Analysis of LTE-A Random Access Procedure: A Foundation to Propose Mechanisms for Managing the M2M Massive Access in Wireless Cellular Networks ⋆

Luis Tello-Oquendo, Israel Leyva-Mayorga, Vicent Pla, Jorge Martinez-Bauset and Vicente Casares-Giner

ITACA Research Institute
Universitat Politècnica de València, Spain

Abstract. Machine-to-machine (M2M) communication or machine-type communication (MTC) is considered an integral part of the so-called Internet of Things. M2M provides ubiquitous connectivity among devices without human intervention. Using cellular networks to provide M2M connectivity presents numerous advantages (coverage, roaming support, interoperability, well developed charging and security solutions among others). However, a large number of devices, that may need to communicate over a short period of time, can cause problematic situations which hugely impacts the radio access and core networks of the cellular system. In this paper, we provide a comprehensive survey on random access procedure in the context of 3GPP Long-Term Evolution (LTE) and LTE-Advanced (LTE-A). Specifically, we give a complete description of the parameters and the information necessary to evaluate the random access procedure when the cellular network is subjected to M2M traffic. It will be useful for designing mechanisms and algorithms to efficiently manage the massive accesses on the air interface.

1 Introduction

The world is moving beyond standalone devices into a new technological age in which everything is connected. Machine-to-Machine (M2M) communication stands for the ubiquitous automated exchange of information between devices on the edge of networks (computers, sensors, actuators, cars or mobile devices) inside a common network, the so-called Internet of Things (IoT).

M2M is what provides the IoT with the connectivity that enables capabilities, which would not be possible without it. A wide area of applications have already started to emerge in several fields, such as health-care, smart robots, smart grids, smart home technologies, intelligent transportation systems and smart cities [13]. One of the major motivating factors of MTC, besides the total automation of devices without involving human effort, is that MTC devices get smaller and more power efficient, while gaining more computing power.

⋆ This work has been supported by the Ministry of Economy and Competitiveness of Spain through the project TIN2013-47272-C2-1-R.

All of these properties would be of no use without a network to link those devices together. Two main approaches have emerged for this purpose. One the one hand are short or medium range wireless networks (IEEE 802.15.4 (ZigBee) and IEEE 802.11 (WiFi) [7, 10]) that can be used in a given area. On the other hand are cellular networks that offer many benefits, such as ubiquitous coverage by widely deployed infrastructure, global connectivity with a number of providers, and well developed charging and security solutions. Therefore, the cellular networks may be the solution for linking MTC devices by providing a solid infrastructure, high mobility, easy deployment and most importantly, wide area of coverage [16, 18].

In recent years, most research focuses on Long Term Evolution Advanced (LTE-A) as the technology enabling M2M communications. LTE-A is the most recent standard of wireless communication proposed by the 3rd Generation Partnership Project (3GPP) to serve its cellular mobile users by providing high data bandwidth. LTE-A cellular networks were not designed for M2M traffic. Instead, they were mainly designed to support human-to-human (H2H) services such as voice and web browsing, and bandwidth demanding services such as video streaming.

The nature of M2M communication is different from that of regular H2H communication and it is mainly characterized by a high device density in a cell, small amounts of payload, low traffic volumes per device, predominance of up-link traffic, low device mobility, low device processing capacity and battery operated devices. Although MTC devices generate small amount of data traffic, the massive amount of MTC devices could create an overload problem which hugely impacts the radio access and core networks of the system, reducing the overall performance of LTE-A [17]. To mitigate this problem, it is important to provide an efficient way for managing the massive access in the radio access network and to minimize the network overload.

In this paper, we study the random access (RA) procedure of LTE-A networks. MTC devices must perform it for initial and periodic RA, in order to request resources to the network before they transmit data. The aim is to provide a complete description of the parameters and the information necessary to evaluate the RA procedure of LTE-A. We consider this work as the first step to propose efficient mechanisms and algorithms to mitigate the overload in the radio access network.

