This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2017.2697420, IEEE Communications Letters

Success Probability of Multiple-Preamble Based Single-Attempt Random Access to Mobile Networks

Serdar Vural, Ning Wang, *Member, IEEE*, Gerard Foster, Rahim Tafazolli, *Senior Member, IEEE*

***Abstract*—In this letter, we analyse the trade-off between collision probability and code-ambiguity, when devices transmit a sequence of preambles as a *codeword*, instead of a single preamble, to reduce collision probability during random access to a mobile network. We point out that the network may not have sufficient resources to allocate to every possible codeword, and if it does, then this results in low utilisation of allocated uplink resources. We derive the optimal preamble set size that maximises the probability of success in a single attempt, for a given number of devices and uplink resources.**

I. Introduction

In Long Term Evolution (LTE) networks, to get connected to the network, each device first indicates its intention to send a message to the network, so that it can be allocated with a distinct "frequency and time" uplink resource. This is achieved by randomly choosing a *preamble* signal out of a set of preambles (typically 54 in LTE). If multiple devices choose the same preamble, then their messages collide as they use the same uplink resource.

The load on the random access (RA) channel is expected to grow, especially with the advent of emerging technologies, such as the Internet-of-Things (IoT) [1]. When more devices attempt network access within a small time interval, this leads to a higher chance for devices to pick the same preamble, leading to more message collisions. This problem is often referred to as the "Massive Access problem", and has been noted by the Third Generation Partnership Project (3GPP) [2], with thousands of devices attempting to perform random access within seconds, which is equivalent to tens of devices at each RA subframe[1]. Various approaches have been proposed to address the massive access problem [1]. However, these methods cause excessive time delay [3], require additional frequency [4] or time [5] resources, are effective only in low load conditions [6], or ban device access all together [7].

A recent approach is called Code-expanded Random Access (CeRA) [8], which is a "game-changing" way of using preamble signals. The idea is to make consecutive preamble transmissions by each device attempting random access. Such a sequence of preambles is then to be interpreted as a *codeword* by the network, where the number of preamble transmissions in a codeword is referred to as the *codeword length*[2]. LTE uses a single preamble transmission, hence a codeword length of 1, and a preamble space size of typically 54. In contrast, CeRA proposes to have a larger codeword length of $n > 1$, resulting in an codeword space size of 54^n instead of a preamble space size of 54. This provides a significant reduction in message *collision probability*, as it is less likely for multiple devices to pick the same codeword, as compared to the probability to pick the same single preamble.

In this letter, this method is referred to as *Multi-preamble RA*. We note that uplink resources are limited, which affects success probability in a single-shot RA attempt. Hence, some devices may not get a resource when Multi-preamble RA is used, even if they do not experience code collision, since the use of multiple preambles causes ambiguity. The letter provides an analysis of the single-shot RA success probability for Multi-preamble RA, as compared to single-shot LTE.

II. Analysis

In this section, the single-shot random access attempt of Multi-preamble RA is mathematically analysed. m devices perform random access, and each device randomly chooses a sequence of n preambles (hence a codeword of length n), and then transmits each preamble of its sequence at an RA subframe, over n consecutive RA subframes.

A. Collision probability

When consecutive preamble transmissions are made for network access, the network assigns an uplink resource to different preamble sequences, but not to individual preambles. Naturally in this case, the same uplink resource is allocated to multiple devices only when these devices choose the same preamble sequence. For instance, for a preamble sequence {A,B} where A and B represent the two preambles of the sequence, for collision to happen between two devices, they must both choose preamble A in the first preamble transmission time, and then preamble B in the second. Hence, the collision probability in multi-preamble RA experienced by a device D is:

$$P_{coll}(a,m,n) = 1 - \left(1 - \frac{1}{a^n}\right)^{m-1}, \tag{1}$$

where $1/a^n$ is the probability that device D chooses a specific codeword, and $\left(1 - 1/a^n\right)^{m-1}$ is the probability that the rest $m-1$ devices do not choose the same codeword that device D chooses, hence no message collision later on. Note that $P_{coll}(1,m,n) = 1$ when there is a single preamble to choose from, i.e. $a = 1$. Furthermore, $P_{coll}(a,1,n) = 0$ when there is a single device i.e. $m = 1$.

