An Early Preamble Collision Detection Scheme Based on Tagged Preambles for Cellular M2M Random Access

Han Seung Jang, *Student Member, IEEE*, Su Min Kim, *Member, IEEE*, Hong-Shik Park, *Member, IEEE*, and Dan Keun Sung, *Fellow, IEEE*

***Abstract*—In this paper, we propose an early preamble collision detection (e-PACD) scheme at the first step of random access (RA) procedure based on tagged preambles (PAs), which consist of both PA and tag Zadoff–Chu sequences using different root numbers, respectively. The proposed e-PACD scheme enables faster PA collision detection and notification, compared with the conventional RA scheme. Accordingly, it can reduce the RA delay and remove resource wastes which occur at the third step of the conventional RA procedure. The reduced RA delay achieved by the proposed e-PACD scheme can contribute to an ultralow latency requirement in fifth-generation (5G) cellular networks. In addition, the proposed e-PACD scheme enables the eNodeB to monitor the number of RA-attempting nodes (RA load) on each physical RA channel slot. The PACD probability and the RA load monitoring accuracy are mathematically analyzed, and the RA performance enhancement of the proposed scheme is evaluated in terms of RA success probability, average RA delay, and RA resource efficiency, compared with the conventional RA scheme.**

***Index Terms*—Machine-to-machine (M2M) communication, preamble (PA), preamble collision detection (PACD), random access (RA), RA load monitoring, tag, Zadoff–Chu (ZC) sequence.**

I. Introduction

CELLUAR machine-to-machine (M2M) communications play a major role in the future infrastructure of Internet of Things. They support a wide range of applications such as e-health, public safety, surveillance, remote maintenance and control, and smart metering [1]. However, different features of M2M communications from those of the conventional human-to-human and human-to-machine communications require the cellular networks to evolve in a different way [2]. The main features of M2M communications come from a massive number of machine nodes and their traffic volumes and patterns. They transmit small-sized data packets for status report and emergency alarm [3] and mostly stay out of connection to reduce energy consumption after data transmission, which makes initial random access (RA) procedure more important in cellular M2M communications.

When a massive number of machine nodes attempt RAs with a limited number of RA preambles (PAs) on physical random access channel (PRACH) and a limited amount of resource on physical uplink shared channel (PUSCH), severe PA collision and PUSCH resource shortage problems may occur. The PRACH is uplink shared resource, which is periodically allocated as part of PUSCH, in order to send and receive PAs at the first step of RA procedure. If PRACH is more frequently allocated within PUSCH due to an increase in RA arrivals, the PUSCH capacity for transmitting user data and RA-step 3 data is proportionally reduced. It implies that just increasing PRACH opportunities is not a good way to accommodate a massive number of RAs. So far, there have been several studies to accommodate a massive number of RA attempts with a limited amount of PRACH resource (e.g., 10 ms PRACH period) in [4]–[7]. In this situation, it is inevitable to suffer from severe PA collisions. These PA collisions can be reduced and treated by using effective methods, such as increasing the number of available PAs [4], [5], reusing PAs [6], and avoiding PUSCH collisions [7]. Moreover, in order to deal with severe RA network (RAN) overload from a huge number of M2M devices, RAN overload control or mitigation schemes were studied in [8]–[11].

The resource allocation for uplink accesses from a large number of machine nodes is another critical issue [12] since most M2M communications require a large number of resource allocations, which may significantly increase signaling overhead[13]–[15]. Since a conventional connection establishment or a conventional scheduling request in LTE system involves four steps in the RA procedure, which increases signalling overhead and delay, a simpler RA procedure reducing part of the steps is required for M2M small-sized data transmission for faster access [16]. Thus far, there have been a few studies which proposed to incorporate small-sized data transmissions into the RA procedure [17]–[19]. A message-embedded RA

Manuscript received April 20, 2016; revised September 6, 2016 and November 28, 2016; accepted December 27, 2016. Date of publication December 30, 2016; date of current version July 14, 2017. This work was supported in part by the Korea government, Ministry of Science, ICT and Future Planning under the National Research Foundation of Korea under Grant 2014R1A2A2A01005192 and in part by Institute for Information and Communications Technology Promotion Grant funded by the Korea government, Ministry of Science, ICT and Future Planning (B0101-16-1270, Research on communication technology using bio-inspired algorithm). The review of this paper was coordinated by Prof. D. B. da Costa.

H. S. Jang, H.-S. Park, and D. K. Sung are with the School of Electrical Engineering, College of Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: jhans@kaist.ac.kr; park1507@kaist.ac.kr; dksung@kaist.ac.kr).

S. M. Kim is with the Department of Electronics Engineering, Korea Polytechnic University, Siheung 15073, South Korea (e-mail: suminkim@kpu.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2016.2646739

scheme enabling machine nodes to quickly transmit small-sized message bits using message-embedded PAs at the first step of the RA procedure has been proposed in [19].

In order to deal with PA collision problems, PA collision detection and notification methods are necessarily required. In the conventional RA procedure, if more than one machine nodes randomly pick the same ZC sequence as their PAs and transmit them on the same PRACH slot at the first step (of the RA procedure), they receive the same uplink resource grant in RA response (RAR) message at the second step. Then, they transmit their desired data on the same uplink resource at the third step, and the data decoding may fail at the eNodeB. Based on this operation, the eNodeB can detect a PA collision at the third step and notify the PA collision to the corresponding machine nodes at the fourth step. However, when a massive number of machine nodes participate in the RA procedure, a longer RA delay is inevitable, and a larger number of resource blocks (RBs) on PUSCH may be wasted at the third step by the conventional PA collision detection and notification methods. In addition, it is hard to monitor the number of RA-attempting nodes on each PRACH slot in the conventional RA scheme, but the average number of RA-attempting nodes can be estimated based on the PA collision probabilities or the number of RA success nodes at the last step of the RA procedure [20]. As a result, it is more difficult to implement fast and adaptive RA overload control in the conventional RA scheme. Recently, Zhang *et al.* [21] proposed a PA collision detection (PACD) scheme by exploiting narrow guard bands on PRACH. In this scheme, nodes transmit their identifiers (IDs) on the guard bands during PA transmissions. If a single ID is not decoded for a specific PA, the eNodeB regards it as a PA collision. However, since the performance of PACD is limited by the bandwidth of the guard bands determined by the system parameters, it is generally hard to be applicable to current LTE system.

In this paper, we propose an early PACD (e-PACD) scheme at the first step of the RA procedure based on the tagged PAs which consist of both PA and tag ZC sequences using different root numbers, respectively. Compared to the message-embedded PAs in [19], the tagged PAs do not have any information but rather they are utilized for the eNodeB to detect PA collisions. More specifically, the proposed e-PACD scheme exploits the received multiple tags attached to the same PA in order to detect the PA collision. The proposed e-PACD scheme enables a faster PACD at the first step and a faster collision notification at the second step, compared to the conventional PACD, which is done at the third step and the conventional collision notification, which is done at the fourth step. As a result, the proposed e-PACD scheme can reduce the RA delay and also remove resource wastes on PUSCH since nodes with collided PAs do not have to proceed to the remaining third and fourth steps of the RA procedure. The reduced RA delay achieved by the proposed e-PACD scheme can contribute to an ultralow latency requirement in fifth-generation (5G) cellular networks [22]. In addition, the proposed e-PACD scheme can monitor RA load on each PRACH slot, and the RA load information can be utilized for fast and adaptive control of RA overload by generating an elaborate access class barring (ACB) factor [23] in a massive M2M RA environment.

The message decoding probability and message throughput on PRACH were analyzed and evaluated in [19], while in this paper, the PACD probability and RA load monitoring accuracy are mathematically analyzed, and from the viewpoint of MAC layer, the RA performance enhancement of the proposed RA scheme is evaluated in terms of RA success probability, average RA delay, and RA resource efficiency, compared with those of the conventional RA scheme.

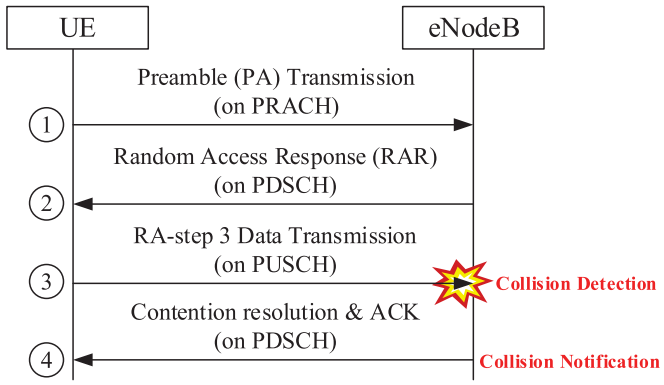

Fig. 1. Conventional RA procedure and its PACD.

The rest of this paper is organized as follows. In Section II, we present background, especially the tagged PA, and system model. In Section III, we propose an e-PACD scheme. The performance of the proposed scheme is mathematically analyzed in Section IV. The performance evaluation is presented in Section V. Finally, conclusive remarks are drawn in Section VI.

II. Background and System Model

A. Conventional RA and Collision Detection Procedure

We briefly introduce the conventional contention-based RA procedure [24], which is shown in Fig. 1. It consists of the following four steps.

