of SNR with maximum number of retransmissions, T, as a parameter.

out packet combining and the case without HARQ are also shown. The proposed method improves the average PER compared to the conventional method, and this becomes more significant when T is set higher. In the conventional method, even if T is increased, the RAIDs continue to collide until one of the users who selected the same RAID succeeds in decoding the packet. On the other hand, in the proposed method, each user randomly reselects one RAID from multiple candidates dedicated to retransmission. So, the packet decoding error due to RAID collision decreases more effectively as the number of retransmissions increases. Furthermore, the proposed method reduces the SNR required to achieve the target PER as T increases thanks to the effect of the packet combining.

Figure 6 shows the average PER as a function of the average total number of available RAIDs per random access time slot. Term T is set to 1 and the SNR is −22 dB. In addition to the proposed method with $R = 4$, the conven-

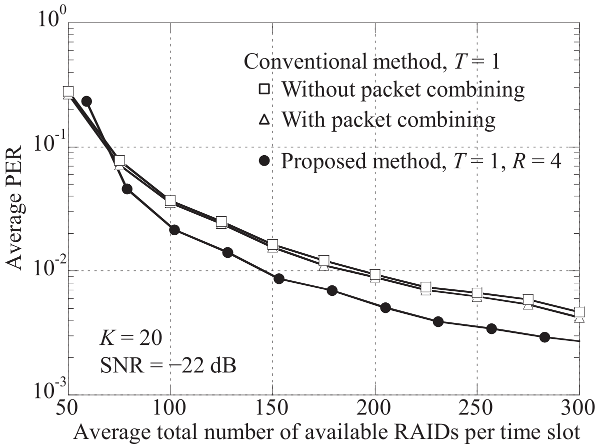


Fig. 6 Average PER as a function of average total number of available RAIDs per random access time slot.

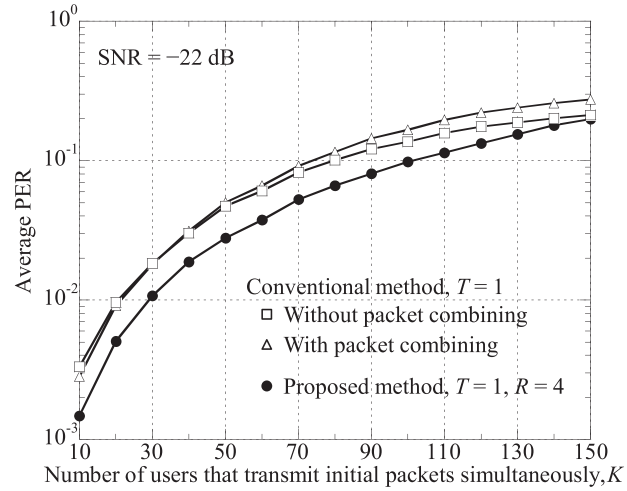


Fig. 7 Average PER as a function of number of users that transmit initial packets simultaneously, K.

tional method with and without packet combining is tested. The proposed and conventional methods are evaluated while changing the value of M. The proposed method, in which the RAID group is prepared separately for the initial transmission and the retransmissions, resolves the RAID collision more efficiently, so that the number of RAIDs required to achieve the same PER is reduced compared to that for the conventional method with and without packet combining. Since the base station tries to detect active users (RAIDs) for all the prepared RAIDs, the proposed method that reduces the effective number of RAIDs also contributes to reducing the computational complexity of the base station receiver.

Figure 7 shows the average PER as a function of the number of users that transmit initial packets simultaneously, K. The SNR is −22 dB and T is set to 1. The proposed method with $R = 4$ and the conventional method with and without packet combining are compared. The proposed method reduces the PER compared to that for the conventional method for a given K. This is because, the proposed method can resolve the RAID collision more efficiently at retransmission than the conventional method although the reduction effect of the PER by using the proposed method is decreased when K is excessively large. The K values of the proposed method are increased by approximately 20 users compared to that for the conventional method for achieving the same PER. This means that the proposed method is effective in accommodating a larger number of random access users, which should contribute to actualizing mMTC.

Figure 8 shows the complementary cumulative distribution function (CCDF) of the number of retransmissions required for correct decoding. The SNR is −22 dB and T is set to 2. The proposed method with $R = 4$ and the conventional method without packet combining are compared. The proposed method has a higher probability of correct decoding with a smaller number of retransmissions than the conventional method. This means that the proposed method is effective in reducing the transmission latency of the random access, which can contribute to actualizing URLLC.

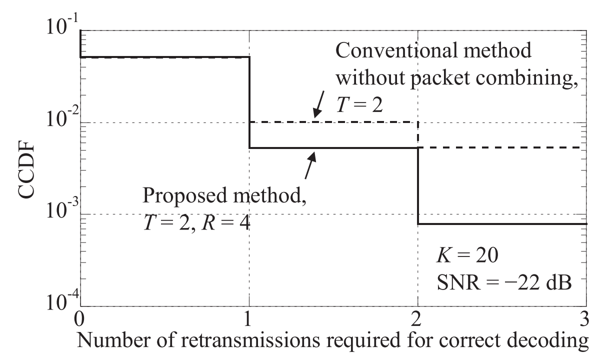


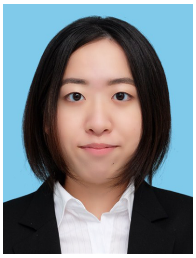
Fig. 8 CCDF of number of retransmissions required for correct decoding.

