RESEARCH ARTICLE

ADVANCED ACCESS CLASS BARRING FOR MASSIVE IOT DEPLOYMENT OVER LTE-A NETWORKS

1Bouba Goni Mahamadou, 2Mbainaibeye Jérôme, 3Ahmed Kora and 1Claude Lishou

1LTI Laboratory, ESP/UCAD, Dakar Senegal
2University of DOBA, Chad
3LTI & RSI Research Group, ESMT, DAKAR Senegal

ABSTRACT

LTE-A is one of the latest generations of communication mobile systems that should support, in addition to traditional voice and data communication services, new communication paradigms, such as the Internet of Things (IoT) and machine-to-machine (M2M) services, which implement communication between machine-type communication devices (MTCs) in a fully automated manner, without human intervention or with minimal human intervention. We will use the term User Equipment (UE) to designate these devices. With the IoT, the network is subject to recurrent congestions, due to an increased solicitation of the uplink channel, during radio access. Idle terminals must compete in a process called Random Access Channel (RACH) to access the network. The collision that causes congestion occurs during this process and it has a negative impact on the quality of service (QoS). The Third Generation Partnership Project (3GPP) has proposed for this purpose, among others solutions, the Access Class Barring (ACB) to alleviate the problem. With the ACB, terminals are divided into classes. An access probability $P_{acb}$ (ac-barring factor) is broadcast by the base station toward the different classes of terminals. The UEs in turn generate a random number $q$, and are allowed to pass the RACH process only in the case where $q \leq P_{acb}$. Otherwise, it must wait for $L_{barring}$ (called barring time) before resuming the process. In this paper we propose the Advanced ACB (A-ACB) for significantly reducing the collisions responsible of the congestion problem in M2M communications for IoT. The originality and the contribution of this paper reside in the fact that our method combines congestion detection method that we have proposed in (Mahamadou, 2018), which consists of determining the threshold of resource utilization ($R_{lim}$) and the amount of used resources ($R_{now}$), in order to activate the basic ACB if the $R_{now}$ reaches the $R_{lim}$. The simulations results show that the proposed A-ACB gives much better results compared to basic ACB. Furthermore, the proposed scheme is less complex and easy to implement in the LTE-A networks. Also, it does not require large investments for network operators.

INTRODUCTION

The concept of smart cities refers to an urban environment where all everyday objects communicate and exchange data and information with each other without human intervention. These objects are used in various fields such as e-health, smart grids, e-learning, e-libraries, telemedicine, traffic management, environmental management, smart transport systems, etc. The establishment of a smart city entails the deployment of a very large number of connected objects. We are talking about 26 (Gartner, 2013) to 50 (Han, 2015) billion connected objects by 2020. Different technologies such as wired networks, local short-range wireless networks, sensor networks and cellular wireless networks have been studied to allow massive MTC applications.

However, because of the mobility of some equipment and for the sake of efficiency and also, to ensure a good quality of service, some equipment must be directly connected to the LTE-A mobile network. The LTE-A is one of the latest mobile network standards that have benefits such as good mobility, good accessibility, good coverage, good security, broadband and many other features that are very important for M2M applications and services. Since the LTE-A network was primarily designed for typical Human-to-Human (H2H) communication, it is clear that adaptations must be made to meet the requirements of M2M communication. In recent years, studies have shown that the Random Access (RA) procedure is not effective for managing massive M2M communication within the LTE network, as the LTE random access physical channel (PRACH) would suffer of overload when a large number of UEs are competing for accessing resources (Ferdouse, 2015; Laya, 2014). The massive deployment of M2M devices leads to congestion of the radio
access and core network of the LTE system. 3GPP, which is the organization responsible for standardizing mobile networks, including LTE, has identified and proposed some solutions to the problems and requirements which may arise when integrating M2M devices into LTE network. Overload and congestion control in the LTE radio access network (RAN) is considered by 3GPP as a high priority issue that must be addressed to enable M2M communication over the LTE network (Ferdouse et al., 2015). Overload and congestion on the network normally occur when a large number of access requests are sent by devices to a single base station during the random access procedure (RACH procedure). The RACH procedure has a very low efficiency when the number of devices increases (Zheng et al., 2014). In this paper, we study congestion problems and we proposed an original method called Advanced Access Class Barring (A-ACB) in the LTE-A network, in a transient scenario of massive access of M2M terminals. The remaining of the paper is organized as following: in section II, we present the context of the M2M communications for IoT and the state of the art in this domain; in section III, we present the model of our proposed system called A-ACB; the section IV presents the simulations results of the proposed method and the discussions. The conclusion is presented in section V.

