Resource-Hopping-Based Grant-Free Multiple Access for 6G-Enabled Massive IoT Networks

Han Seung Jang, *Member, IEEE*, Bang Chul Jung, *Senior Member, IEEE*, Tony Q. S. Quek, *Fellow, IEEE*, and Dan Keun Sung, *Life Fellow, IEEE*

***Abstract*—Grant-free multiple access (GFMA) is an emerging technology to accommodate a massive number of devices for 6G-enabled Internet of Things (IoT) networks. The main advantages of GFMA are to efficiently reduce control signaling overhead for resource scheduling while improving resource efficiency. In this article, we propose a *novel* resource-hopping-based GFMA (RH-GFMA) framework with resource hopping schemes for providing massive connectivity in 6G cellular IoT networks, where each IoT device is allowed to access physical radio resources by using a preassigned resource hopping pattern without not only resource request but also grant procedure, which is the so-called "one-shot" noninteractive multiple access. We exploit three types of resource hopping schemes in the proposed RH-GFMA framework: 1) random hopping; 2) resource group hopping; and 3) Latin-square group hopping. We mathematically analyze the RH-GFMA system performance in terms of the hopping pattern collision probability, maximum allowable packet delay, and interference-over-thermal. Finally, we derive an accommodation capacity of the proposed RH-GFMA framework, which is defined as the expected number of IoT devices accommodated in a cell under a maximum allowable packet-delay requirement and an interference-over-thermal constraint. With the proposed GFMA, massive IoT devices are expected to be efficiently accommodated in 6G wireless networks, while satisfying strict latency and reliability requirements.**

***Index Terms*—Cellular uplink, grant-free multiple access (GFMA), interference-over-thermal, Internet of Things (IoT), low-density parity-check (LDPC) codes, massive connectivity, packet delay.**

I. Introduction

MASSIVE Internet-of-Thing (mIoT) applications, including virtual/augmented reality, autonomous vehicles, wearable sensors, smart home appliances, smart manufacturing, etc., are expected to be one of main services for the sixth (6G) wireless networks [1]. One expects that 6G wireless networks will provide ultimate experience to intelligently combine people, process, data, and things [2]. For instance, Cisco research report reveals that the number of connected devices will reach 500 billion by 2030 [3]. As a representative service scenario of 5G, massive machine-type communications (mMTCs) was considered, which supports a massive number of IoT devices that are sporadically active and send small data payloads. With emerging industrial use cases, the mMTC scenarios will be evolved into several specialized scenarios and 6G will need to serve highly diverse applications for various types of IoT devices [4]. In particular, *critical* mMTC refers to supporting massive connectivity with high reliability and low latency, e.g., critical medical monitoring and factory automation [5]. For example, industrial automation is an example, where a lot of sensors are communicating and generating a huge amount of data, while requiring high reliability and ultralow latency. Industrial IoT automation/orchestration applications require the reliability of 99.9%–99.99999% and latency of 10–50 ms, while autonomous cars require reliability of 99.9%–99.99999% and latency of 2 ms [5]. In the fifth-generation (5G) communication systems, it is difficult to satisfy scalability, reliability, and latency requirements simultaneously.

In such an (critical) mIoT era, multiple access technology will still play an important role by reinventing the conventional multiple access philosophy that was originally designed for human-centric wireless networks [6]. Recently, grant-free multiple access (GFMA) techniques have received much attention from both industry and academia to effectively accommodate a large number of bursty devices transmitting short packets [4], [7]–[11]. The basic principle of GFMA is to allow each device to communicate with the base station (BS) randomly and, thus, multiple devices may share the same physical radio resources (time and frequency) [4]. In other words, once a data packet is generated at a certain device, it is transmitted immediately by using randomly selected or preconfigured radio resources without waiting for the BS's scheduling grant. Hence, GFMA techniques can reduce signaling overhead and latency, and have overloading capability to accommodate massive devices with limited resources. However, the GFMA may induce resource collision issues among active devices and reliability of packet transmission may be degraded. Nonorthogonal multiple access

Manuscript received July 31, 2020; revised November 18, 2020 and January 9, 2021; accepted February 24, 2021. Date of publication March 9, 2021; date of current version October 7, 2021. This work was supported in part by the NRF grant funded by the Korea Government Ministry of Science and ICT under Grant 2019R1F1A1061023; in part by the NRF through the Basic Science Research Program funded by the Ministry of Science and ICT under Grant NRF-2019R1A2B5B01070697; and in part by the MOE ARF Tier 2 under Grant T2EP20120-0006. *(Corresponding author: Bang Chul Jung.)*

Han Seung Jang is with the School of Electrical, Electronic Communication, and Computer Engineering, Chonnam National University, Yeosu 59626, Republic of Korea (e-mail: hsjang@chonnam.ac.kr).

Bang Chul Jung is with the Department of Electronics Engineering, Chungnam National University, Daejeon 34134, Republic of Korea (e-mail: bcjung@cnu.ac.kr).

Tony Q. S. Quek is with the Information Systems Technology and Design Pillar, Singapore University of Technology and Design, Singapore 487372 (e-mail: tonyquek@sutd.edu.sg).

Dan Keun Sung is with the School of Electrical Engineering, KAIST, Daejeon 34141, Republic of Korea (e-mail: dksung@kaist.ac.kr).

Digital Object Identifier 10.1109/JIOT.2021.3064872

(NOMA) techniques have been considered to solve such resource collision issues [9]–[14].

In general, massive devices tend to transmit sporadic or periodic small data packets to BS [15]. Conventional grant-based multiple access (GMA) schemes in which each device sends a scheduling request and waits for a scheduling grant for sending data may not efficiently support massive connections in 5G networks [16]. The significant gap between available radio resources and a massive number of connections motivates us to study GFMA with signature-based NOMA techniques for massive IoT networks [17]. There are two types of resource management schemes in GFMA techniques: 1) random selection type and 2) preconfigured/predetermined allocation type [18]. In the random selection-based GFMA, each device sends data over a randomly selected preamble and time/frequency resources among a certain resource pool, while resource collisions may occur when two or more devices select the same pilot sequence in the random selection-based GFMA [19]. A nonorthogonal pilot design for joint collision detection and channel estimation was proposed for the random selection type of GFMA [20]. On the other hand, in the preconfigured type of GFMA, the BS (semi)persistently configures radio resources, such as preamble and time/frequency resources for each device [21], where there exist no collisions in pilot or preamble transmissions. However, the time/frequency resource may experience collisions as the number of devices accommodated increases. Hence, one of the most challenging issues in GFMA techniques is to identify or detect device activity and to estimate channel state information of active devices.

There have been many studies on device activity detection and channel estimation techniques especially based on the sparse signal processing or compressed sensing framework [22]–[26]. In [24], a compressive random access (RA) framework for active user detection was investigated, where a new channel coding scheme for sparse identification sequences and a low-complexity maximum likelihood decoding techniques were proposed. In [26], a joint user activity detection and channel estimation scheme based on a block sparse signal recovery problem was proposed. The sparse activity characteristics of massive IoT devices makes the compressed sensing approach as a promising solution to user detection problems [23], [24], [27]. Senel and Larsson [27] proposed a noncoherent transmission scheme for GFMA techniques, which can detect device activity and decode embedded information bits.

In GMA schemes, each time/frequency resource block is exclusively assigned to a single device to eliminate interuser interference in each cell, while GFMA may cause the interuser or multiple access interference among UEs due to radio resource collisions, which results in a performance degradation in packet decoding in general. The effect of resource collisions can be alleviated through radio resource hopping and channel coding techniques, such as turbo codes and low-density parity-check (LDPC) codes [28]–[30]. In order to reduce the effect of interuser interference, a tandem spreading mechanism was recently proposed for GFMA schemes [31], [32].

In this article, we propose a novel resource-hopping-based GFMA (RH-GFMA) framework with radio resource hopping to efficiently accommodate massive devices while satisfying latency and reliability requirements, resulting in one-shot uplink access for massive 6G-enabled cellular IoT networks. In the proposed RH-GFMA, each IoT device sends its preassigned preamble to notify its activity and then the BS can obtain its channel state information, while it sends data in the same time-slot with the preassigned resource hopping pattern related to its preamble. We consider three different types of resource hopping schemes in the proposed RH-GFMA framework: 1) random hopping; 2) resource group hopping; and 3) Latin-square group hopping. We mathematically analyze the hopping pattern collision probability, and the accommodation capacity of the RH-GFMA framework under a given maximum packet-delay requirement and an interference-over-thermal constraint. To the best of our knowledge, interference-over-thermal is first considered for designing the GFMA technique even though it is closely related to energy efficiency of UEs or system operation efficiency in uplink cellular networks. We adopt LDPC codes for reducing the effect of the resource collisions among UEs by using a joint log-likelihood ratio (LLR) calculation method.

