LTE-Advanced Random Access Channel Congestion Detection Method for IoT

Goni Mahamadou Bouba[1], Jerôme Mbainaibeye[2],
James Kouawa Tamgno[3](✉), and Claude Lishou[1]

[1] LTI Laboratory, ESP/UCAD, Dakar, Senegal
bouba.mahamadou@ucad.edu.sn
[2] Université de DOBA, Doba, Tchad
mbai_jerome@yahoo.com
[3] LTI & RSI Research Group, ESMT, Dakar, Senegal
james.tamgno@esmt.sn

Abstract. The Long Term Evolution - LTE - is one of the very last evolutions in mobile communication systems that offer a much wider bandwidth than its predecessors. That is why it is very much in demand for a massive deployment of the Internet of Things (IoT) also called Machine to Machine communication or Machine Type Communication (MTC). With the IoT, the network is subject to recurrent congestion when densely charged which is due to increased uplink solicitation. MTC devices must complete the RACH process to access the network. Collisions occur during this process that leads to the congestion which, in turn, has a negative impact on the quality of service. The Third Generation Partnership Project (3GPP) provided some solutions to alleviate the problem. In this paper we propose a congestion detection method since 3GPP only proposed contention resolution methods. We first determine the interval of use of preambles during which the success rate is the highest. By doing so, we determine the maximal preamble utilization threshold (Rlimit) beyond which quality of service is no more guaranteed. The novelty with this method is that once Rlimit threshold is reached, a contention resolution scheme could be activated and will remain so until the threshold drops below Rlimit. Our method can give better results if applied to contention resolution methods. Moreover it is simple, less complex and easy to implement in the LTE. Moreover, it does not require large investments.

Keywords: Machine Type Communication (MTC) · Long Term Evolution (LTE) · Radio Access Network (RAN) overload · Random Access Channel (RACH) · Congestion

1 Introduction

The Internet of Things (IoT) is a recent communication paradigm that envisions a near future, in which everyday objects will be equipped with microcontrollers, transceivers for digital communication, and appropriate protocol stacks that will make them able to communicate with each other, with users or with a remote server, becoming part of the Internet [1]. These objects are able to collect, store, transmit and process data from the

G. Mendy et al. (Eds.): AFRICOMM 2018, LNICST 275, pp. 45–55, 2019.
https://doi.org/10.1007/978-3-030-16042-5_5

physical world. It is a paradigm that finds its application in many different areas, such as home automation, industry, medical aid, mobile health care, help for the elderly, intelligent energy management and smart grids, automobiles, agriculture, traffic management, and many others [2]. Several standardization bodies, among which IEEE and 3GPP, are working to set standards for it [3]. Its deployment on the LTE network is exponentially increasing, which is not without causing enormous challenges when we know that the LTE is designed for basic Human-to-Human (H2H) communications type (large downlink bandwidth and narrow uplink bandwidth). The IoT is meanwhile very greedy in the uplink band as the MTC devices transmit much more packets than they receive. We are talking about 26 [4] to near 50 [5] billion of connected objects by 2020. MTC devices have to compete for resources to get access to network. This is done through RACH process where congestion often happens.

2 Background

2.1 Random Access Procedure

In the LTE system, access to the network is through a RACH process in which UEs use preambles broadcasted at regular time slots by the base station (a total of 64 preambles). The preambles are generated from the sequences of the Zadoff-Chu algorithm include a cyclic prefix CP, a sequence and a guard time as in Fig. 1.

