

Received October 16, 2020, accepted November 3, 2020, date of publication November 6, 2020, date of current version November 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3036398

# Dynamic Backoff Collision Resolution for Massive M2M Random Access in Cellular IoT Networks

HUDA DAKHILALLAH ALTHUMALI<sup>1,2</sup>, (Graduate Student Member, IEEE),

MOHAMED OTHMAN<sup>1,3</sup>, (Senior Member, IEEE),

NOR KAMARIAH NOORDIN<sup>4</sup>, (Senior Member, IEEE),

AND ZURINA MOHD HANAPI<sup>1</sup>, (Member, IEEE)

<sup>1</sup>Department of Communication Technology and Network, Universiti Putra Malaysia (UPM), Serdang 43400, Malaysia

<sup>2</sup>Department of Computer Science, College of Science and Humanities, Imam Abdulrahman Bin Faisal University, Al Jubail 35816, Saudi Arabia

<sup>3</sup>Laboratory of Computational Science and Mathematical Physics, Institute of Mathematical Research (INSPEM), Universiti Putra Malaysia (UPM), Serdang 43400, Malaysia

<sup>4</sup>Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang 43400, Malaysia

Corresponding authors: Huda Dakhilallah Althumali (hdalthemaly@iau.edu.sa) and Mohamed Othman (mothman@upm.edu.my)

This work was supported in part by the Saudi Ministry of Education, in part by the Ministry of Higher Education under the Malaysian International Scholarship Scheme, in part by the Malaysian Ministry of Education through the Research Management Center, Universiti Putra Malaysia, under Grant 9001103, and in part by the Putra Grant with High Impact Factor under Grant UPM/700-2/1/GPB/2017/9557900.

**ABSTRACT** The deployment of machine-to-machine (M2M) communications on cellular networks provides ubiquitous services to Internet-of-Things (IoT) systems. Cellular networks have been chosen as the best infrastructure for M2M communications due to the wide coverage and spectral efficiency. However, with the increased number of devices connecting to the network, massive number of devices are expected to simultaneously access the network resources. This massive access results in excessive congestion and collisions in the random access channel (RACH) which causes major degradation in systems performance. This article focuses on resolving the RACH collisions during the massive access scenarios for cellular M2M communications. We propose a collision resolution scheme using the backoff procedure which dynamically adjusts the backoff indicator ( $BI$ ) based on the number of backlog devices and the available resources. The proposed scheme is integrated with three well-known random access schemes; standard random access (SRA), static access class barring (ACB) and dynamic access class barring (DAB). Furthermore, the paper presents an analysis for access success probability based on the dynamic backoff procedure. The optimal value of  $BI$  that achieves the highest access success probability is derived for the three different schemes. The analysis and simulation results indicate that the dynamic value of  $BI$  achieves approximately 99.9% access success rate with a slight increase in access delay of around 10%, which is considered a reasonable increment for delay-tolerant applications during the massive arrivals scenarios.

**INDEX TERMS** Cellular IoT networks, machine-to-machine, massive random access, backoff procedure, collision resolution.

## I. INTRODUCTION

Machine-to-machine (M2M) communication stands for the automated interaction between machine devices without human intervention. This technology provides significant services for several IoT applications such as e-healthcare, e-transportation, e-commerce and control systems. The Third-Generation Partnership Project (3GPP) has introduced

The associate editor coordinating the review of this manuscript and approving it for publication was Kashif Sharif<sup>1</sup>.