The remainder of this article is organized as follows. In Section II, we describe a system model considered in this article. In Section III, we propose an RH-GMFA framework. In particular, we describe the overall procedure of the RH-GFMA framework with three different types of resource hopping schemes, and analyze their hopping pattern collision probability. We characterize the accommodation capacity of the proposed RH-GFMA framework under a given maximum allowable packet-delay requirement and an interference-over-thermal constraint in Section IV. We evaluate the performance of the proposed RH-GFMA framework through extensive computer simulations in Section V. Finally, we draw conclusions in Section VI.

II. System Model

In this article, we focus on the communication scenario, where massive IoT devices send their packets to the BS intermittently, where the packets have certain delay requirements as we will explain later. Let us first define a frequency–time resource block for RH-GFMA as a GFMA block. We assume that there exist $R \times 1$ resources for preamble part and $R \times R$ resources for data part within the GFMA block. Let T_{block} denote the single GFMA block duration time. We constrict that the maximum number of devices in a single GFMA block is equal to $M = R \cdot L$, where $L \geq 1$ denotes the loading factor, and it is assumed to be an integer number. In other words, the loading factor L represents the ratio of the number of devices in a single GFMA group M to the number of available resources R. For example, $L = 1$ means that the same number of devices as R belong to a single GFMA group. $L = 2$ means that compared to R, twice as many devices belong to a single GFMA group. When U IoT devices potentially attempt GFMA in a cell, there exist $I = \lceil (U/R \cdot L) \rceil$ GFMA groups,

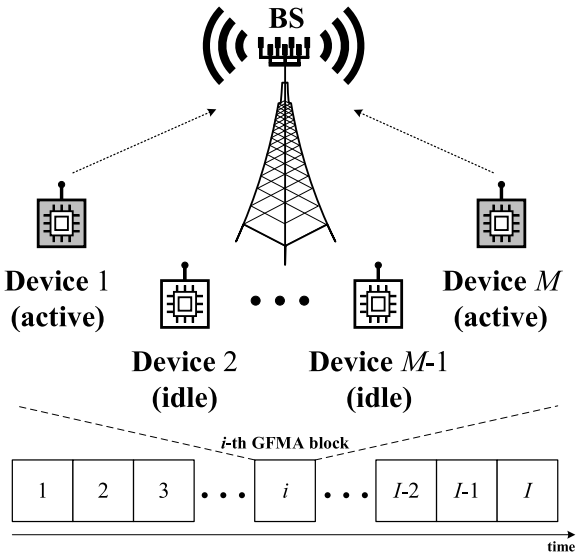


Fig. 1. System model based on RH-GFMA.

and each device belonging to the ith GFMA group[1] can periodically have an opportunity to transmit data on the ith GFMA opportunity every $I \cdot T_{\text{block}}$ time for $i \in \{1, 2, \ldots, I\}$. In what follows, without loss of generality, we describe the proposed RH-GFMA scheme in a single GFMA block.

Fig. 1 shows the system model of the proposed RH-GFMA, which considers a single cell network consisting of a BS with J antennas and shows only M devices belonging to the ith GFMA group. Each device in the ith GFMA group is active if it has at least one packet to transmit; otherwise, it is idle. To express this in a general way, we assume that the activation of a device belonging to the ith GFMA group is determined by the activation probability ν ($0 \leq \nu \leq 1$) on the ith GFMA block. Therefore, depending on ν, the number of active devices varies on each GFMA block. In this article, we assume that the activation probability is the same for all devices utilizing the RH-GFMA. Both of the activation probability ν and the loading factor L affect the number of *active* devices approximately as $\nu \times R \times L$ in a single GFMA group. Hence, depending on the activation probability, the optimal L should be determined.

For multiple access signature allocation, there are two options: one option is that a device selects a random signature and the other option is that each device has a preconfigured or predetermined signature [33]. In the following, we assume that the each device has a preconfigured or predetermined signature. Thus, there are no signature collisions.

The GFMA block mainly consists of *preamble part* and *data resource part*. Each of the active devices, which want to transmit uplink data, sends a predetermined preamble signal as its signature on the preamble part in order for the BS to identify which devices attempted the GFMA, compared to the random preamble selection [34]. We can utilize the Zadoff–Chu (ZC) sequences as preamble signal [35], [36] since the BS can easily detect ZC sequence-based preambles by calculating the correlation values and investigating whether the correlation values exceed a detection threshold on each preamble detection zone. Moreover, based on the preamble detection, the corresponding channels can be subsequently estimated [37]. After identifying the active devices, the BS obtains an active UE list of N devices on a specific GFMA block. On the data resource part, active devices transmit uplink data symbols according to the resource hopping patterns of devices. Here, the resource hopping patterns are one-to-one mapped to predefined preamble indices so that the BS can decode data for the corresponding device.

[1]If the proposed RH-GFMA framework supports total I GFMA groups, for example, each device can determine its GFMA group by dividing the international mobile subscriber identity (IMSI) by I. In other words, the remainder of this division represents the device's GFMA group index.

III. Resource-Hopping-Based Grant-Free Multiple Access

A. Overall Procedure

For massive IoT networks, semipersistent resource allocation for multiple access can significantly reduce signaling and control overhead associated with massive resource requests from a massive number of devices and massive resource grants from the BS. Fig. 2 shows the overall procedure of RH-GFMA, in which for example, devices 2 and 9 in the same GFMA group are active for packet transmissions, and they attempt the GFMA on their dedicated group GFMA block.

1) *Block Transmission and Reception:* Each of devices that want to transmit uplink data first selects its own exclusive preamble index based on its IoT device identifier (device-ID). Then, the device obtains its own resource hopping pattern, which is one-to-one mapped by the preamble index. On its dedicated group GFMA block, it transmits its preamble on the preamble part of the GFMA block, and data symbols on the data part of the GFMA block according to its resource hopping pattern. Afterward, the BS receives multiple preambles on the preamble part and data symbols on the data part.
2) *Device-Activity Detection and Data Decoding:* Upon reception of preambles, the BS verifies all of preamble detection zones whether each device is active or idle. For each of active devices, the BS proceeds data decoding according to its resource hopping pattern.

B. Resource Hopping Methods

In the proposed RH-GFMA framework, each device can transmit data symbols based on its own resource hopping pattern and, thus, the BS does not perform a resource scheduling and allocation procedure to give a resource grant for each device, which is very effective to reduce control and signaling overhead. However, if the number of active devices attempting the GFMA is significantly large, hopping pattern (HP) collisions may occur since the number of available orthogonal resources is limited in practice.

As stated in [38], for URLLC traffic with a Poisson type of sporadic features, it is difficult to avoid collisions in multiple access physical resources in grant-free transmission. Thus, for UL grant-free, NOMA is a good solution to solve the resource collisions while maintaining the reliability at the required level. However, there may exist various resource

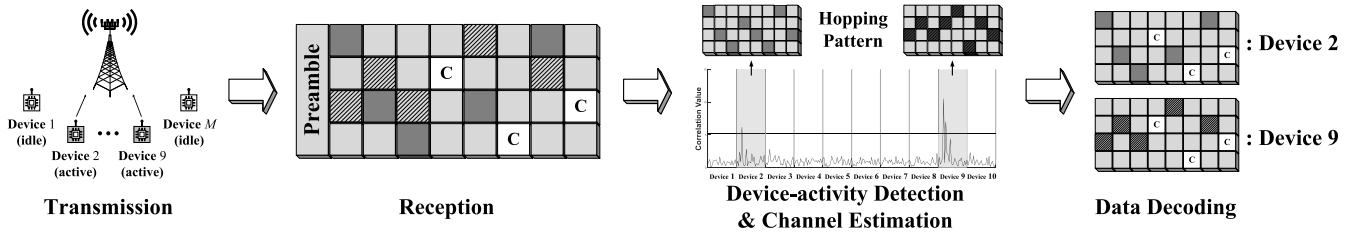


Fig. 2. Procedure of RH-GFMA.

hopping schemes for multiple devices to utilize the shared resources in a nonorthogonal manner. In this section, we deal with three different types of resource hopping schemes: 1) Latin-square group hopping; 2) random resource hopping; and 3) resource group hopping. We assume that the number of devices in a single GFMA group M, the activation probability ν, and the number of available resources R are the same in all three resource hopping schemes. Compared to a typical NOMA technique [9]–[14], our proposed resource hopping schemes can be considered as partially (generalized) NOMA schemes since not all resources are shared with other interfering devices, i.e., the target device utilizes some of resources by orthogonal multiple access (OMA) and the remaining resource by NOMA.

