Universidad Nacional de Chimborazo
NOVASINERGIA, 2018, Vol. 1, No. 2, Diciembre, (38-44)
ISSN: 2631-2654

Extended Access Barring for Handling Massive Machine Type Communication (mMTC) Deployments

Restricción de Acceso Extendida para Manejar Despliegues Masivos de Comunicaciones Tipo Máquina (mMTC)

Luis Tello-Oquendo[1,2]*, José Ramón Vidal[2], Vicent Pla[2], Jorge Martinez-Bauset[2]
[1]*College of Engineering, Universidad Nacional de Chimborazo, Riobamba, Ecuador, 060108*
[2]*ITACA Institute, Universitat Politècnica de València, Valencia, Spain, 46022*
*luis.tello@unach.edu.ec

Abstract: Massive machine type communication (mMTC) has presented a promising moment to generate powerful and ubiquitous connections that face plenty of new challenges. Cellular networks are the potential solution owing to their extensive infrastructure deployment, reliability, security, and efficiency. In cellular-based mMTC networks, the random access channel is used to establish the connection between MTC devices and base stations (eNBs), where the scalable and efficient connectivity for a tremendous number of devices is the primary challenge. To deal with this, the Third Generation Partnership Project (3GPP) has suggested the extended access barring (EAB) as a mechanism for congestion control. The eNBs activate or deactivate EAB using a congestion coefficient. In this paper, an approach to implementing the congestion coefficient is presented so that EAB can operate thus handling congestion episodes in mMTC scenarios. Moreover, the performance of EAB is examined under different MTC traffic loads and paging cycle configurations concerning network key performance indicators (KPIs). Numerical results demonstrate the effectiveness of the proposed method to detect congestion episodes. Also, it is shown that increasing the value of the paging cycle configuration influence on the network behavior under EAB.

Keywords: Extended access barring (EAB), cellular networks, internet of things, performance analysis, random access.

Resumen: *La comunicación masiva tipo máquina (mMTC) ha presentado un momento prometedor para generar conexiones potentes y ubicuas que enfrentan muchos desafíos nuevos. Las redes celulares son la solución potencial debido a su amplio despliegue de infraestructura, confiabilidad, seguridad y eficiencia. En las redes mMTC basadas en comunicación celular, el canal de acceso aleatorio se utiliza para establecer la conexión entre los dispositivos MTC y las estaciones base (eNBs), donde el principal desafío es la conectividad escalable y eficiente para una enorme cantidad de dispositivos. Para hacer frente a esto, el Third Generation Partnership Project (3GPP) ha sugerido la restricción de acceso extendida (EAB) como un mecanismo para el control de la congestión. Las eNBs activan o desactivan EAB utilizando un coeficiente de congestión. En este documento se presenta un enfoque para implementar el coeficiente de congestión de modo que EAB pueda operar y así manejar los episodios de congestión en escenarios de mMTC. También se examina el rendimiento de EAB bajo diferentes cargas de tráfico de MTC y configuraciones de ciclo de paginación en términos de indicadores clave de rendimiento de la red (KPIs). Los resultados numéricos demuestran la efectividad del método propuesto para detectar episodios de congestión. Además se demuestra que el aumento del valor de la configuración del ciclo de paginación influye en el comportamiento de la red bajo EAB.*

Palabras clave: *Restricción de acceso extendida (EAB), redes celulares, internet de las cosas, análisis de rendimiento, acceso aleatorio.*

