Available online at www.sciencedirect.com
ScienceDirect
ICT Express 1 (2015) 97–101
www.elsevier.com/locate/icte

A novel root-index based prioritized random access scheme for 5G cellular networks☆,☆☆

Taehoon Kim, Han Seung Jang, Dan Keun Sung*

School of Electrical Engineering, KAIST, Daejeon, Republic of Korea

Received 31 August 2015; accepted 13 October 2015
Available online 2 December 2015

Abstract

Cellular networks will play an important role in realizing the newly emerging Internet-of-Everything (IoE). One of the challenging issues is to support the quality of service (QoS) during the access phase, while accommodating a massive number of machine nodes. In this paper, we show a new paradigm of multiple access priorities in random access (RA) procedure and propose a novel root-index based prioritized random access (RIPRA) scheme that *implicitly embeds the access priority in the root index of the RA preambles*. The performance evaluation shows that the proposed RIPRA scheme can successfully support differentiated performance for different access priority levels, even though there exist a massive number of machine nodes.

Keywords: Internet-of-Everything; M2M; Random access; Zadoff–Chu sequence; Access priority

1. Introduction

Internet-of-Everything (IoE) is a newly emerging concept that implies that everything including humans, machines, processes, and data are all connected through a network. Due to the huge market potential and wide range of applications of machine-to-machine (M2M) communications such as smart metering, emergency alarm, surveillance and security, and e-health, a number of studies have investigated M2M communications as one of the fundamental elements for realizing IoE in cellular networks [1,2].

In cellular networks, random access (RA) is used for connecting and synchronizing user equipment (UEs) with an eNodeB [3]. Since the RA procedure in cellular networks uses a variation of slotted ALOHA protocol [4], as the number of contending UEs increases, collision probability may significantly increase. There have been a number of studies on efficient RA schemes to mitigate the collision problem during the RA procedure [5–7]. Utilizing these schemes, the network may accommodate more UEs during the access phase with low collision probability; however, there still exists an inherent drawback that they may not support multiple priority levels properly during the access period.

Before the concept of cellular M2M communications emerged, since there were only a relatively small number of UEs for human-to-human (H2H) communications, it did not take much time to access the networks even though the RA procedure could not guarantee a reliable access delay. Therefore, the cellular network considers the QoS after the completion of the access phase (i.e., RA procedure), which seems to be a reasonable approach for supporting QoS. However, this approach should be enhanced for future cellular networks, since most UEs may hardly access the networks when accommodating a massive number of machine nodes for M2M communications, and, thus, it may be hard to support the required QoS. As a result, *a new paradigm will be required in which the future cellular networks should deal with the QoS from the access phase*, not after the access phase. In other words, the future cellular networks should be able to support '***access priority***', which is defined as the *priority during the access phase* (i.e., the RA procedure). Using the concept of access priority, we can support

* Corresponding author.
E-mail addresses: thkimm@kaist.ac.kr (T. Kim), jhans@kaist.ac.kr (H.S. Jang), dksung@kaist.ac.kr (D.K. Sung).

Peer review under responsibility of The Korean Institute of Communications Information Sciences.

☆ This work was supported by the Korea government (MSIP) under the National Research Foundation of Korea (NRF) Grant 2014R1A2A2A01005192.

☆☆ This paper is part of a special issue entitled "Next Generation (5G/6G) Mobile Communications" guest-edited by Prof. Jungwoo Lee, Dr. Sumei Sun, Prof. Huaping Liu, Prof. Seong-Lyun Kim and Prof. Wan Choi.

http://dx.doi.org/10.1016/j.icte.2015.10.003
2405-9595/

different QoSs for various service applications from the access phase.

To our best knowledge, a few studies have dealt with *access priority in cellular networks* [8,9]. The key idea of these studies is to additionally utilize the power domain as well as the time–frequency domain during the access phase. However, there are some limitations in generalizing these schemes to support multiple access priorities due to a limitation in power amplifiers at UEs and machine nodes, and the RA performance may degrade due to uncertainty in wireless channels.