1) Latin Square Group Hopping: When $M \leq R$, the loading factor L is equal to 1, and the BS provides each device with an exclusive resource hopping pattern. Thus, there is no resource collision in this situation. On the other hand, when $M > R$ or equivalently $L > 1$, there can exist $\lceil M/R \rceil = L$ Latin square groups (LSGs) [or collision-free (CF) groups] with a size of R based on the Latin square matrix [39], [40]. For $l = 1, \ldots, R-1$, $R \times R$ Latin square matrix Q^l with the (i, j)th entry is defined as

$$q^l_{i,j} = (l \times i + j) \bmod R \tag{1}$$

where mod represents a modulo operation. The lth LSG has the lth Latin-square matrix Q^l. One virtual channel in the lth LSG experiences only one resource collision with each virtual channel in the different kth LSG based on the orthogonal property of Latin squares. For example, Q^1 Latin square matrix with $R = 5$ is given by

$$Q^1 = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 0 \\ 2 & 3 & 4 & 0 & 1 \\ 3 & 4 & 0 & 1 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix}$$

where there are five exclusive hopping patterns from 0 to 4. In the same LSG, each device utilizes an exclusive hopping pattern and, thus, no resource collisions occur in the same LSG. However, a device may experience a resource collision by up to $(L-1)$ interfering devices from different LSGs when there are L LSGs in total. Thus, the probability that a single resource is collided by $K(\leq L-1)$ interfering devices for a given activation probability ν is given by

$$P_{\text{group}}(K, \nu) = \binom{L-1}{K} \nu^K (1-\nu)^{(L-1-K)}. \tag{2}$$

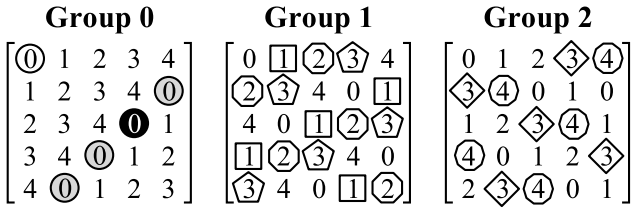


Fig. 3. Example of resource collisions in the LSG hopping when $M = 15$, $R = 5$, and $L = 3$. Device 0 is a target device in group 0. Devices 1, 2, and 3 are active in group 1, and devices 3 and 4 are active in group 2.

From (2), the probability that $(N_0, \ldots, N_{L-1})$ resources are collided by $0, \ldots, L-1$ interfering devices, respectively, is given by

$$\Pr(N_0, \ldots, N_{L-1} | \nu) = \binom{R}{N_0 \cdots N_{L-1}} \prod_{K=0}^{L-1} \{P_{\text{group}}(K, \nu)\}^{N_K} \tag{3}$$

where $N_0 + \cdots + N_{L-1} = R$. The expected number of resources collided by K interfering devices for a given activation probability ν is also given by

$$\mathbb{E}(N_K | \nu) = R \cdot P_{\text{group}}(K, \nu), \quad K = 0, \ldots, L-1. \tag{4}$$

Fig. 3 illustrates an example of resource collisions in the LSGs when $M = 15$, $R = 5$, and $L = 3$. The target device 0 in group 0 experiences three resource collisions in $\{(5, 2), (4, 3), (3, 4), (2, 5)\}$, and there are no collisions in $(1, 1)$. Specifically, resource $(5, 2)$, $(4, 3)$, and $(2, 5)$ are collided by device 3 in group 2, device 3 in group 1, and device 1 in group 1, respectively. Resource $(3, 4)$ is collided by two devices: 1) device 2 in group 1 and 2) device 4 in group 2. The HP collisions may cause bit errors at the receiver. However, we can handle these errors using a strong channel coding method such as LDPC coding utilized for the 5G new radio (NR) standard [41].

2) Random Resource Hopping: The random resource hopping patterns may be based on an ID, such as the electronic serial number [42]. The probability that a target device experiences a resource collision with $K \geq 1$ active devices for a given activation probability ν and the number of available orthogonal resources R on a time slot is expressed as

$$P_{\text{random}}(K, \nu) = \binom{M-1}{K} \left(\frac{\nu}{R}\right)^K \left(1 - \frac{\nu}{R}\right)^{M-1-K} \tag{5}$$

which details are given in (5), shown at the bottom of the next page, where we assume that a target device is active, and it

utilizes a specific orthogonal resource on a given time slot. Thus, we only consider the other $(M-1)$ devices and their activation by the binomial distribution $\binom{M-1}{m}\nu^m(1-\nu)^{M-1-m}$, which is the probability that m devices are active, while $(M-1-m)$ devices are idle. In addition, $\binom{m}{K}(1/R)^K(1-[1/R])^{m-K}$ represents the probability that K active devices out of m active devices utilize the same resource as that of the target device and the remaining $(m-K)$ devices utilize different resources. From (5), the probability that $(N_0, \ldots, N_{M-1})$ resources are collided by $0, \ldots, M-1$ interfering devices, respectively, is given by

$$\Pr(N_0, \ldots, N_{M-1}|\nu) = \binom{R}{N_0 \cdots N_{M-1}} \times \prod_{K=0}^{M-1} \{P_{\mathsf{random}}(K, \nu)\}^{N_K} \quad (6)$$

where $N_0 + \cdots + N_{M-1} = R$. The expected number of resources collided by K interfering devices for a given activation probability ν is also given by

$$\mathbb{E}(N_K|\nu) = R \cdot P_{\mathsf{random}}(K, \nu), \quad K = 0, \ldots, M-1. \quad (7)$$

3) Resource Group Hopping: In a similar manner to LSG hopping, when $M > R$ or equivalently $L > 1$, there can exist $\lceil M/R \rceil = L$ resource groups with a size of R. In the same resource group, each device utilizes an exclusive hopping pattern, and, thus, no resource collisions occur. However, since each group utilizes the identical resource hopping pattern matrix, the probability that a single device attempts a GFMA with $K(\leq L-1)$ interfering devices from the different resource group for a given activation probability ν is the same as (2). It is worth noting that a whole packet with R resources is collided by K interfering devices from different groups with a probability of $P_{\mathsf{group}}(K, \nu)$.

For example, Fig. 4 illustrates an example of resource collisions in the resource group hopping scheme when $M = 15$, $R = 5$, and $L = 3$. As shown in this figure, each resource group has the identical resource hopping matrix. Since devices 0 in each group are all active, each of devices 0 in groups 0, 1, and 2 experiences collisions with two interfering devices during whole data transmission. Device 1 in group 0 is only active and, thus, it does not experience any collisions. Next, since devices 2 in group 1 and 2 are active, each of them experiences collisions with one interfering device during whole data transmission.

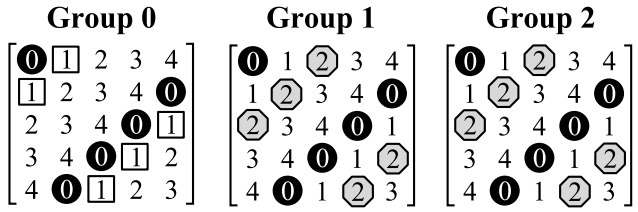


Fig. 4. Example of resource collisions in the resource group hopping scheme when $M = 15$, $R = 5$, and $L = 3$. Devices 0 and 1 are active in group 0. Devices 0 and 2 are active in group 1, and devices 0 and 2 are active in group 2.

C. LLR Computation for Symbol Decoding

Based on the preamble detection, the BS can expect which devices are active for the GFMA and also which resources experience collisions from the resource hopping patterns. Thus, we require the LLR computation for the CF symbol and collided symbol, respectively.

The received signal vector of the mth device with J antennas is expressed as

$$\mathbf{y}_m = \begin{bmatrix} y_{m,1} \\ y_{m,2} \\ \vdots \\ y_{m,J} \end{bmatrix} = \begin{cases} \mathbf{h}_m\sqrt{E_m}s_m + \mathbf{n}, & \text{CF} \\ \mathbf{h}_m\sqrt{E_m}s_m + \mathbf{I}_K + \mathbf{n}, & \text{Collision} \end{cases} \quad (8)$$

where $\mathbf{h}_m$, E_m, and s_m denote the channel coefficient vector, the transmit energy, and the symbol of the mth device, respectively

$$\mathbf{I}_K = \sum_{k=1}^{K} \mathbf{h}_k\sqrt{E_k}s_k$$

represents the received signal vector from K interfering devices, where $\mathbf{h}_k$, E_k, and s_k denote the channel coefficient vector, transmit energy, and the symbol of the kth interfering device experiencing the same resource collision with the mth device, and $\mathbf{n}$ represents the thermal noise vector, which is a collection of J independent identically distributed complex Gaussian random variables with zero-mean and variance of $2\sigma^2$.