1 Introduction

Internet of Things (IoT) is one of the most transformative and disruptive technologies of the upcoming wireless systems that has the potential to change the world radically. It is predicted that billions of heterogeneous IoT devices use cellular connections by 2022 (Ericsson, 2017), which empowers individuals and industries to achieve their full potential. Machine type communication (MTC) is becoming the dominant communication paradigm for a wide range of emerging IoT applications including health-care, smart cities, smart grids, smart transportation, and environmental monitoring. In these applications, a vast number of devices are deployed in a specific area to provide ubiquitous services with minimal (or without) human intervention. Thus, the network has to face an increased load and surges of MTC traffic.
The 5th generation (5G) cellular networks will support this huge number of devices generating sporadic small packets at random times. In this context, the random access channel (RACH) is used to start the communication sessions, aimed at delivering this kind of traffic. The RACH is accessed through a four-message handshake contention-based procedure. First, the devices (named UEs hereafter) wait to the next random access opportunity (RAO) and sends a *Msg1* using a randomly chosen preamble from a pool of preambles. *Msg1* is detected at the eNB if the preamble is chosen by just one UE in the current RAO; if not, a collision occurs. For each detected preamble, the eNB sends a random access response (RAR) message, *Msg2*, which includes one uplink grant, from a few grants available. *Msg2* is used to assign time-frequency resources to the UEs for the transmission of the connection request. UEs that received an uplink grant send their connection request message, *Msg3*, using the resources specified by the eNB. Finally, the eNB responds to each *Msg3* transmission with a contention resolution message, *Msg4*. The interested reader is referred to (Tello-Oquendo *et al.*, 2018; 3GPP, 2017b,d,a) for further details.
A fundamental issue is the efficient management of network resources in overload situations; they are produced when many MTC devices react to the same event generating mass concurrent data and signaling transmission. As result network congestion is engendered including both radio access network congestion and signaling network congestion as defined in (3GPP, 2017f). This may cause intolerable delays, packet loss or even service unavailability.
The 3GPP proposes the extended access barring (EAB) as one mechanism to guarantee network availability and help network to meet performance requirements under such MTC load (3GPP, 2017e). EAB selectively restricts the UEs' access attempts to the RACH. Each UE configured for EAB is allocated an access class (AC) in the range 0–9. When the network determines that it is appropriate to apply EAB (using a congestion coefficient), it barres all UEs except one in a given set of ACs, and broadcasts a system information block type 14 (SIB14) containing a 10-bit barring bitmap. The barring is of simple on/off type, where access to each AC is either allowed or not. EAB may be effective whenever the congestion occurs sparingly and during short periods of time (in the order of several seconds). This fact goes in line with the bursty traffic behavior of MTC described in (3GPP, 2011).
In the literature, several studies address the EAB mechanism. Some of them misinterpret the EAB behavior or do not conform to 3GPP specifications (Kim *et al.*, 2017; Hwang *et al.*, 2016). On the other hand, studies such as (Phuyal *et al.*, 2012; Larmo and Susitaival, 2012; Cheng *et al.*, 2015; Toor and Jin, 2017) analyze EAB mainly in terms of access success probability and access delay. In such studies, a practical way to implement the congestion coefficient remains unclear since the number of preamble transmissions is not known at the eNB.
In this paper, a realistic method to implement the congestion coefficient is proposed for the proper functioning of EAB. Then, a thorough performance analysis of this mechanism is conducted and the impact of the paging timing on the EAB performance is evaluated. The main contributions of this study are summarized as follows:

- The EAB scheme is implemented and evaluated in massive MTC scenarios following the 3GPP directives for this kind of studies.
- A method to estimate the congestion coefficient from the number of used preambles at every RAO is proposed, this number of used preambles is effectively known at the eNB so that our proposed solution conforms to the network specifications (3GPP, 2017b,d, 2014, 2017e) and can be successfully integrated into the system.
- The impact of the paging timing on the EAB performance is evaluated. For doing so, a realistic situation is considered in which the congestion coefficient is estimated as mentioned above.

The rest of the paper is organized as follows. Section 2 presents the EAB operation mode; then, a method to compute the congestion coefficient used by the network to turn on or off this mechanism is proposed. Section 3 analyzes in-depth the perfor-

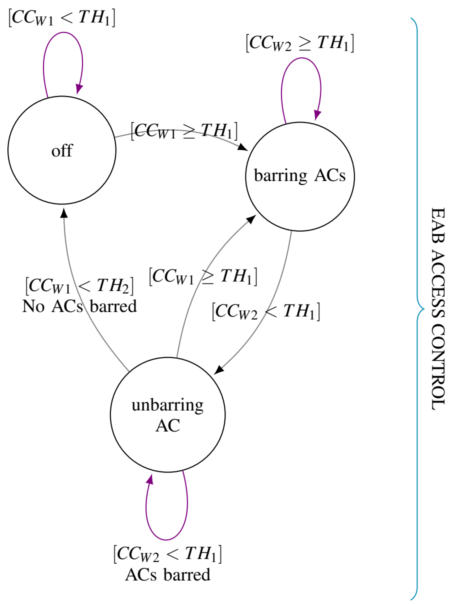

Figure 1: State diagram of EAB operation mode.