1) *(Step 1) PA Selection and Transmission*. Machine nodes randomly select a single PA among N_{PA} available PAs. Then, each node transmits the selected PA on the shared PRACH slot. Due to the random PA selection property, more than one machine nodes can select the same PA.
2) *(Step 2) Random Access Response*. After the eNodeB detects the PAs, it sends RAR messages, each of which conveys the identity of detected PA, timing advance information, and an initial uplink resource grant for RA-step 3 data transmission. Here, the eNodeB only can detect whether a specific PA is transmitted or not, but it cannot recognize that the detected PA is transmitted by a single node or multiple nodes.
3) *(Step 3) RA-Step 3 Data Transmission*. Using the initial uplink resource informed by the received RAR message, the corresponding node transmits the uplink RA-step 3 data which convey a radio resource control connection request, a tracking area update, or a scheduling request on PUSCH. If multiple nodes use the same PA and the corresponding RAR message, all relevant nodes transmit their own RA-step 3 data on the same uplink resource, which makes the eNodeB hard to decode the RA-step 3 data correctly. However, the eNodeB can recognize a PA collision for this case.

4) *(Step 4) Contention Resolution Message Transmission*. If the eNodeB successfully decodes the RA-step 3 data transmitted by a single node, it sends back the contention resolution message including the UE identity (ID) obtained from the decoded RA-step 3 data; otherwise, the eNodeB sends nothing back in order to notify a PA collision. The machine node which receives its own ID in the contention resolution message sends back a positive ACK message, while the other nodes which do not get the ID regard the situation as a PA collision in which they do not send any message back and then retry as next RA.

As shown in Fig. 1, the conventional RA scheme can detect PA collisions and notify them at the third and fourth steps, respectively, which causes a longer RA delay and a waste of RBs for RA-step 3 data due to PA collisions. Throughout this paper, the nth step refers to the nth step of the RA procedure.

B. Conventional PA

ZC sequences [25] have been widely used for various applications in modern wireless communication systems[26]–[28]. In LTE RA system, ZC sequences are used to generate PAs defined as $z_r[n] \triangleq \exp[-j\pi rn(n+1)/N_{\rm ZC}]$ for $n = 0, \ldots, N_{\rm ZC}-1$,

where $N_{\rm ZC}$ denotes the sequence length and $r \in \{1, \ldots, (N_{\rm ZC}-1)\}$ is the root number [25]. The ZC sequences have an ideal cyclic *auto-correlation property* representing that the magnitude of the cyclic correlation with a circularly shifted version of itself becomes a scaled delta function, i.e., $|c_{rr}[\tau]| = \left|\frac{1}{\sqrt{N_{\rm ZC}}}\sum_{n=0}^{N_{\rm ZC}-1} z_r[n]z_r^*[n+\tau]\right| = \sqrt{N_{\rm ZC}}\delta[\tau]$,

where $c_{rr}[\tau]$ is the discrete cyclic auto-correlation function of $z_r[n]$ at lag τ and $(\cdot)^*$ denotes the complex conjugate. From this property, we can observe how much the received sequences are shifted, compared to the reference ZC sequence. These sequences also have a cyclic *cross-correlation property* representing that the magnitude of the cyclic cross-correlation between any two ZC sequences with different root numbers, r and k, is constant, i.e., $|c_{rk}[\tau]| = \left|\frac{1}{\sqrt{N_{\rm ZC}}}\sum_{n=0}^{N_{\rm ZC}-1} z_r[n]z_k^*[n+\tau]\right| = 1$. In principle, multiple PAs can be generated from a ZC sequence by cyclically shifting the sequence by a factor of cyclic shift value, $N_{\rm CS}$. The sequence shifted by $i \in \{0, \ldots, (\lfloor N_{\rm ZC}/N_{\rm CS}\rfloor - 1)\}$th multiple of $N_{\rm CS}$ is $z_r^i[n] = z_r[(n+iN_{\rm CS}) \bmod N_{\rm ZC}]$ which is the ith PA. Conventionally, an eNodeB serves UEs with 64 PAs, and the number of available root numbers is $(N_{\rm ZC}-1) = 838$ in a single LTE cell.

C. Tagged PA

Similarly to the message-embedded PA [19], which incorporates message ZC sequences using difference root numbers in a normal PA, tagged PA consists of both PA and tag ZC sequences, which are transmitted as a mixed sequence

$$X_{r,k_i}^{i,l}[n] = \beta(p_r^i[n] + q_{k_i}^l[n]) \tag{1}$$

where β denotes the transmit power of the tagged PA, and $p_r^i[n]$ and $q_{k_i}^l[n]$ denote the PA sequence and the tag sequence, respectively. The PA sequence is expressed as $p_r^i[n] = z_r[(n+iN_{\rm CS}) \bmod N_{\rm ZC}]$ for $n = 0, \ldots, (N_{\rm ZC}-1)$, where r denotes the PA root number, $i \in \{0, \ldots, (N_{\rm PA}-1)\}$ denotes the randomly selected PA index, $N_{\rm PA}$ denotes the number of available PAs. Similarly, the tag sequence is expressed as $q_{k_i}^l[n] = z_{k_i}[(n+l) \bmod N_{\rm ZC}]$ for $n = 0, \ldots, (N_{\rm ZC}-1)$, where k_i denotes the tag root number of the ith PA, l denotes the tag index randomly selected in $\{0, \ldots, (N_{\rm ZC}-1)\}$, and $N_{\rm ZC}$ denotes the tag sequence length. The tag root number k_i is determined by a mapping function of the PA index i, i.e., $k_i = f(i)$. In other words, the PA index i is mapped into a specific tag root number k_i, which is exclusive of the PA root number r, i.e, $r \neq k_i$. Using the tag root number k_i, the PA index i is tagged with the tag index l.

D. System Model

Let m_i denote the number of RA-attempting nodes using the ith PA index for $i = 0, \ldots, (N_{\rm PA}-1)$, and $M = \sum_{i=0}^{(N_{\rm PA}-1)} m_i$ represents the total number of RA-attempting nodes using the same PA root number r on the same PRACH. Let us define the following three sets:

1) $\mathcal{I} = \{i_d | d = 1, \ldots, M\}$
   : A set of PA indices transmitted by M nodes.
2) $\mathcal{L} = \{l_d | d = 1, \ldots, M\}$
   : A set of tag indices transmitted by M nodes.
3) $\mathcal{K} = f(\mathcal{I}) = \{k_{i_d} = f(i_d) | d = 1, \ldots, M\}$
   : A set of tag root numbers used by M nodes.

Here, subscript d denotes the device (node) index. All M nodes transmit their own tagged PAs as

$$X_{r,k_{i_d}}^{i_d,l_d}[n] = \beta_d(p_r^{i_d}[n] + q_{k_{i_d}}^{l_d}[n]),\ d = 1, \ldots, M \tag{2}$$

where $p_r^{i_d}[n]$ denotes the PA sequence of the dth node using PA root number r and PA index i_d and $q_{k_{i_d}}^{l_d}[n]$ denote the tag sequence of the dth node using tag root number k_{i_d} and tag index l_d, and β_d denotes the transmit power. Then, at the eNodeB, the received sequence from M nodes is obtained by

$$Y_{r,\mathcal{K}}^{\mathcal{I},\mathcal{L}}[n] = \sum_{d=1}^{M}\sum_{e=1}^{E_d} h_d^e X_{r,k_{i_d}}^{i_d,l_d}[(n+t_d^e) \bmod N_{\rm ZC}] + W[n] \tag{3}$$

where h_d^e denotes the eth multipath channel coefficient for the dth node, t_d^e denotes the round-trip delay of the eth multipath for the dth node, E_d is the total number of multipaths for the dth node, and $W[n]$ represents the circular symmetry complex Gaussian noise with zero mean and variance σ^2, i.e., $W[n] = W_{\rm R}[n] + jW_{\rm I}[n] \sim \mathcal{CN}(0, \sigma^2)$.

III. Proposed e-PACD Scheme Based on Tagged PAs

A. Detection of Tagged PAs

To detect PA indices, the eNodeB calculates the correlation value $|c_{\{r,\mathcal{K}\},r}[\tau]|$ between the received sequence $Y_{r,\mathcal{K}}^{\mathcal{I},\mathcal{L}}[n]$ and the PA reference sequence $z_r[n]$ and checks each PA detection zone $\mathcal{D}_i = \{\tau | \tau \in [iN_{\rm CS}, ((i+1)N_{\rm CS}-1)]\} \subset \{\mathcal{D}_0, \ldots, \mathcal{D}_{(N_{\rm PA}-1)}\}$ whether $|c_{\{r,\mathcal{K}\},r}[\tau]| \geq \gamma_{\rm pa}$ or not, where $\gamma_{\rm pa}$ is the PA detection threshold. If at least one among m_i peaks exceeds the threshold $\gamma_{\rm pa}$ on the ith detection zone $\mathcal{D}_i$, the ith PA is detected. We denote the set of detected PA-peak positions for the ith PA as $\mathbf{\Omega}_{\rm pa}^i = \{\Omega_d^i | d = 1, \ldots, m_i\}$. Next, in order to detect tag indices attached to the ith detected PA, the eNodeB

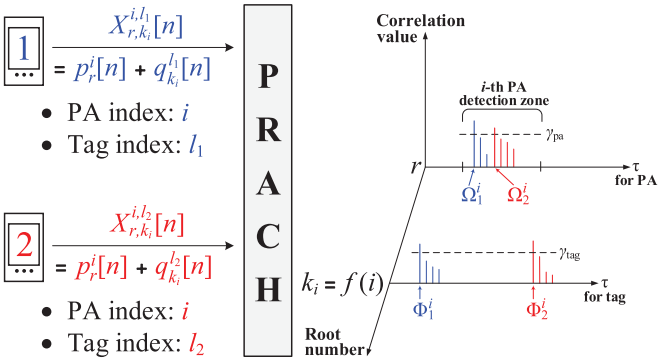

Fig. 2. Two-node example for tagged PA transmission and detection ($i_1 = i_2 = i$ and $l_1 \neq l_2$).

