An Adaptive Access Class Barring Scheme for Handling Massive M2M Communications in LTE-A

Israel Leyva-Mayorga, Miguel A. Rodriguez-Hernandez, Vicent Pla, Jorge Martinez-Bautset and Luis Tello-Oquendo
Instituto Universitario de Tecnologías de la Información y Comunicaciones (ITACA)
Universitat Politècnica de València, Spain
email: {isleyma, marodrig, vpla, jmartinez, luiteloq}@upv.es

Abstract—In the near future, a massive number of machine-to-machine communication devices will provide with ubiquitous information and services, but such a high number of devices can cause severe congestion in relaying networks. This is the case of LTE-A networks, in which the random access channel suffers from congestion whenever a bulk of user equipments (UEs) attempt to access the cellular base stations in a highly synchronized manner. Under these conditions, the access class barring (ACB) scheme can effectively reduce congestion in exchange for a longer access delay. Therefore, the access delay can be greatly affected if the configuration parameters of ACB are not correctly adapted to the traffic intensity. In this paper, we present a novel adaptive ACB scheme that can be directly implemented in the LTE-A system. In this scheme, the configuration parameters are updated by means of an adaptive filter algorithm: namely the least-mean-square algorithm. Results show that our adaptive ACB scheme sharply enhances the access of UEs during periods of high congestion; i.e., the access delay can be reduced up to 50 percent when compared to other ACB schemes. In addition, the access of UEs under normal operating conditions is not affected.

Index Terms—Access class barring (ACB); adaptive access control; least-mean-square (LMS); LTE-A networks; machine-to-machine (M2M) communications.

I. INTRODUCTION

Modern society is in the need for ubiquitous device connectivity, where small devices exchange data autonomously to provide continuous access to information and services. Machine-to-machine (M2M) communication stands for the autonomous exchange of data between devices and is a fundamental component of the Internet of Things (IoT) [1]. Nevertheless, the rapid increase in the number of interconnected devices in M2M applications (the projected number of mobile-connected devices by 2020 is around 11.6 billion [2]) poses serious challenges regarding the signaling capabilities of relaying networks.

LTE-A networks present the best option for the interconnection of devices (user equipments, UEs, in LTE-A) in M2M applications because its infrastructure has already been widely deployed and can meet high QoS requirements [3]. However, the random access channel (RACH) of LTE-A was designed to handle human-to-human communications, where a few UEs (when compared to those expected in M2M communications) attempt to access the cellular base stations (evolved NodeBs, eNBs, in LTE-A) [4]. Consequently, the capacity of the RACH can be exceeded when a bulk of UEs transmit in a highly synchronized manner, which is a typical behavior in M2M applications.

UEs access the eNB by means of the random access (RA) procedure, which comprises a four-message handshake: preamble transmission (only permitted during random access opportunities, RAOs), random access response (RAR), connection request and contention resolution messages; these are described in detail in Section III. Several studies have concluded that the RA procedure of LTE-A is not efficient at handling M2M traffic [5]–[8] and the rapid increase in the number of UEs will undoubtedly increase the frequency and severity of congestion periods. This creates the imperative need for developing efficient access control schemes.

Access class barring (ACB) is an access control scheme included in the LTE-A Radio Resource Control (RRC) specification [9] that redistributes the UE accesses through time. For this, each UE may randomly delay the beginning of its RA procedure according to a barring rate and a mean barring delay; these are broadcast by the eNB through the SystemInformationBlockType2 (SIB2). That is, upon arrival, the UEs are allowed to begin its RA procedure with a probability equal to the barring rate; otherwise, the beginning of the RA procedure is delayed according to the mean barring time. The ACB scheme is further explained in detail in Section III.

Several studies have concluded that when ACB is implemented and correctly configured, the sporadic congestion in the RACH can be relieved in exchange for a longer access delay [4], [5], [10]. Consequently, it is clear that the configuration of the ACB scheme must be continuously adapted to the traffic intensity. By doing so, the performance under highly congested intervals can be optimized without increasing the access delay of UEs under normal operating conditions.