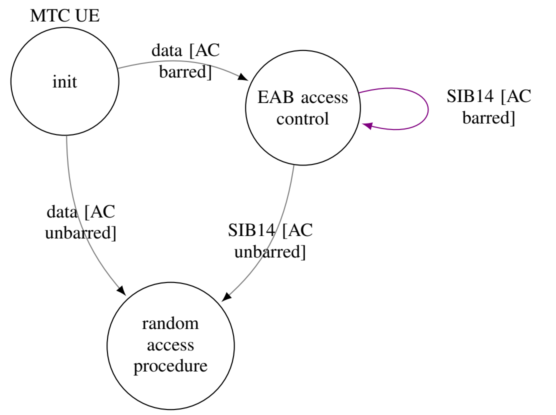

Figure 2: State diagram of the network with EAB access control.

mance of EAB in terms of network key performance indicators (KPIs) and evaluate the impact of the paging timing. Finally, Section 4 draws the conclusions.

2 Extended Access Barring Mechanism

In the following, the operation mode of EAB is presented. Then, a method to implement the EAB congestion coefficient es proposed for its proper functioning.

2.1 EAB operation mode

EAB is activated when congestion is detected. For this purpose, the 3GPP defines a congestion coefficient (CC_W) for a moving time-window of W ms, as detailed in Section 2.2. With a periodicity given by a *modification period* parameter, CC_Ws are used to update the EAB state. If CC_{W1} for $W1 = 1000$ ms exceeds 0.4, EAB is turned on and all ACs except one are barred. From here, every time that CC_{W2} for $W2 = 500$ ms is under 0.4, barring state is released for one AC. The release of ACs proceeds in cyclic order until all ACs are unbarred. Then, if CC_{W1} is under 0.2, EAB is turned off. Fig. 1 illustrate the state diagram of the operation mode of EAB.

The SIB14 contains the bitmap of barred ACs; the eNB broadcasts messages containing SIB14s with a periodicity of $T_{SIB14} \in \{80, 160, 320, 640, 1280, 2560, 5120\}$ ms (3GPP, 2017e). Every time the bitmap has to be changed, the eNB notifies it to the UEs through a *system information change* parameter contained in the paging messages (3GPP, 2017g). Paging messages are sent at specific radio frames and subframes, namely paging frames (PF)s and paging occasions (PO)s, within a paging cycle (T_P). UEs in idle state wake up at their respective PO and read the paging message. UEs calculate their POs from their local identifiers, in order that the POs of the different UEs are distributed homogeneously throughout T_P (3GPP, 2017g). When a UE reads a paging message with *system information change* set to on, it reads the next message containing the SIB14. To make sure that all UEs are notified of all changes and have a chance to update their EAB information, the *modification period* is set to the maximum of T_P and T_{SIB14}, and changes on the bitmap are notified when they are produced but the SIB14 update is delayed up to the next *modification period*. When a UE has to access to the RACH, it checks its barring state from the bitmap contained in the latest updated SIB14 as illustrated in Fig. 2.

2.2 Computing the congestion coefficient

The CC_W is defined as (WG2, 2012)

$$CC_W = 1 - \frac{nRAR_W}{nPT_W}, \quad (1)$$

where $nRAR_W$ is the number of RARs sent during W ms and nPT_W is the number of preamble transmissions during W ms.

To calculate CC_W and therefore update its value at every RAO, the eNB would need to know the number of preamble transmissions at each RAO. But in the commonly assumed collision model defined by the 3GPP, this number is unknown, because those preambles transmitted by more than one UE are not decoded. Therefore, the value of nPT_W defined in (1) should be estimated from the number of preambles used (by at least one UE) at each RAO. This estimation was obtained as follows.

