Received <day> <Month>, <year>; | Revised <day> <Month>, <year>; | Accepted <day> <Month>, <year>

DOI: xxx/xxxx

INTERNET TECHNOLOGY LETTERS

A look at Random Access for Machine-Type Communications in 5G cellular networks

Alessandro Grassi | Giuseppe Piro* | Gennaro Boggia

[1]Dep. of Electrical and Information Engineering (DEI), Politecnico di Bari, Italy, Bari

Correspondence
*Giuseppe Piro, v. Orabona 4, 70125, Bari, Italy. Email: giuseppe.piro@poliba.it

Present Address
v. Orabona 4, 70125, Bari, Italy

Abstract

Machine-Type Communications represent a major challenge for the upcoming 5G technology. Future cellular systems, in fact, will be in charge of supporting a huge number of devices generating sporadic small packets at random times. In this context, the Random Access Channel protocol is generally used to initiate the communication sessions, aimed at delivering this kind of traffic. But, occasional peaks of requests, generated when many devices react to the same event, may severely degrade network performance (i.e., by increasing the collision probability). This letter investigates, through computer simulations, the performance of well known procedures for the Random Access Channel, designed for the current 4G technology and the upcoming 5G system in challenging scenarios never seen before. Specifically, the evaluation targets access peaks caused by emergency situations, including every phase of the protocol from the initial contention to the transmission of the application payload. Obtained results highlights pros and cons of available solutions, while showing challenging issues that should be carefully addressed in future research activities.

KEYWORDS:
Machine-Type Communications ; Random Access; 5G; Computer Simulation

1 | INTRODUCTION

In the next few years, cellular-based Machine-Type Communications (MTCs) will grow at a very fast pace, thus outnumbering the usual human-based communications by orders of magnitude (1). They will enable an entirely new class of services, such as smart grids, intelligent transportation systems, and remote health monitoring.

Conventional broadband services leverage human-type communications which typically produce large bursts of data, e.g., while browsing the internet or downloading files. On the contrary, MTCs introduce new traffic patterns: very often, they only produce very small reports, either periodical or triggered by specific events.

The characteristics and the amount of MTC traffic poses significant challenges from the technological point of view, even for state-of-the-art technologies, such as Long-Term Evolution (LTE) and LTE-Advanced (LTE-A). Indeed, while the number of connected objects is increasing (e.g., in the number of tens of billions worldwide (1)), current cellular networks may not be prepared to support this new trend because they are originally designed for different traffic types (2). A critical bottleneck resides in the Random Access Channel (RACH) procedure. As MTC devices experience long sleep periods between two consecutive communications, they lose synchronization with the network and need to re-establish it every time, going through the RACH. To better support emerging MTC services, the worldwide scientific community is devoting a large interest in the research of RACH schemes with greater capacity and tolerance of traffic peaks, to be employed in 5th Generation (5G) mobile networks. Promising solutions are presented in (3, 4, 5).

Nevertheless, the behavior of available approaches could be compromised in particular scenarios where an extremely large number of MTC devices try to access the network in a very short time. For example, security sensors may respond to large-scale emergency events, such as earthquakes or fires. As the RACH procedure is contention-based by nature, an excessive number of connection attempts would result in a high collision rate and make the network unavailable when it is most needed.

Unfortunately, at the time of this writing, a good performance evaluation of methods presented in (3, 4, 5) is still missing for the case of synchronous transmission from a very large number of devices.

To bridge this gap, this letter investigates, through system-level simulations, the performance of two reference schemes, when used in emergency scenarios. Specifically, the conducted study considers the method currently used in LTE and its improved version suitable for the 5G, as described in (3). It is important to note that the reference contribution discussed in (3) is not sufficient to clarify the advantages provided by the new RACH procedure (which is candidate for the 5G) with respect to the baseline approach already used in LTE, especially in challenging scenario with very high traffic loads. In fact, it investigates the first step of the access protocol, thus reporting a packet loss probability due to preamble collisions only; it considers a uniform probability distribution for incoming requests, which is not compatible with emergency scenarios targeted in this letter; and it evaluates the performance of networks with a limited number of devices. This letter, instead, models the entire access protocol (i.e., including data transmission), considers the possibility that the access protocol can also fail because of insufficient radio resources; evaluates a beta distribution for the arrivals for modeling a synchronized burst of requests; and investigates networks with much higher number of devices comparable to typical MTC scenarios expected for the 5G.