5. Conclusion

In this paper, we proposed an efficient HARQ scheme for NOMA-based random access systems for multi-packet reception to actualize mMTC. Different from the conventional approaches, the proposed method simultaneously actualizes packet combining and resolution of RAID collisions at retransmission in HARQ. To achieve this, we introduce novel hierarchical tree-structured RAID groups in which the RAID for the previous packet transmission has a one-to-one relationship with the set of RAID candidates for retransmission. The proposed method resolves RAID collisions at retransmission through individual and random RAID reselection at each user from the multiple retransmission RAID candidates that are dedicated to the RAID selected at the previous transmission. Furthermore, since the relationships between the RAIDs at the previous transmission and retransmission are known at the base station, packet combining is achieved simultaneously. Based on extensive computer simulation results, we showed that the proposed method reduces the PER floor due to RAID collision and reduces the required SNR by packet combining compared to those for the conventional method. Furthermore, we showed that by using the proposed method, it is possible to reduce the number of RAIDs the system must prepare to obtain the same PER performance, which contributes to reducing the computational complexity of the base station receiver. Finally, the proposed method is also effective in reducing the transmission latency of the random access, which can contribute to actualizing URLLC.

References

[1] S. Lien, S. Shieh, Y. Huang, B. Su, Y. Hsu, and H. Wei, "5G new radio: Waveform, frame structure, multiple access, and initial access," IEEE Commun. Mag., vol.55, no.6, pp.64–71, June 2017.
[2] M. Shirvanimoghaddam, M. Dohler, and S.J. Johnson, "Massive non-orthogonal multiple access for cellular IoT: Potentials and limitations," IEEE Commun. Mag., vol.55, no.9, pp.55–61, Sept. 2017.
[3] Y. Hu, C. Xu, and L. Ping, "NOMA and IDMA in random access systems," Proc. IEEE VTC2018-Spring, Porto, Portugal, June 2018.
[4] M. Kawata, K. Tateishi, and K. Higuchi, "Performance evaluation of IDMA-based random access with various structures of interference canceller," IEICE Trans. Commun., vol.E103-B, no.9, pp.1030–1037, Sept. 2020.
[5] Y. Chen, G. Wang, H. Yi, and W. Zhang, "Priority-based distributed queuing random access mechanism for mMTC/uRLLC terminals coexistence," Proc. IEEE VTC2021-Spring, Virtual Conference, April-May 2021.
[6] J. Metzner, "On improving utilization in ALOHA networks," IEEE Trans. Commun., vol.COM-24, no.4, pp.447–448, April 1976.
[7] K. Higuchi and A. Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," IEICE Trans. Commun., vol.E98-B, no.3, pp.403–414, March 2015.
[8] ITU-R, "IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond," Recommendation M.2083-0, Sept. 2015.
[9] E. Dahlman, S. Parkvall, and J. Sköld, 5G NR: The Next Generation Wireless Access Technology, Academic Press, 2018.
[10] NTT DOCOMO, "White paper: 5G evolution and 6G," Jan. 2023.
[11] 3GPP, RP-182894, "New work item: 2-step RACH for NR," Dec. 2018.
[12] E. Peralta, T. Levanen, F. Frederiksen, and M. Valkama, "Two-step random access in 5G new radio: Channel structure design and performance," Proc. IEEE VTC2021-Spring, Virtual Conference, April-May 2021.
[13] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, 3G Evolution: HSPA and LTE for Mobile Broadband, 2nd ed., Academic Press, 2008.
[14] E. Dahlman, S. Parkvall, and J. Sköld, 4G: LTE/LTE-Advanced for Mobile Broadband, 2nd ed., Academic Press, 2013.
[15] L. Ping, L. Liu, K. Wu, and W.K. Leung, "Interleave division multiple-access," IEEE Trans. Wireless Commun., vol.5, no.4, pp.938–947, April 2006.
[16] R. Hoshyar, F.P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channels," IEEE Trans. Signal Process., vol.56, no.4, pp.1616–1626, April 2008.
[17] K. Au, L. Zhang, H. Nikopour, E. Yi, and A. Bayesteh, "Uplink contention based SCMA for 5G radio access," Proc. IEEE Globecom2014, Austin, USA, Dec. 2014.
[18] N. Abramson, "The ALOHA system — Another alternative for computer communications," Fall Joint Computer Conference, pp.281–285, 1970.
[19] E. Ziouva and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: Throughput and delay analysis," Computer Communications, vol.25, issue 3, pp.313–321, Feb. 2002.
[20] A.C. Cirik, N.M. Balasubramanya, L. Lampe, G. Vos, and S. Bennett, "Toward the standardization of grant-free operation and the associated NOMA strategies in 3GPP," IEEE Commun. Standards Mag., vol.3, no.4, pp.60–66, Dec. 2019.
[21] Y. Shake, Y. Kishiyama, and K. Higuchi, "Investigation on retransmission scheme for IDMA-based random access," Proc. IEEE APWCS2019, Singapore, Aug. 2019.
[22] Q. Zhang and S.A. Kassam, "Hybrid ARQ with selective combining for fading channels," IEEE J. Sel. Areas Commun., vol.17, no.5, pp.867–880, May 1999.
[23] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "On the performance of massive grant-free NOMA," Proc. IEEE PIMRC2017, Montreal, Canada, Oct. 2017.
[24] J. Zhang, P. Pan, L.-L. Yang, and R.G. Maunder, "Channel estimation and user activity identification in massive grant-free multiple-access," IEEE Open J. Veh. Technol., vol.1, pp.296–316, 2020.
[25] T. Tomizawa, Y. Kishiyama, and K. Higuchi, "Performance evaluation of IDMA-based random access considering user detection and channel estimation," Proc. ISPACS 2019, Taipei, Taiwan, Dec. 2019.
[26] Y. Muroki, Y. Murakami, Y. Kishiyama, and K. Higuchi, "Interleaver-linked receiver beamforming with transmitter channel equalization in TDD-IDMA-based random access," Proc. IEEE VTC2020-Fall, Virtual conference, Nov.-Dec. 2020.
[27] Y. Muroki, Y. Murakami, Y. Kishiyama, and K. Higuchi, "Random access identifier-linked receiver beamforming with transmitter filtering in TDD-based random access," IEICE Trans. Commun., vol.E105-B, no.12, pp.1548–1558, Dec. 2022.
[28] Y. Sun and J. Shi, "Retransmission diversity in large CDMA random access systems," IEEE Trans. Signal Process., vol.55, no.7, pp.3471–3483, July 2007.
[29] E. Paolini, G. Liva, and M. Chiani, "High throughput random access via codes on graphs: Coded slotted ALOHA," Proc. IEEE ICC2011, Kyoto, Japan, June 2011.
[30] Y. Iizuka, M. Tanno, and M. Sawahashi, "Efficient random access channel transmission method using packet retransmission according to QoS," IEICE Trans. Commun., vol.E86-A, no.7, pp.1669–1675, July 2003.
[31] N.V. Matveev, "The adaptive retransmission management in random multiple-access system with successive interference cancellation," Proc. WECONF2018, St. Petersburg, Nov. 2018.
[32] J.C. Marinello and T. Abrão, "Collision resolution protocol via soft decision retransmission criterion," IEEE Trans. Veh. Technol., vol.68, no.4, pp.4094–4097, April 2019.
[33] F.J. Block, "Packet combining considerations for random-access networks," Proc. MILCOM2010, San Jose, USA, pp.724–730, Oct.-Nov. 2020.
[34] M. Asada, N. Nonaka, and K. Higuchi, "HARQ using hierarchical tree-structured physical channel identifiers in NOMA-based random access for multi-packet reception," Proc. IEEE VTC2021-Fall, Virtual Conference, Sept.-Oct. 2021.
[35] A.F. Molisch, Wireless Communications, IEEE Press, 2010.