The authors are with the 5G Innovation Centre, Institute for Communication Systems (ICS), Electronic Engineering Department, University of Surrey, Guildford, GU2 7XH, UK. E-mail: {s.vural, g.foster, n.wang, r.tafazolli}@surrey.ac.uk.

Manuscript received Jan 25, 2017; accepted Apr 11, 2017.

[1]In LTE, devices make random access attempt by transmitting a preamble during an RA subframe

[2]Codeword delivery over consecutive RA subframes [8] with x subframe intervals is completed in $(n-1)x + n$ subframes.

1089-7798 (c) 2016 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.

B. Code Ambiguity Problem

When multiple devices transmit a sequence of preambles as codewords, the network does not exactly know which codewords are used by devices. This is because all that the network knows is a set of preambles received in each random access subframe, yet not how to match these received preambles to form the exact codewords chosen by devices. This is called the *code-ambiguity* problem, which is caused by the *false-positives* in the set of deduced codewords at the network side. To quantitatively analyse this problem, the utilisation of uplink resources must first be analysed, as presented next.

C. Utilisation of allocated uplink resources

Let $C(a,m,n)$ denote the number of chosen codewords for m devices choosing codewords of length n using a preamble set of size a. Let $S(a,m,n)$ denote the number of all possible codewords that might have been chosen by devices. In other words, $S(a,m,n)$ is the number of all possible codewords that can be deduced by the network, using the different sets of preambles it receives; e.g. if codeword length is $n=2$, then the network receives 2 consecutive sets of preambles. Note that the number of the deduced (active) codewords is likely to be large as compared to available uplink resources. If the network is able to allocate a resource to each such active codeword, then the utilisation of uplink resources, i.e. goodput, can be defined as follows:

$$U(a,m,n) = C(a,m,n)/S(a,m,n). \quad (2)$$

For instance, if there are $m=3$ devices, choosing codewords of length $n=2$ using a preamble set of size $a=4$, and if these devices choose a total of $C(4,3,2)=3$ different codewords, the network can then deduce a total of $S(4,3,2)=6$ different codewords. Hence, the utilisation in this example is $C(4,3,2)/S(4,3,2)=0.5$. It must be noted that for large preamble set sizes, the codeword space is also large, and the network is likely to be unable to assign a resource to every active codeword. This is later analysed in Section II-F.

To further quantify the uplink resource utilisation term U defined by (2), the numerator and the denominator of this expression must be modelled. This is explained next.

1) The number of chosen codewords: The probability that a specific codeword is chosen by at least one device is:

$$P_{chosen}(a,m,n) = 1 - \left(1 - 1/a^n\right)^m, \quad (3)$$

where $1/a^n$ is the probability for a device to pick the codeword, and $\left(1-1/a^n\right)^m$ denotes the probability that none of the m devices pick that codeword. Note that when there is a single device, i.e. $m=1$, we have $P_{chosen}(a,1,n)=1/a^n$, which is the probability that a device picks a specific codeword.

Multiplying (3) by the codeword space size a^n gives the expected number of codewords $E[C(a,m,n)]$ that are chosen by at least one device, given by:

$$E[C(a,m,n)] = P_{chosen}(a,m,n)a^n = \left[1 - \left(1 - 1/a^n\right)^m\right]a^n, \quad (4)$$

where $E[.]$ is the expected value operator. Note that when there is a single device, i.e. $m=1$, we have $E[C(a,1,n)]=1$, or when $a=1$, we have $E[C(1,m,n)]=1$.

2) Number of all possible codewords: The number of distinct preambles N chosen by devices at a preamble transmission subframe is a random variable. Let $N = N_1, N_2, ..., N_n$ be the corresponding number of distinct preambles chosen by the network at the n consecutive transmission subframes. The network can then deduce the total number of all possible codewords that may have possibly been chosen by devices as $\prod_{i=1}^n N_i$. The expected value of N can be calculated by:

$$E[N] = \overline{N} = a\left[1 - (1-\frac{1}{a})^m\right], \quad (5)$$

where $1-(1-\frac{1}{a})^m$ is the probability that a preamble is chosen by at least one device out of m, and multiplying this by a gives the expected number of distinct preambles chosen by devices.