In this article, we utilize a joint maximum-likelihood (ML) decoding strategy. When a resource collision does not occur for the symbol of the mth device, the LLR for the ith bit $s_{m,i}$

$$\begin{aligned} P_{\mathsf{random}}(K, \nu) &= \sum_{m=K}^{M-1} \binom{M-1}{m} \nu^m (1-\nu)^{M-1-m} \binom{m}{K} \left(\frac{1}{R}\right)^K \left(1-\frac{1}{R}\right)^{m-K} \\ &= \left(\frac{\nu}{R}\right)^K \sum_{m=K}^{M-1} \binom{M-1}{m} \binom{m}{K} \left(\nu - \frac{\nu}{R}\right)^{m-K} (1-\nu)^{M-1-m} \\ &= \binom{M-1}{K} \left(\frac{\nu}{R}\right)^K \sum_{m=K}^{M-1} \binom{M-1-K}{m-K} \left(\nu - \frac{\nu}{R}\right)^{m-K} (1-\nu)^{M-1-m} \\ &= \binom{M-1}{K} \left(\frac{\nu}{R}\right)^K \left(1-\frac{\nu}{R}\right)^{M-1-K} \end{aligned} \quad (5)$$

is calculated as

$$\Lambda(s_{m,i}) = \ln \frac{\sum_{s\in\mathbf{S}_{i0}} \exp\left\{-\frac{\left\|\mathbf{y}_m - s\mathbf{h}_m\sqrt{E_m}\right\|^2}{2\sigma^2}\right\}}{\sum_{s\in\mathbf{S}_{i1}} \exp\left\{-\frac{\left\|\mathbf{y}_m - s\mathbf{h}_m\sqrt{E_m}\right\|^2}{2\sigma^2}\right\}} \tag{9}$$

where $\mathbf{S}_{i0}$ and $\mathbf{S}_{i1}$ represent the ideal constellation points with bits 0 and 1, respectively, in the ith bit. Equation (9) is shown at the bottom of the page in full detail. For example, $\mathbf{S}_{10} = \{(1/\sqrt{2} + j1/\sqrt{2}), (-1/\sqrt{2} + j1/\sqrt{2})\}$ and $\mathbf{S}_{11} = \{(1/\sqrt{2} - j1/\sqrt{2}), (-1/\sqrt{2} - j1/\sqrt{2})\}$ for the first bit of 4-QAM modulation. On the other hand, when a resource collision occurs for $s_{m,i}$ with K interfering devices, the LLR is expressed as (9), where c_k represents the constellation point of the kth interfering device. In order to reduce calculation complexity, if the number of interfering devices to the target device is less than a limit, i.e., $K < \bar{K}$, LLR is computed according to (9), otherwise LLR value is set to 0.

IV. Connection Capacity of the Proposed GFMA

A. Packet-Delay Constraint

When U devices utilizing GFMA exist in a cell, and the number of available resources and the loading factor are R and L, respectively, in a single GFMA opportunity, there exist $I = \lceil U/R \cdot L \rceil$ GFMA groups. It implies that if the number of potential devices U increases, the number of GFMA groups I and the GFMA period $T_{\mathsf{period}} = I \cdot T_{\mathsf{block}}$ increase. T_{period} also represents as *the maximum packet delay* for a device to transmit data using GFMA. In theory, we can set a required packet delay for Quality of Service (QoS), and this requirement may determine the total number of devices accommodated in GFMA system, which we call *"GFMA accommodation capacity"*.

It is worth noting that a larger loading factor L can accommodate a larger number of devices in a single GFMA group and reduce the packet delay. However, a larger loading factor L causes devices to increase their transmission power for satisfying a block error rate (BLER) requirement. Even though the devices can increase their transmission power, the aggregated received power on a resource block affects the interference-over-thermal, which will be described in the next section. Hence, we carefully choose the loading factor L in order to accommodate devices up to the maximum by considering a tradeoff relationship between the maximum allowable packet delay and the interference-over-thermal.

B. Interference-Over-Thermal Constraint

The interference-over-thermal parameter is an uplink design constraint that we need to take into account when designing an uplink system [43]. To ensure that a cellular system operates at a reasonable uplink load, the IMT-2020 requires that uplink interference-over-thermal is kept below 10 dB [44]. In general, the interference-over-thermal parameter is defined as

$$\mathbb{I} = \frac{P_{\mathsf{interference}} + P_{\mathsf{noise}}}{P_{\mathsf{noise}}} = 1 + \frac{P_{\mathsf{interference}}}{P_{\mathsf{noise}}} \tag{10}$$

where $P_{\mathsf{interference}}$ and P_{noise} denote the average interference power and noise power at the receiver. When we assume that the own cell and outer cells operate in the same way, the average interference power is approximately calculated as

$$P_{\mathsf{interference}} = \delta \cdot P_{\mathsf{own}} = \delta \cdot \left(\nu \cdot L \cdot P_{\mathsf{required}}\right) \tag{11}$$

where P_{own} and δ denote the received power of a device in its own cell and the interference ratio of outer cells to the own cell [45], respectively, and specifically, P_{own} consists of ν, L, and P_{required} representing the activation probability, the loading factor, and the required received power for a target BLER on a resource block, respectively. The required received power for the target BER P_{required} is obtained from the required energy per bit to interference plus noise power spectral density ratio γ, which is given by

$$\gamma = \frac{G \cdot P_{\mathsf{required}}}{P_{\mathsf{interference}} + P_{\mathsf{noise}}} = \frac{G \cdot P_{\mathsf{required}}}{\delta \cdot \nu \cdot L \cdot P_{\mathsf{required}} + P_{\mathsf{noise}}} \tag{12}$$

where G represents the effective gain, which is the multiplication of the coding gain and the number of bits per symbol. By rearranging (12) with respect to P_{required}, we have

$$P_{\mathsf{required}} = \frac{\gamma \cdot P_{\mathsf{noise}}}{G - \delta \cdot \nu \cdot L \cdot \gamma}. \tag{13}$$

Finally, substituting (13) into (10) yields

$$\mathbb{I} = 1 + \frac{\delta \cdot \nu \cdot L \cdot \gamma}{G - \delta \cdot \nu \cdot L \cdot \gamma}. \tag{14}$$

The main causes to increase the interference-over-thermal are the loading factor L and the activation probability ν, which is determined by the packet arrival rate λ and the total number of devices U. We will carefully design the proposed RH-GFMA and calculate the number of devices accommodated (GFMA accommodation capacity) under a constraint that the interference-over-thermal is less than 10 dB in the next section.

V. Numerical Results

We compare three different types of hopping schemes in the proposed RH-GFMA framework, such as resource group hopping, random hopping, and LSG hopping and the existing RA-based GFMA scheme with NOMA capability, in terms

$$\Lambda\left(s_{m,i}\right) = \ln \frac{\sum_{s\in\mathbf{S}_{i0}} \sum_{c_1\in\{\mathbf{S}_{i0},\mathbf{S}_{i1}\}} \cdots \sum_{c_K\in\{\mathbf{S}_{i0},\mathbf{S}_{i1}\}} \exp\left(-\frac{\left\|\mathbf{y}_m - s\mathbf{h}_m\sqrt{E_m} - \sum_{k=1}^{K} c_k\mathbf{h}_k\sqrt{E_k}\right\|^2}{2\sigma^2}\right)}{\sum_{s\in\mathbf{S}_{i1}} \sum_{c_1\in\{\mathbf{S}_{i0},\mathbf{S}_{i1}\}} \cdots \sum_{c_K\in\{\mathbf{S}_{i0},\mathbf{S}_{i1}\}} \exp\left(-\frac{\left\|\mathbf{y}_m - s\mathbf{h}_m\sqrt{E_m} - \sum_{k=1}^{K} c_k\mathbf{h}_k\sqrt{E_k}\right\|^2}{2\sigma^2}\right)} \tag{9}$$

TABLE I
SIMULATION PARAMETERS AND VALUES

Parameters	Values
The number of RBs in a GFMA block, $R \times R$	17×17
GFMA block time, T_{block}	18 msec.
GFMA block bandwidth, B_{block}	3.06 MHz
Total number of available preambles in a GFMA block	192
Interfering ratio, δ	0.3 and 0.6
Packet length	2448 bits
Maximum transmittable bytes	198 bytes
Channel coding scheme	LDPC
Coding rate	0.3235
Modulation scheme	QPSK
Number of receive antennas	2

of the block error rate (BLER) and the interference-over-thermal. The existing RA scheme with NOMA capability is based on contention-based RA, and each device in this scheme can transmit data at the third step of RA without resource scheduling [46]. Due to preamble contentions at the first step of RA, each device may transmit data on the same resources at the third step with varying number of interfering devices, and also it may not transmit data since it may not receive RA response (RAR) message including the available resource information when the number of detected preambles D exceeds R. Hence we need to consider two events: E_1) the target device receives an RAR message and E_2) the target device does not receive an RAR message with probabilities of $P[E_1]$ and $P[E_2]$, respectively. Then, the BLER of the existing scheme is calculated by

$$P[\text{Block Error}] = P[\text{Block Error}|E_1]P[E_1] + P[E_2]. \quad (15)$$

It is worth noting that a device may experience collisions during a whole packet transmission with $1 \leq K \leq \lceil M \times \nu \rceil - 1$ interfering devices since they utilize the identical preamble at the first step of RA, and then the identical resources at the third step of RA. With a low number of interfering devices (e.g., $K < 4$), the packet of the target device may be decoded due to NOMA capability while the packet decoding may fail.

Based on the uplink system requirements of packet delay and interference-over-thermal, we approximately calculate the GFMA accommodation capacity, that is, the number of devices accommodated in a single cell for the proposed RH-GFMA framework.