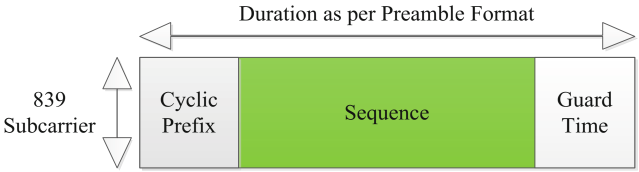

Fig. 1. RA preamble structure [6]

Two types of access exist in LTE:

- **Contention-free access:** Among the 64 preambles, 10 are dedicated to specific uses of high priorities in contention-free access. During contention-free access, the connection is initiated by the base station which, at the same time, provides the UE with the necessary resources. This is applied to priority communications such as emergency alert messages and specific uses.
- **Contention-based access:** In the case of contention-based access, UEs compete for the remaining 54 preambles in the RACH process. The random access request consists of this preamble, which is a digital signature transmitted by the UEs in a time slot. The RACH process is consist of four steps [7]:

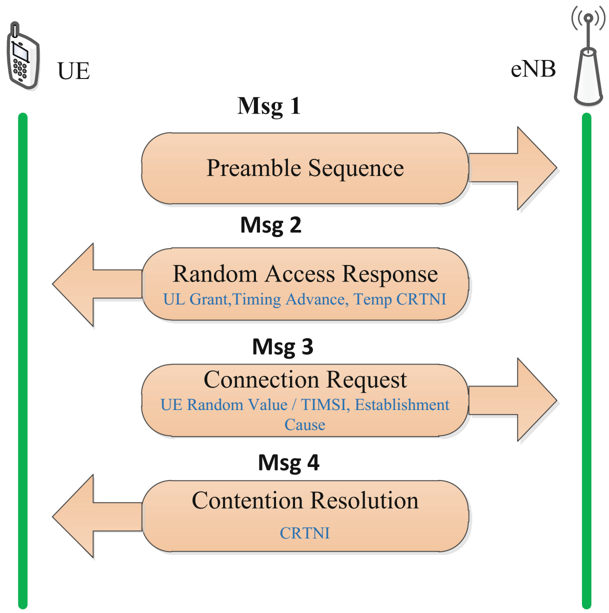

Fig. 2. RACH process

Figure 2 describes the four steps of the Random Access Channel process:

Step 1 (Preamble Send): In this step, each UE sends the access request by sending one of the 54 orthogonal predefined preambles, as well as a temporary identity RA-RNTI (Random Access - Radio Network Temporary Identifier) which is actually based on the time interval in which the preamble is issued.

Step 2 (Random Access Response): In this step, the base station transmits the access response that contains the detected preamble index, the timing for step 3, the time offset (so the UE can modify its schedule to compensate for round-trip delay), and the uplink resources necessary for UE to perform step 3.

Step 3 (Connection Request): After obtaining the resources in Step 2, the UE sends a connection request to the base station. This message contains the identity of the cell (C-RNTI) in which the UE is located and the reason of the request.

Step 4 (Contention Resolution): The base station responds with a contention resolution message. Each device that has received this message compares the identity in the message with the identity passed in the previous step. In case of correspondence between these identities access is granted. In case of non-correspondence, the UE back-off and go back to step 1.

When two or more UEs use same preamble collisions can be detected by the base station, based on the difference in preamble transmission delays. Then it will not send a response for this preamble. The UEs concerned will then be required to resume the operation. However, if these UEs are equidistant from the base station, the collision will not be detected and the response from step 2 will be sent to all UEs having used this preamble. In this case, the collision is resolved by the contention resolution in step 4. Only the retained UE will have access to the network, the others are led to resume the operation for a given number of iterations.

2.2 Congestion Resolution Methods

In order to solve network overload problems due to a very high number of requests from MTC devices and recurrent collisions that occur in these high-load environments causing congestion during access to LTE, 3GPP proposed the following solutions [8]:

(1) Access Class Barring: In the ACB, the UEs are divided into 16 classes. Classes 0–9 are called normal classes; class 10 is dedicated to emergencies while classes 11–15 are dedicated to specific uses of high priority.