3) Utilisation: For a codeword length n and an expected number $\overline{N}$ of received preambles per RA subframe (see (5)), the expected number $E[S]$ of all possible codewords of length n that can be deduced by the network can be calculated by:

$$E[S] = E[\prod_{i=1}^{n} N_i] = E[N_1]E[N_2]\dots E[N_n] = \overline{N}^n, \quad (6)$$

where $N_1 \dots N_n$ are identically distributed random variables with mean $\overline{N}$. Then, using (4), (6), and (2), the expected utilisation of allocated uplink resources in the presence of code ambiguity can be calculated by:

$$\begin{aligned} E[U(a,m,n)] &= E\left[\frac{C(a,m,n)}{S(a,m,n)}\right] \approx \frac{E[C(a,m,n)]}{E[S(a,m,n)]} \\ &= \frac{a^n\left[1-\left(1-1/a^n\right)^m\right]}{a^n\left[1-(1-\frac{1}{a})^m\right]^n} = \frac{\left[1-\left(1-1/a^n\right)^m\right]}{\left[1-(1-\frac{1}{a})^m\right]^n}. \quad (7) \end{aligned}$$

Note that when $n=1$ (LTE case), $E[C(a,m,1)] = aP_{chosen}(a,m,1) = a\left[1-(1-1/a)^m\right] = \overline{N}$ and $E[S]=\overline{N}$; hence the expected utilisation is $E[U(a,m,1)]=1$ for $n=1$.

D. Trade-off between collision probability and code ambiguity

Although Multi-preamble RA can significantly reduce collision probability, this has a trade-off with its code ambiguity. In case of a codeword length of $n=1$ (as in LTE) there is no code-ambiguity, and resource utilisation is 1; however, collision probability quickly reaches 1 for larger number of devices m. In contrast, Multi-preamble RA, i.e. codeword lengths of $n>1$, can significantly reduce collision probability. However, this results in an equivalently significant reduction in utilisation of allocated uplink resources.

In Multi-preamble RA, to guarantee that all devices are assigned with a resource, a significantly high number of codewords $E[S]=\overline{N}^n$ must each be allocated with a resource (see (6)). For instance, for a codeword length of $n=2$, $a=54$ preambles, and $m=50$ devices, this requires resources for $N^2 = a^2(1-(1-1/a)^m)^n > 1000$ codewords. As a result, many devices cannot get allocated, as uplink resources are practically limited. If the network can allocate, say 100 resources a time, it would need to randomly choose 100 out of these 1000 codewords; ≈10% chance to get a resource.

In short, making multiple preamble transmissions reduces collision probability, but creates a more significant problem: *resource non-allocation*. The combined effect of collision and resource non-allocation is analysed in the following.

1089-7798 (c) 2016 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.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2017.2697420, IEEE Communications Letters

E. Modelling collision and resource allocation events

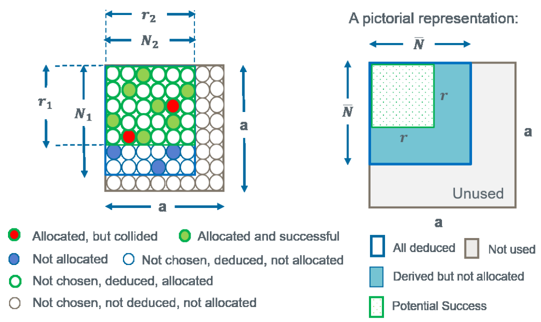


Fig. 1. Collision and non-allocation events. Codeword length is $n = 2$, $N = N_1, N_2$ are the number of the different preambles chosen by devices at the first and second preamble transmission, respectively. $E[N] = \overline{N}$, and $R = r_1 r_2 = r^2$: total number of resources that can be allocated at a time.

A device succeeds in an access attempt, if both of the following conditions are satisfied: (***i***) the device chooses a codeword that no other device chooses (i.e. no collision), and (***ii***) the codeword chosen by a device gets allocated with an uplink resource. In Fig. 1, different event combinations are illustrated for a codeword length of $n = 2$, i.e. 2 dimensions. Success cases are shown to be confined in the inner square area, as shown on the left. The pictorial representation of the event types shown in the left is the generic model on the right, which is to be used for analysis purposes. Here, R denotes the number of resources that the network can allocate to codewords. Note that, $R = r^n$, and the average number of all possible codewords selected by devices is $E[S] = \overline{N}^n$. Since $R = r^n < \overline{N}^n$ in this illustration, the network can only allocate resources to a subset of the set of all possible codewords.