Let $Y_j(i) \in \{0,1\}$ be the random variable that denotes the transmission of preamble j at RAO(i) given that the total number of preamble transmissions at RAO(i) is $n_t(i)$. Then, $Y_j(i) = 0$ when the preamble j has not been transmitted by any UE at RAO(i), and $Y_j(i) = 1$ otherwise. Its probabilities are

$$\begin{cases} \mathbb{P}\{Y_j(i) = 0\} = \left(1 - \frac{1}{R}\right)^{n_t(i)}, \\ \mathbb{P}\{Y_j(j) = 1\} = 1 - \left(1 - \frac{1}{R}\right)^{n_t(i)}, \end{cases} \quad (2)$$

where R is the number of available preambles, and

$$\mathbb{E}\{Y_j(i)\} = 1 - \left(1 - \frac{1}{R}\right)^{n_t(i)}. \quad (3)$$

Then, the number of used preambles at RAO(i), $n_u(i)$, is

$$n_u(i) = \sum_{j=0}^{R} Y_j(i), \quad (4)$$

and its expected value is

$$\mathbb{E}\{n_u(i)\} = R\left[1 - \left(1 - \frac{1}{R}\right)^{n_t(i)}\right]. \quad (5)$$

Since $n_u(i)$ is known at the eNB, and assuming that $\mathbb{E}\{n_u(i)\}$ changes slowly, it can be estimated from a short term time average of $n_u(i)$. Let $\hat{n}_u(i)$ be an estimate of $\mathbb{E}\{n_u(i)\}$ at RAO(i) obtained by exponential smoothing of $n_u(i)$,

$$\hat{n}_u(i) = \alpha \hat{n}_u(i-1) + (1-\alpha)\, n_u(i), \quad (6)$$

with $\alpha < 1$. Finally, from Eq. (5), the estimated value of the number of transmitted preambles used to calculate CC_W is

$$n_t(i) = \frac{\log\left(1 - \frac{\hat{n}_u(i)}{R}\right)}{\log\left(1 - \frac{1}{R}\right)}. \quad (7)$$

Several simulations were conducted to check that the values of CC_W obtained using this estimator are very close to those obtained using the real number of preambles transmitted. Fig. 3 illustrates an example of the values of CC_{W1} for $W1 = 1000$ ms during a congestion episode induced by the MTC traffic benchmark described in Section 3 with $N_M = 30\,000$ MTC UEs arrivals and $(T_P, T_{SIB14}) = (2560, 320)$ ms.

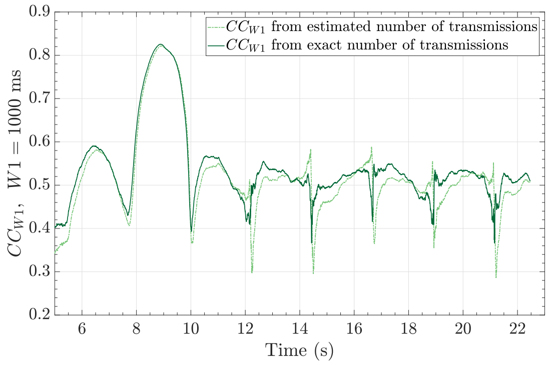

Figure 3: CC_{W1} operation. $N_M = 30\,000; (T_P, T_{SIB14}) = (2560, 320)$ ms.

The CC_{W1} obtained from the estimated number of transmissions is compared with the obtained from the exact value of transmissions. As can be seen, the error in the estimated CC_{W1} is minimal and can be used as real-time congestion coefficient.

3 Performance Analysis

A single cell environment is assumed in which the access requests of MTC UEs follow a $Beta(3,4)$ distribution over a period of 10 s, according to the traffic model 2 specified by the 3GPP in (3GPP, 2011). This traffic model can be seen as an extreme scenario in which a vast number of MTC UE arrivals (ranging from $N_M = 5\,000$ to $N_M = 30\,000$) occur in a highly synchronized manner (e.g., after an alarm that activates them).

Three network KPIs were measured, namely the probability to successfully complete the random access procedure, P_s; the mean number of preamble transmissions needed by the UEs to successfully complete the random access procedure, K; and the access delay (mean and percentiles) of the successful accesses, D. These KPIs are in conformance to the 3GPP directives (3GPP, 2011) to assess the efficiency of the LTE-A random access procedure with MTC.