Table I lists simulation parameter values. We choose the number of resource blocks in a GFMA block as a prime number of 17 in the frequency domain and time domain, respectively, for Latin square matrix, i.e., $R \times R = 17 \times 17$. Since one resource block occupies 180 kHz and 1 ms, the GFMA block time is 18 ms including a preamble time duration of 1 ms, and the GFMA bandwidth is 3.06 MHz. We utilize an LDPC channel coding scheme in the proposed RH-GFMA framework. According to [41], we utilize the base graph 1 and set the LDPC lifting size to 36. From the LDPC lifting size, a parity check matrix is generated, and the corresponding packet length is set to 2448 bits. With a QPSK modulation scheme and a coding rate of 0.3235, a device can transmit up to 1584 bits (198 bytes) over 792 symbols in a GFMA block.

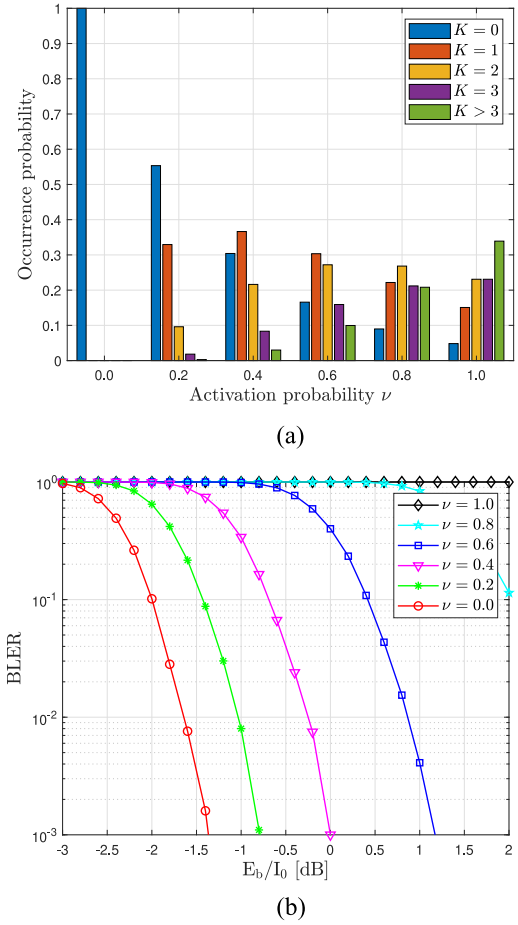


Fig. 5. (a) Occurrence probability of interfering devices and (b) BLER of the random hopping scheme when $L = 3$.

Fig. 5(a) shows the occurrence probabilities of K interfering devices on a single resource according to the activation probability. $K = 0$ represents that only a target device transmits its data without interfering devices. In other words, the number of competing devices (denoted as C) is counted as the number of interfering devices plus a single target device, i.e., $C = K + 1$. As the activation probability increases, the occurrence probability of a larger number of interfering devices increases. Moreover, there is a possibility that the number of competing devices exceeds the loading factor, i.e., $C > L$. Due to decoding complexity, we limit the decoding capability up to $C = 4$; in other words, the decoder computes LLR values up to four competing devices. If the number of competing devices exceeds 4, the decoder sets LLR values to 0.

Fig. 5(b) shows the BLER of the random hopping scheme for varying active probabilities when the loading factor L is equal to 3. As the activation probability increases, the BLER increases due to an increase in the number of competing devices on the same resource. Especially, the BLER significantly increases over an activation probability of 0.6 since the occurrence of $K > 3$ becomes severe.

Fig. 6 shows the occurrence probabilities of interfering devices on a single resource for varying the active probabilities

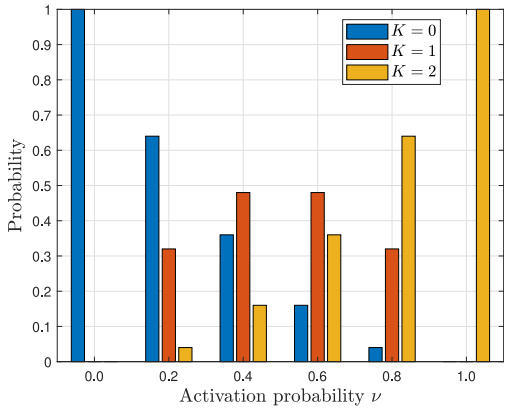


Fig. 6. Occurrence probability of interfering devices for both of resource group hopping and LSG hopping when $L = 3$.

for both of resource group hopping and LSG hopping when $L = 3$. $K = 2$ or equivalently $C = 3$ represents that a target device and two interfering devices (three competing devices) transmit their data on the identical resource. Compared to Fig. 5(a) for the random hopping, the number of competing devices for both resource group hopping and LSG hopping can be limited to the loading factor L.

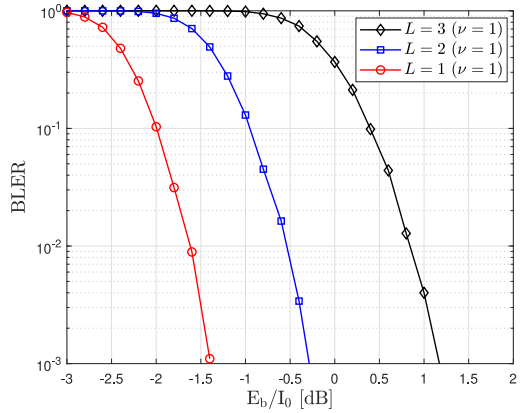
Fig. 7(a) shows the BLER of the resource group hopping scheme with $\nu = 1$ for varying the loading factor L. In particular, $L = 3$ with $\nu = 1$ represents that the whole GFMA block is shared with three competing devices. The performance degrades by 1 dB and 2.5 dB in $L = 2$ and $L = 3$, respectively, compared to $L = 1$.

Fig. 7(b) shows the BLER of the resource group hopping scheme for varying active probabilities when the loading factor L is equal to 3. In the resource group hopping scheme, if the interfering devices in the same resource group are active, a whole data packet is interfered by them. Hence, we can obtain the results by combining the results between the occurrence probabilities of interfering devices in Fig. 6 and the BLER of the resource group hopping scheme with $\nu = 1$ in Fig. 7(a). More specifically, we can calculate the BLER of the resource group hopping scheme as follows:

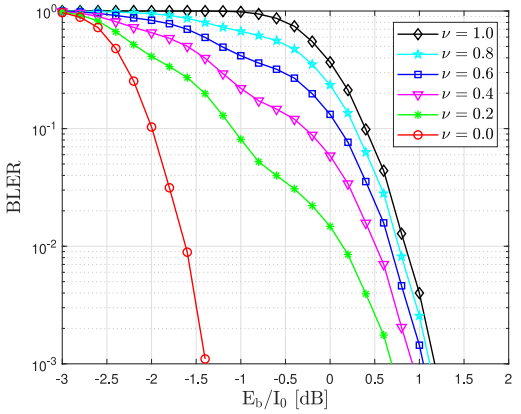
$$\mathsf{BLER}(L, \nu) = \sum_{k=0}^{L-1} \Theta(k, \nu) \times \mathsf{BLER}(k+1, \nu = 1) \quad (16)$$

where $\mathsf{BLER}(L, \nu)$ denotes the BLER of the loading factor L at an activation probability of ν, and $\Theta(k, \nu)$ denotes the occurrence probability of k interfering devices at an activation probability of ν.

Fig. 8 shows the BLER of the LSG hopping scheme for varying active probabilities when the loading factor L is equal to 3. Even though the occurrence probabilities of interfering devices are the same as those of the resource group hopping scheme, the number of interfering devices varies resource by resource due to a Latin square hopping pattern, compared to that of the resource group hopping scheme, in which the number of interfering devices is fixed during packet transmission. As a result, the LSG hopping scheme outperforms the resource group hopping scheme in therms of the BLER. For example, at an activation probability of $\nu = 0.8$, the required $\mathsf{E_b}/\mathsf{I_0}$ for



(a)



(b)

Fig. 7. (a) BLER of the resource group hopping scheme with $\nu = 1$ and (b) BLER of the resource group hopping scheme when $L = 3$.

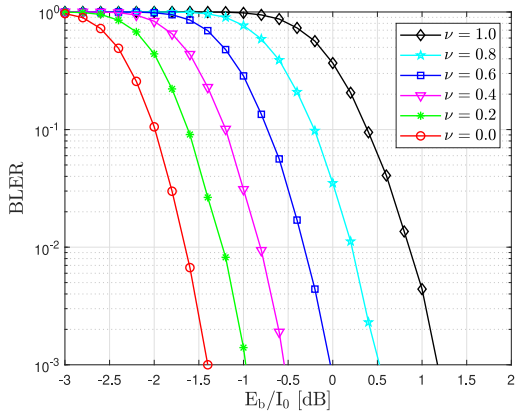


Fig. 8. BLER of the LSG hopping scheme when $L = 3$.

a target BLER of 10^{-2} is 0.2 dB in the LSG hopping scheme, while 0.75 dB in the resource group hopping scheme.