The remainder of the paper is organized as follows. Section 2 describes the LTE-A radio interface. Section 3 presents the RA procedure, key in mobile networks, which enables user equipment (UE) to initiate communications to a base station. Section 4 details the considerations to evaluate the RA performance when the cellular network is subjected to M2M traffic. Section 5 provides numeric results of access success probability. Finally, the conclusions are presented in Section 6.

2 Brief Overview of LTE-A Radio Interface

The LTE-A radio access network consists of the base stations, denoted as enhanced NodeB (eNB), that are connected to each other through the X2 interface and to the evolved packet core (EPC) through the S1 interface. The mobile terminal or MTC device is denoted as user equipment (UE). In this section, we describe briefly the LTE-A radio interface. We focus in the physical layer characteristics which provide the resources for the RA procedure. A more detailed description can be found in [2, 6, 8, 11, 12].

In LTE-A, orthogonal frequency-division multiplexing (OFDM) is the down-link (DL) multiple access scheme, while single-carrier frequency-division multiple access (SC-FDMA) is the up-link (UL) multiple access scheme. LTE-A also supports scalable bandwidth up to 20 MHz per component carrier (a maximum of 100 MHz with carrier aggregation); it uses DL/UL frequency selective and DL frequency diverse scheduling. In the UL, like the DL, frequency domain orthogonality is maintained among intra-cell users, which allows the eNB the ability to efficiently manage the amount of interference seen at the base station.

The UL and DL sub-frame structure is common to both time-division duplex (TDD) and frequency-division duplex (FDD). Each sub-frame consists of two slots of length 0.5 ms (7 OFDM symbols for normal cyclic prefix) with reference symbols located within each slot. UL control signaling such as channel quality indication (CQI) and acknowledgment/negative acknowledgment (ACK/NACK) is located in the system band-edge.

Each portion in the time and frequency resource grid is called a resource element (RE). Each DL sub-frame contains reference signals, control information, and data transmission. The following physical channels provide DL control signaling: physical control format indicator channel (PCFICH), physical hybrid automatic repeat request (HARQ) indicator channel (PHICH), and physical DL control channel (PDCCH). The DL and UL scheduling assignment transmitted on the PDCCH is addressed to a specific user, and contains control information needed for data reception and demodulation. Users are assigned data allocation in amounts of resource blocks (RBs), where an RB is defined as 12 REs by one slot. Table 1 summarizes the available DL and UL physical channels and their purpose.

Table 1. Physical Channels in LTE

Channel		Purpose
DL	PDSCH	Carry user data (DL)
	Physical broadcast channel (PBCH)	Carry broadcast information
	Physical multi-cast channel (PMCH)	Carry multi-cast services
	PCFICH	Indicate the size of the control region in number of OFDM symbols
	PHICH	Carry ACK/NACK associated with UL transmission
	PHICH	Carry DL scheduling assignments and UL scheduling grants
UL	PUSCH	Carry user data (UL)
	PUCCH	Carry ACK/NACK associated with DL transmission, scheduling request, and feedback of DL channel quality and precoding vector
	Physical random access channel (PRACH)	Initiate random access procedure

2.1 LTE-A Physical Random Access Channel

The Physical Random Access Channel (PRACH) is used to signal a connection request when a mobile device desires to access the mobile system (call origination or paging response). The PRACH carries the RA preamble and the messages involved in the RA procedure (represented in Fig. 2); it uses small preamble to lower overhead and it is orthogonal to PUSCH/PUCCH.

Random access attempts are done in predefined time/frequency resources called RA slots herein. In the time domain, the duration of each RA slot depends on the format of the access requests (as shown in Fig. 1). In the frequency domain, each RA slot occupies 1.08 MHz, which corresponds to the bandwidth of 6 RBs. These RA slots are reserved in the UL channel of the network for the transmission of access requests. The eNB broadcasts the periodicity of the RA slot by means of a variable referred to as the PRACH Configuration Index. The periodicity varies between a minimum of 1 RA slot every 2 frames, i.e., every 20 ms, and a maximum of 1 RA slot per 1 sub-frame, i.e., every 1 ms [3, 14].