To obtain the above KPIs, a discrete-event simulator was developed; it fully reproduces the behavior of UEs, eNB, and RACH during the random access procedure. A typical physical RACH configuration, *prach-ConfigIndex 6* (3GPP, 2017c), is assumed where the subframe length is 1 ms and the periodicity of RAOs is 5 ms. $R = 54$ out of the 64 available preambles are used for the contention-based random access procedure and the maximum number of preamble transmissions of each UE, *pream-*

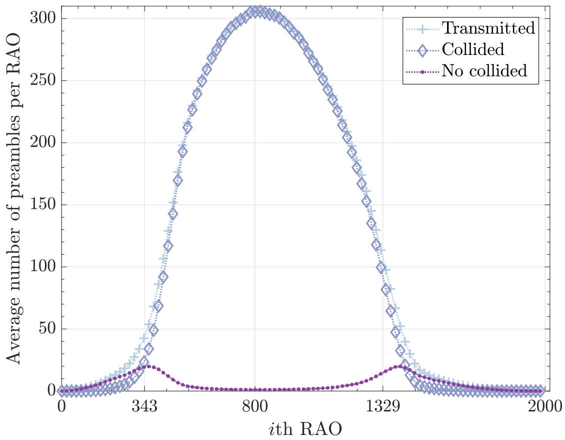

Figure 4: Temporal distribution of preamble transmissions, collided preambles, and successful preambles; traffic model 2, $N_M = 30\,000$.

bleTransMax, is set to 10. Additional system configuration parameters can be found in (Tello-Oquendo *et al.*, 2018, Table III).

First, an overload situation produced when many MTC devices ($N_M = 30\,000$) attempt to access the network is illustrated. Fig. 4 depicts the total preamble transmissions per RAO, preambles with collision (collided), and successful preamble transmissions (no collided); no access control is implemented. It can be seen that when $N_M = 30\,000$, traffic model 2 leads to network congestion, as the $Beta(3,4)$ distribution of UE arrivals exceeds the physical RACH capacity [$c(54) = 20.05$ UE arrivals per RAO as calculated using (Tello-Oquendo *et al.*, 2018, Eq. (4))] from the 343rd to the 1 329th RAO. This massive number of UE arrivals results in a congestion period of $T_c = 4.93$ s, where up to 300 average preamble transmissions per RAO occur at the 800th RAO. As a result, the average number of successful accesses sharply decreases during this period, and the access success probability is severely affected: $P_s = 31.305\,\%$.

In the following, how the standard EAB mechanism handles congestion episodes as the one described above is shown in detail. Fig. 5 plots the temporal distribution of preamble transmissions when EAB is in operation and $(T_P, T_{SIB14}) = (\{640, 1280, 2560\}, 320)$ ms. It can be seen that, when congestion builds up, at $t \approx 4$ s, EAB is enabled. Then, ACs start to be unbarred, one at a time, with a periodicity of *modification period* = T_P. Every time that an AC is unbarred, a number of UEs access the RACH in bursts of periodicity $T_{SIB14} = 320$ ms. Each burst corresponds to those UEs whose *PO*s are between two consecutive broadcasts of the SIB14;