Megumi Asada received the B.E. and M.E. degrees from Tokyo University of Science, Noda, Japan in 2021 and 2023, respectively. Her research interests include wireless communications. She is a member of the IEICE.

Nobuhide Nonaka received the B.E. and M.E. degrees in electronic engineering from Tokyo University of Science, Japan, in 2013 and 2015, respectively. Since April 2015, he has been with NTT DOCOMO, Inc. Since April 2018, he has been engaged in the research of next generation radio access technologies. He received the Young Researcher's Award from IEICE in 2019. He is a member of the IEICE.

Kenichi Higuchi received the B.E. degree from Waseda University, Tokyo, Japan, in 1994, and received the Dr.Eng. degree from Tohoku University, Sendai, Japan in 2002. In 1994, he joined NTT Mobile Communications Network, Inc. (now, NTT DOCOMO, INC.). While with NTT DOCOMO, INC., he was engaged in the research and standardization of wireless access technologies for wideband DS-CDMA mobile radio, HSPA, LTE, and broadband wireless packet access technologies for systems beyond IMT-2000. In 2007, he joined the faculty of the Tokyo University of Science and currently holds the position of Professor. His current research interests are in the areas of wireless technologies and mobile communication systems, including advanced multiple access, radio resource allocation, inter-cell interference coordination, multiple-antenna transmission techniques, signal processing such as interference cancellation and turbo equalization, and issues related to heterogeneous networks using small cells. He was a co-recipient of the Best Paper Award of the International Symposium on Wireless Personal Multimedia Communications in 2004 and 2007, the Best Paper Award from the IEICE in 2021, a recipient of the Young Researcher's Award from the IEICE in 2003, the 5th YRP Award in 2007, the Prime Minister Invention Prize in 2010, and the Invention Prize of Commissioner of the Japan Patent Office in 2015. He is a member of the IEEE.