In practice, the dynamic selection and modification of ACB parameters is hindered by: a) the limited information available at the eNB regarding the number of contending UEs, b) the delay of notification mechanisms and c) the selectivity of the ACB scheme (only the UEs that have not yet begun its RA procedure are subject to the ACB scheme).

In this paper, we propose a novel dynamic ACB scheme that can be directly implemented in the LTE-A system. Our ACB scheme relies on an adaptive filter to continuously adapt its configuration parameters according to the traffic intensity.
The contributions of our adaptive ACB scheme are:

- It effectively operates with minimal information regarding the level of congestion on the RACH. In fact, the only congestion indicator that is needed is the number of successful accesses per RAO, which obviously is known by the eNB.
- It efficiently supports the periodicity of the ACB configuration-update mechanisms; i.e., SIB2. Note that the shortest SIB2 periodicity defined in the specification is 80 ms, which is much longer than the RAO periodicity [9].
- It strictly adheres to the RRC specification [9] unlike the most of the other dynamic ACB schemes [11]–[13]. That is, the ACB scheme only affects the UEs that have not yet begun its RA procedure (performed its first preamble transmission).

In our adaptive ACB scheme, the eNB selects a mean barring time that remains constant throughout the operation of the network and performs the periodic calculation of the barring rate by means of an adaptive filter. The mean barring time and the barring rate are then broadcast by the eNB through the SIB2, as defined in the specification [9]. By implementing our scheme, the access success probability of UEs is sharply increased during sporadic periods of high congestion, while the access delay of UEs during periods of low congestion is not affected.

The rest of the paper is organized as follows. In Section II, we perform a review of the existing dynamic ACB schemes. In Section III we present the model of our adaptive ACB scheme and describe in detail each one of its components. In Section IV, we identify the optimal configuration of our adaptive ACB scheme and showcase its benefits by comparing its performance with that of the following scenarios: a) a static adaptive ACB scheme and b) the time that the UEs are barred is deterministic (of one RAO) and b) no ACB is implemented under normal operating conditions. The article concludes with the discussion of results.

II. RELATED WORK

ACB is one of the most promising access control schemes for the LTE-A RACH [4], [5]. In fact, several studies agree with its efficiency during sporadic periods of congestion, even if its configuration parameters remain constant throughout the congestion period [4], [5], [10]. Under these conditions, the dynamic modification of ACB parameters can, potentially, optimize the access of UEs, but the limitations of the ACB scheme and of the LTE-A system hinder its practical implementation.

A promising approach to implement a dynamic ACB scheme is to estimate the total number of contending UEs, which then enables the selection of an optimal barring rate. Such an approach is proposed in [11], [12], but in both studies the authors assume that: a) all of the UEs are subject to the ACB scheme even after the beginning of its RA procedure; b) the time that the UEs are barred is deterministic (of one RAO) and; c) the barring rate is calculated and broadcast by the eNB at each RAO. While the first two assumptions are simplifications of the ACB scheme, broadcasting a newly calculated barring rate at each RAO is certainly not achievable in a practical implementation. Still, the latter is a common assumption in the literature that is also present in [13].

The lack of information related to the number of contending UEs and collided preamble transmissions is recognized in [14], where the authors propose the use of a state transition diagram for the dynamic activation of a barring scheme. That is, the state of the system depends on the average number of successful preamble transmissions and the barring scheme is activated when the system reaches the state of severe congestion. However, the authors do not consider that the number of available uplink grants is limited and is (in a typical RACH configuration [7]) lower than the number of successful preambles. In addition, the performance of the presented ACB scheme is only assessed in terms of the success probability.

Recognizing the limitations of the ACB scheme and LTE-A system that hinder the implementation of a dynamic ACB scheme, we propose the use of an adaptive filter for the dynamic selection of one of the two ACB parameters: the barring rate.