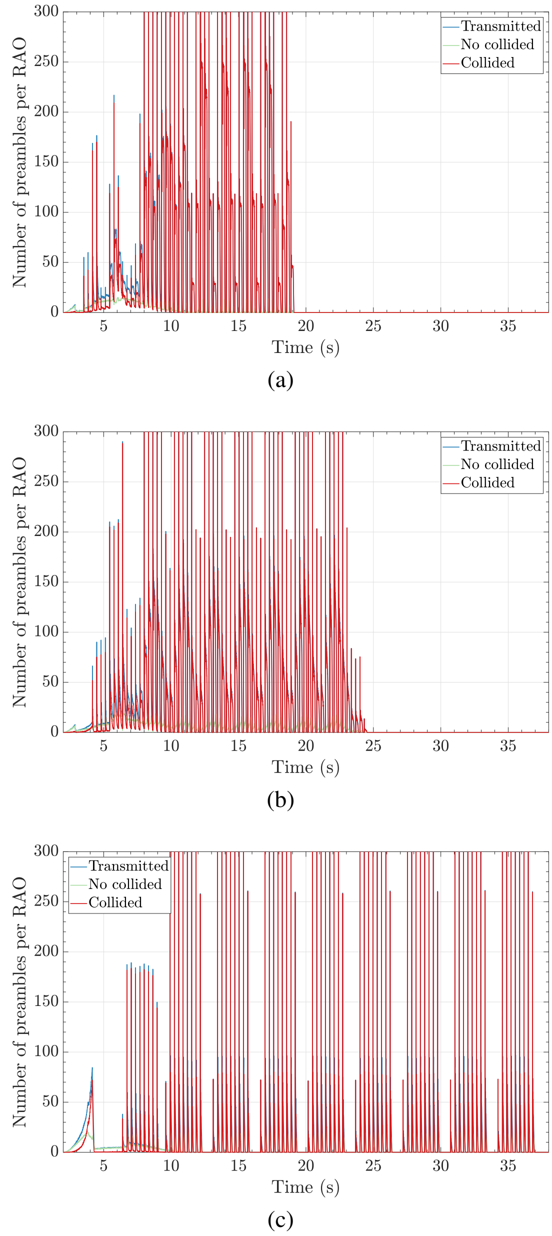

Figure 5: UE access attempts. (a) $T_P = 640$ ms. (b) $T_P = 1280$ ms. (c) $T_P = 2560$ ms.

these UEs access the RACH simultaneously because they update their SIB14 simultaneously. As the T_P configuration value increases, the time required to relieve a congestion episode is longer (e.g., the network is able to handle $N_M = 30\,000$ access requests in ≈ 37 s with $T_P = 2560$ ms, whereas ≈ 19 s are required with $T_P = 640$ ms.)

In Fig. 6, the network KPI results are presented for different configurations of T_P. Considering feasible parameter values in the LTE-A standard, several con-

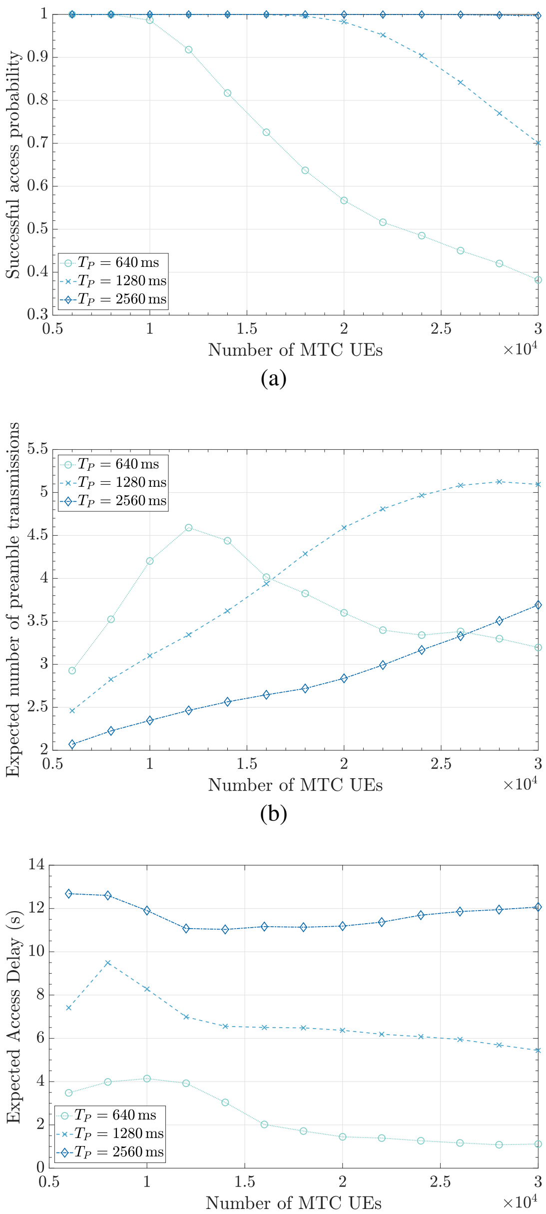

Figure 6: Key performance indicators. (a) Successful access probability. (b) Preamble transmissions. (c) Access Delay.

figurations for the combination of T_P and T_{SIB14} were tested and it was verified that, as in (WG2, 2012; Cheng *et al.*, 2015), combinations in which $T_P < T_{SIB14}$, result in very poor performance in terms of P_s (these results were omitted because of lack of space). Hence, the results for $T_P \in \{640, 1280, 2560\}$ ms and $T_{SIB14} = 320$ ms are shown.

Fig. 6a shows that P_s increases as the T_P duration increases. This is because the greater the T_P, the greater the number of information updates per T_P, which results in lower intensity of the traffic burst after each SIB14 update. However, the cost of increasing T_P is that ACs are unbarred at a slower rate (one AC per T_P), thus increasing D as can be seen in Fig. 6c. In terms of K, Fig. 6b illustrates that increasing T_P reduces this metric particularly in light-loaded MTC scenarios (i.e., $N_M < 16000$) thus decreasing the energy consumption; however, in heavy-loaded MTC scenarios (i.e., $16000 \leq N_M$) this metric increases gradually as T_P increases. This is because the delay caused by the paging mechanism: long paging cycle limits the performance of the barring phase. As a result, some devices cannot receive the EAB updated information on time and their accesses will probably fail; therefore, they should start the access attempt again.