Random Access Preamble If a UE tries to access to a RA slot, it must select a RA preamble (signature). There are 64 orthogonal preambles per cell, which allow up to 64 UEs to get simultaneously access. These preambles are generated by Zadoff-Chu (ZC) sequences due to their good correlation properties [3, 19].

There are altogether 5 formats of RA preambles, format 0 to format 4. From Fig. 1, we can see that the format 0 to 3 have different lengths. These should be chosen specifically according to different cell size. In cells with larger size, longer preamble must be used since the UL timing difference between the farthest and nearest UE is larger. On the other hand, shorter preamble should be used in small cell to avoid unnecessary wastage of the resources. The format 4 preamble is not shown in Fig. 1; it is available only for the TDD system.

Preambles are split into two sets:

- Contention-free: it is used for critical situations such as handover, DL data arrival or positioning, where there is a coordinated assignment of preambles so collision is avoided. eNB alone can assign these preambles for specific slots to specific UEs. UEs can only use these, if assigned by the eNB, and for the specific slots assigned.
- Contention-based: it is the standard mode for network access (there are more preambles in this set). Preambles are selected in a random fashion, so there is risk of collision, i.e., there is a probability that multiple UEs in the cell could pick the same preamble signature and the eNB would assign the same RB to both UEs; therefore contention resolution is needed.

3 Random Access Procedure

This section provides a general overview of the operation of the RACH of LTE-A. We focus on contention-based RA mechanism (as shown in Fig. 2).

A UE can only be selected for UL transmission if it is time-synchronized. The main role of the RA procedure is to request for UL resources, for that it is necessary to assure such time synchronization for a UE which either has not yet acquired, or has lost its UL synchronization [19].

3.1 Contention-Based Random Access Procedure

In order to explain the RA procedure, we use the messages exchanged in Fig. 2. The eNB broadcasts information (System Information Blocks) about which RA

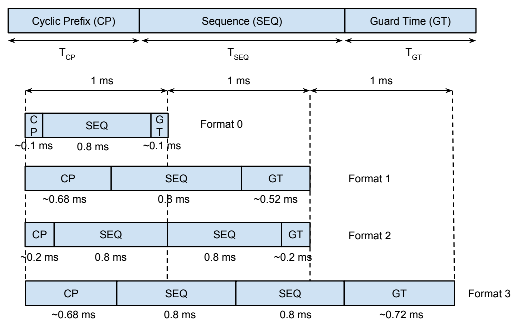


Fig. 1. Structure and Preamble Formats

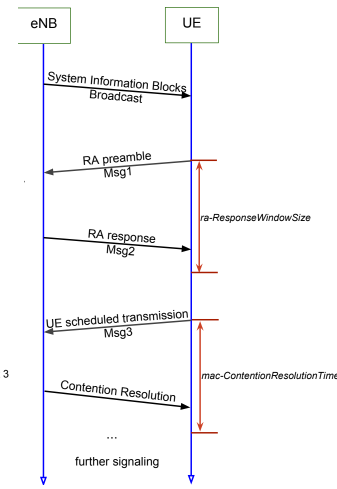


Fig. 2. Contention-based RA Procedure

slot the RA preamble transmission is allowed; in each RA slot, the eNB reserves a number of preambles for UEs to perform RA. A UE sends a randomly chosen preamble in a RA slot (Msg1) and waits for a time window to receive a response from the eNodeB. Then, the eNB sends the RA response (Msg2). If more than two UEs choose the same preamble in the RA slot, a collision will occur in the transmission of Msg3. The eNB can detect non-collided UEs with a probability which mainly depend on the transmission power of the UE. The UEs which are collided or not detected by the eNB are referred as failed UEs. The failed UEs should ramp up its power, and retransmit a new randomly chosen preamble in a new RA slot based on a uniform back-off algorithm; they repeat this process until a transmission limit is reached. The success UEs will exchange Msg3 with eNB through dedicated channels and hybrid automatic repeat request (HARQ) is used to protect the message transmission. The eNB transmits a Contention Resolution message as an answer to Msg3. A device which does not receive the Contention Resolution message declares a failure in the contention resolution and schedules a new access attempt, i.e., a new preamble transmission, starting the process over again. Each device keeps a preamble transmission counter that is increased after each unsuccessful attempt. When the counter reaches the maximum allowed value (informed as system information by the eNB), the network is declared unavailable by the device and a random access problem is indicated to upper layers [2, 4, 5, 9, 14].

