General Model for RACH Procedure Performance Analysis

Osama Arouk, *Student Member, IEEE*, and Adlen Ksentini, *Senior Member, IEEE*

***Abstract*—This letter introduces an analytical model aiming at modeling the performance of the Random Access CHannel (RACH) procedure, with a special focus on the case of highly synchronized machine type communication (MTC) traffic activated according to beta distribution. The accuracy of the proposed model is then verified through computer simulation. Obtained results demonstrate the effectiveness of the model to capture the RACH procedure behavior when MTC devices are activated in a highly synchronized way. Note that 1) the proposed model is not dedicated only to beta distribution, but other distributions could be used and 2) this model could be used with any congestion control method managing the number of new arrivals, such as access class barring (ACB).**

***Index Terms*—3GPP, LTE, LTE-A, MTC, M2M, congestion control, overload control, RACH procedure, beta distribution.**

I. Introduction

MACHINE-Type-Communication (MTC), or Machine-to-Machine (M2M), could be defined as autonomous devices connecting to the network without a human intervention. Although the traffic generated per MTC is somewhat little, the aggregated traffic of tens of thousands of MTCs (in each cell) will put very high pressure not only on the Radio Access Network (RAN) Part, but also on the Core Network (CN) Part. The main problem in the RAN part (the focus of this letter) comes from the fact that the current cellular mobile networks are designed and optimized for Human-to-Human (H2H) communication, featured by their relative low numbers. Therefore, when a large number of MTCs are trying to get access to the network, via the RACH procedure, the congestion at the RACH stage would surely occur. Therefore, the RACH performance may become a bottleneck of the wireless access network. The higher the congestion of the network is, the lower the number of MTCs to get access is, and the lower the resource utilization is.

In order to evaluate the network performance under different access intensities and to show the effectiveness of the congestion control methods for MTC applications, one should first define a good traffic model that characterizes the behavior of MTCs. 3GPP identified two traffic models: Uniform Distribution (over 60 s) as a realistic scenario (non-synchronized traffic), and Beta Distribution (over 10 s) as an extreme scenario (synchronized traffic) [1]. In the literature, there are many works trying to model the traffic of MTC applications. The work [2] provides a real traffic of MTC applications, where it tends to be periodic during the days of the week. In [3], the authors proposed Coupled Markov Modulated Poisson Processes (CMMPP) framework to model the traffic of MTC, which reflects in a more accurate way the behavior of devices. However, this model is more complex and traffic-dependent, as the parameters of the proposed scheme have to be established for each real traffic MTC application [4].

In this letter, we propose a general analytical model for modeling the performance of RACH procedure, and it will be applied to MTC traffic with Beta Distribution, as an example. The choice of MTC with Beta traffic is motivated by the high load resulting in the network, causing high congestion and system overload. Thus, activating MTCs based on Beta distribution represents the worst case for RACH procedure. In the literature there are many works trying to model RACH procedure, such as [5], [6], [7], [8], [9], [10], [11], [12]. Some of them proposed an analytical model for RACH procedure, but either they did not take into consideration the network's constraints, i.e. the number of responses in RACH procedure is limited [12], or they proposed a proprietary solution for certain traffic model [11], [10], [7]. However, it is stated in [10] that the errors could reach up to 200%. The authors in [5] proposed an analytical model to separate the preambles between M2M and H2H, while in [6] the authors introduced an analysis of Slotted Access scheme. In [7], The authors proposed to dynamically allocate resources to accommodate the number of arrivals in the case of Group Paging. The focus of the authors in [8] is on the design of new RACH procedure for MTC communication. The authors in [9] introduced an analytical model of access reservation for LTE. However, it will be shown in the next section that this model does not accurately capture the behavior of RACH procedure. The authors in [13] propose to model the network with MTC application by Beta/M/1 queue model, where the focus is on the overall system performances. However, the aim of our model is to estimate the number of MTCs contending for the channel in the next RACH period. Moreover, the authors assume that the service time (RACH procedure) follows an exponential distribution, which is not realistic and limits the accuracy of the model. In our model, we take into consideration this limitation. However, none of the above mentioned works have introduced a clear analytical model for RACH procedure. The aim of current letter is to help validating new solutions that address the problem of congestion when activating MTCs according to Beta distribution, or other traffic models. The advantage of our proposed analytical model is simple, accurate (by report to the above-cited models), and general one, i.e. it can be applied with any traffic model. Regarding the consistency with congestion control methods, our model can be directly applied with all the methods that control the number of new arrivals, such as ACB, slotted access, and Extended Access Barring (EAB). Therefore, our model is aligned with LTE/MTC standard, and it does not require any changes to the assumptions already taken by the 3GPP on MTC traffic.