CONTEXT

LTE-A Random Access Slot: The random access slot is a radio resource also known as the Physical Random Access Channel (PRACH), through which resources are distributed and then used by the UEs to transmit requests to the eNB. It is the channel used to carry the random access preambles used to initiate the random access procedure. In Frequency Division Duplex (FDD) mode, the RA slot consists of 6 physical resource blocks (RBs) in the frequency domain, while the duration of each RA interval can be 1, 2 or 3 sub-frames, depending on the preamble format (7). There are a total of 864 subcarriers in an RA interval that are at distances of 1.25 KHz. The 64 preambles are distributed over 839 RACH subcarriers, while the remaining 25 subcarriers are used as guard subcarriers. Depending on the distance from the eNB, the UE can choose a generic format of preamble in order to compensate the round trip delay. During a given activation time (\(T_d\)) random access time, the duration of the access channel must be shorter than the time interval between two random access channels. The \(T_d\) is then discretized in several time intervals as shown in figure 1. The \(i^{th}\) random access channel starts at the beginning of the \(i^{th}\) time slot. The \(i^{th}\) time interval starts at \(t_{i-1}\) and ends at \(t_i\). The first time interval starts at \(t_0 = 0\). At the beginning of each time interval, i.e. between \((t_{i-1}; t_i)\), activated UEs proceed to random access.

The number of random access channels in each radio frame is defined by the preamble configuration index. For each preamble format, 16 different indexes are available, where the eNB allocates radio resources in the form of PRACH. Depending on the system bandwidth, some LTE systems may not be able to use some preamble configuration indexes. However, systems using 20 MHz bandwidth can use all indexes (Sesia, 2009). The eNB periodically broadcasts the preamble information in a message called the System Information Block (SIB2), which is used by the UEs to obtain the necessary information, in order to transmit preambles.

In LTE, the Zadoff-Chu (Chu, 1992) sequences are used for transmission of the random access preamble because of its low peak average power ratio (PAPR). Since UEs are limited power equipment, the use of Zadoff-Chu sequences allow power saving, leading to the increase of UEs lifetime. The Zadoff-Chu sequence of odd length \(N_{ZC}\) (Sesia, 2009) is given by equation (1):

\[ x_u(n) = \exp\left[-j \frac{\mu_u(n+1)}{N_{ZC}}\right], \quad n \in [0, N_{ZC} - 1] \quad (1) \]

Where \(u\) is the root index of the ZC and \(N_{ZC}\) is the length of the ZC sequence.

ZC sequences have an ideal cyclic autocorrelation property that can be given by equation (2):

\[ r_{uu}(\sigma) = \sum_{n=0}^{N_{ZC}-1} x_u(n) x_u^*(n + \sigma) = \delta(\sigma) \quad (2) \]

Where \(x_u(n + \sigma)\) is the cyclic shift version of \(x_u(n)\) with an offset \(\sigma\).

The random access preambles are obtained from a ZC sequence with different cyclic offsets. More precisely, the number of preambles per ZC sequence is given by equation (3):

\[ N_p = \frac{N_{ZC}}{N_{CS}} \quad (3) \]

Where \(N_{CS}\) is the size of the cyclic shift.

In FDD-LTE, \(N_{ZC}\) and \(N_{CS}\) are respectively 839 and 13, which means that the number of available preambles per ZC

![Fig 1. Random Access Time Slots](Image)

![Fig 2. Structure of random access preamble](Image)
sequence is 64. The eNB reserves some preambles called $N_{CF}$, for specific contention-free RA accesses, and allocates separate preambles to different UEs. The other remaining preambles ($64 - N_{CF}$) are used for random access with contention. Each UE randomly uses one preamble (7). During contention-free access, the connection is initiated by the base station which, at the same time, provides the UEs with the necessary resources. This is applied to priority communications such as emergency alert messages and specific uses. In the case of contention access, UEs compete for the remaining 54 preambles in the RACH process.