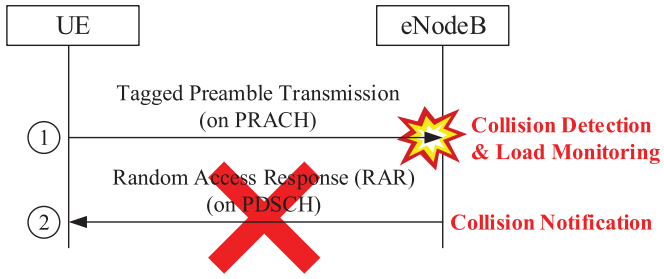

Fig. 3. Proposed e-PACD and RA load monitoring based on tagged PAs.

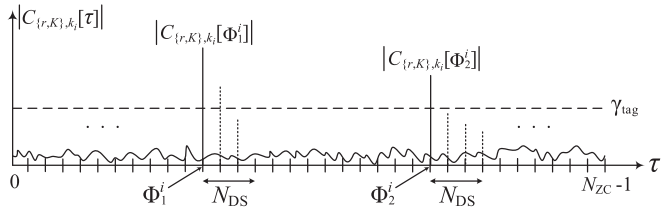

Fig. 4. Detection mechanism for tag spreading zones in case of $N_{\mathrm{SZ}}^{i} = 2$.

again calculates the correlation value $|c_{\{r,\mathcal{K}\},k_i}[\tau]|$ between the received sequence $Y_{r,\mathcal{K}}^{\mathcal{I},\mathcal{L}}[n]$ and the tag reference sequence $z_{k_i}[n]$ using the tag root number determined by $k_i = f(i)$ and checks the overall tag detection zone $\mathcal{D}_{\mathrm{tag}} = \{\tau | \tau \in [0, (N_{\mathrm{ZC}} - 1)]\}$, whether $|c_{\{r,\mathcal{K}\},k_i}[\tau]| \geq \gamma_{\mathrm{tag}}$ or not, where γ_{tag} denotes the tag detection threshold.[1] Here, m_i peaks can be detected on the detection zone $\mathcal{D}_{\mathrm{tag}}$ and we denote the set of detected tag-peak positions of the ith PA as $\mathbf{\Phi}_{\mathrm{tag}}^{i} = \{\Phi_d^i | d = 1, \ldots, m_i\}$. Fig. 2 shows a two-node example for tagged PA transmission and detection. Since nodes 1 and 2 select the same PA index i ($i_1 = i_2 = i$ and $m_i = 2$) but different tag indices l_1 and l_2, they generate the identical PA sequence $p_r^i[n]$ using the same PA root number r, while they generate different tag sequences $q_{k_i}^{l_1}[n]$ and $q_{k_i}^{l_2}[n]$ using the same tag root number $k_i = f(i)$ but different tag indices l_1 and l_2. After they transmit their own tagged PAs $X_{r,k_{i_1}}^{i,l_1}[n]$ and $X_{r,k_{i_2}}^{i,l_2}[n]$, the eNodeB detects the ith PA on the ith PA detection zone by comparing the correlation value $|c_{\{r,\mathcal{K}\},r}[\tau]|$ with the PA detection threshold γ_{pa}. However, it is very hard to recognize the PA collision on the narrow PA detection zone since PA detection is performed only through comparison with a single detection threshold value. In our proposed e-PACD scheme, we again calculate the correlation value between $Y_{r,\mathcal{K}}^{\mathcal{I},\mathcal{L}}[n]$ and the tag reference sequence $z_{k_i}[n]$ using the tag root number $k_i = f(i)$. Then, two tag-peak positions can be obtained sufficiently apart from each other as Φ_1^i and Φ_2^i on the whole tag detection zone $\mathcal{D}_{\mathrm{tag}}$.

B. Proposed e-PACD Scheme

We propose an e-PACD scheme based on tagged PAs. As shown in Fig. 3, at the first step, machine nodes transmit tagged PAs. During the tag detection for the ith detected PA, multiple tags can be captured in a wide range of tag detection zone $\mathcal{D}_{\mathrm{tag}} = \{\tau | \tau \in [0, (N_{\mathrm{ZC}} - 1)]\}$. Basically, in case of a single tag detected, the eNodeB regards the i-th PA as a successful PA transmission from a single node. In this case, the RAR message is transmitted at the second step in order to notify further steps allowed. However, in other cases of multiple tags detected, the eNodeB regards the ith PA as a collided PA. For these cases, based on the number of tags detected, the eNodeB can count the number of RA-attempting nodes using the ith PA index on the same PRACH. If a PA collision is detected, the eNodeB does not transmit an RAR message in order to notify the PA collision to the corresponding machine nodes which transmitted the collided PA, and after a time-out, the nodes recognize the PA collision and attempt new RAs again.

Due to multipath effect of wireless channel, a single tag sequence can be spread across several samples. Let us define a tag spreading zone, which consists of N_{DS} samples. N_{DS} should be set to include significant multipath signals from a single tag sequence [29]. Accordingly, N_{DS} is obtained as $N_{\mathrm{DS}} = \lceil t_{\mathrm{ds}}(N_{\mathrm{ZC}}/T_{\mathrm{ZC}}) \rceil$, where t_{ds}, N_{ZC}, and T_{ZC} denote the delay spread time (μs), ZC sequence length, and ZC sequence time duration (μs), respectively.

Fig. 4 shows the proposed detection mechanism for tag spreading zones. From $\tau = 0$, the eNodeB checks whether the tag correlation value is greater than or equal to the tag detection threshold, i.e., $|c_{\{r,\mathcal{K}\},k_i}[\tau]| \geq \gamma_{\mathrm{tag}}$. Assuming the same situation as Fig. 2 and the first multipath among E_d multipaths has the largest intensity, the first tag-peak position of node 1 is marked as Φ_1^i. Then, the first tag spreading zone is set to $[\Phi_1^i, (\Phi_1^i + N_{\mathrm{DS}} - 1)]$. Even though multiple signals can exceed the tag detection threshold on this tag spreading zone, they are regarded as a single tag. In Fig. 4, the dotted lines following the first tag-peak at $\tau = \Phi_1^i$ are considered as multipaths of the first tag-peak. Similarly, the tag spreading zone of node 2 appears in $[\Phi_2^i, (\Phi_2^i + N_{\mathrm{DS}} - 1)]$. The same operation is repeated until $\tau = (N_{\mathrm{ZC}} - 1)$.

Finally, we can obtain the number of detected tag spreading zones N_{SZ}^i for the detected ith PA. In case of $N_{\mathrm{SZ}}^i = 1$, the eNodeB regards the detected ith PA as the collision-free PA

[1] In the conventional RA scheme, a single correlation task is required, while $(1 + x)$ correlation tasks are required in the proposed e-PACD scheme when the number of detected PAs is x. However, the computational time may not be much longer than that of the conventional scheme since correlation computations can be done in parallel between the received sequence, and each of x tag reference sequences after the PAs are detected. Moreover, $(1 + x) \leq (1 + N_{\mathrm{PA}})$ implies that the correlation tasks do not significantly increase due to a massive number of RA attempts since it is limited by N_{PA}.

and then transmits an RAR message to the corresponding node at the second step. In other cases of $N^i_{\rm SZ} \geq 2$, the eNodeB regards the detected ith PA as the collided PA and the RA load for the ith PA is monitored as $L^i_{\rm mon} = N^i_{\rm SZ}$. Then, it does not transmit an RAR message for the ith PA at the second step in order to notify the PA collision of the ith PA. Nodes, which receive no RAR messages, do not go forward to the remaining RA steps, and attempt a new RA after a back-off time. As a result, since the eNodeB effectively allocates PUSCH resources for RA-step 3 data to only nodes with collision-free PAs, the proposed e-PACD scheme can save PUSCH resources for RA-step 3 channels (S3CHs) and reduce the RA delay, compared to the conventional RA scheme.

Moreover, the monitored RA load can be utilized as an important parameter for adaptive RA overload control in an ACB scheme which is an efficient technique to adjust the RA loads. The probability value called the ACB factor is calculated by the proportion of the number of target RA-attempting nodes to the number of potential RA-attempting nodes. In the existing ACB schemes [10], [30], [31], the number of potential RA-attempting nodes is approximately estimated at the last fourth step by the PA collision probability or the number of RA success nodes, and thus, it may be inaccurate and outdated at the next PRACH slot. However, the proposed RA load monitoring scheme can obtain the latest and more accurate information on the number of RA-attempting nodes at the first step. This information can be directly utilized for generating an elaborate ACB factor for the next PRACH slot.

IV. Performance Analysis

For simplicity, we assume that each tagged PA experiences only a single path and the average channel coefficient is equal to 1. Thus, the received sequence in (3) is rewritten as

$$Y^{\mathcal{I},\mathcal{L}}_{r,\mathcal{K}}[n] = \sum\nolimits_{d=1}^{M} X^{i_d,l_d}_{r,k_{i_d}}[(n+t_d) \bmod N_{\rm ZC}] + W[n] \quad (4)$$

where r, $\mathcal{K}$, k_{i_d}, and t_d denote the PA root number, the set of utilized tag root numbers from M nodes, the tag root number of the dth node, and the delay shift of the dth node, respectively.

