



HAL
open science

Latency Efficient Request Access Rate for Congestion Reduction in LTE MTC

Rémi Bonnefoi, Tarcisio F Maciel, C. Estêvão R. Fernandes

► **To cite this version:**

Rémi Bonnefoi, Tarcisio F Maciel, C. Estêvão R. Fernandes. Latency Efficient Request Access Rate for Congestion Reduction in LTE MTC. ICT 2018 - 25 th International Conference on Telecommunications , Jun 2018, Saint-Malo, France. hal-01769409v2

HAL Id: hal-01769409

<https://hal.science/hal-01769409v2>

Submitted on 29 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Latency Efficient Request Access Rate for Congestion Reduction in LTE MTC

Rémi Bonnefoi*, Tarcisio F. Maciel[†] and C. Estêvão R. Fernandes[†]

*CentraleSupélec/IETR, 35510 Cesson-Sévigné, France, e-mail: remi.bonnefoi@centralesupelec.fr

[†]Wireless Telecommunications Research Group (GTEL), Federal University of Ceará (UFC), Fortaleza, Brazil
maciel@ufc.br, estevao@ufc.br

Abstract—The development of Machine-Type Communication (MTC) in the Long Term Evolution (LTE) standard can lead to an overuse of the Physical Random Access Channel (PRACH). An efficient way to avoid it consists in reducing the number of requests done by end-devices. In this paper, we first study the probability of collision in a congested LTE-MTC network. As a second step, we analyze the latency in the network as a function of the request access rate. Then, we propose a closed-form for a sub-optimal request access rate. We finally conduct numerical simulations in order to assess the performance of the proposed approximation.

I. INTRODUCTION

The development of the Internet of Things (IoT) has led to an increase in the number of devices¹ with wireless connectivity and to the development of Machine-Type Communication (MTC). MTC can be done in dedicated networks or can operate in already deployed cellular networks [1].

This second solution can be limited as cellular networks, such as the Long Term Evolution (LTE) standard developed by the Third Generation Partnership Project (3GPP), are not designed to be used by a very large number of IoT devices. As a consequence, the development of MTC in LTE networks can be restricted by the congestion of the Radio Access Network (RAN) and more precisely of the Physical Random Access Channel (PRACH) [2].

The problem of congestion in LTE MTC can be solved using different techniques. Among others, the evolved Node B (eNB) can prevent devices from accessing the network during predefined periods [3] or can adapt the backoff interval [4]. Another efficient solution to reduce congestion in LTE-MTC lies in adapting the Request Access Rate. This solution is called Access Class Barring (ACB). With it, the eNB broadcasts an ACB factor q . Then, each time a device wants to transmit to the eNB, it computes a uniform random number in $[0; 1]$ and, if this number is below q , transmits the request [5], [6].

Part of the literature about congestion reduction, including [5], [6], does not consider the effect of retransmissions on the probability of collision in the network. In this paper, we consider this effect, where the retransmission mechanism of an end-device can be modeled using a Markov chain. This modeling is inspired by the one proposed in [7] for WiFi (IEEE 802.11) networks. Moreover, a similar modeling has been proposed in [8] for LTE MTC. In [4], this modeling is

used for congestion reduction. However, therein the authors focus on the optimization of the throughput whereas, in the present paper, we focus on the latency.

In this article, we study a solution similar to ACB, we assume that instead of starting a Random Access (RA) procedure each time devices have data to transmit, they start fewer RA procedures and transmit more data in the Physical Uplink Shared Channel (PUSCH) once the access is granted. In other words, suppose a sensor that makes a measure each time there is a change in the measured physical parameter. Instead of making a RA procedure after every measure, the device makes a RA procedure with a probability p_R . The aim of this paper is to find a closed-form approximation for the value of p_R which minimizes the latency. For this purpose, we proceed in two steps: as a first step, we derive an approximation for the probability of collision in congested LTE-MTC networks; then, as a second step, we express the latency in the considered scenario and we use our approximation for the probability of collision in order to derive a closed-form approximation for the value of p_R which minimizes the latency.