Fig. 9 compares the BLER of all schemes when the loading factor L is eqaul to 3. In Fig. 9(a), the activation probability is set to 0.2, representing relatively low activity of devices. In this environment, the BLER of the proposed LSG scheme shows 1.5, 0.5, and 0.2-dB gains compared to resource group, existing RA with NOMA, and random hopping schemes, respectively. On the other hand, in Fig. 9(b), the activation probability is set to 0.4, representing a relatively high activity of devices. In this environment, the BLER of the proposed

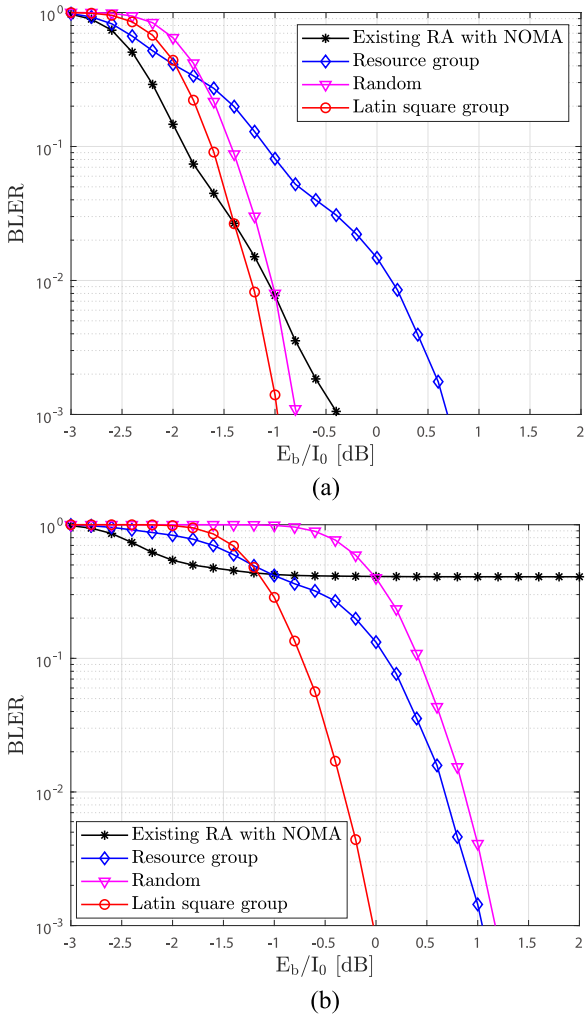


Fig. 9. Comparison of BLER of all schemes when $L = 3$. (a) $\nu = 0.2$. (b) $\nu = 0.6$.

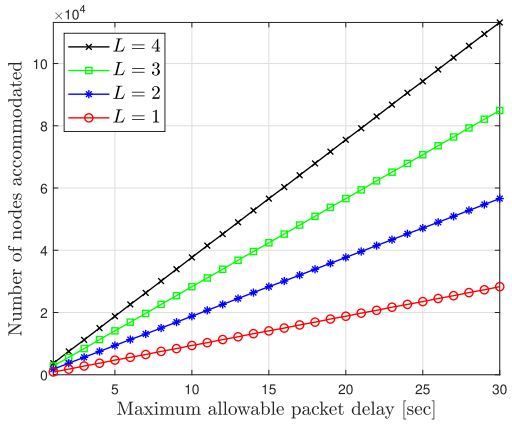


Fig. 10. Number of devices accommodated for varying the maximum allowable packet delay.

LSG scheme shows 1.2 and 1-dB gains compared to random hopping and resource group schemes, respectively. However, the existing RA scheme with NOMA shows the saturated BLER of 0.40 starting at -0.7 dB $\mathsf{E_b/I_0}$.

Fig. 10 shows the number of devices accommodated (GFMA accommodation capacity) for varying the maximum allowable packet-delay time, which is defined as the maximum allowable waiting time for a device to transmit a packet in RH-GFMA schemes. Basically, as the maximum allowable packet delay increases, the number of devices which can be accommodated also increases. In addition, a larger loading factor L causes to accommodate a larger number of devices. At a maximum allowable packet delay of 30 s, the RH-GFMA schemes can accommodate 113 200 devices with $L = 4$. However, it is not guaranteed that the RH-GFMA schemes can accommodate 113 200 devices for all active probabilities. We should note that any uplink systems should be operated under a target interference-over-thermal value, which is affected by the number of competing devices and their activation probability.

TABLE II
MAXIMUM SUPPORTABLE ACTIVATION PROBABILITY FOR $\delta = 0.6$ (DENSE NETWORK)

	Existing RA	RG	Random	LSG
$L = 4$	0.24	0.16	0.25	0.28
$L = 3$	0.30	0.27	0.34	0.37
$L = 2$	0.47	0.51	0.52	0.57
$L = 1$	1.00	1.00	1.00	1.00

Fig. 11 shows the interference-over-thermal for varying active probabilities. Here, we set the interference ratio to 0.3 and 0.6 for sparse and dense networks, respectively. When $L = 1$, each device can utilize an exclusive resource without interfering devices. Thus, we can observe that lower interference-over-thermal values even at an activation probability of 1. We also observe that the larger loading factor, the larger interference-over-thermal values. The LSG hopping scheme yields lower interference-over-thermal values compared to the other hopping schemes, while the resource group hopping scheme yields the highest interference-over-thermal values except $L = 2$. It is worth noting that the existing RA scheme with NOMA capability shows lower interference-over-thermal values compared to the resource group hopping scheme except $L = 2$. In this scheme, higher loading factor ($L > 2$) causes more devices not to transmit data at the third step of the RA procedure due to the lack of available resources to be scheduled by RAR messages. As suggested in [44], we set the target interference-over-thermal value to 10 dB. Especially, with $L = 4$ and $\delta = 0.6$, the interference-over-thermal value exceeds 10 dB at the active probabilities of 0.16, 0.24, 0.25, and 0.28 in the resource group, existing RA, random, and LSG hopping schemes, respectively.

Tables II and III summarize the maximum supportable active probabilities for each hopping scheme by considering both of the required BLER and the 10-dB interference-over-thermal constraints for a dense network with an interference ratio δ of 0.6 and a sparse network with an interference ratio δ of 0.3, respectively. Among all schemes, the LSG scheme shows the highest supportable active probabilities for all loading factors. In addition, the maximum supportable active probabilities of the existing RA scheme with NOMA capability are mostly affected by the required BLER and, thus, in both cases, it shows similar values for each loading factor.

Finally, Tables IV and V summarize the number of accommodating active devices for each hopping scheme by considering both of the maximum packet delay and the maximum

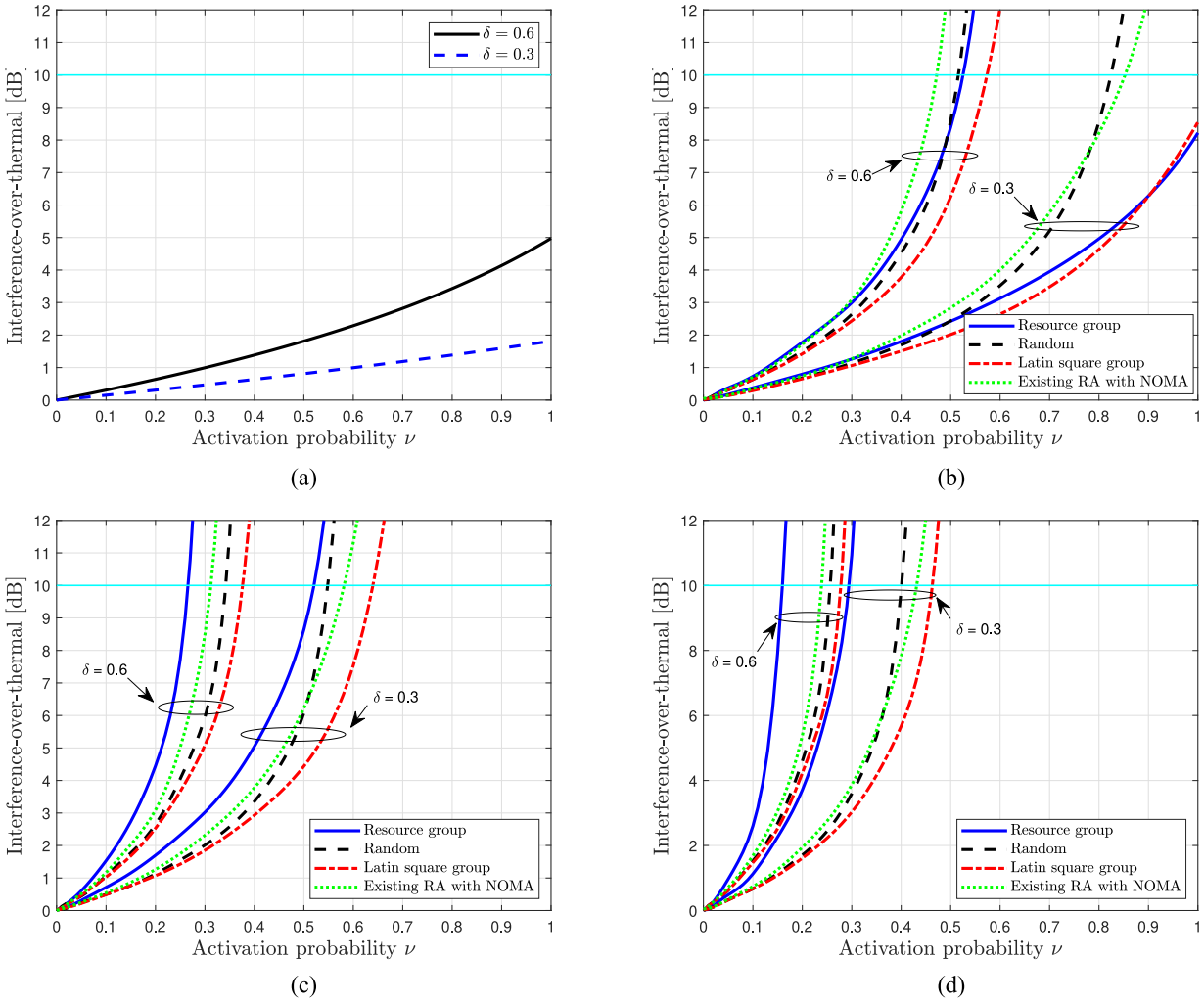


Fig. 11. Interference-over-thermal for varying loading factors. (a) $L = 1$. (b) $L = 2$. (c) $L = 3$. (d) $L = 4$.