III. ADAPTIVE ACCESS CLASS BARRING SCHEME

In this section we describe in detail the operation of our adaptive ACB scheme. As a baseline, we show its block diagram in Fig. 1, from which two main blocks can be clearly identified: the random access (RA) and the adaptive ACB configuration. In the RA, depicted in the upper part of Fig. 1, the UEs are subject to the ACB scheme in the first random access opportunity (RAO) after arrival; once a UE is no more subject to the ACB scheme, it proceeds to perform the RA procedure. Note that the RA, depicted in the upper part of Fig. 1, is performed in each and every RAO; i.e., with a periodicity of $T_{RAO}$ as defined in the specification [9], [15].

In the adaptive ACB configuration, depicted in the lower part of Fig. 1, the eNB calculates and broadcasts the ACB parameters, i.e., the mean barring time and the barring rate, through the SIB2. Hence, the periodicity of the adaptive ACB configuration process is determined by the SIB2 periodicity, $T_{SIB2}$. Consequently, each part in Fig. 1 (upper/lower) operates at a different time scale, which is represented by the following notation. The discrete unit $t$ stands for the epoch number when the epoch duration is $T_{RAO}$, whereas the discrete unit $j$ stands for the epoch number when the epoch duration is $T_{SIB2} \times T_{RAO}$. The processes performed in each of the blocks shown in Fig. 1 are now described in detail.

A. Random access (RA)

In the ACB scheme, the UEs are divided into access classes (ACs) 0 to 15 according to its traffic characteristics. Each UE belongs to one out of the normal ACs (from ACs 0 to 9) and can also belong to one or more out of the five high-priority categories (ACs 11 to 15) [9]. The eNB calculates and broadcasts a barring rate, $p_{AC}(j)$, and a mean barring time, $t_{ACB}$, through the SIB2, that are applied to ACs 0 to 9
and to one or more of the ACs 11 to 15. In a typical RACH configuration (see Table I in Section IV), the SIB2 is broadcast every $T_{\text{SIB2}} \in \{16, 32, 64, \ldots, 1024\}$ RAOs. At the $i$th RAO, the $a(i)$ UEs that are about to begin its RA procedure and subject to the ACB scheme must first perform the barring check as follows [9], [15].

repeat
  Select the latest mean barring time, $t_{\text{ACB}}$, and barring rate, $p_{\text{ACB}}(j)$, broadcast by the eNB.
  Generate $r \equiv U[0, 1]$ a random number with uniform distribution between 0 and 1.
  if $r \leq p_{\text{ACB}}(j)$ then
    Initiate the RA procedure.
  else
    Generate $r_2 \equiv U[0, 1]$ and select the barring time as
    \[ t = [0.7 + 0.6 \times r_2] \times t_{\text{ACB}}. \] (1)
  end if
end if

Let $n(i)$ be the number of UEs that are allowed to begin its RA procedure at the $i$th RAO. Also let $b(i)$ be the number of UEs that are set to delay the beginning of its RA procedure; each of these UEs waits for a random time, $t$, and continues to be subject to the ACB scheme in the next RAO.

The UEs that are no more subject to the ACB scheme proceed to perform the RA procedure, illustrated in Fig. 2, as follows.

Preamble (Msg1): At the first step of the RA procedure, each UE randomly selects one out of the available preambles and sends it towards the eNB in a RAO. Due to the orthogonality of the different preambles, multiple UEs can access the eNB in the same RAO, using different preambles. The eNB decodes the preambles transmitted (with sufficient power) by exactly one UE. In this study we assume that a collision occurs whenever two or more UEs transmit the same preamble at the same RAO. This goes in line with the 3GPP recommendations for the performance analysis of the RACH [7] and with most of the literature [10], [12], [14], [16]–[18].

Exactly two subframes after the preamble transmission has ended (this is the time needed at the eNB to process the received preambles), the UE begins to wait for a time window, RA response (RAR) window, to receive a RAR message from the eNB.

RAR (Msg2): The eNB computes an identifier for each successfully decoded preamble and sends the RAR message. It includes, among other data, time alignment, uplink grants for the transmission of Msg3, the backoff indicator, and the assignment of a temporary identifier. There can be up to one RAR message in each subframe, but it may contain several uplink grants; each of which is associated to a successfully decoded preamble.