The rest of this paper is organized as follows. The system model is introduced in section II. The closed-form approximation for the probability of collision is proposed in section III. In section IV, we derive the expression of the latency in the network and then an approximation for the value of p_R which minimizes it. Numerical simulations are conducted in section V and section VI concludes this paper.

II. SYSTEM MODEL

We assume an LTE eNB which is shared by N end-devices (e.g. sensors). These devices make uplink communications to the eNB. When one of them wants to transmit data to the LTE eNB, it makes a request into the PRACH. If the request is successful, some resource blocks are granted to this device in the PUSCH. Then, these resource blocks are used by the end-device to transmit its packets. We assume that we have one PRACH slot per frame and we denote by T_f the frame duration.

We model the behavior of an end-device with the Markov chain of Fig. 1, which is inspired by the one proposed in [4]. It is also similar to the model used in [8], but where we have neglected the probability to reach the maximum number of repetitions. In idle mode, the device collects data (i.e., by making some measurements). The probability that a device makes a new measure follows a Poisson process of intensity

¹Herein, we use the term *device* to refer to User Equipment (UEs).

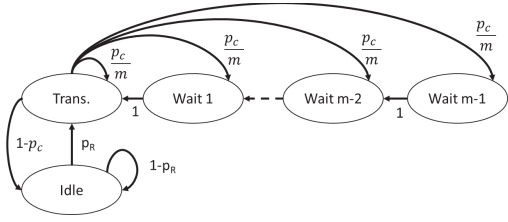


Fig. 1. Markovian modeling of the behavior of an end-device.

λ . Thus, there is a probability $p = \lambda T_f$ that a device has new data to transmit between two PRACH slots. We assume that λ is small enough so that at most one measurement is done between two PRACH slots. As to reduce contention in the network, instead of requesting access to the eNB each time it has a packet to transmit, a device makes a request with a probability $p_R < p$. Moreover, in our analysis of the latency, we assume that p_R is small compared to p .

For the sake of simplicity, we suppose that, once the device starts the transmission procedure, it stops collecting data. The request access is a four-step procedure [9]. We consider a contention-based RA and, in that case, during the first step the end-device that needs to access the eNB chooses one of the $N_{zc} = 64$ possible preambles. If this preamble is chosen by only one device, it can be decoded and the request can be granted by the eNB. If this preamble is chosen by more than one device, there is a collision and the request fails. If the request is successful, we assume that the end-device can transmit its data to the eNB. The first step of this procedure is the focus of this paper, in which we assume that the PRACH channel is the limiting factor. In case of failure, the end-device waits for a random number of PRACH slots uniformly distributed between 0 and $m - 1$, where m denotes the length of the backoff interval. The value of m depends on the Backoff Indicator (BI) transmitted by the eNB [10]. In all this paper, we assume that we have one PRACH slot per frame and we denote $T_f = 10$ ms as the frame duration, which thus corresponds to the delay between two PRACH slots.

In case of success of the RA procedure, all the data collected since the last transmission is transmitted to the eNB in the PUSCH. After the end of the transmission process, the device switches to idle mode and collects new data for the next transmission. In the following, p_c denotes the probability that a request transmitted by a device collides with another request sent by another device. In other words, p_c is the probability that the preamble chosen by the end-device is chosen by at least one other end-device.