TABLE III
MAXIMUM SUPPORTABLE ACTIVATION PROBABILITY FOR $\delta = 0.3$ (SPARSE NETWORK)

	Existing RA	RG	Random	LSG
$L = 4$	0.24	0.29	0.40	0.46
$L = 3$	0.32	0.52	0.54	0.64
$L = 2$	0.49	1.00	0.82	1.00
$L = 1$	1.00	1.00	1.00	1.00

supportable activation probability constraints for a dense network with an interference ratio δ of 0.6 and a sparse network with an interference ratio δ of 0.3, respectively. In both networks, all schemes in the proposed RH-GFMA framework accommodate the highest number of active devices with a loading factor of $L = 2$ under each packet-delay constraint. In particular, for the dense network ($\delta = 0.6$), the LSG, resource group, and random hopping schemes can accommodate 32 262, 29 432, and 28 866 devices, respectively, under the packet-delay constraint of 30 s. However, even though the existing RA scheme is equipped with NOMA capability, the result shows that the loading factor of $L = 1$, i.e., ($M = R$), is the best way to maximally accommodate devices in both of sparse and dense networks. In addition, the existing RA scheme with NOMA capability shows approximately half the number of accommodating active devices compared to that of the LSG scheme in the proposed SG-GFMA framework for the sparse network.

VI. CONCLUSION

In this article, we proposed an RH-GFMA framework for simultaneously satisfying connectivity, reliability, and latency requirements in 6G-enabled massive IoT networks. In particular, we considered three different types of resource hopping schemes for the proposed RH-GFMA framework and the existing RA scheme with NOMA capability, and compared their BLER performance as a reliability metric. Based on the BLER performance, we derived interference-over-thermal values for these hopping schemes and the existing scheme. As a result, we calculated the expected number of devices accommodated for the proposed RH-GFMA framework by considering the maximum-allowable packet delay and the interference-over-thermal with the required target BLER. Simulation results showed that the Latin-square group hopping scheme yields the best performance among the proposed RH-GFMA techniques over a large range of activation probability for a given number of IoT devices. Therefore, the proposed RH-GFMA can be adopted as a promising technique for critical mMTC service scenarios in 6G-enabled IoT networks.

TABLE IV
NUMBER OF ACCOMMODATING ACTIVE DEVICES FOR $\delta = 0.6$ (DENSE NETWORK)

L	Maximum allowable packet delay											
	5 sec.	15 sec.	30 sec.	5 sec.	15 sec.	30 sec.	5 sec.	15 sec.	30 sec.	5 sec.	15 sec.	30 sec.
	Resource group			Random			Latin square group			Existing RA with NOMA		
$L = 4$	3008	9056	18112	4700	14150	28300	5264	15848	31696	4512	13584	27168
$L = 3$	3807	11448	22923	4794	14416	28866	5217	15688	31413	4230	12720	25470
$L = 2$	**4888**	**14716**	**29432**	**4888**	**14716**	**28866**	**5358**	**16131**	**32262**	4418	13301	26602
$L = 1$	4700	14100	28300	4700	14100	28300	4700	14100	28300	**4700**	**14100**	**28300**

TABLE V
NUMBER OF ACCOMMODATING ACTIVE DEVICES FOR $\delta = 0.3$ (SPARSE NETWORK)

L	Maximum allowable packet delay											
	5 sec.	15 sec.	30 sec.	5 sec.	15 sec.	30 sec.	5 sec.	15 sec.	30 sec.	5 sec.	15 sec.	30 sec.
	Resource group			Random			Latin square group			Existing RA with NOMA		
$L = 4$	5452	16414	32828	7520	22640	45280	8648	26036	52072	4512	13584	27168
$L = 3$	7332	22048	44148	7614	22896	45845	9204	27136	54336	4512	13586	27168
$L = 2$	**9400**	**28300**	**56600**	**7708**	**23206**	**46412**	**9400**	**28300**	**56600**	4606	13867	27734
$L = 1$	4700	14100	28300	4700	14100	28300	4700	14100	28300	**4700**	**14100**	**28300**

REFERENCES

[1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May/Jun. 2020.

[2] "6G: The next hyper-connected experience for all," Bengaluru, Karnataka, Samsung Res., White paper, Jul. 2020.

[3] Cisco, *Cisco Edge-to-Enterprise IoT Analytics for Electric Utilities Solution Overview*. 2018. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/big-data/solution-overview-c22–740248.html

[4] N. H. Mahmood *et al.*, *White Paper on Critical and Massive Machine Type Communication Towards 6G*. Apr. 2020. [Online]. Available: https://arxiv.org/pdf/2007.08793.pdf.

[5] S. R. Pokhrel, J. Ding, J. Park, O.-S. Park, and J. Choi, "Towards enabling critical mMTC: A review of URLLC within mMTC," vol. 8, pp. 131796–131813, Jul. 2020.

[6] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy, and G. Caire, "Massive access for future wireless communication systems," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 148–156, Aug. 2020.

[7] S.-Y. Lien, S.-L. Shieh, Y. Huang, B. Su, Y.-L. Hsu, and H.-Y. Wei, "5G new radio: Waveform, frame structure, multiple access, and initial access," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 64–71, Jun. 2017.

[8] *Study on New Radio Access Technology Physical Layer Aspects*, 3GPP, Sophia Antipolis, France, 3GPP Rep. TR 38.802, Sep. 2017.

[9] Y. Chen *et al.*, "Toward the standardization of non-orthogonal multiple access for next generation wireless networks," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 19–27, Mar. 2018.

[10] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic, and E. de Carvalho, "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 88–99, Sep. 2018.

[11] R. Abbas, T. Huang, B. Shahab, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, *Grant-Free Non-Orthogonal Mulitiple Access: A Key Enabler for 6G-IoT*. Mar. 2020. [Online]. Available: https://arxiv.org/abs/2003.10257.

[12] M. T. P. Le, G. C. Ferrante, T. Q. S. Quek, and M. Di Benedetto, "Fundamental limits of low-density spreading NOMA with fading," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4648–4659, Jul. 2018.

[13] Z. Chen, Q. Yao, H. H. Yang, and T. Q. S. Quek, "Massive wireless random access with successive decoding: Delay analysis and optimization," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 457–471, Jan. 2019.

[14] H. S. Jang, H. Lee, and T. Q. S. Quek, "Deep learning-based power control for non-orthogonal random access," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 2004–2007, Nov. 2019.

[15] A. Azari, P. Popovski, G. Miao, and C. Stefanovic, "Grant-free radio access for short-packet communications over 5G networks," in *Proc. IEEE GLOBECOM*, Dec. 2017, pp. 1–7.

[16] Y. Cai, Z. Qin, F. Cui, G. Y. Li, and J. A. McCann, "Modulation and multiple access for 5G networks," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 629–646, 1st Quart., 2018.

[17] M. Mohammadkarimi, M. A. Raza, and O. A. Dobre, "Signature-based nonorthogonal massive multiple access for future wireless networks: Uplink massive connectivity for machine-type communications," *IEEE Veh. Technol. Mag.*, vol. 13, no. 4, pp. 40–50, Dec. 2018.

[18] Z. Yuan, Y. Hu, W. Li, and J. Dai, "Blind multi-user detection for autonomous grant-free high-overloading multiple-access without reference signal," in *Proc. IEEE 87th VTC Spring*, Jun. 2018, pp. 1–7.

[19] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "A novel analytical framework for massive grant-free NOMA," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2436–2449, Mar. 2019.

[20] A. Quayum, H. Minn, and Y. Kakishima, "Non-orthogonal pilot designs for joint channel estimation and collision detection in grant-free access systems," *IEEE Access*, vol. 6, pp. 55186–55201, 2018.

[21] "Semi-persistent scheduling for 5G new radio URLLC," 3GPP, Sophia Antipolis, France, Nokia, Espoo, Finland, Alcatel-Lucent Shanghai Bell, Shanghai, China, Rep. R1-1609748, Oct. 2016.