Since the downlink resources are limited, a maximum of $N_{\text{UE}}$ uplink grants can be sent within a RAR window. Let $s(i)$ be the number of UEs that receive an uplink grant at the $i$th RAO; these UEs proceed with the transmission of Msg3.

Connection request (Msg3): After receiving the corresponding uplink grant, the UEs adjust their uplink transmission time according to the received time alignment and transmit a scheduled connection-request message, Msg3, to the eNB through dedicated resources.

Contention resolution (Msg4): The eNB transmits a contention resolution message as an answer to Msg3. If a UE does not receive Msg4 within the Contention Resolution Timer, then it declares a failure in the contention resolution and schedules a new access attempt.

The maximum number of allowed preamble transmissions for each UE is broadcast by the eNB through the SIB2 [9]. Whenever a collision occurs and, if the maximum number of preamble transmissions has not been reached, the UE waits for a random backoff time (determined by the backoff indicator),

![Figure 1. Block diagram for the implementation of our adaptive ACB scheme.](image)

![Figure 2. Four-message handshake performed in the RA procedure of LTE-A.](image)
then selects and transmits a preamble (Msg1) at the next RAO.

B. Adaptive ACB configuration

In our adaptive ACB scheme, the eNB calculates a normalized resource utilization indicator for the \( j \)th SIB2 broadcast interval immediately before broadcasting the SIB2; i.e., from the \( (i - T_{\text{SIB2}} + 1) \)th to the \( i \)th RAO. For this, let \( s(j) = [s(i), s(i - 1), \ldots, s(i - T_{\text{SIB2}} + 1)] \) be the vector that contains the number of successful UE accesses at each RAO within the \( j \)th SIB2 broadcast interval; the normalized resource utilization indicator can be easily calculated as

\[
\hat{s}(j) = \frac{1}{T_{\text{SIB2}} \times N_{\text{UG}}} \sum_{m=0}^{T_{\text{SIB2}}-1} s(i - m),
\]

(2)

since \( T_{\text{SIB2}} \times N_{\text{UG}} \) is the maximum number of uplink grants that can be sent within the SIB2 broadcast interval. Next, a suggested barring rate for the \( j \)th SIB2 broadcast interval is calculated as

\[
u(j) = 1 - \hat{s}(j).
\]

(3)

That is, the suggested barring rate is the complement of the normalized resource utilization indicator for the \( j \)th SIB2 broadcast interval. Then, the vector that contains the suggested barring rates for the last \( M \) SIB2 broadcast intervals, \( u(j) = [u(j), u(j - 1), \ldots, u(j - M + 1)] \), serves as the input to an adaptive filter.

In this study, the least-mean-square (LMS) adaptive algorithm is used, though any adaptive algorithm can be selected. The LMS is an adaptive filter algorithm that is widely used because of its simplicity [19]. Specifically, the complexity of the LMS algorithm scales linearly with the input length, \( M \); i.e., \( 2M+1 \) multiplications and \( 2M+1 \) additions per iteration (adaptive ACB configuration process) are performed [19]. Since the eNBs possess great computational power, they can efficiently implement the LMS algorithm.

It consists of two processes: the filtering and the adaptive process, which result in a feedback loop as it can be observed in the lower part of Fig. 1.

In the filtering process, the output of the filter, \( y(j) \), is computed from the input vector (of length \( M \) ), \( u \), and then compared to a desired response, \( d(j) \). In our adaptive ACB, the desired response, \( d(j) \), represents the desired barring factor at each RAO under normal operating conditions; i.e., in the case of no congestion. In other words, \( d(j) \) is the barring rate that will be applied during periods in which \( \hat{s}(j) \approx 0 \). Hereafter, we select \( d(j) = 1 \) in order to minimize the access delay of UEs under normal operating conditions, for which the ACB scheme is not needed. The error, \( e(j) \), is the difference between the desired response and the filter output.