III. CLOSED-FORM APPROXIMATION FOR THE PROBABILITY OF COLLISION IN HEAVY TRAFFIC

In this section, we derive a closed-form approximation for the probability of collision in the studied network. As a first step, we compute the probability that a device is transmitting. In the case where $p_R > 0$ and $p_c > 0$ the Markov chain of Fig. 1 is ergodic, assuming a steady state, we can compute

the probability that a device is in each state by finding the expression of the vector:

$$\mathbf{x} = [x_i \quad x_t \quad x_{w_{m-1}} \quad \dots \quad x_{w_1}]^T, \quad (1)$$

where x_i is the probability that the device is in idle mode, x_t the probability that the device is transmitting a request and x_{w_k} is the probability that the device is in waiting state k (see Fig. 1). This vector verifies:

$$\mathbf{M}\mathbf{x} = \mathbf{x}, \quad (2)$$

where \mathbf{M} is the transition matrix of the Markov chain introduced in Fig. 1. This matrix can be expressed as:

$$\mathbf{M} = \begin{bmatrix} 1 - p_R & 1 - p_c & 0 & \dots & 0 & 0 \\ p_R & \frac{p_c}{m} & 0 & \dots & 0 & 1 \\ 0 & \frac{p_c}{m} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \frac{p_c}{m} & 0 & \dots & 1 & 0 \end{bmatrix}. \quad (3)$$

Solving the linear system of (2) and using the normalization equation of the Markov chain, we derive the expression of x_t , the probability that the device is transmitting a request, as:

$$x_t = \frac{1}{1 + \frac{(1-p_c)}{p_R} + p_c \frac{(m-1)}{2}}. \quad (4)$$

Moreover, assuming that the probability of collision follows a Bernoulli process, we can express it as:

$$p_c = 1 - \left(1 - \frac{x_t}{N_{zc}}\right)^{N-1} \underset{N \gg 1}{\approx} 1 - e^{-\frac{Nx_t}{N_{zc}}}. \quad (5)$$

As x_t is a function of p_c , equation (5) is an implicit expression of p_c . In the following, we use a first order Taylor expansion in order to derive a closed-form approximation for the probability of collision when p_c is close to 1. For that purpose, we rewrite x_t as:

$$x_t = \frac{2}{m+1} \frac{1}{1 + \frac{2(1-p_c)}{m+1} \left(\frac{1}{p_R} - \frac{(m-1)}{2}\right)}. \quad (6)$$

In case where p_c is close to 1 or where $\frac{m-1}{2} \approx \frac{1}{p_R}$, we can use the first order Taylor expansion: $\frac{1}{1+x} \underset{x \approx 0}{\approx} 1 - x$, in order to provide an approximation for the probability of transmission:

$$x_t \approx \frac{2}{m+1} \left[1 - \frac{2(1-p_c)}{m+1} \left(\frac{1}{p_R} - \frac{(m-1)}{2}\right)\right]. \quad (7)$$

Inserting the approximation of equation (7) in (5), we obtain:

$$p_c \approx 1 - e^{-\frac{N}{N_{zc}} \frac{2}{m+1} \left[1 - \frac{2(1-p_c)}{m+1} \left(\frac{1}{p_R} - \frac{(m-1)}{2}\right)\right]}. \quad (8)$$

This equality can be rewritten:

$$\begin{aligned} & \frac{-N}{N_{zc}} \frac{4(1-p_c)}{(m+1)^2} \left(\frac{1}{p_R} - \frac{m-1}{2}\right) e^{\frac{-N}{N_{zc}} \frac{4}{(m+1)^2} (1-p_c) \left(\frac{1}{p_R} - \frac{m-1}{2}\right)} \\ & \approx \frac{-N}{N_{zc}} \frac{4}{(m+1)^2} \left(\frac{1}{p_R} - \frac{m-1}{2}\right) e^{-\frac{N}{N_{zc}} \frac{2}{m+1}}. \end{aligned} \quad (9)$$

Moreover, equation (9) has a solution if and only if:

$$p_R \geq \frac{1}{e^{\frac{N}{N_{zc}} \frac{2}{m+1} - 1} \frac{N_{zc}}{N} \frac{(m+1)^2}{4} + \frac{m-1}{2}}. \quad (10)$$

In order to find a closed-form approximation for p_c , we apply the W-Lambert function in equation (9). This function is