The remainder of this letter is as follows. Section II introduces system model and our proposed analytical model. Computer simulation-based evaluation of our model is presented in section III. Finally, conclusions are introduced in section IV.

Manuscript received June 28, 2015; revised November 19, 2015; accepted November 20, 2015. Date of publication December 3, 2015; date of current version February 12, 2016. The associate editor coordinating the review of this paper and approving it for publication was A. Lioumpas.

The authors are with the IRISA Laboratory, University of Rennes 1, Cedex 35042, France (e-mail: firstname.lastname@irisa.fr).

Digital Object Identifier 10.1109/LCOMM.2015.2505280

1558-2558 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

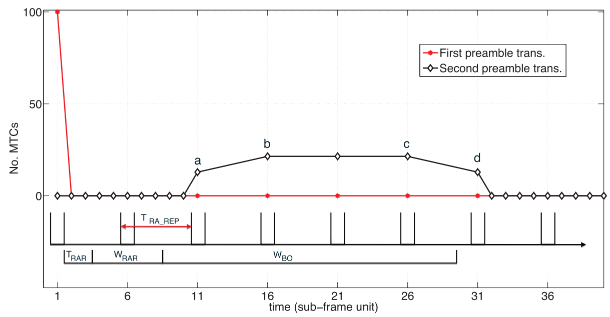

Fig. 1. Number of MTCs at each RA slot for the first and second preamble transmission for $R = 54$, and $M = 100$ [15].

II. The Proposed Recursive Estimation

A. System Model

In our study, we assume that MTCs are activated according to Beta distribution [1], i.e. there will be just one type of traffic. All MTCs will be activated during certain interval I_β. It is assumed that all the MTCs fall within the coverage of just one eNB. Regarding the channel resources, the eNB reserves R random access preambles. Further, we assume that the cell is large enough, which allows the eNB to differentiate the preambles collision when they are chosen by more than one MTC device (you may refer to [8]), i.e. collisions in the preamble space can be detected. Generally, for each preamble transmission the MTC device could take up to $\lceil \Psi/T_{RA\_REP} \rceil T_{RA\_REP} = \lceil (T_{RAR} + W_{RAR} + W_{BO})/T_{RA\_REP} \rceil T_{RA\_REP}$ sub-frames before retrying the transmission of preamble, where T_{RAR}, W_{RAR}, W_{BO}, and T_{RA_REP} are the waiting time before the start of Random Access Response (RAR) window, the size of RAR window, the size of backoff window, and the interval between two consecutive RA slots, respectively, as illustrated in Fig. 1. Therefore, the number of RA slots required in our study, i.e. for Beta Distribution, will be equal to:

$$I_{ra} = \left\lceil I_\beta / T_{RA\_REP} \right\rceil + \left(N_{PT_{max}} - 1 \right) \left\lceil \Psi / T_{RA\_REP} \right\rceil \quad (1)$$

where I_β is the interval of Beta Distribution, in a sub-frame unit, $N_{PT_{max}}$ is the maximum number of preamble transmission. Note that this interval can be adapted to any other traffic by changing the value I_β. In our study, MTCs will be considered as successful ones when receiving the message Msg2 of RACH procedure, ignoring Msg3 and Msg4 [11].

B. Analytical Model

In order to calculate the success, idle, and collision probabilities of the preambles transmission, we use the balls and bins problem [14]. Let R and M be the number of bins/preambles and balls/MTCs, respectively. Let p be the probability a ball falls in a bin, which is equal to $(\frac{1}{R})$ as all the bins are considered to have the same probability. Generally, the probability that k balls fall in a bin, let w, is equal to $Pr(N_w = k) = \binom{M}{k} p^k (1-p)^{M-k}$, where $\binom{m}{k}$ is k-combinations. The probabilities that (none of the balls fall/one ball falls) in the bin w represent the idle and success probabilities, respectively, and they are equal to $P_S(i) = Pr(N_w = 1) \approx (M_i/R)e^{-M_i/R}$, and $P_I(i) = Pr(N_w = 0) \approx e^{-M_i/R}$. M_i is the number of MTCs at the time (i), or the RA slot (i) for our problem. Moreover, the collision probability is the complement of success and idle probabilities, i.e. $P_C(i) = 1 - P_I(i) - P_S(i)$, where this equation is not well positioned in [9]. The number of successful MTCs, which is equal to the number of preambles/bins chosen by only one MTC/ball, is equal to $M_S(i) = R \times P_S(i) = M_i e^{-M_i/R}$. However, the total number of MTCs M_i includes the ones whose preambles are transmitted for the first, second,..., and the $N_{PT_{max}}$-th time, and thus the precedent equation becomes $M_S(i) = \sum_{n=1}^{N_{PT_{max}}} M_i[n] e^{-M_i/R}$. As the probability to detect the n^{th} preamble transmission by the eNB is equal to $(p_n = 1 - e^{-n})$ rather than (1) [1], the precedent equation will be written by:

$$M_S(i) = \sum_{n=1}^{N_{PT_{max}}} M_{S,n}(i) = \sum_{n=1}^{N_{PT_{max}}} M_i[n] p_n e^{-\frac{M_i}{R}} \quad (2)$$

One of the main constraints imposed by the network consists in the fact that no more than N_{ACK} responses can be sent, after the preamble transmission, even if the number of successful preambles is more than N_{ACK}. Note that $N_{ACK} = N_{RAR} \times W_{RAR}$, where N_{RAR} is the number of responses per a RAR message. Therefore, the number of successful MTCs for the nth preamble transmission is:

$$M_{S,n}(i) = \begin{cases} M_i[n] p_n e^{-\frac{M_i}{R}} & ; \text{if } \chi_i \leq N_{ACK} \\ \frac{M_i[n] p_n e^{-\frac{M_i}{R}}}{\chi_i} N_{ACK} & ; \text{otherwise} \end{cases} \quad (3)$$

where $\chi_i = \sum_{n=1}^{N_{PT_{max}}} M_i[n] p_n e^{-\frac{M_i}{R}}$. Based on the analysis in [15], we have:

$$\begin{aligned} x_a(i) &= i + \left\lceil \frac{T_{RAR} + W_{RAR}}{T_{RA\_REP}} \right\rceil = i + \left\lceil \frac{\Gamma}{T_{RA\_REP}} \right\rceil \\ x_{bc}(i) &= i + \left\lceil \frac{\Gamma}{T_{RA\_REP}} \right\rceil + k \\ x_d(i) &= i + \left\lfloor \frac{\Psi}{T_{RA\_REP}} \right\rfloor + 1 \end{aligned} \quad (4)$$

where $x_a(i)$, $x_{bc}(i)$, and $x_d(i)$ are the order of the RA slots (a), (bc), and (d), respectively, within the backoff interval W_{BO} relative to the preamble transmission at the RA slot (i), as illustrated in Fig. 1, $k = 1, 2, \ldots, K_{max}$ and $K_{max} = \lfloor (W_{BO} - \alpha_a W_{BO})/T_{RA\_REP} \rfloor$. Alternatively, the proportions of collided MTCs whose backoff timers expire and retransmit their preambles at the RA slots (a), (bc), and (d) are equal to:

$$\begin{aligned} \alpha_a &= \left(\left\lceil \Gamma / T_{RA\_REP} \right\rceil T_{RA\_REP} - \Gamma \right) / W_{BO} \\ \alpha_{bc} &= \left(T_{RA\_REP} / W_{BO} \right) \\ \alpha_d &= \left(\Psi - T_{RA\_REP} \left\lfloor \Psi / T_{RA\_REP} \right\rfloor \right) / W_{BO} \end{aligned} \quad (5)$$

From the equations (4) and (5) and Fig. 1, we can conclude that the number of MTCs retransmitting their preambles for the n^{th} time is equal to:

$$M_i[n] = \sum_{j=i-k_2}^{i-k_1} \alpha_j M_{C,n-1}(j) \text{ ; for } n = 2 : N_{PT_{max}} \quad (6)$$

where $M_{C,k}(j)$ is the number of collided MTCs corresponding to the preamble transmission at the RA slot (j) for the k^{th} time,

α_j can be α_a, α_{bc}, or α_d, while $k_1 = \lceil \Gamma / T_{RA\_REP} \rceil$ and $k_2 = 1 + \lfloor \Psi / T_{RA\_REP} \rfloor$. Note that k_1 and k_2 are directly obtained from $x_a(i)$ and $x_d(i)$, respectively. Regarding the equation (6), it can be written as follows:

$$M_i[n] = \alpha_a M_{C,n-1}(i-k_1) + \alpha_d M_{C,n-1}(i-k_2) + \sum_{k=i-k_2+1}^{i-k_1-1} \alpha_{bc} M_{C,n-1}(k) \text{ ; for } n = 2 : N_{PT_{max}} \quad (7)$$

For $n = 1$, the number of MTCs will be the value determined by Beta distribution (for our study) or by any traffic model.