In this paper, we show a new paradigm of multiple access priorities in RA procedure, where each RA preamble should be able to *indicate its own access priority level* during the RA procedure. We propose a novel root-index based prioritized random access (RIPRA) scheme that *implicitly embeds the access priority in the root index of the RA preambles*. The performance evaluation shows that the proposed RIPRA scheme can successfully support differentiated performance for different access priority levels, even though there exist a massive number of machine nodes.

2. Conventional random access scheme

The conventional RA scheme consists of four steps [3]. In the first step, each node randomly generates a preamble signature and transmits it on a physical random access channel (PRACH). A Zadoff–Chu (ZC) sequence is used to generate RA preamble signatures, defined as $z_r[n] \triangleq \exp[-j\pi \cdot r \cdot n \cdot (n+1)/N_{ZC}]$ for $n = 0, \ldots, N_{ZC}-1$, where $r \in \{1, \ldots, N_{ZC}-1\}$ is the root index and N_{ZC} is the sequence length [10]. The ZC sequence is one of the constant amplitude zero auto-correlation (CAZAC) sequences, which mainly show the following two properties: One is an auto-correlation property between two ZC sequences with the same root index r, which yields a delta function, i.e., $|c_{rr}[\sigma]| = \left|\sum_{n=0}^{N_{ZC}-1} z_r[n] z_r^*[n+\sigma]\right| = N_{ZC} \cdot \delta[\sigma]$, where $c_{rr}[\sigma]$ is the discrete cyclic auto-correlation function of $z_r[n]$ at lag σ and $(\cdot)^*$ denotes the complex conjugate. The other is that the magnitude of the cross-correlation between any two ZC sequences with different root indices, r and s, is constant, i.e., $|c_{rs}[\sigma]| = \left|\sum_{n=0}^{N_{ZC}-1} z_r[n] z_s^*[n+\sigma]\right| = \sqrt{N_{ZC}},\ r \neq s$. In principle, $N_{PA}^{root}(d) (= \lfloor N_{ZC}/N_{CS}(d) \rfloor)$ preambles can be generated from a single ZC sequence with a given root index, where $N_{CS}(d)$ is the size of the cyclic shift, which highly relies on the cell radius, d. In other words, multiple RA preambles can be generated from a single ZC sequence by cyclically shifting an integer multiple of $N_{CS}(d)$. For example, the sequence shifted by $i \cdot N_{CS}(d)$, where $i \in \{0, \ldots, \lfloor N_{ZC}/N_{CS}(d) \rfloor - 1\}$, can be represented as $z_{r,i}[n] = z_r[(n + i \cdot N_{CS}(d)) \bmod N_{ZC}]$. Conventionally, there is a required number of RA preambles in a single cell, $N_{PA,req}$. If $N_{PA}^{root}(d) < N_{PA,req}$, the eNodeB needs to utilize more root indices in order to provide $N_{PA,req}$. $N_{root}^{conv}(d) (\triangleq \lceil N_{PA,req}/N_{PA}^{root}(d) \rceil)$ denotes the required number of root indices for generating $N_{PA,req}$ preambles. Finally, it is worth noting that $N_{root}^{conv}(d)$ only relies on d. In short, in this step, each node randomly selects one RA preamble among $N_{PA,req}$ preambles.

In the second step, if the preamble has been received correctly, the eNodeB sends an RA response (RAR) message through the physical downlink control channel (PDCCH). The RAR message includes the RA preamble identifier (RAPID), timing alignment (TA), uplink grant (UG), and cell-radio network temporary identifier (C-RNTI). In the third step, the node transmits a scheduled message on the physical uplink shared channel (PUSCH). Once the message is transmitted, the node starts a contention resolution (CR) timer for determining a collision for the previous transmission of a preamble. In the fourth step, the eNodeB echoes the identifiers of the node through the physical downlink shared channel (PDSCH). If the node does not receive any messages within the CR timer, then it regards the RA as a failure and retries the RA after performing a back-off mechanism. Otherwise, it regards the RA as a success.