Reported results show that the candidate RACH procedure for the 5G can reduce the collision rate and support a larger number of users than the LTE protocol, in the same operating conditions. But, a significant number of users is still unable to transmit their data within the required time.

To conclude, the rest of this letter is organized as follows: section 2 explains the basic RACH procedure of LTE and some of the improved methods presented in the literature, while section 3 describes the simulation setup and presents the attained results. Finally, section 4 draws the conclusion and reports future research activities.

2 | TECHNICAL DESCRIPTION OF COMPARED RANDOM ACCESS SCHEMES

The main purpose of the RACH procedure is to allow a mobile device to establish a connection with the base station, without assuming any previously shared information(6). This process is intrinsically contention-based: since multiple mobile users can access the channel at the same time, collisions are possible. Thus, a key aspect characterizing a RACH procedure is the collision resolution.

This section presents the current LTE random access procedure and some novel techniques proposed in the literature.

2.1 | LTE Random Access

The standardized random access procedure for LTE is based on a four-message handshake initiated by the mobile terminal (see Figure 1) (2). It allows the mobile terminal to achieve tight synchronization with the base station and to receive an allocation of uplink resources.

The first message can only be sent during RACH opportunities, periodically scheduled by the base station. It consists of a preamble sequence, randomly chosen from a set of 64 orthogonal sequences. The purpose of the preamble message is to indicate the presence of an access request, and to allow the base station to estimate the distance of the mobile terminal for the Timing Advance procedure (message 1 in Figure 1). If two or more devices pick the same preamble during the same RACH opportunity, there is a collision and the procedure will fail immediately or at a later stage. More details on RACH opportunities and preamble sequences are given in (7).

After the preambles are detected, the base station sends back a Random Access Response (RAR) message (message 2 in Figure 1). It contains a set of relevant information for each detected preamble, the most important being the allocation of un uplink resource for sending the third message.

If a collision is detected for a specific preamble, then the corresponding information is not sent in the RAR and the devices retry the procedure after a waiting time. On the other hand, if a collision goes unnoticed, then two or more mobile terminals will be assigned the same uplink resource and they will collide again on the third message.

After receiving a resource allocation through the RAR, the mobile terminal can send the Connection Request (message 3 in Figure 1). If there was an undetected preamble collision, two or more devices will send this message over the same resource, i.e. they will collide again and their messages will be lost.

Finally, after a successful delivery of the Connection Request, the base station replies with the last message, that is the Contention Resolution (message 4 in Figure 1). A device that receives a Contention Resolution addressed to him assumes that the random access procedure is completed. Therefore, it can now have a reliable, collision-free communication with the base station. On the other end, if the Contention

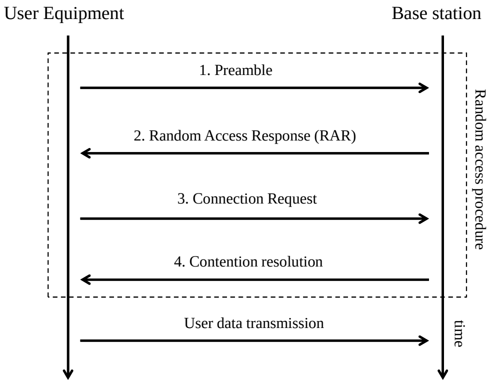

FIGURE 1 Baseline RACH procedure for LTE

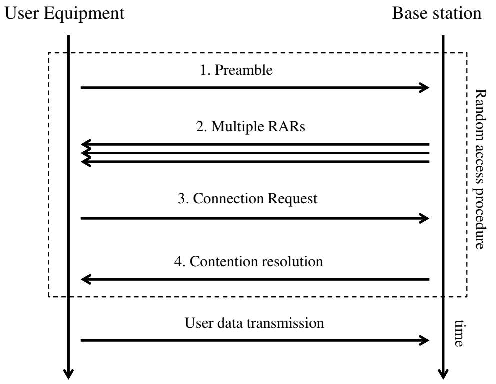

FIGURE 2 Enhanced RACH procedure

Resolution is not received at the proper time, the random access procedure has to be restarted.

In a typical configuration, 54 preambles are dedicated to the contention-based access, while the other 10 are reserved for contention-free access, and the Physical Random Access Channel (PRACH) is scheduled every 5 ms. This gives a theoretical capacity of 10800 preambles per second, which could be sufficient for most MTC scenarios. But, the real capacity is much lower because of the collisions occurring at moderate and high loads, especially in scenarios enabling event-triggered reports (like those considered in this letter).