the standards for M2M communications on cellular networks [1]. Cellular networks such as Long-Term Evolution (LTE), LTE-Advanced (LTE-A) and fifth generation (5G) networks have been considered an appropriate infrastructure for M2M communications due to the widespread coverage, scalability, low latency, large capacity, spectral efficiency and Quality-of-Service (QoS) guarantees. However, LTE/LTE-A and 5G networks are optimised for Human-to-human (H2H) communication, which has different characteristics from M2M communications. Usually, M2M communications include short

data transmissions with higher random access requests compared to H2H communications. Therefore, the current cellular networks require some improvements to cope with the special characteristics of M2M communications. The most challenging issue for M2M communications is the massive number of devices that simultaneously try to access the base station or the evolved node B (eNodeB), which results in excessive congestion and collisions in the random access channel (RACH). There are many studies that have attempted to reduce the congestion in the RACH during the massive access scenarios. However, only a few researchers have conducted a study of the backoff procedure for RACH collisions resolution. It is important to optimise the standard backoff procedure of cellular networks to tolerate the massive access of M2M communications. For this reason, this article aims to analyse the performance of the RACH under the usage of the backoff procedure for collision resolution in massive M2M access scenarios. A dynamic backoff collision resolution (DBCR) scheme is proposed to improve the performance of the RACH by modifying the backoff parameters according to the system load. Then, the paper studies the combination of the dynamic backoff with three different random access schemes: standard random access (SRA) [2], access class barring (ACB) [3] and dynamic access class barring (DAB) [4]. Specifically, the main contributions of this study are:

- 1) To investigate the impact of modifying the backoff indicator (BI) on different performance metrics such as access success probability, access delay, collision rate and the number of access retrials.
- 2) To propose a DBCR scheme that optimises the value of BI according to the number of contending devices and the available resources for SRA, ACB and DAB.
- 3) To analyse the access success probability for SRA, ACB and DAB under the proposed DBCR scheme.

PP [2] for SRA, ACB and DAB. It has been observed that most studies evaluating the RACH performance assume a fixed BI, usually equal to 20 ms, whereas the 3GPP specified that this parameter is randomly chosen for each random access opportunity (RAO). Therefore, this research aims to find the best BI value that achieves the highest success probability based on the total number of M2M devices and the available resources. Furthermore, the proposed scheme can be applied for both LTE/LTE-A and 5G networks since they have the same standard random access procedure [5]. The implementation of the proposed scheme in the real world only requires a simple modification for the system settings related to the computations of BI value. It is considered a costless solution because no need for any infrastructure modifications.

The rest of this article is organised as follows. Section II gives a brief background about the random access procedure, access class barring and the backoff procedure. Section III presents related work on random access procedure improvements. The system model is shown in section IV. After that, the proposed DBCR scheme is explained, in section V. The performance evaluation is presented in section VI, while the

evaluation results are shown in section VII. Finally, there is a brief conclusion in section VIII.

## II. BACKGROUND

This section provides a brief demonstration of the main concepts related to this work.

### A. RACH PROCEDURE OVERVIEW

In cellular networks, the standard mode for network access is contention-based random access [2]. The user equipment (UE) must perform the random access procedure as a first step to acquire initial access to the network resources. This procedure is performed through the RACH and consists of four steps:

- 1) Preamble Transmission: in this step, the UE randomly selects one of the available preambles in the current time slot and transmits it to the eNodeB. Normally, there are 54 preambles for random access purposes in each time slot. These preambles are used as an initial identification for the UE. Thus, each preamble must be selected by only one UE to guarantee successful access. If a preamble is selected by more than one UE, a collision occurs and the access is considered failed.
- 2) Random Access Response (RAR): the eNodeB transmits the random access response message to all detected preambles. This message contains a temporary identification for the UE and the uplink resources that must be used by the UE in the next step. The RAR message also includes the value of BI that can be used for the backoff procedure.
- 3) Connection Request: the UE sends a connection request to the eNodeB using the uplink resources assigned in the previous step. This message contains the UE identifier and the access request purpose.
- 4) Contention Resolution: the eNodeB transmits this message as a response to the previous message to inform the UE that it guarantees successful access to the network resources. The UE that does not receive this message has to try another access attempt in an upcoming time slot. There is a maximum number of transmission attempts after which the UE request is dropped.