F. Resource allocation probability

If $R < S$, then the network has insufficient resources to allocate to each possible codeword. All that the network can do is to randomly assign its available R resources to a subset of R possible codewords out of the S deduced (see Fig. 1), since the network does not know which codewords are actually chosen by devices. Hence, for $R \leq \overline{N}^n$, the probability P_{alloc} that a codeword in the set of deduced codewords gets allocated with an uplink resource is:

$$P_{alloc}(a,m,n,R) = \frac{R}{E[S]} = \frac{R}{\overline{N}^n} = \frac{R}{a^n\left[1-(1-\frac{1}{a})^m\right]^n}. \quad (8)$$

If $R \geq S$, then all possible codewords get assigned a different resource, which means all devices are allocated, i.e. $P_{alloc} = 1$.

G. Probability of random access success in a single attempt

Since codewords are chosen with equal probability by devices, and given that a device's chosen one is among the set of S codewords deduced by the network, the probability that a specific codeword is chosen by the device is $1/S$. Then, the probability that the codeword is also allocated with a resource is P_{alloc}, as in (8). Since there are S such possible codewords that the device might have chosen, the probability that the device gets allocated with a resource is then $S \times 1/S \times P_{alloc} = P_{alloc}$. Therefore, P_{alloc} in (8) also represents the probability that a device gets allocated with a resource. To succeed in random access, a device's codeword must not be chosen by another device, i.e. no collision, and the device must get allocated with a resource. Hence:

$$P_s(a,m,n,R) = [1 - P_{coll}(a,m,n,R)]\, P_{alloc}(a,m,n,R). \quad (9)$$

Using (1) and (8), the probability of success in a single attempt is then given by:

$$P_s(a,m,n,R) = \begin{cases} f_1 = \left(1-\frac{1}{a^n}\right)^{m-1}, & \text{if } R \geq \overline{N}^n \\ f_2 = \left(1-\frac{1}{a^n}\right)^{m-1}\frac{R}{\overline{N}^n}, & \text{if } R \leq \overline{N}^n \end{cases} . \quad (10)$$

The probability of success expression in (10) is a piecewise probability distribution function with two sub-functions f_1 and f_2, meeting at a point $a = a^*$, such that $\sqrt[n]{R} = \overline{N}^* = a^*(1-(1-1/a^*)^m)$. For $1 \leq a \leq a^*$, resources are sufficient, i.e. $R \geq \overline{N}^n$, hence we have only potential collision events, and no non-allocation events. For $a > a^*$ however, since $R < \overline{N}^n$, non-allocation events may also occur.

H. Maximising success probability

To determine the preamble subset size a that maximises the single-shot RA success probability given by (10), it is necessary to study the properties of functions f_1 and f_2, which meet at $a = a^*$. This meeting point is related with the term $R/\overline{N}^n$, where $R = \overline{N}^n$. Here, the denominator $\overline{N}^n$ has the following properties[3]: $\overline{N}^n_{|a=1} = 1$, $\mathrm{d}\overline{N}^n/\mathrm{d}a > 0$ for $m > 1, a \geq 1$, i.e. this term strictly increases for increasing a. Hence, for a given number of uplink resources R, i.e. constant R value, $R = \overline{N}^n$ can occur at a single value of $a = a^*$. For $a \leq a^*$, i.e. where $P_s(a,m,n,R) = f_1$, we have $\mathrm{d}f_1/\mathrm{d}a > 0$ for $m > 1,\ a > 1$; and for $a \geq a^*$, i.e. where $P_s(a,m,n,R) = f_2$, $\mathrm{d}f_2/\mathrm{d}a < 0$ holds[4]. Hence, P_s reaches its maximum value at $a = a^*$, i.e. when $R = \overline{N}^n$. This leads to the conclusion that probability of success has its maximum value at $a = a^*$, which is the common end-point of functions f_1 and f_2. Note that since (1) a is a positive integer in practice, (2) the solution a^* may not be an integer value, and (3) f_1 is strictly increasing and f_2 is strictly decreasing, then the maximum value that $P_s(a,m,n,R)$ can claim is at either $a^- = \lfloor a^* \rfloor$ or $a^+ = \lceil a^* \rceil$.