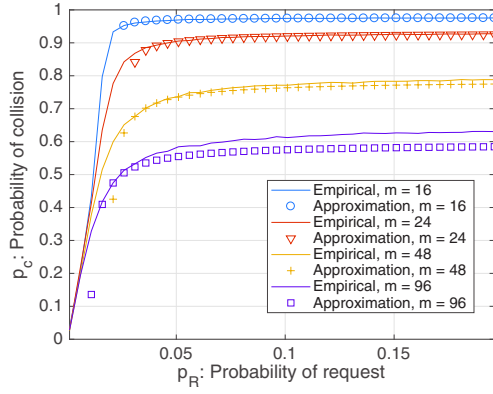


Fig. 2. Probability of collision with 2000 users and for different values of m , the length of the backoff interval.

the reciprocal bijection of the function $x \mapsto xe^x$ [11]. In order to obtain realistic values for the probability of collision, i.e., that belong in $[0; 1]$, we apply the main branch of this function denoted \mathcal{W}_0 . We finally obtain a closed-form approximation for the probability of collision:

$$p_c \approx 1 + \frac{\mathcal{W}_0 \left(\frac{-N}{N_{zc}} \frac{4}{(m+1)^2} \left(\frac{1}{p_R} - \frac{m-1}{2} \right) e^{-\frac{N}{N_{zc}} \frac{2}{m+1}} \right)}{\frac{N}{N_{zc}} \left(\frac{1}{p_R} - \frac{m-1}{2} \right) \frac{4}{(m+1)^2}}. \quad (11)$$

We now conduct some numerical simulations in order to validate the proposed approximation. In Fig. 2, we display the probability of collision versus the probability of access in the studied network with $N = 2000$ users and for different values of m . These values of m are compliant with the LTE standard [10]. The probability of collision is either computed using empirical simulations or using the approximation of equation (11). We can see that the proposed approximation is very precise when the probability of collision is high. Moreover, we can see that the precision of the proposed approximation is reduced when the backoff interval is larger and the probability of collision lower.

In the following, we use the approximation derived here to find the value of p_R which minimizes the latency.

IV. VALUE OF p_R WHICH MINIMIZES THE LATENCY

A. Latency

In this article, we focus on the latency of the access to the eNB which is the delay between the first PRACH slot in which the request for a packet can be transmitted (i.e., the first PRACH slot after the measure) and the successful reception of the PRACH preamble for this packet. We recall here that, instead of transmitting a request each time it has a packet to transmit, i.e., with a probability p , each device transmits a request to the eNB with a probability p_R . And, in case of success of the RA procedure, transmits more data in the PUSCH. Moreover, the device stops collecting data once it starts the RA procedure. In such case, the average latency of the access can be decomposed as:

$$\mathbb{E}[\mathcal{L}] = \mathbb{E}[\mathcal{L}_w] + \mathbb{E}[\mathcal{L}_a], \quad (12)$$

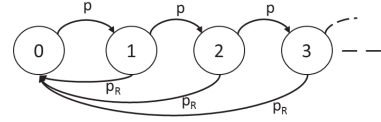


Fig. 3. The behavior of a device in idle mode can be modeled by a queue. We have a probability p to have a new packet in the queue and a probability p_R that the device starts the transmission process and transmits all the packets.

where $\mathbb{E}[\mathcal{L}_w]$ is the queueing delay which is the time between the first PRACH slot during which the device could make its request for this packet and the PRACH slot during which it starts the RA procedure. $\mathbb{E}[\mathcal{L}_a]$ denotes the delay between the beginning of the random access procedure and the successful decoding of the preamble by the eNB.