### B. ACCESS CLASS BARRING (ACB)

The 3GPP has proposed the ACB scheme to control congestion that may occur in the RACH due to the massive number of arrival access requests [6]. In this scheme, the M2M devices are classified into 16 access classes. Any of these classes may be prevented by the network operator from accessing the eNodeB at any time. This process depends on the barring parameters that are broadcast by the eNodeB through the system information block type 2 (SIB2) message. The barring parameters are the access barring rate  $P_{ACB}$  and the barring time  $T_{ACB}$ . The value of  $P_{ACB}$  is uniformly randomly generated by the eNodeB in the range of 0 and 1. The value of  $T_{ACB}$  ranges from 4 to 512 seconds. Under

the ACB access, the UE required to access the network must generate a random number  $q$  where  $0 \leq q \leq 1$ . The UE is allowed to start the random access procedure only if  $q \leq P_{ACB}$ . Otherwise, the UE must wait for  $T_{ACB}$  seconds before it retries the ACB process. This process allows the system to reduce the massive access congestion by redistributing the access requests over time. The main drawback for the ACB is that it is not appropriate for delay-sensitive devices due to the barring time that increases the access delay. The ACB scheme has been studied and improved by many researchers. For example, the extended ACB scheme has been proposed to consider the priority of devices [7]. In this scheme, low priority devices are prevented from accessing the network during massive access scenarios. Another approach is the adaptive ACB where the ACB parameter is dynamically updated according to the network load [8]. The multiple ACB is another scheme where different ACB parameters are assigned to different access classes and updated based on the network traffic [9]. The performance of the ACB has been optimised using timing advance information to determine the optimal ACB parameter that increases the number of successful devices per RAO [10]. This scheme has reduced the access delay by reducing the number of RAOs required to serve all devices up to 50%. Moreover, the DAB scheme has improved the ACB by deriving a dynamic optimal value of  $P_{ACB}$  to achieve the minimum total service time for massive M2M devices [2]. However, this article combines the ACB and the DAB schemes with the proposed dynamic backoff scheme for collision resolution which enhances the overall performance of the ACB and the DAB.

### C. BACKOFF PROCEDURE

In the 3GPP standards, a backoff procedure is proposed in the case of access collision or access failure [2]. In this procedure, the devices that failed in the first access attempt must wait for a predetermined period of time before retrying the access process. This time interval is called the backoff time  $T_{BO}$  and is uniformly randomly chosen by the backlog device based on  $BI$  where

$$T_{BO} = \mathcal{U}(0, BI), \quad (1)$$

where  $\mathcal{U}(\cdot)$  stands for uniform distribution. In fact, the value of  $BI$  has a great influence on the overall performance of the RACH. For example, a small  $BI$  allows devices to retry access after a short period of time, which may increase the probability of collisions during burst arrivals. However, the large value of  $BI$  may increase the access success probability at the expense of increasing the access delay. Most previous research has considered a fixed value for the  $BI$ , equal to 20 ms, which does not follow the standards of the 3GPP which determine that the value of  $BI$  is randomly chosen by the eNB in the range of 0 to 960 ms. Therefore, it is important to find the optimal value of  $BI$  that increases the access success probability with a reasonable access delay. For this reason, this research focuses on the optimisation of the  $BI$  value with respect to the access success probability.