A. Analysis of PA Detection Probability

The ith PA can be detected by calculating the correlation value $|c_{\{r,\mathcal{K}\},r}[\tau]|$ between the received sequence $Y^{\mathcal{I},\mathcal{L}}_{r,\mathcal{K}}[n]$ and the PA reference sequence $z_r[n]$ and then verifying the ith PA detection zone $\mathcal{D}_i$. The correlation value is calculated as

$$\begin{aligned}|c_{\{r,\mathcal{K}\},r}[\tau]| &= \left| \frac{1}{\sqrt{N_{\rm ZC}}} \sum_{n=0}^{N_{\rm ZC}-1} Y^{\mathcal{I},\mathcal{L}}_{r,\mathcal{K}}[n] z^*_r[n+\tau] \right| \\ &= \left| \sum_{d=1}^{M} (c^d_{r,r}[\tau] + c_{k_{i_d},r}[\tau]) + W[\tau] \right|\end{aligned}$$

where

$$\begin{aligned}c^d_{r,r}[\tau] &= \frac{1}{\sqrt{N_{\rm ZC}}} \sum_{n=0}^{N_{\rm ZC}-1} \beta_d p^{i_d}_r[n+t_d] z^*_r[n+\tau] \\ &= \sqrt{N_{\rm ZC}} \beta_d \delta[\{(i_d N_{\rm CS} + t_d) \bmod N_{\rm ZC}\} - \tau]\end{aligned}$$

and

$$c_{k_{i_d},r}[\tau] = \frac{1}{\sqrt{N_{\rm ZC}}} \sum_{n=0}^{N_{\rm ZC}-1} \beta_d q^{l_d}_{k_{i_d}}[n+t_d] z^*_r[n+\tau].$$

Since there are m_i nodes using $i_d = i$, m_i peaks can be detected at each of $\tau \in \mathbf{\Omega}^i_{\rm pa} = \{\Omega^i_d | d = 1, \ldots, m_i\}$, where $\Omega^i_d = \{(iN_{\rm CS} + t_d) \bmod N_{\rm ZC}\}$. At time instant $\tau = \Omega^i_d$

$$\left|c_{\{r,\mathcal{K}\},r}[\Omega^i_d]\right| = \sqrt{(\theta^i_d \cos\alpha + W_{\rm R})^2 + (\theta^i_d \sin\alpha + W_{\rm I})^2}$$

becomes a random variable (RV) $G^i_d \sim {\rm Rice}(\theta^i_d, \sigma)$ following a Rician distribution [32], whose probability density function (pdf) is given by

$$f_{G^i_d}(g; \theta^i_d, \sigma) = \frac{g}{\sigma^2} I_0\left(\frac{g\theta^i_d}{\sigma^2}\right) \exp\left[-\frac{g^2 + (\theta^i_d)^2}{2\sigma^2}\right] \quad (5)$$

where $\theta^i_d = |\sqrt{N_{\rm ZC}}\beta_d + \sum_{s=1}^{M} c_{k_{i_s},r}[\Omega^i_d]|$

$$\alpha = \cos^{-1}\left\{ {\rm Re}(\sqrt{N_{\rm ZC}}\beta_d + \sum_{s=1}^{M} c_{k_{i_s},r}[\Omega^i_d])/\theta^i_d \right\}$$

and $I_0(\cdot)$ is the modified Bessel function of the first kind with the zeroth order. Therefore, we have m_i RVs $\{G^i_d | d = 1, \ldots, m_i\}$ at each of $\mathbf{\Omega}^i_{\rm pa}$ in $\mathcal{D}_i$. Let us define the maximum RV $G^i_{\rm pa} = \max\{G^i_d | d = 1, \ldots, m_i\}$. Finally, the PA detection probability for the ith PA is defined as

$$P^i_{\rm PA} \triangleq \Pr[G^i_{\rm pa} \geq \gamma_{\rm pa}] = 1 - \prod_{d=1}^{m_i} \left\{ 1 - Q_1\left(\frac{\theta^i_d}{\sigma}, \frac{\gamma_{\rm pa}}{\sigma}\right) \right\} \quad (6)$$

where $Q_1(a,b)$ and $\gamma_{\rm pa}$ denote the Marcum Q function and the PA detection threshold, respectively.

B. Analysis of PACD and False Collision Detection Probabilities

After the ith PA is detected, the eNodeB again calculates the correlation value $|c_{\{r,\mathcal{K}\},k_i}[\tau]|$ between the received sequence $Y^{\mathcal{I},\mathcal{L}}_{r,\mathcal{K}}[n]$ and the tag reference sequence $z_{k_i}[n]$ using the tag root number $k_i = f(i)$ and checks the overall tag detection zone $\mathcal{D}_{\rm tag} = \{\tau | \tau \in [0, (N_{\rm ZC} - 1)]\}$. The correlation value is calculated as

$$\begin{aligned}|c_{\{r,\mathcal{K}\},k_i}[\tau]| &= \left| \frac{1}{\sqrt{N_{\rm ZC}}} \sum_{n=0}^{N_{\rm ZC}-1} Y_{r,\mathcal{K}}[n] z^*_{k_i}[n+\tau] \right| \\ &= \left| \sum_{d=1}^{M} (c^d_{r,k_i}[\tau] + c_{k_{i_d},k_i}[\tau]) + W[\tau] \right|\end{aligned}$$

where

$$c^d_{r,k_i}[\tau] = \frac{1}{\sqrt{N_{\rm ZC}}} \sum_{n=0}^{N_{\rm ZC}-1} \beta_d p_{r,i_d}[n+t_d] z^*_{k_i}[n+\tau]$$

and

$$c_{k_{i_d},k_i}[\tau] = \frac{1}{\sqrt{N_{\rm ZC}}} \sum_{n=0}^{N_{\rm ZC}-1} \beta_d q_{k_{i_d},l_d}[n+t_d] z^*_{k_i}[n+\tau].$$

Since there are m_i tags attached to the ith PA using $k_{i_d} = k_i$, m_i peaks can be detected at each of $\tau \in \mathbf{\Phi}^i_{\rm tag} = \{\Phi^i_d | d = 1, \ldots, m_i\}$, where $\Phi^i_d = \{(l_d + t_d) \bmod N_{\rm ZC}\}$. At time instant $\tau = \Phi^i_d$

$$|c_{\{r,\mathcal{K}\},k_i}[\Phi^i_d]| = \sqrt{(\phi^i_d \cos\varphi + W_{\rm R})^2 + (\phi^i_d \sin\varphi + W_{\rm I})^2}$$

becomes an RV $H^i_d \sim {\rm Rice}(\phi^i_d, \sigma)$ following a Rician distribution, whose pdf is given in (5) with the different parameters

$$\phi^i_d = \left| \sqrt{N_{\rm ZC}}\beta_d + \sum_{s=1}^{M} c^s_{r,k_i}[\Phi^i_d] + \sum_{s=1, k_{i_s} \neq k_i}^{M} c_{k_{i_s},k_i}[\Phi^i_d] \right|$$

and

$$\varphi = \cos^{-1}\left\{ {\rm Re}(\sqrt{N_{\rm ZC}}\beta_d + \sum_{s=1}^{M} c^s_{r,k_i}[\Phi^i_d] + \sum_{s=1, k_{i_s} \neq k_i}^{M} c_{k_{i_s},k_i}[\Phi^i_d])/\phi^i_d \right\}.$$

Therefore, we have m_i RVs $\{H^i_d | d = 1, \ldots, m_i\}$ at each of $\mathbf{\Phi}^i_{\rm tag}$ in the tag detection zone.

There exist 2^{m_i} detection cases whether each of m_i RVs exceeds the tag detection threshold $\gamma_{\rm tag}$ or not. The jth detection vector is defined as $\boldsymbol{v}^i_j = [\mathbb{I}^i_{j,1}, \mathbb{I}^i_{j,2}, \ldots, \mathbb{I}^i_{j,m_i}]$ for $j = 1, \ldots, 2^{m_i}$, where $\mathbb{I}^i_{j,d}$ represents the dth indicator function in the jth detection vector, which is 0 if $H^i_d \geq \gamma_{\rm tag}$, 1 otherwise. The probability mass function of the detection vector $\boldsymbol{v}^i_j$ is expressed as

$$\Pr[\boldsymbol{v}^i_j] = \prod_{d=1}^{m_i} \left\{ \mathbb{I}^i_{j,d} + (-1)^{\mathbb{I}^i_{j,d}} Q_1\left(\frac{\phi^i_d}{\sigma}, \frac{\gamma_{\rm tag}}{\sigma}\right) \right\}.$$

In addition, the total number of detectable tag spreading zones is obtained by $N_{\rm tag} = \lfloor N_{\rm ZC}/N_{\rm DS} \rfloor$. As a result, the PACD probability for the ith PA index is defined as

$$P^i_{\rm CD} \triangleq \Pr[L^i_{\rm mon} \geq 2] \tag{7}$$

$$= P^i_{\rm PA}\left(1 - \frac{1}{N_{\rm tag}}\right) \sum_{j=1, f_{\rm sum}(\boldsymbol{v}^i_j) \geq 2}^{2^{m_i}} \Pr[\boldsymbol{v}^i_j] \tag{8}$$

where $L^i_{\rm mon}$ represents the monitored number of RA-attempting nodes using the ith PA index, $f_{\rm sum}(\boldsymbol{v}^i_j) = \sum_{d=1}^{m_i}(1 - \mathbb{I}^i_{j,d})$, and $\sum_{j=1, f_{\rm sum}(\boldsymbol{v}^i_j)\geq 2}^{2^{m_i}} \Pr[\boldsymbol{v}^i_j]$ indicates the sum of all probabilities that more than two tag spreading zones are detected. The term $(1 - 1/N_{\rm tag})$ represents the probability that at least one tag is transmitted on another tag spreading zone.