C. Beta Distribution

The expected number of arrivals in the RA opportunity (i) is given by the following equation [1]:

$$M_i[1] = M \int_{t_{i-1}}^{t_i} p(t) dt \quad (8)$$

where t_i is the time of the RA opportunity (i) and M is the total number of MTCs. The function $p(t)$ follows Beta distribution:

$$p(t) = \frac{t^{\alpha-1}(T-t)^{\beta-1}}{T^{\alpha+\beta-1} Beta(\alpha, \beta)} \text{ ; } \alpha > 0, \beta > 0 \quad (9)$$

where $Beta(\alpha, \beta)$ is Beta function. In order to find the expected number of arrivals at each RA slot, the authors in [3] propose to use modulated Poisson process to find the number of new arrivals at each time. However, we use simpler way by approximating the integration in the equation (8) by using the trapezoidal rule [16]. Thus, we obtain:

$$M_i[1] = M(t_i - t_{i-1})\,(p(t_{i-1}) + p(t_i))\,/2 \quad (10)$$

The accuracy of this integration depends on the difference $(t_i - t_{i-1})$. The smaller the difference is, the more the accuracy is. As the interval between two consecutive RA slots is equal to T_{RA_REP}, we set $t_i - t_{i-1} = T_{RA\_REP}$, and therefore:

$$M_i[1] = \frac{M * T_{RA\_REP}}{2T^{\alpha+\beta-1} Beta(\alpha, \beta)} \left[t_{i-1}^{\alpha-1}(T - t_{i-1})^{(\beta-1)} + t_i^{\alpha-1}(T - t_i)^{(\beta-1)} \right] \quad (11)$$

As suggested in [1], α and β are set to be 3 and 4, respectively. Equation (11) represents the number of new arrivals according to Beta distribution. After determining the number of MTCs for each preamble transmission at the RA slot (i) (equations 7 and 11), we calculate the number of successful MTCs by the equation (3) and the number of collided MTCs is $M_{C,n}(i) = M_i[n] - M_{S,n}(i)$. It is worth noting that our analytical model can be applied for another traffic models, where it is sufficient to change the number of new arrivals, i.e. $M_i[1]$, according to the traffic model whereas the rest of the model remains unchanged. Therefore, the proposed model is complied with the LTE/MTC standard, without requiring any changes to the assumption already taken by the 3GPP on MTC traffic.

III. Performance Evaluation

The proposed model implemented using C++ language. The parameters of RACH procedure are taken as specified in [1], where the number of MTC devices is $(10000 \sim 50000)$ activated during $(10\ s)$. Note that the choice of high number of MTCs in one cell may represent the case of a large cell. For the sake of simplicity, the detection probability of the transmitted preamble is set to be one. Table I summarizes the more important parameters used in the simulations.

TABLE I
BASIC SIMULATION PARAMETERS

Notations	Definition	Values
α, β	The parameters of Beta Distribution	$3, 4$
I_β	The interval of Beta Distribution	$10 * 1000$
M	Average number of MTCs in the cell	$10000 - 50000$
R	No. preambles in a random access slot	54
BI	Backoff indicator in a sub-frame unit	20
$N_{PT_{max}}$	Maximum number of preamble transmission	10
N_{RAR}	Max. No. RARs that can be carried in one response message	3
W_{RAR}	The size of the random access response window in a sub-frame unit	5
N_{ACK}	Max. No. MTCs that can be acknowledged within the RAR window	$N_{ACK} = N_{RAR} \times W_{RAR}$
$PRACH_{config\_indx}$	PRACH configuration index	$PRACH_{confic\_indx} = 6$
T_{RA_REP}	The interval between two consecutive Random Access (RA) slots	5