In the adaptive process, the coefficients of the transversal filter, \( w(j) = [w_0(j), w_1(j), \ldots, w_{M-1}(j)] \), are adjusted automatically according to the obtained error.

The LMS adaptive filter algorithm is as follows.

**Require:** the number of coefficients of the transversal filter, \( M \).

**Require:** the adaptation step size, \( \mu \).

Initialize the vector of filter coefficients, \( w \), for all \( j \) do

Filtering process:

\[
y(j) = w^T(j) u(j)
\]

Adaptive process:

\[
e(j) = d(j) - y(j)
\]

\[
w(j + 1) = w(j) + \mu e(j) u(j)
\]

end for

In this study we select the initial values of the filter coefficients, \( w \), to be 0, but a determined value can be selected if prior knowledge of their values is available.

The barring rate for the \( j \)th SIB broadcast interval is given as

\[
p_{\text{ACB}}(j) = \min \{ y(j), 1 \},
\]

(6)

and will remain constant until the next SIB2 is broadcast, \( T_{\text{SIB2}} \) RAOs later. On the other hand, the mean barring time, \( t_{\text{ACB}} \), is selected once and remains constant throughout the operation of the network.

In the remainder of the paper, we strive to identify the parameter settings, i.e., mean barring time, \( t_{\text{ACB}} \), the number of filter coefficients, \( M \), and adaptation size, \( \mu \), that optimize the performance under a massive M2M scenario and also under normal operating conditions.

IV. PERFORMANCE EVALUATION

In this section we evaluate the efficiency of our adaptive ACB scheme in terms of three key performance indicators (KPIs): the probability to successfully complete the RA procedure (success probability), \( P_s \); the average number of preambles transmitted by the successfully accessed UEs, \( \mathbb{E}[k] \); and the 95th percentile of access delay, \( D_{95} \), i.e., the access delay experienced by 95 percent of the successfully accessed UEs is lower than or equal to \( D_{95} \).

To obtain our results, we have developed a C-based simulator. Each simulation begins 4000 RAOs before the first UE arrival occurs in order to allow for the stabilization of the filter coefficients, \( w(j) \). Then, \( N = 30000 \) UE arrivals are distributed according to the selected Traffic model; the simulation ends when the \( N \) UEs have completed their RA procedure. Simulations are performed \( x \) times until the cumulative results obtained up to the \( x \)th simulation differ from the ones obtained up to the \((x - 1)\)th simulation by less than 0.01 percent for each of the KPIs of interest.

Throughout this study, we assume a typical RACH configuration, prach-ConfigIndex 6 as suggested in [7] and select the shortest SIB2 periodicity defined in the specification [9], \( T_{\text{SIB2}} = 16 \) RAOs (80 ms); other configuration parameters are shown in Table I.

As a baseline, we identify the configuration of our adaptive ACB that minimizes the access delay of UEs while achieving \( P_s \geq 0.95 \) during a high congestion interval, in which \( N = 30000 \) UE arrivals follow a Beta(3,4) distribution over 10 seconds (2000 RAOs) as described by Traffic model.
2 in [7]. Specifically, we investigate the impact that the mean barring time, \( t_{ACB} \), the number of filter coefficients, \( M \), and the adaptation step size, \( \mu \), have on the system performance. For the latter, we have observed that selecting \( \mu = 0.0025 T_{SIB2}/M \) results in an adequate response from the LMS adaptive filter; as such, we use this value hereafter.

Fig. 3 shows the \( P_s \) and \( D_{95} \) obtained by implementing our adaptive ACB scheme with \( M \in \{1, 8, 32, 64\} \) and \( t_{ACB} \in [0.1, 3] \). We also show these KPIs for the case in which no LMS adaptive filter is implemented; hence \( p_{ACB}(j) = \mu(j) \). By doing so, we are able to observe the benefits of the LMS adaptive filter algorithm.