3. Proposed random access scheme

We propose a novel root-index based prioritized random access (RIPRA) scheme that *implicitly embeds the access priority in the root index of the RA preambles*. In principle, the proposed RIPRA scheme cannot only support access priority, but also allow some access priority levels to follow a different access procedure, for example, directly supporting *connection-less data transmissions*[1] during the RA procedure. However, here, we mainly focus on the *access priority mechanism* of the proposed RIPRA scheme, where the objective is to connect with the eNodeB while supporting access priority.

In the proposed RIPRA scheme, the eNodeB can support up to K access priorities, if $\sum_{k=1}^{K} N_{root}^{k}(d) \leq N_{ZC}$ satisfies, where $N_{root}^{k}(d)$ denotes the number of root indices assigned for the kth access priority level, which also highly relies on cell radius, d. Note that $N_{root}^{k}(d)$ may be different from $N_{root}^{l}(d)$, where $k \neq l$, since $N_{root}^{k}(d)$ is determined by considering $N_{PA,req}^{k}$, which is the number of RA preambles required to satisfy the target collision probability of the kth access priority level. If $N_{PA}^{root}(d) < N_{PA,req}^{k}$, then more root indices should be allocated to the kth access priority level until $N_{PA}^{k}(d) \geq N_{PA,req}^{k}$ satisfies, where $N_{PA}^{k}(d) (= N_{PA}^{root}(d) \cdot N_{root}^{k}(d))$ is the number of RA preambles assigned for the kth access priority level. $N_{root}^{prop}(d, K) (\triangleq \sum_{k=1}^{K} N_{root}^{k}(d))$ denotes the total number of used root indices in a single cell and it is worth noting that the $N_{root}^{prop}(d, K)$ relies on K as well as d. According to $N_{root}^{prop}(d, K)$, the preamble detection threshold needs to be optimized in order to achieve a reasonably low false-alarm probability and mis-detection probability during the preamble detection procedure; however, this is out of scope of this paper.

Fig. 1 shows the overall procedure of the proposed RIPRA scheme. The eNodeB broadcasts the configuration parameters such as RA parameters and access priority information. Especially, the access priority information includes the total number of access priorities, the set of root indices

[1] Using the connection-less data transmissions, each node can transmit a small packet during the RA procedure (in step 3). To support this mode, the eNodeB should assign a number of root indices to the corresponding access mode.

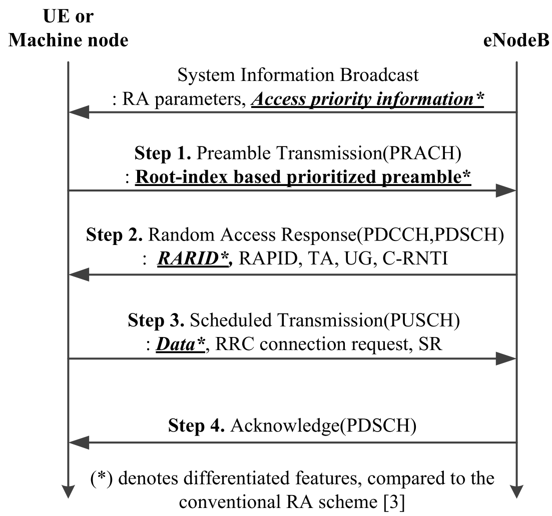

Fig. 1. Procedure of the proposed RIPRA scheme.

corresponding to each of the access priorities, and the "application-access priority" mapping table.

In the first step, based on the broadcast mapping table, each node determines its access priority according to its application. If the access priority level is determined as k, the node randomly selects a root index among $\mathbf{r}^k = \{r_1^k, r_2^k, \ldots, r_{N_{\text{root}}^k(d)}^k\}$, where $\mathbf{r}^k$ denotes a set of root indices assigned to the kth access priority level, and randomly generates a preamble signature by cyclically shifting a ZC sequence from the selected root index. If the node selects r_i^k and j as a root index for the ZC sequence and an integer multiple of cyclic shift, respectively, the generated preamble signature can be expressed as $z_{r_i^k,j}[n] = z_{r_i^k}[(n + j \cdot N_{\text{CS}}(d)) \bmod N_{\text{ZC}}]$ for $n = 0, \ldots, N_{\text{ZC}} - 1$, where $j \in \{0, \ldots, \lfloor N_{\text{ZC}}/N_{\text{CS}}(d) \rfloor - 1\}$.