If only a single node transmits the ith PA ($m_i = 1$), the monitored RA load for the ith PA index, $L^i_{\rm mon}$, should be equal to 1. However, due to the additive noise, a false collision detection (FCD) may occur, i.e., $L^i_{\rm mon} \geq 2$. At any time instants $\tau \notin \mathbf{\Phi}^i_{\rm tag}$ on the tag detection zone $\mathcal{D}_{\rm tag}$, $|c_{\{r,\mathcal{K}\},k_i}[\tau]|$ becomes an RV $V^i[\tau] \sim {\rm Rice}(\eta^i[\tau], \sigma)$ following a Rician distribution, whose pdf is given in (5) with the different parameter

$$\eta^i[\tau] = \left| \sum_{d=1}^{M} c^d_{r,k_i}[\tau] + \sum_{d=1, k_{i_d} \neq k_i}^{M} c_{k_{i_d},k_i}[\tau] \right|.$$

Here, let us define the maximum noise RV $Z^i_{\rm tag} \triangleq \max\{V^i[\tau] | \forall \tau \notin \mathbf{\Phi}^i_{\rm tag}\}$, whose cumulative density function is given by

$$\Pr[Z^i_{\rm tag} \leq z; \boldsymbol{\eta}^i, \sigma] = \prod_{\tau \notin \mathbf{\Phi}^i_{\rm tag}} \left\{ 1 - Q_1\left(\frac{\eta^i[\tau]}{\sigma}, \frac{z}{\sigma}\right) \right\}$$

where $\boldsymbol{\eta}^i = [\eta^i[\tau] | \forall \tau \notin \mathbf{\Phi}^i_{\rm tag}]$. Therefore, the FCD probability for the ith PA is obtained by

$$P^i_{\rm FCD} \triangleq P^i_{\rm PA} \Pr[Z^i_{\rm tag} \geq \gamma_{\rm tag}] \Pr[H^i_1 \geq \gamma_{\rm tag}]. \tag{9}$$

C. Analysis of RA Load Monitoring Accuracy

We define RA load monitoring accuracy as the probability that the eNodeB accurately estimates the number of RA-attempting nodes transmitting the same PA. When multiple clusters of tag peaks are placed within the same tag spreading zone, the eNodeB detects them as a single tag. Thus, in order to correctly monitor the number of tags attached to the specific PA, tags should be spaced sufficiently apart from each other. We consider the tag nonoverlapping probability that m_i tags from the m_i nodes transmitting the same ith PA are not overlapped as

$$P^i_{\rm NO} = \prod_{d=0}^{m_i - 1} \frac{(N_{\rm tag} - d)}{(N_{\rm tag})^{m_i}}.$$

In addition, all these tags should be correctly detected while noise values should not exceed the tag detection threshold $\gamma_{\rm tag}$. Hence, the RA load monitoring accuracy of the ith PA is defined as

$$P^i_{\rm LM} \triangleq \Pr[L^i_{\rm mon} = m_i]$$

$$= P^i_{\rm PA} P^i_{\rm NO} \Pr[Z^i_{\rm tag} < \gamma_{\rm tag}] \prod_{d=1}^{m_i} \Pr[H^i_d \geq \gamma_{\rm tag}] \tag{10}$$

where $\Pr[H^i_d \geq \gamma_{\rm tag}] = Q_1(\phi^i_d/\sigma, \gamma_{\rm tag}/\sigma)$.

D. Analysis of RA Performance

We analyze the RA performance in terms of the RA collision probability, the successful S3CH allocation rate, RA success probability, RA failure probability, and RA delay. Here, we assume that M machine nodes attempt RAs with $N_{\rm PA}$ PAs on the same PRACH, and $N_{\rm S3CH}$ S3CHs are available. The probability that no one utilizes the ith PA is derived as

$$p_0 \triangleq \Pr[m_i = 0] = \binom{M}{0}\left(\frac{1}{N_{\rm PA}}\right)^0 \left(1 - \frac{1}{N_{\rm PA}}\right)^M$$

where $\binom{a}{b}$ represents the binomial coefficient. The probability that a single node uses the ith PA is derived as

$$p_1 \triangleq \Pr[m_i = 1] = \binom{M}{1}\left(\frac{1}{N_{\rm PA}}\right)^1 \left(1 - \frac{1}{N_{\rm PA}}\right)^{(M-1)}.$$

The PA detection probability that at least one node uses the ith PA is derived as $\bar{p}_0 \triangleq \Pr[m_i \geq 1] = 1 - p_0$. In addition, the PA collision probability that more than two nodes use the ith PA is derived as $p_{\rm c} \triangleq \Pr[m_i \geq 2] = 1 - (p_0 + p_1)$.

The proposed e-PACD scheme can save S3CHs by exactly allocating S3CHs to only the nodes with collision-free PAs, and

thus, the required number of S3CHs in the proposed scheme becomes $p_1 N_{\mathrm{PA}}$. The successful S3CH allocation rate of the proposed RA scheme is obtained by $\min\left[1, N_{\mathrm{S3CH}}/(p_1 N_{\mathrm{PA}})\right]$. However, the conventional RA scheme should allocate S3CHs to all detected PAs, and thus, the required number of S3CHs becomes $\bar{p}_0 N_{\mathrm{PA}}$. The successful S3CH allocation rate of the conventional RA scheme is obtained by $\min\left[1, N_{\mathrm{S3CH}}/(\bar{p}_0 N_{\mathrm{PA}})\right]$. The RA success probability is defined as the probability that a single node uses an exclusive PA, and one S3CH is allocated to this node. Thus, the RA success probabilities in the proposed and the conventional RA schemes are expressed, respectively, as

$$p_{\mathrm{s}}^{\mathrm{prop}} = \left(1 - \frac{1}{N_{\mathrm{PA}}}\right)^{(M-1)} \times \min\left(1, \frac{N_{\mathrm{S3CH}}}{p_1 N_{\mathrm{PA}}}\right)$$

and

$$p_{\mathrm{s}}^{\mathrm{conv}} = \left(1 - \frac{1}{N_{\mathrm{PA}}}\right)^{(M-1)} \times \min\left(1, \frac{N_{\mathrm{S3CH}}}{\bar{p}_0 N_{\mathrm{PA}}}\right)$$

where N_{S3CH} denotes the number of available S3CHs.

To calculate the average RA delay, there exist four different cases for PA collision and S3CH allocation in a single RA trial

1) $P_1 = \Pr$[PA collision, S3CH allocated];
2) $P_2 = \Pr$[PA collision, S3CH unallocated];
3) $P_3 = \Pr$[no PA collision, S3CH unallocated];
4) $P_4 = \Pr$[no PA collision, S3CH allocated];

where P_4 represents the RA success probability. According to the above four cases, we can calculate each of the average delays T_1, T_2, T_3, and T_4. Especially, the conventional and the proposed RA schemes have different T_1 values (denoted by T_1^{conv} and T_1^{prop}), since the proposed RA scheme can early detect the PA collision and notify it to the corresponding nodes at the second step. This can reduce the average RA delay in the proposed scheme. N_{RT} denotes the total number of RA retrials to succeed, and N_k^j denotes the number of occurences for delay case k in the jth instance before an RA success, where $N_{\mathrm{RT}} = \sum_{k=1}^{3} N_k^j$, $0 \le N_k^j \le N_{\mathrm{RT}}$, $j = 1, \ldots, 3^{N_{\mathrm{RT}}}$, and N_{RT} is limited by the maximum number of allowable RA retrials $N_{\mathrm{RT}}^{\max}$. Therefore, the average RA delay is obtained by

$$T_{\mathrm{delay}} = \frac{P_4 T_4 + \sum_{n=1}^{N_{\mathrm{RT}}^{\max}} \sum_{j=1}^{3^n} A_j B_j}{P_4 + \sum_{n=1}^{N_{\mathrm{RT}}^{\max}} \sum_{j=1}^{3^n} A_j} \tag{11}$$

where $A_j = P_4 \prod_{k=1}^{3} P_k^{N_k^j}$ and $B_j = T_4 + \sum_{k=1}^{3} T_k N_k^j$. Finally, the RA failure probability is defined as the probability that a node fails to access even after the maximum number of RA reattempts, $N_{\mathrm{RT}}^{\max}$, and it is derived as

$$P_{\mathrm{failure}} = 1 - \sum_{n=0}^{N_{\mathrm{RT}}^{\max}} P_4 \left(1 - P_4\right)^{(N_{\mathrm{RT}}^{\max} - n)}. \tag{12}$$

V. Performance Evaluation

In Section V-A, we evaluate the performance of the proposed e-PACD scheme in terms of PA detection probability, PACD probability, and RA load monitoring accuracy. From the viewpoint of MAC layer, we show the RA performance enhancement of the proposed scheme in terms of RA success probability, average RA delay, and S3CH resource efficiency, compared to those of the conventional RA scheme in Section V-B.