We start by calculating $\mathbb{E}[\mathcal{L}_w]$. Assuming that p_R is small compared to p , we can model the packet queue in each device, in idle mode, with the queue of Fig. 3. In the steady state, we can compute the probability that there is k packets in the queue which is the probability that the queue is in state k . For that purpose, we can deduce, from Fig. 3, a system of linear equations that represents the state of the system. Solving this system and using the normalization equation of the Markov chain, we derive the probability that the queue is in state k :

$$x_k = \left(\frac{p}{p + p_R} \right)^k \frac{p_R}{p + p_R}, \quad (13)$$

where x_k is the probability that the device has k packets waiting in its queue. We can deduce from equation (13) the average number of packets in the queue:

$$\mathbb{E}[N_p] = \frac{p_R}{p + p_R} \sum_{k=1}^{\infty} k \left(\frac{p}{p + p_R} \right)^k = \frac{p}{p_R}. \quad (14)$$

Equation (14) is derived using the formula of the derivative of a geometric series. We can now use the Little's formula [12] to derive the average delay in the queue:

$$\mathbb{E}[\mathcal{L}_w] = \frac{T_f}{p_R}. \quad (15)$$

We now compute the latency between the beginning of the RA procedure and the successful reception of the preamble. For that purpose, we use the law of total expectation:

$$\mathbb{E}[\mathcal{L}_a] = \sum_{k=1}^{+\infty} P(N_t = k) \mathbb{E}[\mathcal{L}_a | N_t = k], \quad (16)$$

where N_t denotes the number of times the device transmits the preamble until a transmission without collision. Its expression can be written as:

$$P(N_t = k) = p_c^{k-1} (1 - p_c). \quad (17)$$

Moreover, the average time required to transmit N_t times a preamble is:

$$\mathbb{E}[\mathcal{L}_a | N_t = k] = (k - 1) \mathbb{E}[T_{\text{fail}}] = \frac{(k - 1)mT_f}{2}, \quad (18)$$

where $\mathbb{E}[T_{\text{fail}}]$ is the average time between a collision and the next transmission of a preamble by the end-device. Then,

$$\begin{aligned} \mathbb{E}[\mathcal{L}_a] &= \frac{mT_f}{2} (1 - p_c) p_c \sum_{k=1}^{+\infty} (k - 1) p_c^{k-2} \\ &= \frac{mT_f}{2} \frac{1}{1 - p_c} - \frac{mT_f}{2} \end{aligned} \quad (19)$$

Equation (19) is obtained using the derivative of a geometric series. Finally, the total latency can be expressed as:

$$\mathbb{E}[\mathcal{L}] = \frac{mT_f}{2} \frac{1}{1-p_c} - \frac{mT_f}{2} + \frac{T_f}{p_R}. \quad (20)$$

We can see in equation (20) that $\mathbb{E}[\mathcal{L}_w]$ decreases with p_R whereas $\mathbb{E}[\mathcal{L}_a]$ increases with p_R . In the following, we derive a closed-form approximation for the value of p_R which minimizes the delay. This optimum is a trade-off between $\mathbb{E}[\mathcal{L}_w]$ and $\mathbb{E}[\mathcal{L}_a]$.

B. An approximation for the optimal Request Access Rate

In the following, we denote $\gamma = \frac{1}{p_R}$. The function $\mathbb{E}[\mathcal{L}](\gamma)$ has a linear asymptote for $\gamma \mapsto +\infty$ and an asymptotic behavior for $\gamma \mapsto 1$. This can be explained as follows:

- When γ is large, the probability of collision becomes negligible and the latency is equal to $\mathbb{E}[\mathcal{L}_w]$ which is a linear function of γ .
- When γ is low, the probability of collision is high and, p_c can be approximated by equation (11). We will see that, in that case, the latency can be approximated by a linear function.

A consequence of these observations is that the optimum can be approximated by the intersection of these two asymptotes.