The principle of the ACB consists in the fact that the base station (eNodeB) broadcasts at regular time slots a probability **p** ($p \in [0-1]$) called ACB factor towards all the UEs belonging to classes 0–9. The UEs accessing the network generate a number q ($q \in [0-1]$). If the q generated number by the UE is less than p ($q < p$), UE is allowed to proceed with RACH process. Otherwise, it has to wait a $T_{barring}$ time (barring time) before resuming the process. By this way it is possible for the base station to control the collisions and network overloads by assigning an optimal value to p. $T_{barring}$ can be calculated as follows [8]:

$$T_{\mathrm{barring}} = (0.7 + 0.6 * \mathrm{rand}) * \mathrm{ac\_BarringTime}$$

Where rand is a random number generated by the MTC device after passing a first failed ACB check and before a second attempt. The values of ***ac_BarringTime*** can range from 4 s to 512 s.

Several improved versions of the ACB have been performed to increase its performance. The separate approach of ACB for M2M and H2H [9, 10]. Improvements of Extended Access Barring (EAB) have also been proposed in [11, 12]. For UEs under cover of several base stations, a cooperative approach has been proposed [13] to allow an optimal choice of E-NodeB for the EU. [14, 15] provide priority random access joined to the dynamic ACB mechanism to improve the performance of the random access channel. The improvement of the ACB in most cases leads to an increase in the access time on which, once a certain threshold is reached can be a real problem. In order to overcome this problem, the authors of [16] have developed a scalable ACB system based on the game is proposed.

(2) Separate RACH Resources for MTC: This scheme separates resources for H2H and M2M. When resources are not shared, the network is subject to recurring congestion. The separation of RACH resources between H2H and M2M reduces the impact of each other. A study of the separation of resources is done in [17]. It proposes 2 methods: First method, called “Method 1”, consists of completely dividing all available preambles into two disjoint subsets. The other method, called “Method 2”, also consists of dividing the set into two subsets, but one of them is shared by the H2H and MTC clients, meaning one is reserved for the customers H2H and the other shared between H2H and MTC.

However, the division of RACH resources into 2 groups does not seem efficient. This is a method that can very quickly become ineffective if the M2M traffic becomes excessively high and the H2H traffic remains low and vice versa.

(3) Dynamic Allocation of RACH Resources: In this scheme, resources are allocated to M2M and H2H devices. The network can predict in advance whether the network will be overloaded by excessive access attempts caused by the large number of MTC devices. The network then dynamically allocates additional resources for the RACH procedure. As proposed in [18], M2M devices are categorized by types. In this approach, when the base station accepts the access request of an M2M device, in addition to granting access, it also allocates certain resources to devices of the same type, in which M2M devices of the same type can be content with access resources. Compared to the regular static RACH allocation, the dynamic resource allocation provided a big improvement [19] in the probability of successful access as well as time access. It is a solution that can be effective to some extent. However, it is limited by the unavailability of additional resources.

(4) Backoff Specific Scheme: In this scheme, a lower backoff time is assigned to conventional UEs than to MTC devices. This reduces the collision and congestion in the access network. Unfortunately, this pattern causes considerable delays, which negatively impacts high priority applications that are very sensitive to delays. The authors of [20] suggested a pure back-off scheme as well as a mixed back-off and ACB scheme based on cell load information. This system can provide performance improvements when the network is experiencing a low level of congestion in the RACH. It cannot, however, solve very high congestion levels [21].

(5) Slotted Access: This is an approach that defines access slots for MTC devices, so that the MTC device can perform RACH process only at the beginning of its dedicated time slot. This means that an MTC device cannot access the network when it wants, but only in its predefined time slot. This solution also reduces access level congestion. In [22], the authors have shown that randomly assigning slots to H2H and M2M devices may reduce the performance of this approach, while pre-assigning resources can increase efficiency up to three times.

(6) Pull based Access (Paging): In this scheme RACH process is not initiate by the M2M devices but rather by the base station. M2M devices are in idle mode. M2M servers trigger the random access process via the network (Paging) to collect data from M2M devices. This is a useful mechanism when it comes to, for example, reading smart meter data [23] in the smart gird network. Although it can simply mitigate overload problems at RACH, it can create an overload in the paging channel. Overloading of the paging channel is discussed in [24, 25]. The authors of [26] developed an analytical model for evaluating the performance of group paging in LTE.