2.2 | Candidate approach for the 5G

The procedure described in subsection 2.1 was extended in (3) through a simple modification. That is, after the base station performs the detection of preambles, it sends back a RAR containing multiple responses for each identified preamble, with different uplink resources assigned. Every mobile terminal which receives the RAR can randomly choose one of the uplink resources reserved for the preamble, selected during the first step. Figure 2 shows the modified message sequence chart.

Thanks to this additional randomness, a collision over the selection of the preamble does not translate to a failure in the access procedure. This way, if two or more mobile terminals use the same preamble, they still have a chance to select a different resource assignment in the RAR and thus avoid the collision at the third step of the protocol. Basically, the multiple RAR responses act as multipliers for the number of preambles.

The downside of this technique is that a correspondingly larger amount of uplink resources must be reserved for the transmission of Connection Request messages, which shrinks the resources available for actual user data.

2.3 | Other proposed methods

With Coded Random Access (CRA) (8), each packet is transmitted during multiple slots of a frame. The base station first decodes slots with only one packet, then subtracts their duplicate copies from other slots. This way, more slots become decodable, and the process is repeated until all slots are decoded.

Physical Layer Network Coding (PLNC) takes this idea one step further: it tries to decode all the messages even if all the slots experience collisions, by combining them into a system of linear equations(4).

Compressive Sensing Multi-User Detection (CS-MUD) (9) is an application of the compressed sensing framework: it exploits the sparsity of the active users vector to perform combined user detection and data decoding at the physical layer.

In (5) a signature-based method is described, where in the first step each user sends a specific sequence of preambles (i.e. its "signature") rather than just one preamble. Even if different signatures collide on some preambles, this is usually not problematic as the preamble are still detected as active.

3 | PERFORMANCE EVALUATION

The baseline scheme and the proposed candidate for 5Gwere implemented in LTE-Sim (10) and extensively tested under high-load conditions via numerical simulations.

The simulated environment consists of multiple cells, managed by base stations having an omnidirectional antenna. Since the RACH protocol works independently in each cell, we collected the results from a specific cell only: the same results can be generated by also considering large-scale simulation environments.

Conducted tests assume a simple single-input single-output communication scheme, in accordance with the low-cost requirements of MTC devices. The device density varies from 10000 to 1000000 devices/km^2. The inter-site distance is set to 500 m, thus the resulting radius of the cell is equal to 290 m. Users were positioned with a uniform random distribution over the simulation area. The allocated bandwidth is 10 MHz, with a center frequency of 2 GHz.

To reproduce an event-driven transmission burst, the activation time of the devices in the simulations follows a beta distribution with parameters (3,4) over a time interval of 10 seconds (11). The application payload has a small size (5 bytes), so that it only requires a single LTE Resource Block (RB). Also, it should be delivered within 10 seconds from the generation instant, otherwise it is dropped. Each test was repeated for 50 times with a different seed for the random quantities.

According to the most common configuration in LTE networks, RACH opportunities occur every 5 ms and 54 different preambles are available, while the remaining 10 are reserved for contention-free access (2). In case of collision, the procedure fails for all the involved devices, and can be repeated for a maximum of 3 times. For the candidate 5G approach, the number of RARs transmitted for each preamble is set to 2 and 4.

We assumed the same modulation and coding as LTE, so as to isolate the effect of the proposed extension of the RACH scheme. This choice ensures that any gain observed is actually due to the new extended procedure rather than difference in the modulation and coding. Moreover, we expect that 5G will be deployed through incremental steps, and the application services will be quickly deployed while re-using the existing hardware. For this reason, we chose to compare different RACH procedures, implemented on top of the current LTE technology.

The propagation loss is modeled with the urban macro-cell model (12) $L_{db} = 128.1 + 36.7 \log_{10}(d) + S$, where d is the distance in km and S the large-scale shadowing, with 0 mean and 8 dB standard deviation. The power level employed at the mobile terminals is 23 dB. As frequency and time selectivity are not relevant in this scenario, fast fading was not used.

The performance of investigated approaches was evaluated in terms of success probability of the transmission, average delay, and collision rate of the random access process.