TABLE I
Simulation Parameters and Values for Physical Layer Performance

Parameters	Values
Length of ZC sequence, N_{ZC}	839
Size of tag spreading zone, N_{DS}	3
PA detection threshold, γ_{pa}	-16 dB
Tag detection threshold, γ_{tag}	-19 dB ~ -16 dB
Signal strength (SNR), β	-20 dB ~ -10 dB
Number of nodes transmitting PA 1, m_1	$1 \sim 5$

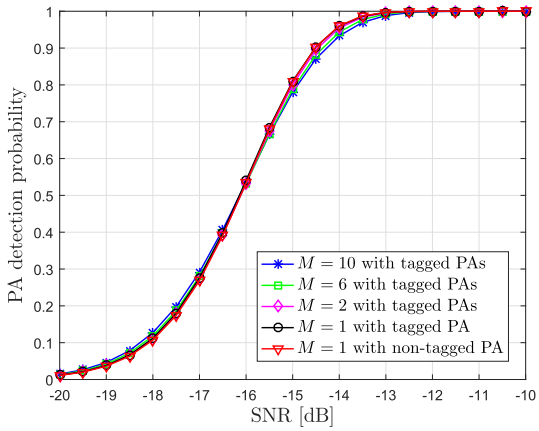

Fig. 5. Comparison of the average PA detection probabilities of both tagged PAs and nontagged PA when total M nodes choose exclusive PAs on a PRACH slot.

A. PA Detection Probability, PACD Probability, and RA Load Monitoring Accuracy

Table I lists the simulation parameter set. Here, the detection thresholds and signal strengths (SNR) are represented in dB scale over N_{ZC} AWGN noises. First, we evaluate the impact of the additional tag sequence on the PA detection probability. The additional tag sequences may degrade the PA detection performance since the additional tag sequences with distinct root numbers from that of the PA sequence have a certain cross-correlation value. This may cause an expense for sending additional tag sequences in the proposed scheme. This negative effect is examined in Fig. 5. Since different setups of M, $\mathcal{I}$, and $\mathcal{L}$ may yield different results, we randomly choose $\mathcal{I}$ and $\mathcal{L}$ 1000 times for varying M to obtain the average performance. Fig. 5 shows the average PA detection probabilities for varying M values. As shown in the figure, the difference in the average PA detection probabilities of the proposed e-PACD scheme using the tagged PAs and the conventional RA scheme using the nontagged PAs is marginal. It implies that the additional tags used in the proposed e-PACD scheme do not significantly degrade the PA detection probability.

Fig. 6 shows the PA detection probability, the PACD probability, and the RA load monitoring accuracy, when the number of RA-attempting nodes transmitting PA 1 among five nodes

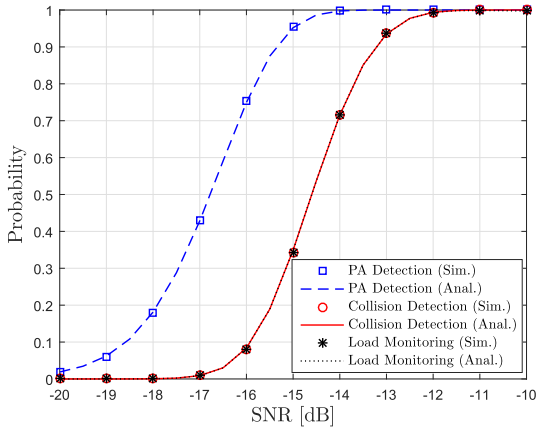

Fig. 6. PA detection probability, PACD probability, and RA load monitoring accuracy when $m_1 = 2$ and $\gamma_{\rm pa} = \gamma_{\rm tag} = -16$ dB.

$(M = 5)^2$ is equal to 2 ($m_1 = 2$) and the PA and tag detection thresholds are set to the same value as $\gamma_{\rm pa} = \gamma_{\rm tag} = -16$ dB. In this experiment, the set of transmitted PA and tag indices from five nodes are set to $\mathcal{I} = \{1, 1, 2, 3, 4\}$ and $\mathcal{L} = \{10, 20, 30, 40, 50\}$, respectively. As shown in the figure, the simulation results agree well with the analytical results. Since the number of RA-attempting nodes using PA 1 is equal to 2 ($m_1 = 2$), the PACD probability ($\Pr[L^1_{\rm mon} \geq 2]$) and the RA load monitoring accuracy ($\Pr[L^1_{\rm mon} = 2]$) show almost the similar performance, but they are apart approximately 2 dB from the PA detection probability since both the PACD and the RA load monitoring require first the successful PA detection.

Fig. 7 shows the PA detection probability, the PACD probability, and the RA load monitoring accuracy in case of $m_1 = 2$, and the FCD probability in case of $m_1 = 1$ when $\gamma_{\rm pa} = -16$ dB and $\gamma_{\rm tag} = -19$ dB ~ -16 dB. In this experiment, the set of transmitted PA indices $\mathcal{I}$ is fixed to $\{1, 2, 3, 4, 5\}$ in case of $m_1 = 1$ and $\{1, 1, 2, 3, 4\}$ in case of $m_1 = 2$, and the set of five tag indices $\mathcal{L}$ is randomly picked 1000 times in order to obtain the average performance. Since the PA detection does not depend on the tag detection threshold, the PA detection probabilities show the similar performance for all cases.

The lower tag detection threshold is set, the higher PACD probability is obtained. However, the RA load monitoring accuracies with $\gamma_{\rm tag} = -19$ dB, -18 dB, -17 dB decrease after SNR $= -14.5$ dB, -13.5 dB, and -11.5 dB, respectively, while that with $\gamma_{\rm tag} = -16$ dB achieves the similar curve to the PACD probability. In addition, in case of $m_1 = 1$, the FCD probabilities with $\gamma_{\rm tag} = -19$ dB and -18 dB increase as the SNR value increases from -17 dB and -15 dB, respectively, but the FCD probabilities with $\gamma_{\rm tag} = -17$ dB and -16 dB are almost zero. The results indicate that the difference between $\gamma_{\rm pa}$ and $\gamma_{\rm tag}$ needs to be less than 2 dB in order to achieve high RA load monitoring accuracy and low FCD probabilities in the SNR region over -15 dB.

Fig. 8 shows the PA detection probability, the PACD probability, and the RA load monitoring accuracy with various m_1 values

[2]Assuming 50 000 nodes in a cell with an RA arrival rate of 0.5 (1 arrival/2 min), the corresponding M values become approximately 5, which is a much higher RA arrival rate than the conventional M2M applications such as smart metering report (1 arrival/15 min).

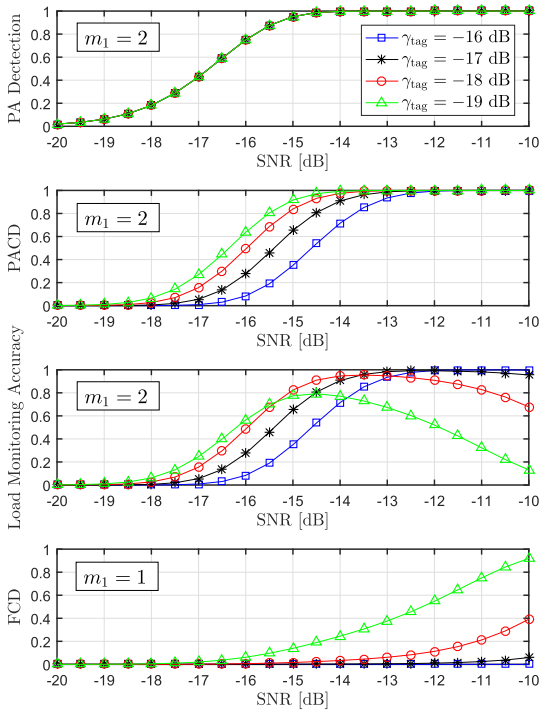

Fig. 7. PA detection probability, PACD probability, RA load monitoring accuracy ($m_1 = 2$), and FCD probability ($m_1 = 1$) when $\gamma_{\rm pa} = -16$ dB, and $\gamma_{\rm tag} = -19$ dB ~ -16 dB.

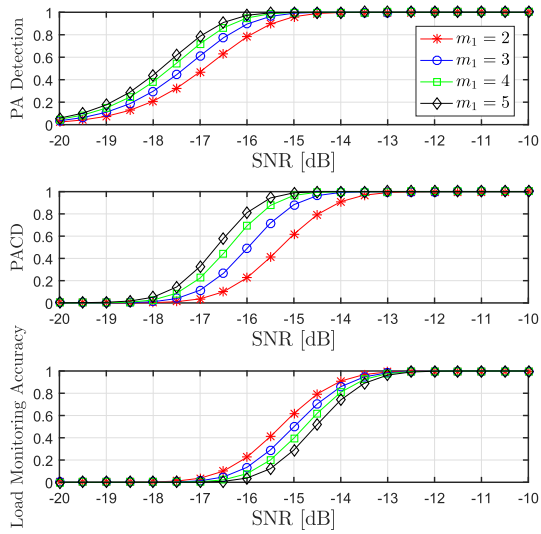

Fig. 8. PA detection probability, PACD probability, and RA load monitoring accuracy when $m_1 = 2 \sim 5$ and $\gamma_{\rm pa} = \gamma_{\rm tag} = -16$ dB.

when $\gamma_{\rm pa} = \gamma_{\rm tag} = -16$ dB. In this experiment, the number of RA-attempting nodes using PA 1 (m_1) increases from 2 with $\mathcal{I} = \{1, 1, 2, 3, 4\}$ to 5 with $\mathcal{I} = \{1, 1, 1, 1, 1\}$, and the set of five tag indices $\mathcal{L}$ is randomly picked 1000 times. For a certain SNR value, as the number of RA-attempting nodes using PA index 1 (m_1) increases, the PA detection and PACD probabilities increase, while the RA load monitoring accuracy decreases. For a target probability of 0.9, as the number of nodes increases from $m_1 = 2$ to $m_1 = 5$, the PACD probability degrades by 1.53, 1.13, 1.00, and 0.93 dB, and the RA load monitoring ac-