**Random Access Procedure**

The LTE’s Medium Access Control (MAC) function is divided into two parts: contention access known as RACH and contention-free access. Contention access is a process initiated by the idle terminals (Radio Resource Control standby mode RRC) to request access to the uplink channel, necessary for the transmission of packets. However, contention-free access is designed for active UEs (RRC connected mode). Only the active UE can transmit data; if a UE is in idle mode, it must first proceed to the RACH process for an initial connection to the network. The contention based random access is explained by the fact that at each access time interval (RAI) granted by the base station, the UEs must compete for a random access preamble to perform network access request. The random access process is divided into four stages:

- **Preamble Emission:** In this first step, the UE selects and transmits a random access preamble via the PRACH, using a temporary identifier, and waits for $T_{RAR}$ (random access response) time necessary for the eNB to process the preamble. This allows the eNodeB to estimate the timing of the terminal transmission. It should be noted that the network broadcasts information in which the time-frequency resources (ie the PRACH) are located, through the SIB-2 message to all terminals authorized to transmit.
- **Random Access Response:** It consists of the eNodeB sending a synchronization advance command to adjust the transmission timing of the terminal, according to the timing estimated in the previous step. This is done through the Physical Downlink Shared Channel (PDSCH). The location of the UE and the response time on the PDSCH can be calculated from the time and location from which the random access message was sent. In addition to synchronization, the second step also assigns the Physical Uplink Shared Channel (PUSCH) to the terminal. This message also contains the random identity sent by the device in step 1, and the temporary identifier of the Cellular Radio Network Temporary Identifier (T_C-RNTI) network cell that will be used in the next step.
- **Radio Resource Control (RRC) connection request:** The UE transmits its identity to the eNodeB using the PUSCH uplink channel allocated in step 2. This message also contains the previously allocated identity of the cell (C-RNTI) in which the EU is located and the reason for the request.
- **RRC connection:** This step consists of the transmission of a contention resolution message by the eNodeB to the UE using the PDSCH channel. This step resolves the conflict from the multiple terminals attempting to access the network using the same preamble.

**Rach Congestion:** The RACH procedure has been identified by 3GPP as a difficult task (Lo, 2011) for M2M communications, due to the sudden increase in signaling traffic. This is caused by a sudden increase of the number of UEs trying to simultaneously access to the same base station. This situation may be for example a result of smart meters suddenly becoming active at the same time, after a period of power failure or alerts, following any kind of disaster. As a result, massive access requests from UEs can overload the PRACH channel, leading to an increase of contention probability which, in turn, will increase access time and failure rate. So the network will be congested. One of possible solutions to reduce PRACH load is to increase the number of scheduled access opportunities per frame. But this produces a reduction of the amount of resources available for data transmission, and therefore a contraction of the uplink channel data transport capacity occurs. Furthermore, it is to be known that the total amount of RA resources which can be allocated in an LTE-A frame is limited. Finally, the processing of the Zadoff-Chu sequence, which is used to generate preambles, is more expansive time computationally and may be an additional source of problems for resource-limited devices, such as MTCs. When two or more UEs send the same preamble during the same RAO, a collision occurs. The eNB can detect a collision less preamble transmission with a certain probability, $P_{col}$, which depends on several parameters, such as the transmission power of the UE.

UEs which do not receive Msg2 in the Random Access Response Window (Random Access Response Window) should increase their power and retransmit a new preamble randomly selected in a new RAO based on a uniform back-off policy (see figure 4). The base station can detect collisions based on the difference in transmission delays of these preambles in step 1, then it will not send any response for this preamble (step 2); the UEs concerned will then be led to resume the operation. However, if the UEs concerned are equidistant from the base station, the collision will not be detected; therefore a copy of the response from step 2 will be sent to them. In this case, the collision is resolved by the contention resolution message in step 4.