First of all, Figure 3 shows the probability that each device completes the transmission of its small data packet within the required delay. It is always equal to 100%, except with the maximum density of 1000000 devices/km^2. In this situation LTE-A performs poorly (6.6%), while the proposed solution shows a clear advantage (21.6% and 31.5% with 2 or 4 RAR responses, respectively). There are two possible causes of failure: (1) a device may not be detected by the base station because of recurring collisions, and (2) transmission can be exceedingly delayed because of the limited amount of physical uplink shared channel (PUSCH) resources. Both of them occur in this work, depending of which technique is adopted, as will be further discussed below.

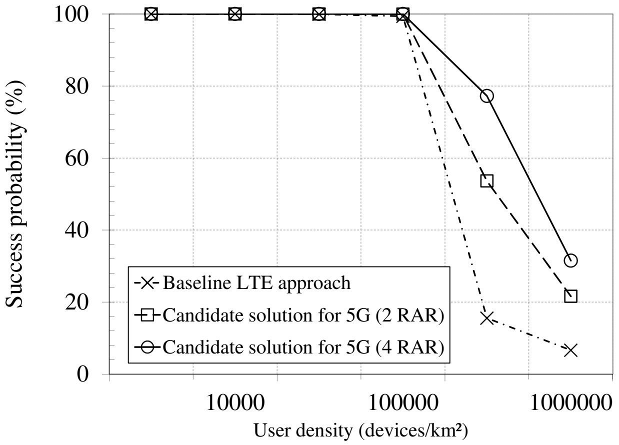

FIGURE 3 Success probability of the complete transmission

Figure 4 shows the average delay of the transmitted packets. This metric refers to application layer data. Therefore, it is calculated by considering both the latency due to the random access procedure and the queuing time at the packet scheduler. Moreover, the delay is only measured when the device completes the random access procedure and it is scheduled for the transmission of the application packet before it expires. For all the investigated techniques, it increases almost linearly with the number of users, and it progressively lower when upgrading from LTE-A to the proposed technique with 2 RARs and finally to 4 RARs. The only exception is with 4 RARs and the highest device density, where the proposed technique grows faster and exceeds the other two cases. This suggests that the PUSCH resources are being saturated for most of the time, due to both the high number of devices handled and the overhead of sending 4 RAR messages.

Finally, Figure 5 shows a temporal trace of the collision rate, measured in the highest-loaded scenario with 1000000 devices/km^2. The most significant result is the very high collision rate of LTE from 3 to 13 seconds, which happens during the peak of the arrivals and for some more seconds. During this period, almost no devices can be detected, and they retry connection until they succeed or fail. Therefore, they produce a longer tail than the other curves. It is thus clear that LTE-A fails to handle 93.4% of the devices due to collisions. By contrast, the candidate solution for 5G with 4 RARs experiences only 63% collision rate at most, meaning that eventually all