Algorithm 1 presents the procedure used by the UE to perform the RA to the network as defined by [2, 4]. For message definitions, field description of information elements (master information blocks, system information blocks and radio resource control) refer to Sections 6.2.2, 6.3.1 and 6.3.2 in [5]. The MAC layer parameters are borrowed from [2] and for the most part detail the RA procedure, which is considered in Algorithm 1.

Algorithm 1 LTE-A Random Access Procedure

1: UE gets system information blocks from eNB.
2: UE needs to access eNB.
3: (a) Flush the Msg3 buffer; set the $Preamble\_Transmission\_Counter$ to 1; set the $backoffParameterValue$ to 0 ms;
4: Is $Preamble\_Transmission\_Counter = preambleTransMax + 1$?:
   – If yes: ACCESS FAILURE; indicate a RA problem to upper layers.
   – Otherwise, go to step 5.
5: Select a random back-off time according to $unif\{0, backoffParameterValue\}$ and delay the subsequent RA transmission by the back-off time;
6: Select a RA preamble sequence, determine the next available sub-frame containing PRACH and proceed to the transmission of the preamble.
7: Monitor the PDCCH for RA Response(s) identified by the $RA\text{-}RNTI$ within the $ra\text{-}ResponseWindowSize$.
8: Did UE receive RAR within $ra\text{-}ResponseWindowSize$ and $RAPID$ matches $RA\text{-}RNTI$ on PDCCH?
   – If yes:
      (a) set the $backoffParameterValue$ as indicated by the BI field of the Backoff Indicator subheader; set the $Temporary\_C\text{-}RNTI$ to the value received in the Msg2; process the received TAC; process the received UL grant value and indicate it to the lower layers;
      (b) indicate the $preambleInitialReceivedTargetPower$ and the amount of power ramping applied to the latest preamble transmission to lower layers (i.e., $(Preamble\_Transmission\_Counter - 1) *$ powerRampingStep);
      (c) UE transmits its initial Msg3 four sub-frames later in PUSCH location assigned in the UL grant. If $C\text{-}RNTI$ is not yet assigned, UE uses the $Temporary\_C\text{-}RNTI$ and CCCH SDU is included in Msg3. Otherwise, UE sends message in PDCCH.
      (d) start $mac\text{-}ContentionResolutionTimer$ and restart it at each HARQ retransmission.
      (e) go to 10
   – Otherwise, go to step 9.
9: Did the UE receive a notification of power ramping suspension?
   – If yes: increment $Preamble\_Transmission\_Counter$ by 1 and suspend power ramping; go to step 4.
   – Otherwise: increment $Preamble\_Transmission\_Counter$ by 1 and boost power by $powerRampingStep$ factor in dB; go to step 4.
10: Did $mac\text{-}ContentionResolutionTimer$ expire?
   – If yes: discard the $Temporary\_C\text{-}RNTI$ and flush the HARQ buffer; go to step 9.
   – Otherwise, go to step 11.
11: Did UE receive notification of a reception of a PDCCH transmission from lower layers?
   – If yes: Is the PDCCH transmission addressed to the $C\text{-}RNTI$?
      • If yes: Contention Resolution successful; stop $mac\text{-}ContentionResolutionTimer$; discard the $Temporary\_C\text{-}RNTI$; RA procedure successfully completed.
      • Otherwise: Does the decoded MAC PDU contain a UE Contention Resolution Identity MAC control element; and matches the CCCH SDU transmitted in Msg3?
         * If yes: Contention Resolution successful; stop $mac\text{-}ContentionResolutionTimer$; set the $C\text{-}RNTI$ to the value of the $Temporary\_C\text{-}RNTI$ and discard it; RA procedure successfully completed.