![Fig. 3. Random Access Process](image)
Only the selected UE will receive the answer in step 4 and will have access to the network. The UEs which have not received the response of the Step 4 declares access failure and plans a new access attempt, meaning restarting the preamble transmission process. Each UE has a preamble transmission counter which is incremented after each unsuccessful attempt. When the counter reaches the preamble TransMax (broadcasted by the eNB in the SIB2 message), the UE declares the network unavailable and then, the failure of the RA is signaled to the upper layers and terminates the random access process (Laya, 2014), (7), (11), (Cheng, 2015), (13).

Several studies have been conducted to improve 3GPP native solutions (15)- (Nan Jiang, 2018). These studies have proposed different approaches to reduce the network congestion of massive M2M communications. In (35), Qinghe Du and al. have proposed an efficient mechanism for massive access control over differentiated M2M services, including delay-sensitive and delay-tolerant services. Specifically, based on the traffic loads of the two types of services, their proposed scheme dynamically partitions and allocates the random access channel (RACH) resource to each type of services. The RACH partition strategy is thoroughly optimized to increase the access performances of M2M networks. The simulation results demonstrate the effectiveness of their design. They conclude that their scheme can outperform the baseline access class barring (ACB) scheme, which ignores service types in access control, in terms of access success probability and the average access delay. However, their results do not allow highlighting the probability of success for simultaneous requests for connection. In (Luis Tello-Oquendo, 2018), Luis Tello-Oquendo and al. have proposed a dynamic ACB scheme in which an estimate of the current number of M2M devices in backoff state is used to adjust in real-time the backoff rate parameter. They evaluate the key performance indicators (KPIs) of dynamic ACB in several scenarios with different degrees of traffic load and compare them with those of a static ACB with optimal parameters. They show that the dynamic ACB outperforms the static one offering shorter access delay and higher successful access probability, while its impact on H2H communications KPIs is negligible. In (Israel Leyva-Mayorgaa et al., 2018), Israel Leyva-Mayorga and al. have presented a novel ACB configuration (ACBC) scheme that can be directly implemented at the cellular base stations. In their ACBC scheme, they calculate the ratio of idle to total available resources, which then serves as the input to an adaptive filtering algorithm. Their main objective of the latter is to enhance the selection of the barring parameters by reducing the effect of the inherent randomness of the system. Their results show that ACBC scheme greatly enhances the performance of the system during periods of high congestion. In addition, the increase in the access delay during periods of light traffic load is minimal. In (Shree Krishna Sharma, ?), Shree Krishna Sharma has presented a detailed overview of the existing and emerging solutions towards addressing RAN congestion problem, and then identifies potential advantages, challenges and use cases for the applications of emerging Machine Learning (ML) techniques in ultra-dense cellular networks. He discusses some open research challenges and promising future research directions. In (Israel Leyva-Mayorga, 2019), Israel Leyva-Mayorga and al. compare the benefits of using different filtering methods to configure an access control scheme included in the 5G standards: the access class barring (ACB), according to the intensity of access requests. These filtering methods are a key component of their proposed ACB configuration scheme, which can lead to more than a three-fold increase in the probability of successfully completing the random access procedure under the most typical network configuration and mMTC scenario.

**Back-Off Process**: According to LTE-A (7), if the RA procedure fails, whatever the cause, the UE must start all over again. For this, it waits a random time, $U(0, BI) ms$, until he can try again. The BI (Back-off Indicator) is the standby indicator defined by the eNB and its value is between 0 and 960 ms. In this paper, we consider an exponential back-off, in which the waiting time depends on the number of preamble transmissions, $P_r$ previously performed. The waiting time is obtained by $T_{Bo} = U(0.10 * 2^{P_r-1})$ (7), knowing that $P_r \leq preambuleTransMax$.