3 System Model

During the RACH access process in a low-load network environment, the success rate of requests is very high with low delay. But as soon as a strong load is felt this rate decreases very quickly to reach a critical threshold and the delay with to reach inconceivable levels. Since the 3GPP has not proposed a method of detection or prevention of overload, we propose an overload detection method so that the network can anticipate an overload situation to enable it to set up a suitable overload resolution solution on optimal time. Our method is first to determine the resource utilization interval (preambles) in which the success rate is highest $\boldsymbol{R_1} \leq \boldsymbol{R_{used}} \leq \boldsymbol{R_{limit}}$ (see Sect. 4). Assuming the network not loaded at the beginning, eNodeB monitors the number of preambles used. If the preamble utilization threshold reaches $\boldsymbol{R_{limit}}$, we assume that the threshold is reached; meaning that the collision rate is going very high in contrary of success rate which drop down drastically. On that point of view, we envisage activation of contention resolution method, which will remain so until the number of used preambles (R_{used}) drops below the $\boldsymbol{R_{limit}}$ threshold (Fig. 3).

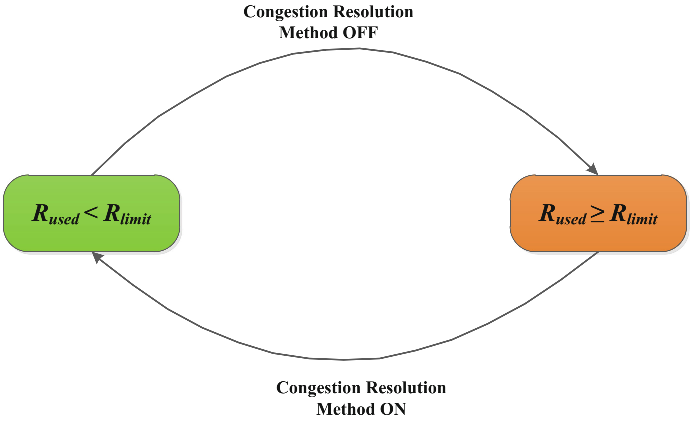

Fig. 3. Proposed congestion detection scheme

We assume a stable power for the MTC and the base station, we also assume that the base station is not able to successfully decode any type of transmission where a preamble is selected by more than one MTC in the same interval access time and therefore does not send any response to the corresponding MTCs. Note that random access can only take place in a frequency block specified by the base station. That is, the Physical Random Access Channel (PRACH) which is the physical layer responsible for mapping the RACH. In our work the RACH configuration index is 6. This means that the RACH occurs every 5 ms in a frequency band of 180 kHz for duration of 1 ms to 3 ms. We denote by R the number of available preambles and N the number of MTCs transmitting the preambles in a time interval $\boldsymbol{T_A}$ called activation time

$0 \leq t \leq T_A$, with a probability p(t) during which p(t) follows a beta distribution with parameters α = 3, β = 4 as in [9].

$$p(t) = \frac{t^{(\alpha-1)}(T_A - t)^{\beta-1}}{T_A^{\alpha+\beta-1} Beta(\alpha, \beta)} \tag{1}$$

Where **Beta (α, β)** is a beta function.

It is considered that there is $\mathbf{I_A}$ access in the activation time interval and that the access time is smaller than the interval between two access channels. $\mathbf{I_A}$ is divided into several activation times where the first activation time starts at $\boldsymbol{t_{(i-1)}}$ and ends at $\boldsymbol{t_i}$.

The number of supposed new arrivals is given by [8]:

$$\lambda_i = N \int_{t_{i-1}}^{t_i} p(t)dt \tag{2}$$

where $i = 1, 2, 3, \ldots, I_A$.