A. Performance Metrics

In order to show the performance of our analytical model, we will consider the following metrics: i) the success and collision probabilities, ii) the average number of preamble transmission, iii) and the average access delay. Success probability is defined by the total number of successful MTCs within the maximum number of preamble transmission to the total number of MTCs. Collision probability is the ratio of total number of collided preambles, during the interval I_{ra}, to the total available resources $I_{ra} \times R$. Regarding average delay, it is the total access delay for all the MTCs successfully finished the RACH procedure divided by the number of successful MTCs, and it can be given by the following equation:

$$T = \frac{\sum_{i=1}^{I_{ra}} \sum_{n=1}^{N_{PT_{max}}} D_n M_{S,n}(i)}{\sum_{i=1}^{I_{ra}} \sum_{n=1}^{N_{PT_{max}}} M_{S,n}(i)} \quad (12)$$

where

$$D_n = \frac{T_{RA\_REP}}{2} + (n-1) \left\lceil \frac{T_{RAR} + W_{RAR} + W_{BO}/2}{T_{RA\_REP}} \right\rceil \times T_{RA\_REP} + T_{RAR} + \frac{W_{RAR}}{2}$$

Note that the first term $T_{RA\_REP}/2$ is the average waiting time for the next RA slot, while $W_{RAR}/2$ is the average waiting time for the RAR message. It is sufficient to replace D_n by (n), in the equation (12), to obtain the average number of preamble transmission. It is worth noting that only the successful MTC devices are considered for determining the average number of preamble transmission.

B. Results

Fig. 2 illustrates the success and collision probabilities, and also the Relative Errors (RE) which is $RE = 100 \times (sim - ana)/sim$ where sim and ana are the simulations and analytical results, respectively. From this figure, we clearly see that the analytical model, generally, gives a good approximation with low RE. However, this RE becomes pretty large for certain values of the total number of MTCs, where it reaches its maximum for $M = 15000$. This can be argued by the fact that the network cannot send back responses to more than N_{ACK} MTCs

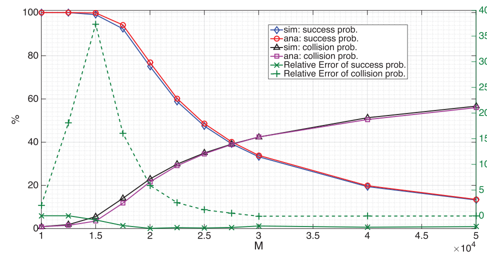

Fig. 2. Success and collision probabilities.

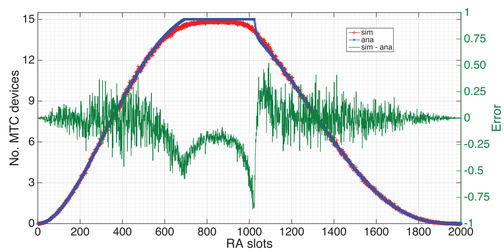

Fig. 3. No. successful MTC devices: M = 15000.

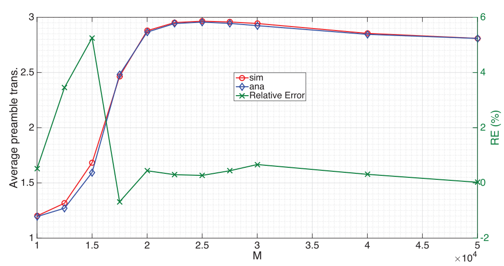

Fig. 4. Average number of preamble transmission.

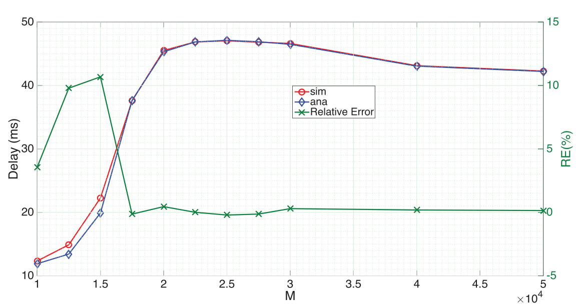

Fig. 5. Average access delay.

even if the number of successful preambles exceeds this value. Therefore, when several experiments are run, the mean value will never reach the constraint (i.e., N_{ACK}) as it is the maximum allowed value, while in the analytical model we used the value N_{ACK} directly if the number of successful preambles is more than N_{ACK} (see equation 3). This case is clearer in Fig. 3, where the upper bound of the analytical results looks as it is cut. Accordingly, solving this problem is one of our future works.

Fig. 4 and 5 show the average number of preamble transmission and the average access delay, respectively. From these figures, we observe that our analytical model is very accurate when the number of MTCs is relatively large (more than 15000 MTCs), where the RE is less than one percent. The worst RE value in our model reaches 40%, which is far from the RE obtained in [10] (i.e., 200%).