**Congestion Resolution Methods** In order to solve network overload problems due to recurrent collisions which occur in high-load environments, 3GPP has advocated the following approaches (14):

- **Access Class Barring**
- **Separate RACH Resources for MTC**
- **Dynamic Allocation of RACH Resources**
- **Backoff Specific Scheme**
- **Slotted Access**
- **Pull based Access (Paging)**
In (Nan Jiang, 2018), Nan Jiang and al., for relieving the congestion during the RA in mIoT networks, they model RA procedure, and analyze as well as evaluate the performance improvement due to different RA schemes, including power ramping (PR), back-off (BO), access class barring (ACB), hybrid ACB and back-off schemes (ACB&BO), and hybrid power ramping and back-off (PR&BO). To do so, they develop a traffic-aware spatio-temporal model for the contention-based RA analysis in the mIoT network, where the signal-to-noise plus-interference ratio (SINR) outage and collision events jointly determine the traffic evolution and the RA success probability. Based on this analytical model, they derive the analytical expression for the RA success probability. Based on this analytical model, they derive the analytical expression for the RA success probability. Based on this analytical model, they derive the analytical expression for the RA success probability.

**Access Class Barring:** With ACB, the UEs are divided into 16 classes. Classes 0-9 are so-called "normal" classes, class 10 is dedicated to emergencies while those between 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 network information at regular time intervals to all the UEs in a message block called SIB2. In this message are among others, the probability $P_{ACB}$ ($P_{ACB} \in (0-1)$) called ACB factor (acBarringFactor) intended for the normal classes, the coefficient ac_BarringTime which enters the calculation of the barring time $T_{barring}$ as well as other necessary information to the UE. The UEs wishing to access the network generate a number $q$ between 0 and 1. If the number $q$ generated by the UE is less than or equal to $P_{ACB}(q \leq P_{ACB})$, then the terminal is authorized to initiate the random access procedure RACH. Otherwise, it has to wait a $T_{barring}$ time before resuming the process. In this way, it is possible for the eNodeB to control network overloads by assigning an optimal value to $P_{ACB}$. $T_{barring}$ can be calculated as follows (7):

$$T_{barring} = (0.7 + 0.6 * rand) * ac\_BarringTime$$  \hspace{1cm} (4)

**SYSTEM MODEL**

**Acb Performance Mesures:** During the RACH access process under a low-load network environment, the success rate of requests is very high with low delay. But, as soon as a heavy load is felt, this rate decreases very quickly to reach a critical threshold, and the delay with. This can continue to reach levels that no longer allow the network to guarantee a good quality of service. Since 3GPP has not proposed any method of overload detection or prevention, we propose in this paper a model for overload detection that allows the network to anticipate this situation, in order to enable it to set up the overload resolution solution. The detection method consists in providing the network (base station) with the information on the state of the use of the resources i.e. the number of preambles used that we name $R_{used}$ and the threshold from which the collisions become important as calculated in (Mahamadou, 2018) and named $R_{limit}$. Then, when $R_{limit}$ is reached, i.e. $R_{used} \geq R_{limit}$ the ACB is then activated in order to decongest the network. The ACB will remain active until the number of preambles used falls below the $R_{limit}$ i.e. $R_{used} < R_{limit}$. At that point ACB is disabled.

$$R_{used} = R \left(1 - \left(1 - \frac{1}{R}\right)^{N}\right)$$  \hspace{1cm} (5)

![Fig. 7. Non Empty Preambles ($R_{used}$) Vs Successful Preambles](image)

![Fig. 6. Access Class Barring Flowchart](image)

![Fig. 8. Proposed A-Acb Flowchart](image)

Where $rand$ is a random number generated by the UEs (i.e.$q$) after passing a first ACB check failed and before a second attempt. The values of the $ac\_BarringTime$ can range from 4s, 8s, 16 s to 512 s.
offers 200 RAO per se for contention opportunity every 5 ms, and the number of preambles is given in Table 2. There is one random access after a network failure. M2M devices access the network simultaneously, for example, in our work, we assume the traffic model 2 which is allows a fair comparison of new congestion solution proposals.

We have proposed and implemented an algorithm to anticipate this situation i.e. in order to enable it to set up the overload detection or prevention and then resolve the congestion of the networks. This algorithm, described above is showed in figure 7.