TABLE II
SIMULATION PARAMETERS AND VALUES FOR MAC LAYER PERFORMANCE

Parameters	Values
Number of PAs, N_{PA}	20
Number of S3CHs, N_{S3CH}	8
Number of tag spreading zones, N_{tag}	279
Number of RA nodes on a PRACH, M	$1 \sim 100$
Number of RA attempts of a device, λ	$1 \sim 12$ per 10 min
Delay time, $\{T_1^{\mathrm{prop}}, T_1^{\mathrm{conv}}, T_2, T_3, T_4\}$	{20, 50, 20, 20, 30} ms

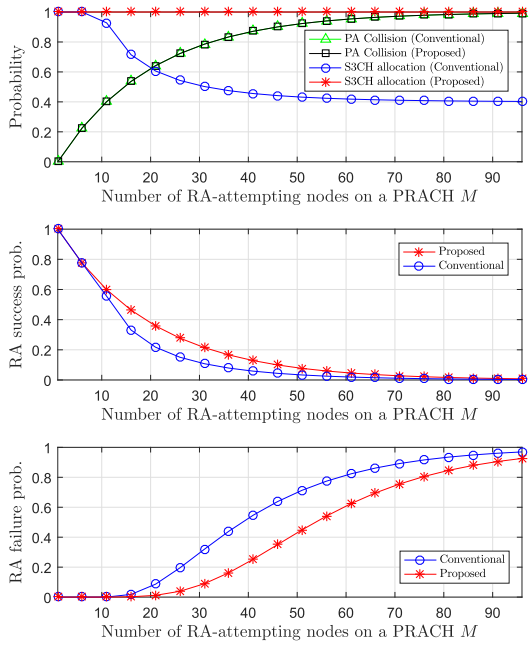

Fig. 9. PA collision probability, successful S3CH allocation rate, RA success probability, and RA failure probability for varying M.

curacy degrades by 1.53, 2.33, 2.80, and 3.26 dB, compared to each of the PA detection probabilities of 0.9. This implies that as m_1 increases, the PACD performance is enhanced while the RA load monitoring accuracy is degraded.

B. RA Performance Enhancement

In this section, from the viewpoint of MAC layer, we compare the performance of the proposed RA scheme with that of the conventional RA scheme in terms of the PA collision probability, the successful S3CH allocation rate, the RA success probability, and the RA failure probability. The average RA delay and the S3CH resource efficiency of the proposed RA scheme are also compared with those of the conventional RA scheme. Table II lists the simulation parameter set. $\{T_1^{\mathrm{prop}}, T_1^{\mathrm{conv}}, T_2, T_3, T_4\}$ and N_{DS} are set to the values specified in Table II (the same values used in [3] and [29]), respectively. The number of S3CHs is fixed to $N_{\mathrm{S3CH}} = 8$, which represents a very scarce uplink resource environment.

Fig. 9 shows the PA collision probability, the successful S3CH allocation rate, the RA success probability, and the RA failure probability of the proposed and conventional RA schemes for varying the number of RA-attempting nodes on a single PRACH M when $N_{\mathrm{PA}} = 20$ and $N_{\mathrm{S3CH}} = 8$. As shown in the figure,

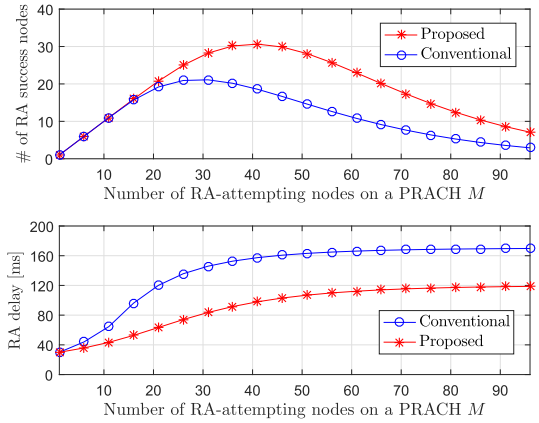

Fig. 10. Average number of RA-success nodes and the average RA delay for varying M.

the PA collision probabilities of both the proposed and conventional schemes are identical. On the other hand, the successful S3CH allocation rate of the conventional RA scheme starts to decrease from $M = 8$. It is because the given eight S3CHs are not sufficient to be allocated to all detected PAs, and nodes with nonscheduled PAs should reattempt RAs in the conventional RA scheme. On the contrary, the proposed RA scheme can always allocate S3CHs successfully with $N_{\mathrm{S3CH}} = 8$ since it does not have to allocate them to the collided PAs, but allocate them only to collision-free PAs. As a result, the RA success probability of the proposed RA scheme is higher than that of the conventional RA scheme, and the proposed RA scheme shows the lower RA failure probability than that of the conventional RA scheme.

Fig. 10 shows the number of RA-success nodes and the average RA delay. The number of RA-success nodes linearly increases until $M = 20$, which implies that almost every RA-attempting node succeeds in RA, but when $M > 20$, the number of RA-success nodes becomes less than M gradually. Accordingly, for the proposed RA scheme, 30 nodes can maximally succeed in RA when $M = 40$, while in the conventional RA scheme, the maximum number of RA-success nodes is 21 when $M = 30$. This result comes from the fact that the proposed e-PACD scheme can accurately allocate the limited S3CHs only to nodes with collision-free PAs. Next, the average RA delay is calculated from the RA delays of the RA-success nodes. The average RA delay of the conventional RA scheme rapidly increases so that it is much longer than that of the proposed RA scheme as M increases. Since the number of RA retrials is limited by 10, the average RA delays of the proposed and conventional RA schemes converge to 120 and 170 ms, respectively.

To show the long-term average performance, we assume that each of 50 000 machine nodes does λ RA attempts per 10 min on the average. Fig. 11 shows the RA failure probability and the RA delay, respectively. In those figures, λ varies from 1 to 12. The RA failure probability of the conventional RA scheme sharply increases from $\lambda = 5.5$ while the RA failure probability of the proposed scheme sharply increases from $\lambda = 7.5$. It implies that the proposed scheme can accommodate more frequent RA attempts from a massive number of machine nodes. In terms of RA delay, the proposed e-PACD scheme achieves significantly lower RA delay for more frequent RA attempts.

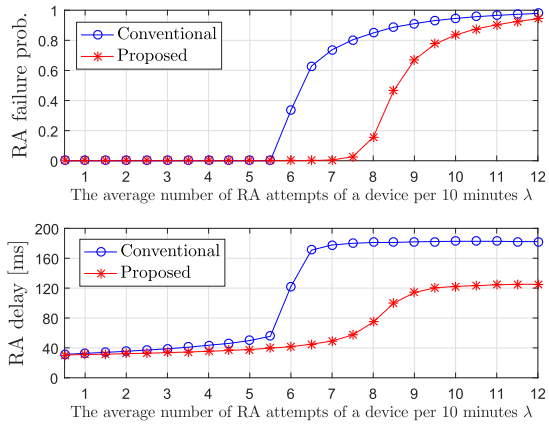

Fig. 11. Long-term average performance of RA-failure probability and RA delay for varying λ when total 50 000 machine nodes are within a cell.

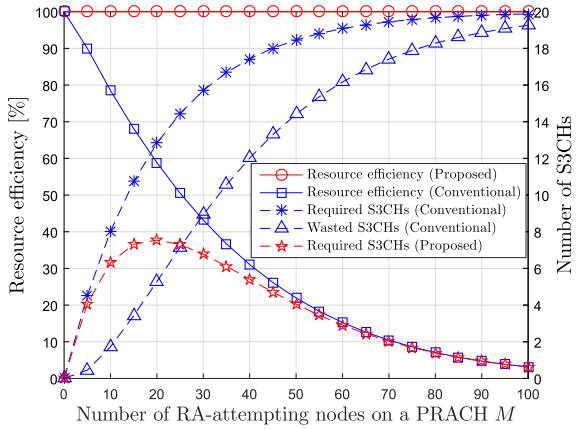

Fig. 12. S3CH resource efficiency, the total number of required S3CHs, and the number of wasted S3CHs for varying M.

For instance, 100 ms delay is reduced compared to that of the conventional scheme when $\lambda = 7$. Note that due to the accurate S3CH allocation (to only collision-free PAs) of the proposed RA scheme, the number of RA-reattempting nodes is much lower than that of the conventional RA scheme, which result in lower RA congestion in the proposed RA scheme.

Finally, Fig. 12 shows the S3CH resource efficiency and the required number of S3CHs in both conventional and proposed RA schemes, and the number of wasted S3CHs in the conventional RA scheme. The S3CH resource efficiency (%) is defined as

$$\Gamma \triangleq \frac{\text{\# of S3CHs allocated to collision-free PAs}}{\text{Total \# of required S3CHs}} \times 100.$$

The proposed RA scheme can achieve up to 100% resource efficiency through the proposed e-PACD scheme, while the conventional RA scheme degrades the resource efficiency as M increases due to an increase in the number of wasted S3CHs allocated to collided PAs. The proposed RA scheme requires maximum eight S3CHs (16 RBs) on a single RA procedure when $M = 20$ since it accurately allocates S3CHs only to the collision-free PAs. It implies that in the proposed RA scheme, the eNodeB is required to reserve a small number of S3CHs for the RA procedure since the required number of S3CHs is limited as shown by the red line with star markers. However, in the conventional RA scheme, the required number of S3CHs increases up to $20 (= N_{\mathrm{PA}})$ as M increases since it has to allocate S3CHs to all detected PAs. It implies that in the conventional RA scheme, the eNodeB is required to reserve a large number of S3CHs for the RA procedure. For example, when $M = 20$, approximately 13 S3CHs (26 RBs) are required in the conventional RA scheme, and among these, six S3CHs (12 RBs) are wasted for the collided PAs, which shows the very inefficient resource utilization of approximately 60%.