    - ∗ Otherwise: discard the successfully decoded MAC PDU; go to step 10.
  - Otherwise: go to step 10.

C-RNTI, Cell Radio Network Temporary Identifier; RA-RNTI, Random Access Radio Network Temporary Identifier; RAPID, Random Access Preamble ID; MAC, Medium Access Control; SDU, Segment Data Unit; PDU, Packet Data Unit; TAC, Timing Advance Command.

4 Evaluation of LTE FDD RACH for M2M communication

In order to evaluate the network performance under M2M traffic, a single cell environment is assumed. The network is subjected to different access intensities; for that, 3GPP TR 37.868 [1] defined two different traffic models for the evaluation of network performance, these are listed in Table 2. Traffic model 1 can be considered as a realistic scenario in which MTC devices access the network uniformly over a period of time, i.e. in a non-synchronized manner. Traffic model 2 can be considered as an extreme scenario in which a large amount of MTC devices access the network in a highly synchronized manner, e.g. after a power outage.

Table 2. Traffic model for RACH evaluation [1]

Characteristics	Traffic model 1	Traffic Model 2
Number of MTC devices (N)	1000, 3000, 5000, 10000, 30000	1000, 3000, 5000, 10000, 30000
Arrival distribution	Uniform distribution over T	Beta distribution over T
Distribution period (T)	60 seconds	10 seconds

Table 3. Basic Simulation Parameters for RACH capacity evaluation for LTE FDD

Parameter	Setting
Cell bandwidth	5 MHz
PRACH Configuration Index	6
Total number of preambles (M)	54
preambleTransMax	10
Number of UL grants per RAR	3
Number of CCEs allocated for PDCCH	16
Number of CCEs per PDCCH	4
ra-ResponseWindowSize	5 subframes
mac-ContentionResolutionTimer	48 subframes
Backoff Indicator (BI)	20 ms
HARQ retransmission probability for Msg3 andMsg4 (non-adaptive HARQ)	10%
Maximum number of HARQ TX for Msg3 andMsg4 (non-adaptive HARQ)	5
Periodicity of PRACH opportunities	5 ms
Preamble transmission time	1 ms

If two (or more) MTC devices select the same preamble at the same time, it is assumed that the eNB will not be able to decode any of the preambles; hence, the eNodeB will not send the Random Access Response (Msg2). MTC devices will only detect a collision if Msg2 is not received in the *ra-ResponseWindowSize*. In case of no collision, $1-(1/e^i)$ preamble detection probability is assumed, where i indicates the $i-th$ preamble transmission to take into account the effects of radio channels, for example path-loss, fading, inter-cell interference, etc [1].

The following measures could be taken into account for the purpose of RACH capacity evaluation [1]:

1. Access success probability, defined as the probability to successfully complete the RA procedure within the maximum number of preamble transmissions.
2. Statistics of number of preamble transmissions, defined as the cumulative distribution function (CDF) of the number of preamble transmissions to perform a RA procedure, for the successfully accessed MTC devices.
3. Statistics of access delay, defined as the CDF of the delay for each RA procedure between the first RA attempt and the completion of the RA procedure, for the successfully accessed MTC devices.

5 Numerical Results

In this section we show some results according the methodology proposed in [1] and detailed in Section 4. In order to understand the limits of the RACH of LTE, we analyze the overloaded scenario, i.e., more than 1.000 devices that need to access the network simultaneously.

The system setup and simulation parameters follow Tables 2 and 3. There is a RA slot every 5 ms and 54 out of the 64 available preambles are used for contention-based access, while the remaining 10 preambles are reserved for contention-free access. Under these conditions, the system offers 200 RA slots per second; the maximum number of preamble transmission is set to 10 [14, 15, 20].

Figure 3 illustrates the total, failed, re-transmitted and successful average number of RA attempts per slot. The reason for access failure is because the difficulty of selecting a unique preamble when a large number of access request happen simultaneously. Note that the contention-based operation of the RACH is based on ALOHA-type access, i.e., transmit the request in the first available opportunity [14]. In case of collision, the UE tries again after the back-off time.