IV. Conclusion

The choice of Beta distribution has been motivated by its ability to model a highly synchronized MTC connection to the network, which puts high pressure on the network in terms of congestion. Any solution addressing the congestion problem raised by MTC should be tested with Beta-based traffic model. Accordingly, in this letter an analytical model has been proposed to model the RACH procedure, where this model is tested on Beta Distribution-activated MTC traffic. To demonstrate the effectiveness of our analytical model, many metrics have been considered, such as the success and collision probabilities and the average access delay. Simulation results show the accuracy of the proposed model when the number of MTCs is large (more than 15000 MTCs), where the relative error is less than one percent on the average preamble transmission and average access delay and less than two percent on the success probability. Moreover, our analytical model could be used to model the performance of RACH procedure with any traffic model since the only required change concerns the number of new arrivals whereas the rest remains unchanged.

References

[1] 3GPP, "Study on RAN improvement for machine-type-communications," 3GPP TR 37.868 V11.0.0, Sep. 2011.

[2] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, and J. Wang, "A first look at cellular machine-to-machine traffic," *SIGMETRICS Perform. Eval. Rev.*, vol. 40, no. 1, pp. 65–76, Jun. 2012.

[3] M. Laner, P. Svoboda, N. Nikaein, and M. Rupp, "Traffic models for machine type communications," in *Proc. Wireless Commun. Syst. (ISWCS'13)*, Aug. 2013, pp. 1–5.

[4] M. Laner, N. Nikaein, P. Svoboda, M. Popovic, D. Drajic, and S. Krco, *8-Traffic Models for Machine-to-Machine (M2M) Communications: Types and Applications*. Oxford, U.K.: Woodhead Publishing, 2015, pp. 133–154. doi: http://dx.doi.org/10.1016/B978-1-78242-102-3.00008-3

[5] A. Pourmoghadas and P. G. Poonacha, "Performance analysis of split preamble RAN over-load protocol for M2M communications in cellular networks," *CoRR*, arXiv: 1508.07742, 2015.

[6] C.-H. Wei, R.-G. Cheng, and Y.-S. Lin, "Analysis of slotted-access-based channel access control protocol for LTE-advanced networks," *Wireless Pers. Commun.*, vol. 85, no. 1, pp. 9–27, Nov. 2015.

[7] R. Cheng, F. Al-Taee, J. Chen, and C. Wei, "A dynamic resource allocation scheme for group paging in LTE-advanced networks," *IEEE Internet Things J.*, vol. 2, no. 5, pp. 427–434, Oct. 2015.

[8] D. Wiriaatmadja and K. W. Choi, "Hybrid random access and data transmission protocol for machine-to-machine communications in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 33–46, Jan. 2015.

[9] J. J. Nielsen, D. Kim, G. C. Madueño, N. K. Pratas, and P. Popovski, "A tractable model of the LTE access reservation procedure for machine-type communications," *IEEE Global Commun. Conf. (GC 2015): Wireless Networks Symposium (GC' 15 - Wireless Networks)*, Dec. 2015.

[10] C.-H. Wei, R.-G. Cheng, and S.-L. Tsao, "Modeling and estimation of one-shot random access for finite-user multichannel slotted ALOHA systems," *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1196–1199, Aug. 2012.

[11] W. Chia-Hung, C. Ray-Guang, and T. Shiao-Li, "Performance analysis of group paging for machine-type communications in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 7, pp. 3371–3382, Sep. 2013.

[12] A. Pourmoghadas and P. Poonacha, "Performance analysis of a machine-to-machine friendly MAC algorithm in LTE-Advanced," in *Proc. Int. Conf. (ICACCI'14)*, Sep. 2014, pp. 99–105.

[13] X. Jian, X. Zeng, Y. Jia, L. Zhang, and Y. He, "Beta/M/1 model for machine type communication," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 584–587, Mar. 2013.

[14] R. Motwani, and P. Raghavan, *Randomized Algorithms*. London, U.K.: Chapman & Hall/CRC, 2010.

[15] O. Arouk, A. Ksentini, and T. Taleb, "Group paging optimization for machine-type-communications," in *Proc. IEEE Int. Conf. Commun. (ICC'15)*, Jun. 2015, pp. 6500–6505.

[16] Wikipedia. *Trapezoidal Rule* [Online]. Available: http://en.wikipedia.org/wiki/Trapezoidal_rule