Equiprobable preambles (1/R) are considered. The probability that one of the N MTCs chooses one and only one preamble successfully is given by the binomial law:

$$\begin{aligned} P_{success} &= \binom{N}{1}\left(\frac{1}{R}\right)^1\left(1-\frac{1}{R}\right)^{N-1} \\ P_{success} &= \frac{N}{R}\left(1-\frac{1}{R}\right)^{N-1} \end{aligned} \tag{3}$$

The probability that one of the N MTCs does not transmit any preamble (Idle) is given by the binomial law:

$$\begin{aligned} P_{Idle} &= \binom{N}{0}\left(\frac{1}{R}\right)^0\left(1-\frac{1}{R}\right)^{N-0} \\ P_{Idle} &= \left(1-\frac{1}{R}\right)^{N} \end{aligned} \tag{4}$$

Success (successfully transmitted queries) can be achieved by multiplying the probability of success (Eq. 1) by the amount of available resources:

$$\begin{aligned} Success &= R * P_{success} \\ Success &= R * \frac{N}{R}\left(1-\frac{1}{R}\right)^{N-1} \\ Success &= N\left(1-\frac{1}{R}\right)^{N-1} \end{aligned} \tag{5}$$

Probability of collision will then be:

$$\begin{aligned} P_{Collision} &= 1 - P_{success} - P_{Idle} \\ &= 1 - \frac{N}{R}\left(1 - \frac{1}{R}\right)^{N-1} - \left(1 - \frac{1}{R}\right)^{N} \\ &= 1 - \frac{N}{R}\left(1 - \frac{1}{R}\right)^{N-1} - \left(\frac{R-1}{R}\right)\left(1 - \frac{1}{R}\right)^{N-1} \\ &= 1 - \left(1 - \frac{1}{R}\right)^{N-1}\left(\frac{N}{R} + \frac{R-1}{R}\right)^{N-1} \\ &= 1 - \left(\frac{N+R-1}{R}\right)\left(1 - \frac{1}{R}\right)^{N-1} \end{aligned} \quad (6)$$

From previous equations, it is clear that the quantity of resources used is the product of the total number of resources by the probability of use $(1 - \mathrm{P}_{\mathrm{Idle}})$:

$$\begin{aligned} R_{used} &= R * (1 - P_{Idle}) \\ R_{used} &= R\left(1 - \left(1 - \frac{1}{R}\right)^{N}\right) \end{aligned} \quad (7)$$

4 Discussion and Performance Analysis

In this section we determine the resource (preambles) utilization interval in term of the success rate of RA transmissions. Simulations (Fig. 4) from Eqs. (5) and (7) show that this interval is located between $\mathrm{R}_{\mathrm{used}} = 18.93$ and $\mathrm{R}_{\mathrm{used}} = 46,5$ gives a higher success rate when considering this rate above 15 and when taking N = 300 (MTC devices) and R = 54 as in [8]. From that interval it is clear that beyond $\mathrm{R}_{\mathrm{used}} = 18.93$ (the threshold called Rlimit in our work) the success rate falls down badly and considered non-viable for the network. So we consider that threshold the point from which congestion resolution method can be applied. By doing so, performances of system can be improved when comparing to the case when congestion resolution method is applied without taking in account the Rlimit threshold. In this section we determine the resource (preambles) utilization interval in term of the success rate of random access transmissions. Simulations (Fig. 4) from Eqs. (5) and (7) show that this interval is located between $R_{used} = 18.93$ (Y = 18.9) and $R_{used} = 46,5$ (Y = 46.5) gives a higher success rate when considering this rate above 15 and when taking N = 300 (MTC devices) and R = 54 as in [8]. From that interval, and considering $R_{used} = 18.93$ as $\boldsymbol{R_1}$, it is clear that beyond $R_{used} = 46,5$ (the threshold called $\boldsymbol{R_{limit}}$ in our work) the success rate falls down badly and considered non-viable for the network. So we consider that threshold the point from which congestion resolution method can be applied. By doing so, performances of system can be improved when comparing to the case when congestion resolution method is applied without taking in account the $\boldsymbol{R_{limit}}$ threshold.