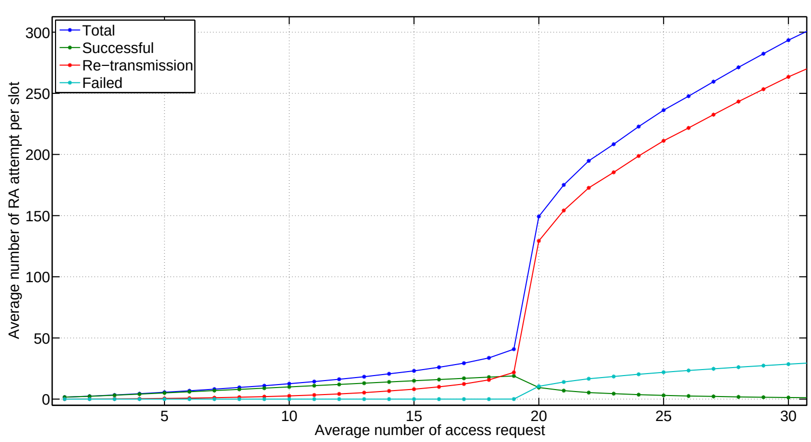
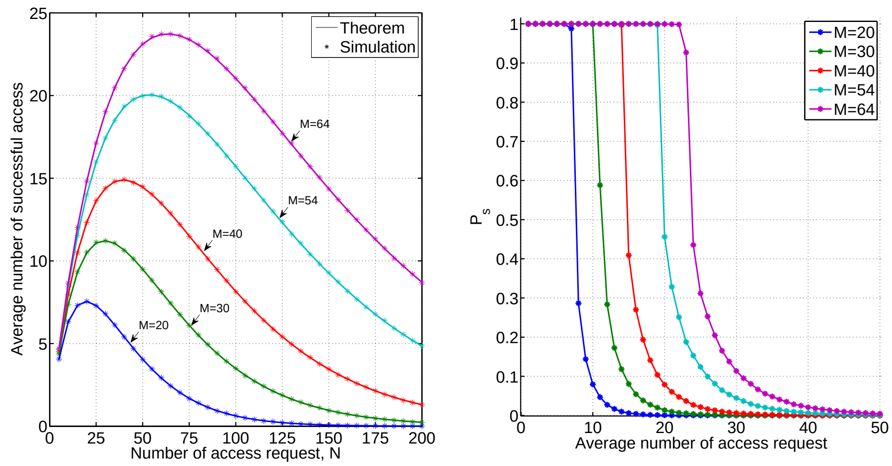


Fig. 3. Random Access Attempts per Slot. N=30000; M=54.

According to [15], the expected number of successful RA requests per PRACH slot is $N(1-1/M)^{N-1}$ for N RA requests and M preambles; Fig. 4(a) illustrates this computation for different values of M.



(a) Capacity (b) Access Success Probability

Fig. 4. System Capacity

Let P_s be the access success probability per device (including the back-off mechanism), it can be computed as $P_s = (N_r \cdot P_r)/N_a$, where P_r is the success probability per slot, N_a is the average number of new accesses per slot, and N_r the average number of total accesses (news+retx) per slot. If we define the mean number of trials as $N_i = N_r/N_a$; $P_s = N_i \cdot P_r$. The system capacity per PRACH slot is computed as $c = \left[log(\frac{M}{M-1})\right]^{-1}(1-\frac{1}{M})^{\left[log(\frac{M}{M-1})\right]^{-1}}$ [15]. Figure 4(b) shows that in the current random access mechanism, an access attempt can be successful when the total number of access requests is lower than the system capacity. Additionally, the maximum capacity increases slightly as the number of preambles increases.

6 Conclusions

The random access procedure is key in mobile networks since it enables the user equipment to initiate communications and time synchronization with a base station. We show that the current mechanism to request access to the cellular network system suffer from congestion and overloading in the presence of a huge number of devices. Therefore, more efficient methods for managing the access to these networks in such circumstances are necessary. Finally, further work on LTE-A and next generations of cellular networks in order to make them capable and efficient to provide M2M services is still ongoing, and a follow-up to this paper will be published at a later date.