III. Simulation Results

In this section, the case of single-shot random access with a single preamble (where codeword length is $n = 1$, as in LTE) and the case of single-shot random access performed by Multi-preamble RA (referred to as Multi(n), with $n = 2, 3, 4$) are compared for their success probability. The analytical expression in (10) has been verified by Monte-Carlo simulations, providing average results of 10000 repetitions. The arrival rate to the system is m devices per RA subframe.

Multi-preamble RA necessitates a convention that all devices must follow. Based on the codeword length n, all RA subframes must be labelled, such as First (F), Second (S), Third (T), and so on, forming an *RA superframe*. An example is shown in Fig. 2 for $n = 2$. This is necessary, because otherwise different devices would use the same RA subframe to transmit a different part of their sequences, e.g. while

[3]These are numerically verified for the parameter ranges $a = 1, 2, \ldots, 54$, $m = 1, 2, \ldots, 1000$, and $n = 1, 2, 3, 4$.

[4]It has been numerically verified that $\mathrm{d}f_2/\mathrm{d}a < 0$ for all $a^* \leq a \leq 54$, $m = 1, \ldots, 1000$, and $n = 1, 2, 3, 4$.

1089-7798 (c) 2016 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.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCOMM.2017.2697420, IEEE Communications Letters

one device transmits its first preamble, another might be transmitting its second. As shown in the figure, this convention results in a load of nm devices per RA subframe for Multi-preamble RA. When $n = 1$ (LTE), for each set of m devices attempting RA access at an RA subframe, the network needs $R = m$ new resources associated to each RA subframe. In contrast in Multi-preamble RA, for each RA superframe, the network needs to allocate resources for this cumulative set of nm devices in bulk, i.e. nm resources.

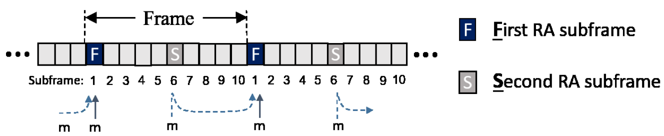
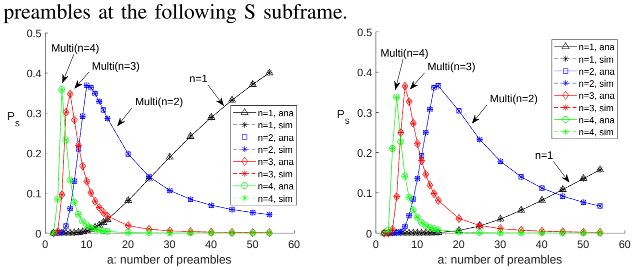


Fig. 2. Multi-preamble RA with codelength $n = 2$: the RA superframe structure. m devices arrive at each RA subframe. Devices arriving at an S subframe must defer to the upcoming F subframe. As a result, a set of $2m$ devices send their first preambles at the F subframe, and then send their second preambles at the following S subframe.



(a) $m = 50$. (b) $m = 100$.

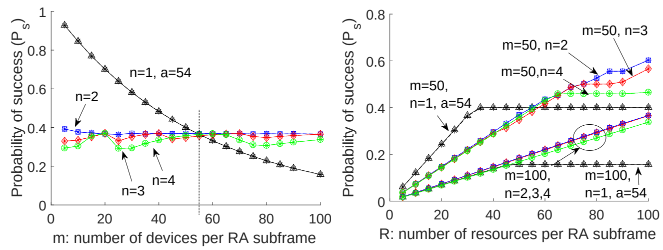
Fig. 3. Success probability (P_s) in a single-shot random access with multi-preamble RA, for varying preamble set size.

A. Parameter effects: a, m and R

Fig. 3 shows that success probability with Multi-preamble RA reaches a maximum level at a certain preamble subset size, which varies with codeword length n. Multi-preamble RA provides higher single-shot RA success probability than the $n = 1$ case when preambles set size is limited, depending on number of devices attempting RA. The figure also shows that analytical (ana) results overlap with simulation (sim) curves, demonstrating their accuracy.