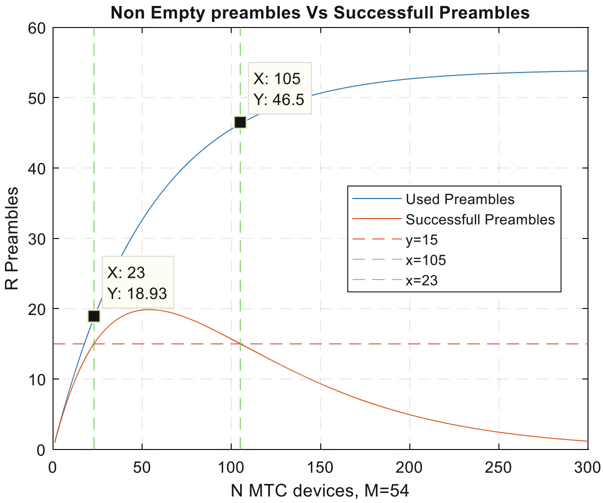

Fig. 4. Rused vs Success

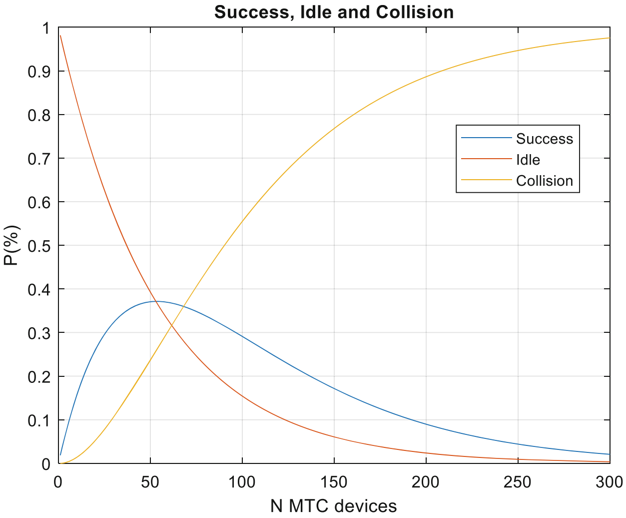

Fig. 5. Success-Collisions-Idle

Figure 5 shows that success probability (Eq. (3)) reaches its highest value at N = 50 and decreases when the number of MTC devices is going high. At that point, congestion is detected. So if no congestion resolution is applied, the collision probability (Eq. (6)) goes increasingly high from reaching its highest value when N increasing towards 300. At the same time Idle probability (Eq. (4)) decreases from higher value to 0 when N is going high.

5 Conclusion

Considered to be the first line of defense against LTE congestion, the LTE RACH procedure is prone to congestion when a large number of MTC devices attempt to access the network simultaneously. 3GPP pays particular attention to the resolution of congestion. Thus it proposed several methods of congestion resolution from which research was conducted for their improvement. However 3GPP did not provide any congestion detection method. Our work corrects this by providing a congestion detection method which, if applied to one of 3GPP's congestion resolution methods, will greatly improve its performance. The simulations allowed us to determine the preamble utilization interval during which the RA success is the highest. It also allowed us to determine the threshold beyond which the quality of service is no longer guaranteed. The method gets the eNodeB informed about the state of use of preamble resources, which enable the network to anticipate congestion states by activating a congestion resolution scheme. In our future work, we envisage the implementation of our detection method with 3GPP methods such as ACB, EAB, Separation of Resources or Backoff Scheme to improve its performance.