References

1. 3GPP: Study on RAN Improvements for Machine Type Communications. TR 37.868, 3rd Generation Partnership Project (3GPP) (Sep 2011)
2. 3GPP: Medium access control (MAC) protocol specification. TS 36.321, 3rd Generation Partnership Project (3GPP) (Sep 2012)
3. 3GPP: Physical channels and modulation. TS 36.211, 3rd Generation Partnership Project (3GPP) (Dec 2014)
4. 3GPP: Physical layer procedures. TS 36.213, 3rd Generation Partnership Project (3GPP) (Dec 2014)
5. 3GPP: Radio Resource Control (RRC), Protocol specification. TS 36.331, 3rd Generation Partnership Project (3GPP) (Dec 2014)
6. 3GPP: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN), Overall Description, Stage 2. TS 36.300, 3rd Generation Partnership Project (3GPP) (Dec 2014)
7. Ahmad, F., M.S.Z.Y., Görg, C.: Tailoring lte-advanced for m2m communication using wireless inband relay node. In: World Telecommunications Congress (WTC) (2014)
8. Akyildiz, I.F., Gutierrez-Estevez, D.M., Reyes, E.C.: The evolution to 4g cellular systems: Lte-advanced. Phys. Commun. 3(4), 217–244 (Dec 2010)
9. Cheng, R., Chen, J., Chen, D., Wei, C.: Modeling and analysis of an extended access barring scheme for machine-type communications in lte networks. Wireless Communications, IEEE Transactions on PP(99), 1–1 (2015)
10. Ghavimi, F., Chen, H.: M2m communications in 3gpp lte/lte-a networks: Architectures, service requirements, challenges, and applications. Communications Surveys Tutorials, IEEE 17(2), 525–549 (Secondquarter 2015)
11. Ghosh, A., Ratasuk, R., Mondal, B., Mangalvedhe, N., Thomas, T.: Lte-advanced: next-generation wireless broadband technology [invited paper]. Wireless Communications, IEEE 17(3), 10–22 (June 2010)
12. Larmo, A., Lindstrom, M., Meyer, M., Pelletier, G., Torsner, J., Wiemann, H.: The lte link-layer design. Communications Magazine, IEEE 47(4), 52–59 (April 2009)
13. Lawton, G.: Machine-to-machine technology gears up for growth. Computer 37(9), 12–15 (Sept 2004)
14. Laya, A., Alonso, L., Alonso-Zarate, J.: Is the random access channel of lte and lte-a suitable for m2m communications? a survey of alternatives. Communications Surveys Tutorials, IEEE 16(1), 4–16 (First 2014)
15. Lin, T.M., Lee, C.H., Cheng, J.P., Chen, W.T.: Prada: Prioritized random access with dynamic access barring for mtc in 3gpp lte-a networks. Vehicular Technology, IEEE Transactions on 63(5), 2467–2472 (Jun 2014)
16. Lo, A., Law, Y., Jacobsson, M.: A cellular-centric service architecture for machine-to-machine (m2m) communications. Wireless Communications, IEEE 20(5), 143–151 (October 2013)
17. Pötsch, T., Khan Marwat, S., Zaki, Y., Görg, C.: Influence of future m2m communication on the lte system. In: Wireless and Mobile Networking Conference (WMNC), 2013 6th Joint IFIP. pp. 1–4 (April 2013)
18. Russler, N.: Networks of the future: Ideas and concepts (August 2014), http://www.whitebyte.info/wp-content/uploads/2014/08/
19. Ubeda, C., Pedraza, S., Regueira, M., Romero, J.: Lte fdd physical random access channel dimensioning and planning. In: Vehicular Technology Conference (VTC Fall), 2012 IEEE. pp. 1–5. IEEE (2012)
20. WG2, G.T.R.: Rach overload solutions. TR N. 70bis R2-103742, 3rd Generation Partnership Project (3GPP) (june 2010)