### Table 1. M2m Traffic Models For Rach Evaluation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>TrafficModel1</th>
<th>TrafficModel2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of M2M devices</td>
<td>1000,3000,5000,10000,30000</td>
<td>1000,3000,5000,10000,30000</td>
</tr>
<tr>
<td>Arrival distribution</td>
<td>Uniform distribution over T</td>
<td>Beta distribution over T</td>
</tr>
<tr>
<td>Distribution period (T)</td>
<td>60 seconds</td>
<td>10 seconds</td>
</tr>
</tbody>
</table>

### Table 2. Basic simulation parameters for rach evaluation (14)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellb and width</td>
<td>5 MHz</td>
</tr>
<tr>
<td>PRACH Configuration Index</td>
<td>6</td>
</tr>
<tr>
<td>Total number of preamble(R)</td>
<td>54</td>
</tr>
<tr>
<td>Preamble Trans Max</td>
<td>10</td>
</tr>
<tr>
<td>Number of uplink grants per RAR</td>
<td>3</td>
</tr>
<tr>
<td>Preamble detection probability for the i(^{th}) preamble transmission</td>
<td>( P_d = 1 - \frac{1}{2^i} )</td>
</tr>
<tr>
<td>W(_{Max})</td>
<td>5 sub-frames</td>
</tr>
<tr>
<td>mac-Contention Resolution Timer</td>
<td>48 sub-frames</td>
</tr>
<tr>
<td>Back off Indicator(BI)</td>
<td>20 ms</td>
</tr>
<tr>
<td>HARQ(_) transmission probability for Msg3 and Msg4 (non-adaptive HARQ)</td>
<td>10%</td>
</tr>
<tr>
<td>Maximum number of HARQ TX for Msg3 and Msg4 (non-adaptive HARQ)</td>
<td>5</td>
</tr>
<tr>
<td>Periodicity of RAOs</td>
<td>5ms</td>
</tr>
<tr>
<td>Preamble trans mission time</td>
<td>1ms</td>
</tr>
</tbody>
</table>

### Table 3. Distribution of UES in the success and the collisions of connection

<table>
<thead>
<tr>
<th>1(^{st}) computation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success distribution</td>
<td>171</td>
<td>172</td>
<td>170</td>
<td>172</td>
<td>170</td>
<td>171</td>
<td>171</td>
<td>170</td>
<td>171</td>
<td>172</td>
</tr>
<tr>
<td>Collisions distribution</td>
<td>829</td>
<td>828</td>
<td>830</td>
<td>828</td>
<td>830</td>
<td>829</td>
<td>829</td>
<td>830</td>
<td>829</td>
<td>828</td>
</tr>
</tbody>
</table>

We have proposed and implemented an algorithm to anticipate this situation i.e. in order to enable it to set up the overload detection or prevention and then resolve the congestion of the networks. This algorithm, described above is showed in figure 7.

### Rach Performance Measurements:

We assume a stable power for the UEs 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 UE in the same access time interval. Therefore, it does not send any response to the concerned UEs. Note that random access can occur only in a frequency block, called RA (RAO) opportunities, specified by the base station, using a variable called the PRACH Configuration Index. That is, the PRACH is the physical layer responsible for the RACH scheduling. In our work, the configuration index of the RACH 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.

Since comparison of new congestion control methods is not straightforward due to the large number of variables and test scenarios, 3GPP TR 37.868 (14) defines two different traffic models (see table 1) and four Key performance indicators to assess network performance with M2M communications. This allows a fair comparison of new congestion solution proposals. In our work, we assume the traffic model 2 which is considered as an extreme scenario in which a large number of M2M devices access the network simultaneously, for example, after a network failure. The performance parameters of the RACH are given in table 2. There is one random access opportunity every 5 ms, and the number of preambles is \( R = 54 \) for contention-based RA. Under these conditions, the system offers 200 RAO per second, the maximum number of preamble transmission per UE \((\text{preambleTransMax})\) is 10.

If two or more UEs select the same preamble in the same RAO, we assume that the eNB will not be able to decode any of the preambles. Therefore, the eNB will not send the random access response (step 2). UEs will only detect a collision if Msg2 is not received in the RAR window.