We start by calculating the asymptote when $\gamma \mapsto +\infty$. In that case, $p_c \mapsto 0$ and the asymptote of the latency is:

$$\mathbb{E}[\mathcal{L}](\gamma) \underset{\gamma \mapsto +\infty}{\approx} \mathbb{E}[\mathcal{L}_w](\gamma) = y(\gamma) = T_f \gamma. \quad (21)$$

We now analyze the asymptotic behavior of the latency for low values of γ . When γ is low, the network is congested and p_c can be approximated by equation (11). As a consequence, the latency can be approximated by:

$$\mathbb{E}[\mathcal{L}] = T_f \gamma - \frac{mT_f}{2} + \frac{-\frac{mT_f}{2} \frac{N}{N_{zc}} (\gamma - \frac{m-1}{2}) \frac{4}{(m+1)^2}}{\mathcal{W}_0\left(\frac{-N}{N_{zc}} \frac{4}{(m+1)^2} (\gamma - \frac{m-1}{2}) e^{-\frac{N}{N_{zc}} \frac{2}{m+1}}\right)}. \quad (22)$$

Moreover, we have $\frac{x}{\mathcal{W}_0(x)} = e^{\mathcal{W}_0(x)}$, $\forall x \geq -e^{-1}$ [11]. Using this property, the latency of equation (22) can be rewritten:

$$\mathbb{E}[\mathcal{L}] = T_f \gamma - \frac{mT_f}{2} + m \frac{T_f}{2} e^{\mathcal{W}_0\left(\frac{-N}{N_{zc}} \frac{4}{(m+1)^2} (\gamma - \frac{m-1}{2}) e^{-\frac{N}{N_{zc}} \frac{2}{m+1}}\right) + \frac{N}{N_{zc}} \frac{2}{m+1}}. \quad (23)$$

The function $x \mapsto e^{\mathcal{W}_0(x)}$ can be approximated by a linear function [13]. We can consequently approximate (23) by one of the tangents of the latency. A good choice is to approximate the latency by its tangent at the point $\gamma = \frac{m-1}{2}$. Indeed, at this point, the approximation of equation (7) is always valid. The derivative of $\mathbb{E}[\mathcal{L}]$ at $\gamma = \frac{m-1}{2}$ is:

$$\frac{\partial \mathbb{E}[\mathcal{L}]}{\partial \gamma} \left(\frac{m-1}{2} \right) = T_f - \frac{mT_f N}{2N_{zc}} \frac{4}{(m+1)^2}. \quad (24)$$

We can use this derivative in order to approximate the latency by its tangent at the point $\gamma = \frac{m-1}{2}$. The equation of this tangent is:

$$z(\gamma) = \frac{\partial \mathbb{E}[\mathcal{L}]}{\partial \gamma} \left(\frac{m-1}{2} \right) \times \left(\gamma - \frac{m-1}{2} \right) + \mathbb{E}[\mathcal{L}] \left(\frac{m-1}{2} \right). \quad (25)$$

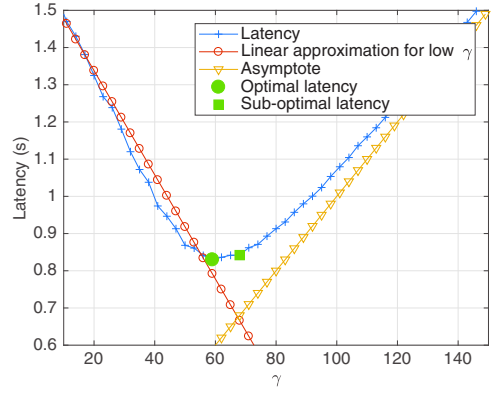


Fig. 4. Evolution of the latency versus γ with $N = 2000$, $m = 24$ and with different values of p . The proposed approximation provides sub-optimal results.

And the optimal value of γ can be approximated by the value of γ denoted γ_i which verifies $y(\gamma_i) = z(\gamma_i)$. We finally derive,

$$\gamma_i = \frac{\frac{N}{N_{zc}} \frac{2(m-1)}{(m+1)^2} + e^{\frac{N}{N_{zc}} \frac{4}{(m+1)^2}} - 1}{\frac{N}{N_{zc}} \frac{4}{(m+1)^2}}. \quad (26)$$

The optimal value of p_R is the inverse of γ_i .