It is worth noting that, since we have selected \( P_s \geq 0.95 \) as the minimum acceptable success probability, \( D_{95} \) is only shown in Fig. 3b for the combinations of \( t_{ACB} \) and \( M \) that lead to \( P_s \geq 0.95 \). It can be clearly seen that the access delay obtained with the dynamic ACB with no adaptive filter is much higher than the one obtained with the adaptive ACB, assuming that an adequate \( M \) is selected. For the latter, the worst performance is obtained by selecting \( M = 1 \). On the other hand, the optimal configuration of the adaptive ACB for Traffic model 2 is \( M = 32 \) and \( t_{ACB} = 0.6 \text{ s} \), because it results in the lowest \( D_{95} \), given that \( P_s \geq 0.95 \); hence, it is selected for the remainder of this paper.

To compare the impact that a static ACB scheme (in which the configuration parameters remain constant throughout the whole congestion interval) and our adaptive ACB scheme have on the UE arrivals, we show the average number of first preamble transmissions and the average number of successful accesses per RAO for three configurations in Fig. 4. In the first one, no ACB scheme has been implemented; therefore, the first preamble transmission of each UE occurs at the first RAO after its arrival. In the second one, a static ACB scheme with the optimal configuration for Traffic model 2: \( t_{ACB} = 4 \text{ s} \) and \( p_{ACB} = 0.5 \) is implemented [10]. In the third one, our adaptive ACB scheme is implemented with \( M = 32 \) and \( t_{ACB} = 0.6 \text{ s} \).

It is evident that the number of successful accesses per RAO is extremely low if no ACB scheme is implemented; in fact, only a \( P_s = 0.313 \) is achieved in this case. If a static ACB scheme is implemented, the UE accesses are effectively distributed through time, but continue to be delayed even after the congestion has been mitigated, which sharply increases the access delay. If our adaptive ACB scheme is implemented, the UE arrivals are effectively delayed during the period of highest congestion and, as the congestion drops, the UEs are allowed to access the RACH. Specifically, the average number of successful accesses in the RAOs with highest traffic intensity is very close to the number of available uplink grants per RAR window, \( N_{UG} = 15 \). This allows for an efficient use of resources and also reduces access delay.

To conclude the performance evaluation under Traffic model
2. We compare the KPIs obtained with our adaptive ACB scheme with those obtained with a static ACB scheme (the optimal configuration for both schemes is selected). In Fig. 5a, we show that our adaptive ACB scheme results in a slightly higher success probability, $P_s$, a slightly lower mean number of preamble transmissions (energy consumption), $E[k]$, and an access delay that is more than 50 percent shorter than the one obtained with the static ACB.

Next, we compare the KPIs obtained with our adaptive ACB scheme with those obtained when no ACB scheme is implemented under normal operating conditions; i.e., a period with no congestion, in which $N = 30000$ UE arrivals are uniformly distributed over 60 seconds (12000 RAOs) as described by Traffic model 1 in [7]. The performance evaluation under Traffic model 1 is essential because the UE arrivals should not be affected by access control schemes under normal operating conditions. In Fig. 5b it is clearly seen that the performance of LTE-A with no ACB scheme is adequate under normal operating conditions and that neither of the KPIs are affected by implementing our adaptive ACB scheme.

V. CONCLUSION

In this paper, we presented a novel adaptive ACB scheme that takes advantage of an LMS algorithm to enhance the selection of the barring rate. Our adaptive ACB scheme can be directly implemented in LTE-A because it operates according to the behavior of the ACB scheme as defined in the specifications. Results show that our adaptive ACB scheme clearly outperforms the optimal configuration of a static ACB during periods of high congestion; i.e., a higher success probability can be obtained while reducing the access delay up to a 50 percent. In addition, the UE accesses under normal operating conditions are not affected.

ACKNOWLEDGMENT

This work has been supported by the Ministry of Economy and Competitiveness of Spain through the projects TIN2013-

47272-C2-1-R and TEC2015-71932-REDT. The research of L. Leyva-Mayorga was partially funded by grant 383936 CONACYT-Gobierno del Estado de México 2014. The research of L. Tello-Oquendo was supported in part by Programa de Ayudas de Investigación y Desarrollo (PAID) of the Universitat Politècnica de València.

REFERENCES