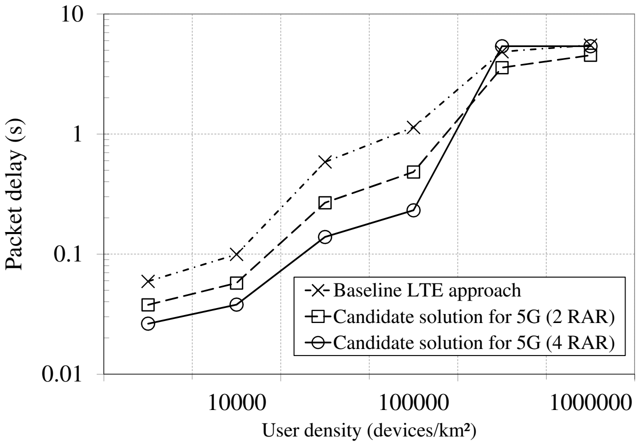

FIGURE 4 Average delay from initial request to successful data transmission

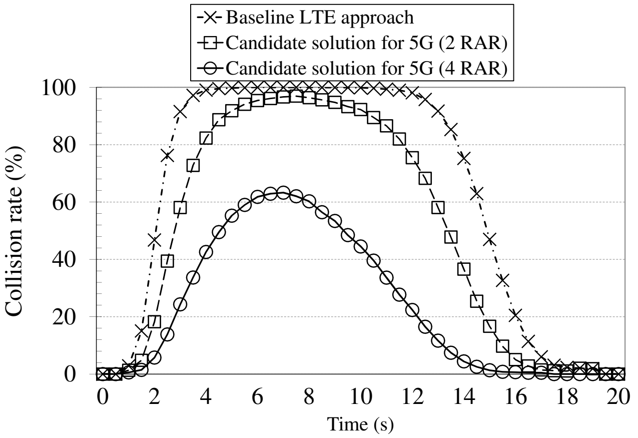

FIGURE 5 Profile of the collision rate occurring during simulations

users should be detected thanks to retries and backoff. In this case, the 78.4% of lost transmission attempts is due to PUSCH saturation. The candidate solution for 5G with 2 RARs lies somewhere in-between, as it experiences a peak of 97% collision rate where part of the users are left out, but recovers quicker than LTE-A. The collision rate for lower user densities is not shown for space reasons, but it is not a concern as it reaches a maximum value of 5%.

4 | CONCLUSION

In this work, two RACH procedures were compared with the goal of supporting massive connectivity in emergency scenarios. The first one is the LTE RACH protocol, and the other one is an improved version suitable for the 5G. The comparison was performed via system-level simulations with a modified version of LTE-Sim. Results show that, with the highest device density, the baseline procedure is strongly limited by the large amount of collisions in the contention phase. On the other hand, the candidate solution for the 5G can support a much larger number of users. Despite the overhead due to the multiple RAR responses, it is still very useful to tolerate an occasional traffic peak from MTC devices. Such traffic peaks can be caused by e.g. monitoring of large-scale events, such as earthquakes, or periodic reporting around a given time of the day. Possible follow-ups of this work include better exploitation of PUSCH resources (e.g. through RBs partitioning or non-orthogonal access) and evaluation of different RACH methods.

ACKNOWLEDGEMENTS

This work is supported by the FANTASTIC-5G project, which receives funding from the European Union’s Horizon 2020 research and innovation programme under the grant agreement ICT-671660. The authors would like to acknowledge Cosimo Damiano Di Pace and Francesco Chianura for their contribution.

References

[1] Ericsson . *White Paper: More than 50 Billion Connected Devices.* : Ericsson; 2011.

[2] Laya Andres, Alonso Luis, Alonso-Zarate Jesus. Is the random access channel of LTE and LTE-A suitable for M2M communications? A survey of alternatives. *IEEE Communications Surveys & Tutorials.* 2014;16(1):4–16.

[3] Kim JS, Munir D, Hasan SF, Chung MY. Enhancement of LTE RACH through extended random access process. *Electronics Letters.* 2014;50(19):1399–1400.

[4] Cocco Giuseppe, Alagha Nader, Ibars Christian, Cioni Stefano. Network-coded diversity protocol for collision recovery in slotted ALOHA networks. *International Journal of Satellite Communications and Networking.* 2014;32(3):225–241.

[5] Pratas Nuno K, Stefanovic Cedomir, Madueno German Corrales, Popovski Petar. Random Access for Machine-Type Communication based on Bloom Filtering. *arXiv preprint arXiv:1511.04930.* 2015;.

[6] Larmo Anna, Lindström Magnus, Meyer Michael, Pelletier Ghyslain, Torsner Johan, Wiemann Henning. The LTE link-layer design. *IEEE Communications magazine.* 2009;47(4):52–59.

[7] 3GPP . *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation.* TS 36.211: 3rd Generation Partnership Project (3GPP); 2008.

[8] Stefanovic Cedomir, Popovski Petar. ALOHA random access that operates as a rateless code. *IEEE Transactions on Communications.* 2013;61(11):4653–4662.

[9] Shim Byonghyo, Song Byungkwen. Multiuser detection via compressive sensing. *IEEE Communications Letters.* 2012;16(7):972–974.

[10] Piro Giuseppe, Grieco Luigi Alfredo, Boggia Gennaro, Capozzi Francesco, Camarda Pietro. Simulating LTE cellular systems: an open-source framework. *IEEE Transactions on Vehicular Technology.* 2011;60(2):498–513.

[11] 3GPP . *RAN Improvements for Machine-type Communications.* TR 37.868: 3rd Generation Partnership Project (3GPP); 2014.

[12] ITU-R . Guidelines for evaluation of radio interface technologies for IMT-Advanced. *Tech. Rep. M.2135.* 2009;.

How cite this article: A Grassi, G. Piro, and G. Boggia (2017), A look at Random Access for Machine-Type Communications in 5G cellular networks, *Internet Technology Letters, Wiley*, *2017;00:1–6*.