V. SIMULATION RESULTS

We now use numerical simulations in order to validate our approach. Besides, we assess the performance of the proposed approximation. For that purpose, we compute the latency with (20) for different values of p_R (or γ) and we compare the latency obtained with our approximation of γ with the optimal value of the latency.

As a first step and in order to validate our approach, we display in Fig. 4 the two asymptotes with $N = 2000$ and $m = 24$. We can see that the value γ_i at which the asymptotes intersect is near the optimal value of γ .

In order to analyze the performance of the proposed solution, we display in Fig. 5 the evolution of the latency versus N , the number of users, for different values of m , the backoff interval. In these figures, we compare four strategies: the proposed one, the optimal one, and two strategies with a fixed value of p_R . We can see that, for a given value of m , the proposed policy provide near optimal performance over a wide range of values of N . The performance of the proposed solution is reduced when N is large or low. This result can be explained by analyzing the probability of collision p_c . In the case where N is low, the probability of collision is lower when the network is saturated. In that case, the proposed approximations for p_c is less precise. This causes errors and reduces the performance of the proposed solution. Moreover, when N is large, p_c is near 1 and a small error on p_c causes large error on the latency which depends on $\frac{1}{1-p_c}$. The proposed solution is consequently less accurate.

We now compare our strategy with a situation in which p_R does not vary with the number of users. As an example, we analyze in Fig. 5c that choosing $p_R = 0.01$ provides near optimal performance when the number of devices is between 3000 and 4000. However, this value of p_R is no longer optimal