VI. Conclusion

In this paper, we proposed an e-PACD scheme which can detect PA collisions at the first step of the RA procedure based on tagged PAs. Basically, in case of a single tag detected for a PA, the eNodeB regards this PA as the successful PA transmission without a PA collision. However, the other cases, in which multiple tags are detected for a PA, can be exploited in order to detect a PA collision and monitor RA load (the number of RA-attempting nodes). The proposed e-PACD scheme reduces the average RA delay and removes resource wastes which occur at the third step of the conventional RA scheme. The reduced RA delay achieved by the proposed e-PACD scheme can contribute to an ultralow latency requirement in 5G cellular networks. Moreover, the proposed RA load monitoring can be further utilized to effectively control the RA load for future cellular M2M communications.

References

[1] *Service Requirements for Machine-Type Communications*, 3GPP TS 22.368 V13.1.0, Dec. 2014.
[2] *M2M service requirements*, ETSI TS 102 689 V2.1.1, Jul. 2013.
[3] *Study on RAN Improvements for Machine-type Communications*, 3GPP TR 37.868 V11.0.0, Sep. 2011.
[4] H. S. Jang, S. M. Kim, K. S. Ko, J. Cha, and D. K. Sung, "Spatial group based random access for M2M communications," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 961–964, Jun. 2014.
[5] H. S. Jang, S. M. Kim, H.-S. Park, and D. K. Sung, "Enhanced spatial group based random access for cellualr M2M communications," in *Proc. IEEE Int. Conf. Commun. Workshop*, London, U.K., Jun. 2015, pp. 2102–2107.
[6] T. Kim, H. S. Jang, and D. K. Sung, "An enhanced random access scheme with spatial group based reusable preamble allocation in cellular M2M networks," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1714–1717, Oct. 2015.
[7] K. S. Ko, M. J. Kim, K. Y. Bae, D. K. Sung, J. H. Kim, and J. Y. Ahn, "A novel random access for fixed-location machine-to-machine communications in OFDMA based systems," *IEEE Commun. Lett.*, vol. 16, no. 9, pp. 1428–1431, Sep. 2012.
[8] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, Jun. 2013.
[9] M.-Y. Cheng, G.-Y. Lin, H.-Y. Wei, and A.-C. Hsu, "Overload control for machine-type-communications in LTE-Advanced system," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 38–45, Jun. 2012.
[10] T.-M. Lin, C.-H. Lee, J.-P. Cheng, and W.-T. Chen, "PRADA: Prioritized random access with dynamic access barring for MTC in 3GPP LTE-A networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2467–2472, Jun. 2014.
[11] C.-H. Wei, R.-G. Cheng, and S.-L. Tsao, "Performance analysis of group paging for machine-type communications in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3371–3382, Sep. 2013.
[12] H. S. Jang, H.-S Park, and D. K. Sung, "A non-orthogonal resource allocation scheme in spatial group based random access for cellular M2M communications," *IEEE Trans. Veh. Technol.*, to be published, doi: 10.1109/TVT.2016.2606103.

[13] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, "Radio resource allocation in LTE-advanced cellular networks with M2M communications," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 184–192, Jul. 2012.
[14] T. Kwon and J.-W. Choi, "Multi-group random access resource allocation for M2M devices in multicell systems," *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 834–837, Jun. 2012.
[15] C. Y. Ho and C.-Y. Huang, "Energy-saving massive access control and resource allocation schemes for M2M communications in OFDMA cellular networks," *IEEE Wireless Commun. Lett.*, vol. 1, no. 3, pp. 209–212, Jun. 2012.
[16] A. Gotsis, A. Lioumpas, and A. Alexiou, "M2M scheduling over LTE: Challenges and new perspectives," *IEEE Veh. Technol. Mag.*, vol. 7, no. 3, pp. 34–39, Sep. 2012.
[17] D. T. Wiriaatmadja and K. 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.
[18] S. Andreev *et al.*, "Efficient small data access for machine-type communications in LTE," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2013, pp. 3569–3574.
[19] H. S. Jang, S. M. Kim, H.-S Park, and D. K. Sung, "Message-embedded random access for cellular M2M communications," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 902–905, May 2016.
[20] C.-H. Wei, G. Bianchi, and R.-G. Cheng, "Modeling and analysis of random access channels with bursty arrivals in OFDMA wireless networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 1940–1953, Apr. 2015.
[21] N. Zhang, G. Kang, J. Wang, Y. Guo, and F. Labeau, "Resource allocation in a new random access for M2M communications," *IEEE Commun. Lett.*, vol. 19, no. 5, pp. 843–846, May 2015.
[22] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
[23] *Study on RAN Improvements for Machine-Type Communications*, 3GPP TS 37.868 V11.0.0, Oct. 2011.
[24] S. Sesia, I. Toufik, and M. Baker, *LTE–The UMTS Long Term Evolution From Theory to Practice*. New York, NY, USA: Wiley, 2009.
[25] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.
[26] D. Katselis, "Some preamble design aspects in CP-OFDM systems," *IEEE Commun. Lett.*, vol. 16, no. 3, pp. 356–359, Mar. 2012.
[27] M. M. U. Gul, X. Ma, and S. Lee, "Timing and frequency synchronization for OFDM downlink transmissions using Zadoff-Chu sequences," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1716–1729, Mar. 2015.
[28] 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.
[29] R. Jain, "Channel models—A tutorial," in *Proc. WiMAX Forum AATG*, 2007.
[30] S. Y. Lien, T. H. Liau, C. Y. Kao, and K. C. Chen, "Cooperative access class barring for machine-to-machine communications," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 27–32, Jan. 2012.
[31] S. Duan, V. Shah-Mansouri, and V. W. S. Wong, "Dynamic access class barring for M2M communications in LTE networks," in *Proc. IEEE Global Commun. Conf.*, Atlanta, GA, USA, Dec. 2013, pp. 4747–4752.
[32] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. J.*, vol. 24, no. 1, pp. 46–156, 1945.

Han Seung Jang (S'14) received the B.S degree in electronics and computer engineering from Chonnam National University, Gwangju, South Korea, in 2012 and the M.S. degree in electrical engineering from Korea Advanced Institute for Science and Technology (KAIST), Daejoen, South Korea, in 2014. Since March 2014, he has been working toward the Ph.D. degree with the School of Electrical Engineering, KAIST.

His research interests include cellular machine-to-machine communications, smart grid, renewable energy forecasting, and electric vehicle charging management.

Su Min Kim (S'05–M'13) received the B.S. degree in electronics engineering from Inha University, Incheon, South Korea, in 2005 and the M.S. and Ph.D. degrees in electrical engineering and computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejoen, Korea, in 2007 and 2012, respectively.

Since Mach 2015, he has been an Assistant Professor with the Department of Electronics Engineering, Korea Polytechnic University, Siheung, South Korea. He was a Postdoctoral Follow with the Department of Electrical Engineering, KAIST, in 2012 and with the School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden, from 2012 to 2014. Between 2014 and 2015, he was an Experienced Researcher with Radio Access Technology, Ericsson Research, Kista, Sweden. His research interests include next-generation mobile communication systems, hybrid automatic repeat request protocols, radio resource management, interference management, cooperative and buffer-aided relaying communications, cognitive radio communications, machine type communications, physical layer security, and statistical signal processing.

Dr. Kim received the Paper Award from the Next Generation PC International Conference in 2005 and the Silver and Bronze Awards from the 17th and 18th Samsung Humantech Thesis Prizes in 2011 and 2012, respectively.

Hong-Shik Park (M'98) received the B.S. degree from Seoul National University, Seoul, South Korea, in 1977 and the M.S. and Ph.D. degrees from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 1986 and 1995, respectively, all in electrical engineering.

In 1977, he joined the Electronics and Telecommunications Research Institute and had been engaged in development of the TDX digital switching system family, including TDX-1, TDX-1A, TDX-1B, TDX-10, and ATM switching systems. In 1998, he moved to Information and Communications University, Daejeon, as a member of Faculty. He is currently a Professor with the School of Electrical and Electronics Engineering, KAIST. From 2004 to 2012, he was the Director of BcN Engineering Research Center sponsored by KEIT, Korea. His research interests are network architecture and protocols, traffic engineering, and performance analysis of telecommunication systems.

Dr. Park is a member of the IEEK and KICS, South Korea.

Dan Keun Sung (S'80–M'86–SM'00–F'15) received the B.S. degree in electronics engineering from Seoul National University, 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 with the School of Electrical Engineering. From 1996 to 1999, he was the Director of the Satellite Technology Research Center, KAIST. He had served as Division Editor of the *Journal of Communications and Networks* during 1998–2007. He also had served as Editor of IEEE COMMUNICATIONS MAGAZINE during 2002–2011. His research interests include mobile communication systems and networks with special interest in resource management, cellular machine-to-machine communications, smart grid, energy networks, energy ICT, wireless local area networks, wireless personal area networks, traffic control in wireless and wired networks, performance and reliability of communication systems, and microsatellites.

Dr. Sung is a member of the National Academy of Engineering of Korea. He received 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, and the 2013 Haedong Academic Grand Award from the Korean Institute of Communications and Information Sciences.