In the second step, the eNodeB detects which preambles are transmitted. During the detection, the eNodeB should utilize $N_{\text{root}}^{\text{prop}}(d, K)$ root indices. For example, by using all reference ZC sequences generated from each of the root indices in the set $\mathbf{r}^k$, the eNodeB can detect the preambles indicating the kth access priority. Especially, when the eNodeB detects the preambles with a reference ZC sequence generated from the root index, r_i^k, if there exist any peaks in the jth preamble detection zone, then it regards that a node has performed the RA procedure with $z_{r_i^k,j}[n]$ to indicate its kth access priority. The eNodeB does not transmit any RAR messages until all detection procedures are completed. After detecting preamble signatures with $N_{\text{root}}^{\text{prop}}(d, K)$ root indices, the eNodeB sorts the detected preambles according to the access priority. Thereafter, the eNodeB generates RAR messages in sequence from the detected preambles with the highest access priority level to the detected preambles with the lowest access priority level, and transmits them on the PDCCH. Based on this *ordered-response mechanism*, the RAR messages for the detected preambles with higher access priority levels are likely to be served in advance, even though there are insufficient downlink resources for transmitting RAR messages and uplink resources for transmitting scheduled messages in the subsequent step. Note that there may be multiple RAR messages including the same preamble identifier because multiple root indices are utilized. In order to distinguish these RAR messages, we need to modify the RAR message format as follows: Even though the preamble identifiers are the same, since the root indices may be different, the RAR message format should be reconfigured to contain the **RA root index (RARID)** as well as the RAPID, TA, UG, and C-RNTI.

In the third step, the node transmits a *scheduled message*[2] on the PUSCH. Once the message is transmitted, the node starts a CR timer to determine whether a collision occurs or not. When two or more nodes select an identical preamble with the same root index in the first step, they utilize the same uplink resources in this step, which results in a collision.

In the fourth step, the eNodeB acknowledges the identifiers of the nodes, which successfully transmit their messages in the third step, on the PDSCH. If the node does not receive any messages within the CR timer, then it regards the RA as a failure and reattempts the RA after performing a back-off procedure. Otherwise, it regards the RA as a *success*.[3]

4. Performance evaluation

4.1. Numerical analysis

If two or more nodes transmit an identical preamble with the same root index in the same RA slot, they receive the same RAR message, and, thus, they utilize the same uplink resources for the transmission of the scheduled message, which eventually results in a collision. The collision probability is defined as the probability that two or more nodes select an identical preamble with the same root index in the same RA slot. The collision probability of the RA attempt with the kth access priority can be derived as [6]:

$$p_{\text{col}}^k = 1 - \exp\left[W\left(\ln\left(1 - \frac{1}{N_{\text{PA}}^k}\right) \cdot \lambda_k \cdot N_k \cdot T_{\text{PRACH}}\right)\right], \tag{1}$$

where $W(\cdot)$ denotes the Lambert W function and N_{PA}^k, λ_k, N_k, and T_{PRACH} represent the number of preambles assigned to the kth access priority, the arrival rate of RA with the kth access priority, the number of nodes attempting RAs with the kth access priority, and the period of PRACH, respectively.

Furthermore, the access delay can be defined as the time duration from a new packet generation to the successful completion of the RA procedure. The average access delay of the RA attempt with the kth access priority, $\overline{T^k}$, can be derived as follows:

$$\overline{T^k} \simeq \overline{T_{\text{Q}}^k} + \overline{T_{\text{RA}}^k} \cdot \left\{\sum_{l=1}^{\text{maxTx}} l \cdot (p_{\text{col}}^k)^{(l-1)} \cdot (1 - (p_{\text{col}}^k)^l)\right\}, \tag{2}$$

2 A radio resource control (RRC) connection request message or scheduling request (SR) message can be transmitted. Especially for connection-less data transmissions, actual data can be delivered through this step.