[22] B. Wang, L. Dai, Y. Zhang, T. Mir, and J. Li, "Dynamic compressive sensing-based multi-user detection for uplink grant-free NOMA," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2320–2323, Nov. 2016.

[23] Y. Du *et al.*, "Efficient multi-user detection for uplink grant-free NOMA: Prior-information aided adaptive compressive sensing perspective," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2812–2828, Dec. 2017.

[24] J. Choi, "Compressive random access with coded sparse identification vectors for MTC," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 819–829, Feb. 2018.

[25] Y. Du *et al.*, "Block-compressed-sensing-based multiuser detection for uplink grant-free NOMA systems," in *Proc. IEEE ICC*, May 2018, pp. 1–7.

[26] Y. Zhang, Q. Guo, Z. Wang, J. Xi, and N. Wu, "Block sparse Bayesian learning based joint user activity detection and channel estimation for grant-free NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9631–9640, Oct. 2018.

[27] K. Senel and E. G. Larsson, "Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6164–6175, Dec. 2018.

[28] B. C. Jung and D. K. Sung, "Performance analysis of orthogonal code hopping multiplexing systems with repetition, convolutional, and turbo codes," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 932–944, Mar. 2008.

[29] B. C. Jung, S. S. Cho, and D. K. Sung, "Performance comparison of downlink capacity improvement schemes: Orthogonal code-hopping multiplexing versus multiple scrambling codes," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 670–681, Feb. 2009.

[30] B. C. Jung, S. S. Cho, and D. K. Sung, "Uplink capacity improvement through orthogonal code hopping in uplink-synchronized CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5404–5410, Nov. 2009.

[31] G. Ma, B. Ai, F. Wang, and Z. Zhong, "Tandem spreading network-coded division multiple access," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 390–398, Feb. 2017.

[32] G. Ma *et al.*, "Coded tandem spreading multiple access for massive machine-type communications," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 75–81, Apr. 2018.
[33] *Consideration on Grant-Free Transmission*, 3GPP, Sophia Antipolis, France, 3GPP CATT R1–1608757, Oct. 2016.
[34] J. Ding, D. Qu, H. Jiang, and T. Jiang, "Success probability of grant-free random access with massive MIMO," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 506–516, Feb. 2019.
[35] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.
[36] K. Lee, J. Kim, J. Jung, and I. Lee, "Zadoff-Chu sequence based signature identification for OFDM," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4932–4942, Oct. 2013.
[37] J. Tao and L. Yang, "Improved Zadoff-Chu sequence detection in the presence of unknown multipath and carrier frequency offset," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 922–925, May 2018.
[38] *Grant-Free Non-Orthogonal MA for Uplink URLLC*, 3GPP, Huawei, Shenzhen, China, HiSilicon R1–1608869, Oct. 2016.
[39] R. Mandl, "Orthogonal Latin squares: An application of experiment design to compiler testing," *Commun. ACM*, vol. 28, no. 10, pp. 1054–1058, Oct. 1985.
[40] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge, U.K.: Cambridge Univ. Press, 2005.
[41] *Multiplexing and Channel Coding*, 3GPP, Sophia Antipolis, France, 3GPP Rep. TS 38.212 V15.1.1, Apr. 2018.
[42] S. Park and D. K. Sung, "Orthogonal code hopping multiplexing," *IEEE Commun. Lett.*, vol. 6, no. 12, pp. 529–531, Dec. 2002.
[43] J. Liu, J. Wang, G. Shen, and Y. Yang, "IoT control based on uplink inter-cell power control in LTE system," in *Proc. IEEE PIMRC*, Sep. 2014, pp. 717–721.
[44] "Guidelines for evaluation of radio interface technologies for IMT-2020," ITU-R, Geneva, Switzerland, ITU-R Rep. ITU-R M.2412–0, Oct. 2017.
[45] T. Holma and A. Toskala, *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*. New York, NY, USA: Wiley, 2000.
[46] D. T. Wiriaatmadja and K. W. Choi, "Hybrid random access and data transmission protocol for machine-to-machine communications in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 33–46, Jan. 2015.

Han Seung Jang (Member, IEEE) received the B.S. degree in electronics and computer engineering from Chonnam National University, Gwangju, South Korea, in 2012, and the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute for Science and Technology, Daejeon, South Korea, in 2014 and 2017, respectively.

From May 2018 to February 2019, he was a Postdoctoral Fellow with the Information Systems Technology and Design Pillar, Singapore University of Technology and Design, Singapore. From September 2017 to April 2018, he was also a Postdoctoral Fellow with Chungnam National University, Daejeon, where he is currently an Assistant Professor with the School of Electrical, Electronic Communication, and Computer Engineering. His research interests include cellular Internet of Things/machine-to-machine communications, machine learning, smart grid, and energy ICT.

Bang Chul Jung (Senior Member, IEEE) received the B.S. degree in electronics engineering from Ajou University, Suwon, South Korea, in 2002, and the M.S. and Ph.D. degrees in electrical and computer engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2004 and 2008, respectively.

He was a Senior Researcher/Research Professor with the KAIST Institute for Information Technology Convergence, Daejeon, from January 2009 to February 2010. From March 2010 to August 2015, he was a Faculty of Gyeongsang National University, Tongyeong, South Korea. He is currently a Professor of the Department of EE, Chungnam National University, Daejeon. His research interests include 6G wireless communications, wireless IoT communications, statistical signal processing, information theory, wireless localization, interference management, radar signal processing, spectrum sharing, multiple antennas, multiple access techniques, radio resource management, machine learning, and GNSS receiver signal processing.

Prof. Jung was the recipient of the Fifth IEEE Communication Society Asia–Pacific Outstanding Young Researcher Award in 2011, the KICS Haedong Young Scholar Award in 2015, and the 29th KOFST Science and Technology Best Paper Award in 2019. He has been serving as an Associate Editor for *IEEE Vehicular Technology Magazine* since May 2020, and also served as an Associate Editor for the *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Sciences* since 2018.

Tony Q. S. Quek (Fellow, IEEE) received the B.E. and M.E. degrees in electrical and electronics engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 1998 and 2000, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2008.

He is currently the Cheng Tsang Man Chair Professor with the Singapore University of Technology and Design (SUTD), Singapore. He also serves as the Head of ISTD Pillar, the Sector Lead of the SUTD AI Program, and the Deputy Director of the SUTD-ZJU IDEA. His current research topics include wireless communications and networking, network intelligence, Internet of Things, URLLC, and big data processing.

Dr. Quek was honored with the 2008 Philip Yeo Prize for Outstanding Achievement in Research, the 2012 IEEE William R. Bennett Prize, the 2015 SUTD Outstanding Education Awards—Excellence in Research, the 2016 IEEE Signal Processing Society Young Author Best Paper Award, the 2017 CTTC Early Achievement Award, the 2017 IEEE ComSoc AP Outstanding Paper Award, the 2020 IEEE Communications Society Young Author Best Paper Award, the 2020 IEEE Stephen O. Rice Prize, the 2020 Nokia Visiting Professor, and the 2016–2020 Clarivate Analytics Highly Cited Researcher. He has actively involved in organizing and chairing sessions, and has served as a member of the Technical Program Committee, as well as the symposium chair in a number of international conferences. He is currently serving as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and an elected member of the IEEE Signal Processing Society SPCOM Technical Committee. He was an Executive Editorial Committee Member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, and the IEEE WIRELESS COMMUNICATIONS LETTERS. He is a Distinguished Lecturer of the IEEE Communications Society.

Dan Keun Sung (Life Fellow, IEEE) received the B.S. degree in electronics engineering from Seoul National University, Seoul, South Korea, in 1975, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Texas at Austin, Austin, TX, USA, in 1982 and 1986, respectively.

Since 1986, he has been with the Faculty of the Korea Advanced Institute of Science and Technology (KAIST), Daejon, South Korea, where he is currently a Professor Emeritus with the School of Electrical Engineering. From 1996 to 1999, he was the Director of the Satellite Technology Research Center, KAIST. His research interests include mobile communication systems and networks with special interest in resource management, cellular M2M, smart grid, energy networks, energy ICT, WLANs, WPANs, traffic control in wireless and wired networks, performance and reliability of communication systems, and microsatellites.

Prof. Sung was a recipient of the 1992 National Order of Merits, the Dongbaek, the 1997 Research Achievement Award, the 2000 Academic Excellence Award, the 2004 Scientist of the Month from the Ministry of Science and Technology and the Korea Science and Engineering Foundation, the 2013 Haedong Academic Grand Award from the Korean Institute of Communications and Information Sciences, and the 2017 National Order of Merits, the Okjo Medal. He had served as the Division Editor for the *Journal of Communications and Networks* from 1998 to 2007. He also had served as an Editor for *IEEE Communications Magazine* from 2002 to 2011. He is a fellow of the Korean Academy of Science Technology and a Life Fellow of the National Academy of Engineering of Korea.