Four performance indicators for the evaluation of RACH capacity for LTE were defined by (33):

- **Collision Probability** (\( P_c \)), defined as the ratio of the number of occurrences when two or more MTC devices send a random access attempt using exactly the same preamble and the total number of opportunities (with or without access) during that period. So, the number of collision preambles is divided by the total number of preambles.
- **Probability of access success** (\( P_s \)), the probability of success is defined as the number of preambles successfully received by the eNB divided by the total number of RACH preamble transmissions in a time interval.
- **Statistics of the number of preamble transmissions** (\( \text{PT} \)) per access attempt is defined as the Cumulative Distribution Function (CDF) of the number of preamble transmissions to perform a random access procedure, for MTC devices for which random access has been successfully performed.
- **Access delay statistics**, defined as the delay CDF for each random access procedure between the first access attempt and the completion of the random access procedure, for MTC devices for which random access has been successfully completed.

We denote by \( R \) the number of available preambles and \( N \) the number of UEs allowed to perform random access by transmitting preambles in a time interval \( T_d \) called activation
time \((0 \leq t \leq T_A)\), with a probability \(p(t)\) during which \(p(t)\) follows a beta distribution with parameters \(\alpha = 3, \beta = 4\) as in equation (6):
\[
p(t) = \frac{t^{(\alpha-1)}(T_A-t)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}
\]
Where \(\text{Beta}(\alpha, \beta)\) is a beta function.

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

The number of devices generating access requests in the \(i^{th}\) access opportunity, meaning access intensity (14), is defined as in equation (7):
\[
\lambda_i = N \int_{t_{i-1}}^{t_i} p(t)dt
\]
Where \(i = 1, 2, 3, ..., I_A\)

Equiprobable preambles \((1/R)\) are considered. The probability that one of the N MTC chooses one and only one preamble successfully is given by the binomial law in equation (8):
\[
P_{\text{success}} = \binom{N}{1} \left( \frac{1}{R} \right) \left( 1 - \frac{1}{R} \right)^{N-1}
\]
\[
P_{\text{success}} = \frac{N}{R} \left( 1 - \frac{1}{R} \right)^{N-1}
\]
(8)

The probability for which one of the N MTC does not transmit any preamble (Idle) is given by the binomial law in equation (9):
\[
P_{\text{idle}} = \binom{N}{0} \left( \frac{1}{R} \right)^0 \left( 1 - \frac{1}{R} \right)^{N-0}
\]
\[
P_{\text{idle}} = \left( 1 - \frac{1}{R} \right)^N
\]
(9)

Success rate (successfully transmitted queries) can be achieved by multiplying the probability of success (equation 1) by the amount of available resources as in equation (10):
\[
\text{Success} = R \times P_{\text{success}}
\]
\[
\text{Success} = R \times \frac{N}{R} \left( 1 - \frac{1}{R} \right)^{N-1}
\]
(10)

Probability of collision can then be calculated as in equation (11):
\[
P_{\text{collision}} = 1 - P_{\text{success}} - P_{\text{idle}}
\]
\[
P_{\text{collision}} = 1 - \frac{N}{R} \left( 1 - \frac{1}{R} \right)^{N-1} - \left( 1 - \frac{1}{R} \right)^N
\]
(11)

**SIMULATIONS RESULTS AND DISCUSSIONS**

In this section, we first evaluate the performance of collisions detection method algorithm. We set the number of available preambles to be used for estimation \(R\) in each RACH slot to 54, the probability factor \(p\) is set to 0.6, and the overload detection parameter \(R_{\text{limit}}\) is 47. The RACH system model is as described in section III-B, and the arrival of UEs follows the time-limited beta distribution as indicated in (14).
We propose a RACH contention resolution method for LTE systems. We investigate the issue about both overload occurrence, so it can anticipate the congestion state by activating ACB. The use of $R_{\text{limit}}$ alleviates the effect of collision over the network. We observe in figure 9 that for the small number of UEs, the success probability in almost equal to 1 and the collision probability is almost equal to 0. We consider in this simulation that all UEs transmit preamble i.e. $P_{\text{idle}} = 0$. So, the sum of the success probability and the collisions probability is always equal to 1. This is verified by the probability theory expressed by equation (12):

$$\sum_{i=1}^{n} p(e_i) = 1$$  \hspace{1cm} (12)