References

1. Atzori, L., Iera, A., Morabito, G.: The Internet of Things: a survey. J. Comput. Netw. **54**(15), 2787–2805 (2010)
2. Xia, N., Yang, C.-S.: Recent advances in machine-to-machine communications. J. Comput. Commun. **4**, 107–111 (2016)
3. IEEE: Machine to Machine (M2M) communications technical report. IEEE 802.16p-10/0005, November 2010
4. Gartner: Gartner says the internet of things installed base will grow to 26 billion units by 2020 (2013)
5. Han, X., Lim, T.J., Xu, J.: Heterogeneous access class barring with QoS guarantee in machine-type communications. Trans. Emerg. Telecommun. Technol. **28**, e2959 (2015)
6. http://www.rfwireless-world.com/Terminology/LTE-PRACH-Physical-Random-Access-Channel.html. Accessed 12 Apr 2017
7. Sesia, S., Toufik, I., Baker, M.: LTE–The UMTS Long Term Evolution: From Theory to Practice. Wiley, Hoboken (2009)
8. GPP TR 37.868 V11.0.0: Study on RAN Improvements for Machine-type Communications, September 2011
9. GPP:R2-100182: Access control of MTC devices. 3GPP TSG RAN WG2 Meeting 68bis, Valencia, Spain (2010)
10. GPP:R2-103143: Discussion on separating RACH resources for MTC. Alcatel-lucent Shanghai Bell, Alcatel-lucent (2010)
11. Larmo, A., Susitaival, R.: RAN overload control for machine type communications in LTE. In: 2012 IEEE GLOBECOM Workshops (GC Workshops), pp. 1626–163 (2012)
12. Cheng, R., Chen, J., Chen, D., Wei, C.: Modeling and analysis of an extended access barring scheme for machine-type communications in LTE-A networks. IEEE Trans. Wirel. Commun. **14**(6), 2956–2968 (2015)
13. Lien, S.-Y., Liau, T.-H., Kao, C.-Y., Chen, K.-C.: Cooperative access class barring for machine-to-machine communications. IEEE Trans. Wirel. Commun. **11**(1), 27–32 (2012)

14. Jiang, T., Tan, X., Luan, X., Zhang, X., Wu, J.: Evolutionary game based access class barring for machine-to-machine communications. In: 2014 16th International Conference on Advanced Communication Technology (ICACT), pp. 832–835, February 2014
15. Cheng, J.-P., Lee, C., Lin, T.-M.: Prioritized random access with dynamic access barring for RAN overload in 3GPP LTE-a networks. In: 2011 IEEE GLOBECOM Workshops (GC Workshops), pp. 368–372 (2011)
16. Lin, G.-Y., Chang, S.-R., Wei, H.-Y.: Estimation and adaptation for bursty LTE random access. IEEE Trans. Veh. Technol. **65**, 2560–2577 (2015)
17. Lee, K.-D., Kim, S., Yi, B.: Throughput comparison of random access methods for M2M service over LTE networks. In: 2011 GLOBECOM Workshops (GC Workshops), December, pp. 373–377
18. GPP:R2-113328: Dynamic separate RACH resources for MTC. 3GPP TSG RAN WG2 74. Institute for Information Industry (III), Coiler Corporation (2011)
19. Pang, Y.-C., Chao, S.-L., Lin, G.-Y., Wei, H.-Y.: Network access for m2m/h2h hybrid systems: a game theoretic approach. Commun. Lett. **18**(5), 845–848 (2014)
20. Jian, X., Jia, Y., Zeng, X., Yang, J.: A novel class-dependent back-off scheme for machine type communication in LTE systems. In: 2013 22nd Wireless and Optical Communication Conference (WOCC), pp. 135–140 (2013)
21. Lien, S.-Y., Chen, K.-C., Lin, Y.: Toward ubiquitous massive accesses in 3GPP machine-to-machine communications. IEEE Commun. Mag. **49**(4), 66–74 (2011)
22. GPP:R2-112247: Merits of the slotted access methods for MTC. Alcatel-lucent Shanghai Bell, Alcatel-lucent (2011)
23. G. R. 104873: Comparing push and pull based approaches for MTC. 3rd Generation Partnership Project (2010)
24. GPP:R2-104007: Pull vs push approach for MTC. 3GPP TSG RAN WG2 70bis, Stockholm, Sweden (2010)
25. GPP:R2-102781: Paging and downlink transmission for MTC. 3GPP TSG RAN WG2 Meeting 70, Montreal, Canada (2010)
26. Wei, C.-H., Cheng, R.-G., Tsao, S.-L.: Performance analysis of group paging for machine-type communications in LTE networks. IEEE Trans. Veh. Technol. **62**(7), 3371–3382 (2013)