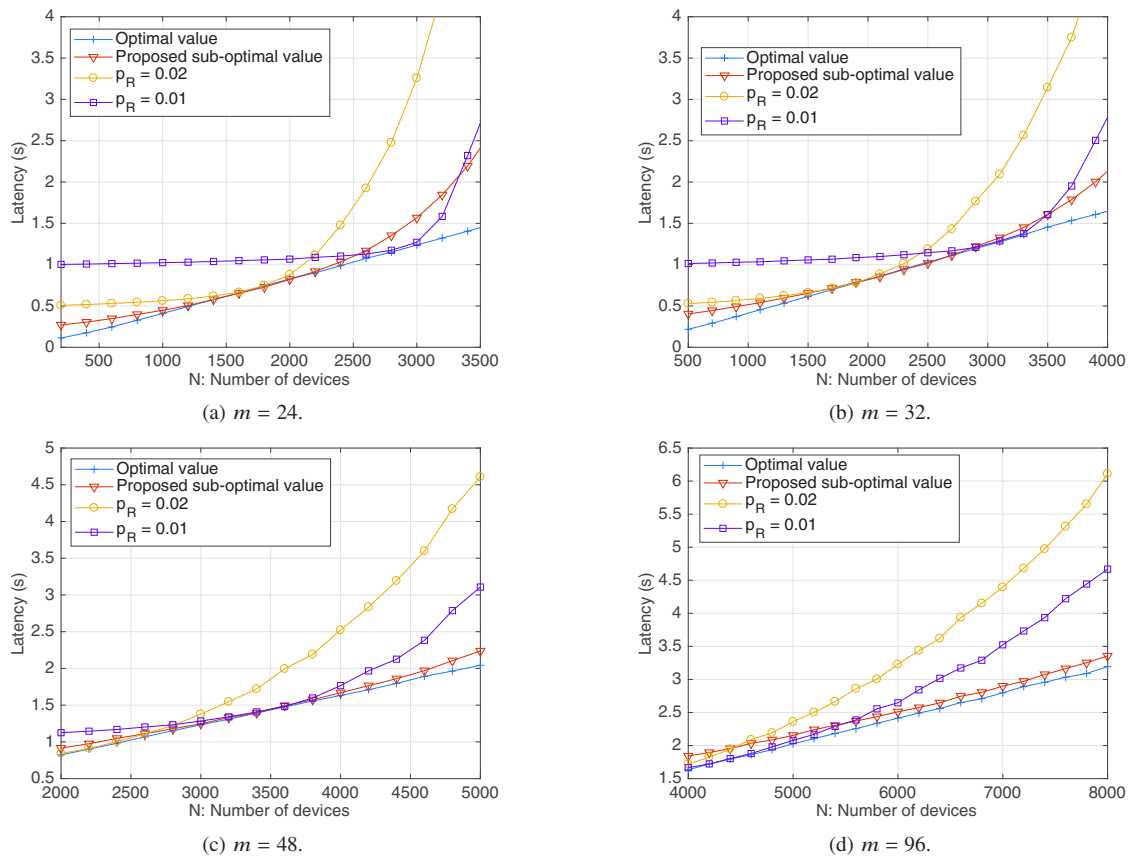


Fig. 5. Evolution of the latency as function of the number of devices N for different values of m : comparison of the proposed approach versus the optimum one and strategies with fixed p_R .

when N increases. In the case where $N = 5000$, the proposed policy provides a 33% gain compared to the use of a fixed value of p_R equal to 0.01.

VI. CONCLUSION

In this paper, we first derive a closed-form approximation for the probability of collision in the PRACH of a congested LTE MTC network. Then, we derive the expression of the latency in the studied network. Afterwards, we use the approximation previously derived to derive a closed-form approximation for the value of the request access rate which minimizes the latency. Finally, we conduct numerical simulations in order to validate our approach. In our future work, we will use the results provided in this paper to optimize other parameters such as the backoff interval. We also envision to analyze the effect of clustering in LTE MTC.

ACKNOWLEDGEMENT

This work has been funded by Université Bretagne Loire (UBL) and by Région Bretagne, France. T. F. Maciel would like to acknowledge CNPq for its financial support under the grants 426385/2016-0 and 308398/2015-7.

REFERENCES

[1] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low Power Wide Area Networks: An Overview," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 855–873, Secondquarter 2017.

[2] M. Y. Cheng, G. Y. Lin, H. Y. Wei, and A. C. C. Hsu, "Overload control for Machine-Type-Communications in LTE-Advanced system," *IEEE Communications Magazine*, vol. 50, no. 6, pp. 38–45, June 2012.

[3] C. W. Chang, Y. H. Lin, Y. Ren, and J. C. Chen, "Congestion Control for Machine-Type Communications in LTE-A Networks," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.

[4] W. Zhan and L. Dai, "Throughput optimization for massive random access of M2M communications in LTE networks," in *2017 IEEE International Conference on Communications (ICC)*, May 2017.

[5] S. Duan, V. Shah-Mansouri, and V. W. S. Wong, "Dynamic access class barring for M2M communications in LTE networks," in *2013 IEEE GLOBECOM*, Dec 2013, pp. 4747–4752.

[6] M. Tavana, V. Shah-Mansouri, and V. W. S. Wong, "Congestion control for bursty M2M traffic in LTE networks," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 5815–5820.

[7] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, March 2000.

[8] X. Yang, A. Fapojuwo, and E. Egbogah, "Performance Analysis and Parameter Optimization of Random Access Backoff Algorithm in LTE," in *2012 IEEE Vehicular Technology Conference*, Sept 2012, pp. 1–5.

[9] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches," *IEEE Comm. Mag.*, vol. 51, no. 6, pp. 86–93, June 2013.

[10] 3GPP, "Technical Specification Group Radio Access Network Evolved Universal Terrestrial Radio Access (E-UTRA) Medium Access Control (MAC) protocol specification," Tech. Rep., Sept. 2017.

[11] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Adv. Comput. Math.*, vol. 5, 1996.

[12] J. D. C. Little, "A Proof for the Queuing Formula: $L = \lambda W$," *Operations Research*, vol. 9, no. 3, pp. 383–387, 1961.

[13] A. Hoorfar and M. Hassani, "Inequalities on the Lambert W function and hyperpower function," *J. Inequal. Pure and Appl. Math.*, vol. 9, 2008.