To sum up, increasing T_P rises P_s at the cost of longer access delay while diminishing the number of preamble transmissions in light-loaded MTC scenarios which translates in energy savings for power-constrained MTC devices.

4 Conclusion

In this paper, a practical method to implement the congestion coefficient used by extended access barring (EAB) was proposed. It allows EAB functioning properly for handling congestion in massive machine-type communication (mMTC) scenarios. Through extensive discrete-event simulations, the EAB scheme was analyzed using our proposed congestion coefficient implementation. Then, the impact of the paging timing on the performance of EAB was studied. Numerical results show that a limiting factor of EAB as defined by the 3GPP is that, when a barred AC is released, its UEs initiate their access procedure in bursts of periodicity T_{SIB14}. These bursts cause many preamble collisions during the first RAOs, deteriorating overall performance. On the other hand, increasing the paging cycle rises the successful access probability at the cost of longer access delay while diminishing the number of preamble transmissions in light-loaded MTC scenarios. This results in energy savings useful for the rather power-constrained devices in MTC applications.

Interest Conflict

Authors declare that there is no conflict of interest in this research.

Acknowledgment

This work was supported in part by the Ministry of Economy and Competitiveness of Spain under Grants TIN2013-47272-C2-1-R and TEC2015-71932-REDT.

References

3GPP (2011). *Study on RAN Improvements for Machine Type Communications*. 3GPP, TR 37.868.

3GPP (2014). *Physical layer procedures*. 3GPP, TS 36.213.

3GPP (2017a). *Feasibility Study for Further Advancements for E-UTRA*. 3GPP, TR 36.912.

3GPP (2017b). *Medium Access Control (MAC) Protocol Specification*. 3GPP, TS 36.321.

3GPP (2017c). *Physical Channels and Modulation*. 3GPP, TS 36.211.

3GPP (2017d). *Radio Resource Control (RRC), Protocol specification*. 3GPP, TS 36.331.

3GPP (2017e). *Service Accessibility*. 3GPP, TS 22.011.

3GPP (2017f). *Service Requirements for Machine-Type Communications*. 3GPP, TS 22.368.

3GPP (2017g). *TS 36.304, V14.4.0, User Equipment (UE) procedures in idle mode*.

Cheng, R. G., Chen, J., Chen, D. W., and Wei, C. H. (2015). Modeling and analysis of an extended access barring algorithm for machine-type communications in LTE-A networks. *14*(6), 2956–2968.

Ericsson (2017). Ericsson mobility report/internet of things forecast.

Hwang, R.-H., Huang, C.-F., Lin, H.-W., and Wu, J.-J. (2016). Uplink access control for machine-type communications in LTE-A networks. *Personal and Ubiquitous Computing*, *20*(6), 851–862.

Kim, H., Lee, S. S., and Lee, S. Dynamic extended access barring for improved M2M communication in LTE-A networks. In *Proc. IEEE SMC*, pp. 2742–2747.

Larmo, A. and Susitaival, R. RAN overload control for Machine Type Communications in LTE. In *IEEE Globecom Workshops*, pp. 1626–1631.

Phuyal, U., Koc, A. T., Fong, M. H., and Vannithamby, R. Controlling access overload and signaling congestion in M2M networks. In *Proc. IEEE ASILOMAR*, pp. 591–595.

Tello-Oquendo, L., Leyva-Mayorga, I., Pla, V., Martinez-Bauset, J., Vidal, J. R., Casares-Giner, V., and Guijarro, L. (2018). Performance Analysis and Optimal Access Class Barring Parameter Configuration in LTE-A Networks With Massive M2M Traffic. *67*(4), 3505–3520.

Toor, W. T. and Jin, H. Comparative study of access class barring and extended access barring for machine type communications. In *Proc. IEEE ICTC*, pp. 604–609.

WG2, G. T. R. (2012). Further performance evaluation of EAB information update mechanisms. Meeting N. 77 R2-120270, 3rd Generation Partnership Project (3GPP).