Where $p(e_i)$ is the probability of the event $e_i$ in a set of $n$ events. Figure 9 shows that for $n = 1000$ UEs whose tried to connect to the network simultaneously, 54 UEs are connected successfully and 946 have reached collisions. This is logic true and coherent because the available resources (preamble) are $R = 54$; so only 54 among 1000 UEs have the luck to be connected with success. When no limitation is set, meaning ACB always is active, ACB shows high success rate and less collisions when the requests begin to be sent to the eNodeB. But success rate drops rapidly as UEs numbers going high, inversely collisions increase rapidly as in figure 10. So, in $R_{\text{limit}}$ free scheme, overload occurs earlier than when A-ACB $R_{\text{limit}}$ method is applied. We have analyzed the access delay in function of the maximum number of the collisions per UE; the results are presented in figures 11 and 12. We may observe that access delay goes higher when $R_{\text{limit}}$ is applied, it the case of figure 11. In contrary the $R_{\text{limit}}$ free A-ACB scheme shows a little reduction of access delay as showed in Figure 12; this is due to the fact that collisions occur before the $R_{\text{limit}}$ value is reached. When $R_{\text{limit}}$ free scheme is applied, all devices undergo directly access barring class by generating access factor $q$ before accessing the network as explained previously. As well in figure 11 as in figure 12, we observe that the maximum delay per collision is between 4 and 4.5 seconds. We may conclude that A-ACB shows a waiting delay acceptable for the network. We are also interested in the random distribution of the UEs concerned by the success of connection and the UEs concerned by the collisions. So, we have computed this distribution for $n = 1000$ UEs. The objective is to quantitatively estimate the distribution of the UEs likely to be connected to the network and the UEs likely to cause the collisions. Figure 13 presents the obtained results. The distribution presented in figure show that for $n = 1000$ UEs whose tried to connect to the network simultaneously, 171 UEs belong to the group likely to be connected successfully and 829 UEs belong to the group likely to cause the collisions. To highlight the randomness of the distribution of the UEs, we have repeated the computation of the distribution and the results are presented in table 3. We observe that the distribution of the UEs likely to be successful connected and the UEs likely to cause the collision is random. However, we remark that the average number of the UEs likely to be successfully connected is 171 and the average number of the UEs likely to cause the collisions is 829. The standard deviation calculated for the each of the two groups of UEs is 0.8165. We may conclude that these distributions have the same standard deviation which may be taken as 1 UE.

**Conclusion**

In this paper we investigate the issue about both overload detection and contention resolution in LTE systems. We propose a RACH contention resolution method for LTE systems. We assume that all UEs are activate between $t = 0$ and $t = 10$ second, and the random access intensity is given (14). We set the maximum number of transmitters (UEs) $n$ to 1000. The algorithm showed in figure 7 is executed and the result is presented in figure 9. We can see in figure 9, that as the number of transmitters becomes larger, the eNodeB observes the increase of collisions rate and the decrease of success rate. When $R_{\text{limit}}$ is applied, collisions are attenuated in the range between 1 and 100 first UE requests. This prevents the network overload occurrence, so it can anticipate the congestion state.
systems, including an overload detection method. Our method called Advanced ACB (A-ACB) combine the 3GPP ACB and our overload detection method in which a preamble utilization threshold Rlimit is calculated. The method gets the eNodeB informed about the state of use of preamble resources, which enable the network to anticipate congestion states by activating The ACB. Simulation results show that A-ACB can provide high RACH success probability while guarantee a fair packet delay in comparison to the native 3GPP ACB scheme described in 3GPP TR37.868. It is simple, less complex and easy to implement in the LTE. Moreover, it does not require large investments.

REFERENCES


Gartner, “Gartner says the internet of things installed base will grow to 26 billion units by 2020”, 2013.

X. Han, T. J. Lim and J. Xu, “Heterogeneous access class barring with QoS guarantee in machine-type communications”, Trans. Emerging Tel. Tech. 2015


3GPP TS 36.331 version 12.7.0 Release 12, “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification”, Oct. 2015


3GPP TS 36.213 version 12.4.0 Release 12, “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures”, Feb. 2015


3GPP TSG RAN WG2 #71bis, “Email discussion output - (71#52) UMTS/LTE: MTC Simulations assumptions and output”, Huawei, October 11-15, 2010, Xian, China


*******