To further evaluate the effect of the access load m, i.e. the number of devices per RA subframe, Fig. 4(a) shows comparison results for device access load conditions of up to $m = 100$ devices per RA subframe. The number of resources R allocated to each RA subframe is $R = m$, so that there is one resource for each device in the $n = 1$ case (LTE); in the Multi-preamble RA case, the nm resources for the n RA subframes of an RA superframe are allocated in bulk, i.e. $R = nm$ resources after each RA superframe. Results show that LTE provides a higher single-shot RA success probability for loads up to a certain value of m.

As observed in this figure, the maximum success probability Multi-preamble RA can achieve stabilises at a certain level, for increasing m. For nm devices over n RA subframes, the network allocates $R = nm$ resources. For a sufficiently large load m, $\overline{N} = a^*(1 - (1 - 1/a^*)^{nm}) \approx a^*$, which leads to $\sqrt[n]{R} \approx a^*$. Hence, we have $\lim_{m\to\infty} P_s(a^*, m, n, m) = \lim_{m\to\infty} (1 - \frac{1}{a^{*n}})^{m-1} = 1/e = 0.3679$, as observed in Fig. 4(a).



(a) Number of devices m per RA subframe. (b) Number of resources R per RA subframe. $m = 50, 100$.

Fig. 4. Single-shot RA success probability. LTE ($n = 1, a = 54$) vs Multi-preamble RA ($n > 1$, a: the subset size providing the highest P_s).

Fig. 4(b) shows results for varying number of resources R per RA subframe, for $m = 50$ and $m = 100$ cases. LTE's single-shot success probability does not change after a certain R, as more resources do not provide further benefit. This point represents $R = \overline{N}$, i.e. the number of active preambles observed, for $a = 54, n = 1$, which occurs for a larger R when m is larger ($m = 100$). In contrast, Multi-preamble RA benefits from increasing number of resources, as this helps the network to allocate more deduced codewords with a resource, which is the combined effect of code-ambiguity and limited resources, as mentioned in Section II-E.

IV. Conclusion

In this letter, we provide mathematical expressions to analyse the single-shot success probability of random access (RA) using a sequence of preamble transmissions in an RA attempt (referred as Multi-preamble RA), instead of a single preamble transmission as performed in LTE. We demonstrate the effects of resource limitation and varying access load, and analytically find the preamble subset size that maximises the single-shot RA success probability, for a given codeword length, number of devices, and available uplink resources. We verify the accuracy of the expressions with Monte-Carlo simulations. It is shown that Multi-preamble RA can achieve better single-shot RA success than LTE when the preamble set size is much smaller than 54.

References

[1] A. Laya *et al.*, "Is the random access channel of LTE and LTE-A suitable for M2M communications? A survey of alternatives," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 4–16, 2014.
[2] 3GPP, "Study on RAN improvements for machine-type communications," Radio Access Network Technical Specification Group, 3rd Generation Partnership Project, Tech. Rep. TR 37868, September 2012.
[3] X. Yang, A. Fapojuwo, and E. Egbogah, "Performance analysis and parameter optimization for random access backoff algorithm in LTE," in *IEEE Vehicular Technology Conference*, 2012, pp. 1–5.
[4] S.-Y. Lien and K.-C. Chen, "Massive access management for QoS guarantees in 3GPP machine-to-machine communications," *IEEE Communications Letters*, vol. 15, no. 3, pp. 311–313, March 2011.
[5] M. J. A. Lo, Y. W. Law and M. Kucharzak, "Enhanced LTE-advanced random access mechanism for massive machine-to-machine (M2M) communications," in *Wireless World Research Forum*, 2011.
[6] L. Alonso *et al.*, "A near-optimum MAC protocol based on the distributed queueing random access protocol (DQRAP) for CDMA mobile communication system," *IEEE JSAC*, vol. 18, no. 9, pp. 1701–1718, September 2000.
[7] T.-M. Lin *et al.*, "PRADA: Prioritized random access with dynamic access barring for MTC in 3GPP LTE-A networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2467–2472, 2014.
[8] N. K. Pratas *et al.*, "Code-expanded random access for machine-type communications," in *IEEE Globecom Workshop on Machine-to-Machine Communications Key to the Future Internet of Things*. IEEE, 2012.

1089-7798 (c) 2016 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.