3 For connection-less data transmissions, after receiving the corresponding acknowledgement message in step 4, each node terminates its RA. Furthermore, the format of acknowledgement for the connection-less data transmission mode may be modified.

Table 1
System parameters and values.

Parameters	Values
Cell radius (d)	0.7 (km)
N_{ZC}, $N_{CS}(d = 0.7)$	839, 13
The number of UEs (N_{UE})	100
The number of machine nodes (N_M)	0 ∼ 50, 000
RA arrival rate of UEs (λ_{UE})	1/300 (s^{-1})
RA arrival rate of machine nodes (λ_M)	1/60 (s^{-1})
The number of access priorities (K)	3
The number of root indices (N^k_{root}, $\forall k$)	1
The number of available preambles (N^k_{PA}, $\forall k$)	64
Maximum number of preamble transmissions (maxTx)	10
PRACH period (T_{PRACH})	10 (ms)
Back-off Indicator (W)	20 (ms)
RA Response Windows	5 (ms)
Contention resolution timer	48 (ms)

where $\overline{T^k_Q}$, $\overline{T^k_{RA}}$, and maxTx represent the average queuing delay, the time required to attempt or reattempt the RA procedure considering the back-off procedure and the maximum number of preamble transmissions, respectively.

4.2. Numerical results

Table 1 lists the system parameters based on [3]. We assume that the location of nodes follows a uniform distribution and the RA inter-arrival time follows an exponential distribution. The number of allocated RA slots is one sub-frame of 1 ms within a radio-frame of 10 ms. Moreover, we assume that there are no mis-detections and false alarms during the preamble detection. We consider a group of UEs for H2H communications and two groups of machine nodes for M2M communications: a group of machine nodes with a high priority level and a group of machine nodes with a low priority level, which correspond to 70% and 30% of the total machine nodes, respectively. We assume that the first, second, and third access priority levels correspond to a group of UEs, a group of machine nodes with a high priority level, and a group of machine nodes with a low priority level, respectively. We mainly observe the effect of machine nodes on the performance of both UEs and machine nodes themselves.

Fig. 2 shows the collision probability for varying the number of machine nodes in a cell, N_M. With the conventional RA scheme, as N_M increases, the collision probabilities of both UEs and machine nodes increase and they are identical to each other. Since both the UEs and machine nodes utilize the RA preambles generated from the same root indices, they have the highest collision probability regardless of the type of nodes, UEs or machine nodes, as N_M increases. On the other hands, with the proposed RIPRA scheme, even though N_M increases, the collision probabilities of the UEs are not affected by the machine nodes due to the good cross-correlation property of the ZC sequences with different root indices (in the figure, collision probabilities of the UEs are depicted as a constant line, since it is constant regardless of N_M).

Fig. 3 shows the average access delay for varying N_M. With the conventional RA scheme, as N_M increases, the average access delay of both the UEs and machine nodes simultaneously increases due to the effect of a massive number of machine nodes, since both UEs and machine nodes utilize the RA preambles generated from the same root indices. However, with the proposed RA scheme, we can observe a differentiated average access delay performance for different access priority levels as N_M increases.

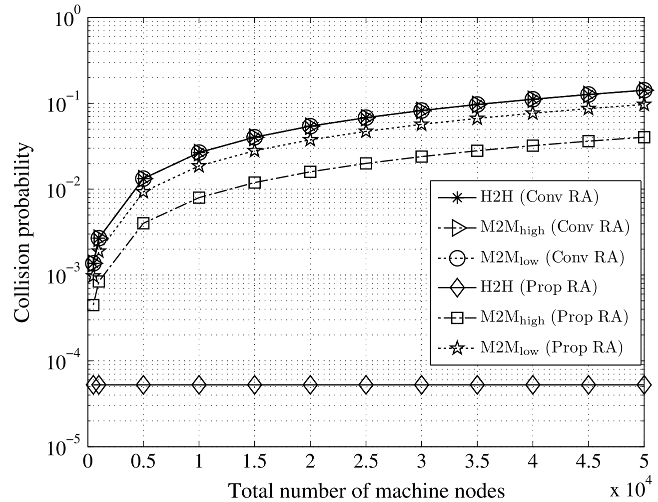

Fig. 2. Total number of machine nodes vs. collision probability when $N_{UE} = 100$, $\lambda_{UE} = 1/60$ (s^{-1}), $\lambda_M = 1/300$ (s^{-1}) and $N^k_{PA} = 64$, $\forall k$.

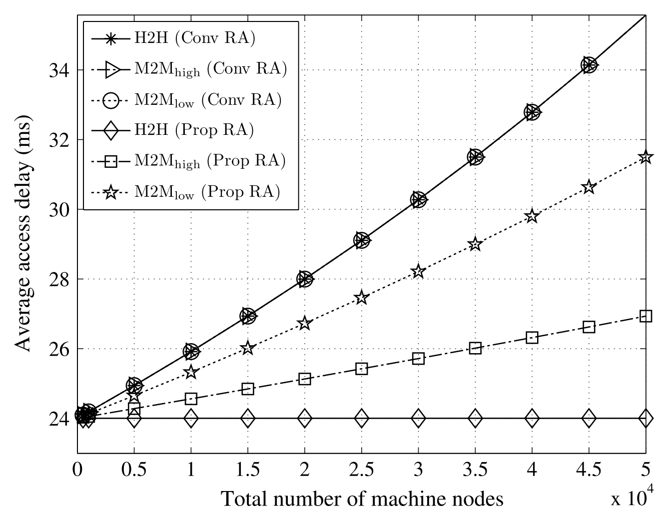

Fig. 3. Total number of machine nodes vs. average access delay when $N_{UE} = 100$, $\lambda_{UE} = 1/60$ (s^{-1}), $\lambda_M = 1/300$ (s^{-1}) and $N^k_{PA} = 64$, $\forall k$.

5. Conclusions

In this paper, we showed a new paradigm of multiple access priorities in the RA procedure, where each RA preamble *indicates its access priority* during the RA procedure. We proposed a novel root-index based prioritized random access (RIPRA) scheme that *implicitly embeds the access priority in the root index of the RA preambles*. The performance evaluation shows that the proposed RIPRA scheme can successfully support differentiated performance for different access priority levels, even though there exist a massive number of machine nodes.

References

[1] Service Requirements for Machine-Type Communications, 3GPP TS 22.368 V10.1.0, June 2010.

[2] RAN Improvements for Machine-type Communications, 3GPP TR 37.868 V11.1.0, October 2011.

[3] Evolved Universal Terrestrial Radio Access Network(E-UTRAN); Medium Access Control (MAC) protocol specification, 3GPP TS 36.321 V11.2.0, March 2013.

[4] A. Laya, L. Alonso, J. Alonso-Zarate, Is the random access channel of LTE and LTE-A suitable for M2M communications? A survey of alternatives, IEEE Trans. Commun. Surv. Tuts 16 (1) (2014) 4–16.

[5] K.S. Ko, M.J. Kim, K.Y. Bae, D.K. Sung, J.H. Kim, J.Y. Ahn, A novel random access for fixed-location machine-to-machine communications in OFDMA based systems, IEEE Commun. Lett. 16 (9) (2012) 1428–1431.

[6] H.S. Jang, S.M. Kim, K.S. Ko, J. Cha, D.K. Sung, Spatial group based random access for M2M communications, IEEE Commun. Lett. 18 (6) (2014) 961–964.

[7] T. Kim, H.S. Jang, D.K. Sung, An enhanced random access scheme with spatial group based reusable preamble allocation in cellular M2M networks, IEEE Commun. Lett. 19 (10) (2015) 1714–1717.

[8] T. Kim, K.S. Ko, I. Bang, D.K. Sung, A random access scheme based on a special preamble for supporting emergency alarms, in: Proc. IEEE WCNC, April 2014.

[9] T. Kim, K.S. Ko, D.K. Sung, Prioritized random access for machine-to-machine communications in OFDMA based systems, in: Proc. IEEE ICC, June 2015.

[10] D. Chu, Polyphase codes with good periodic correlation properties, IEEE Trans. Inform. Theory 18 (4) (1972) 531–532.