### III. RELATED WORK

The contention-based random access control techniques for M2M communication in cellular networks have been classified into congestion control techniques and collision resolution techniques [11]. Congestion control techniques aim to reduce the massive number of access requests by redistributing the arrival requests over time while the collision resolution techniques resolve the RACH collisions after occurrence. There are many studies that have evaluated and improved the performance of the RACH for M2M traffic. For example, Tello-Oquendo *et al.* [3] analysed the capacity of the LTE RACH according to the SRA procedure defined by the 3GPP. They define the combination of parameters that increase the access success probability for the RACH under massive access. They also studied the performance of ACB and identified the optimal parameters for ACB during most congestion scenarios. For collision resolution, they compared the standard uniform backoff procedure with an exponential backoff that depends on the number of previous access attempts performed by UE. They found that the exponential backoff increased the maximum number of devices that leads to access success  $\geq 0.95$  of approximately 18%. Furthermore, Leyva *et al.* proposed a novel analysis model to accurately evaluate the performance of the random access procedure in LTE networks [12]. They obtained the key performance indicators defined by the 3GPP with minimal error compared to simulation results. They also proposed an analytical model for the static ACB scheme. The numerical results show that this model is very accurate and can be easily integrated with other analytical models. However, the model must be optimised to eliminate the error observed in the CDF of access delay in the first few subframes.

Several schemes have been proposed to control and reduce the RACH congestion caused by massive arrivals of M2M traffic. For instance, Harwahy *et al.* proposed a pre-backoff scheme to avoid RACH collisions in group paging random access [13]. In this scheme, all devices are forced to perform a pre-backoff process before the first preamble transmission. The pre-backoff time is uniformly randomly chosen within a predetermined range. This scheme has improved the access success probability and reduced the collision probability as well. However, it increases the access delay due to the extra delay imposed by the pre-backoff procedure. To reduce the access delay, an opportunistic splitting algorithm has been proposed to avoid the congestion caused by the four-step random access [14]. In this algorithm, uplink resources are assigned to the connected devices in a distributed way. A new frame structure was introduced for this algorithm where the contention slot is followed immediately with a data transmission slot. Hence, the device which wins in the contention slot starts the data transmission until the end of the current time slot. Simulation results demonstrate that this algorithm reduced the average access delay as well as the collision rate with optimal resource utilisation. Another solution to control the RACH congestion is the Q-learning based algorithms in which the reinforcement learning is used

to optimise the RACH resources separation for heterogeneous systems during the massive access scenarios [15], [16]. The simulation results for Q-learning based algorithms show their effectiveness in increasing the access success probability and reducing the total service time. Furthermore, a grant-free random access scheme has been proposed to increase the access success probability and the RACH throughput in massive MIMO systems [17]. This scheme allows the massive M2M devices to transmit their access requests through the same time-frequency resource. Then, the base station uses several classifiers to separate the mixed preambles and input them into an ensemble classifier which jointly detects and decodes the preambles. The simulation results demonstrate that the grant-free scheme improves the RACH throughput even in the massive access scenarios. However, the grant-free schemes suffer from interferences which may cause some performance degradations. This issue has been solved by proposing a hybrid grant random access scheme which aims to improve the spectral efficiency by adding a broadcast message after the preamble transmission message [18]. In details, the base station transmits a broadcast message that can be resolved only by the devices that have successful in preamble transmission. The successful devices transmit their data in the next step while the collided devices remain silent. Therefore, the interferences are removed without costing extra service time. The simulation results show that the Hybrid-grant scheme achieve a significant improvement on the spectral efficiency compared to grant-free scheme.

RACH collision is another issue that reduces the overall performance of M2M systems. To overcome this issue, various solutions have been proposed such as the  $q$ -array splitting algorithm [19]. In this algorithm, the colliding devices are directed to a predetermined set of uplink resources to be used in the next access attempt. This process allows the colliding devices to compete separately within several groups of preambles which reduces the collision rate and in turn improves the access success probability for the massive access scenarios. The same concept of preamble grouping is used in the distributed queueing algorithm [20]. In this algorithm, the devices that are colliding in a given preamble are directed to the same queue for the next access attempt. Therefore, for each colliding group, there is a corresponding queue. The splitting process allows the colliding devices to compete in a parallel manner which helps reduce the collision rate and the access delay. The distributed queueing idea has been integrated with multiple-input multiple-output (MIMO) technology to design an efficient system for M2M communications [21]. This article proposes a new frame structure consisting of four parts, contention window, contention feedback, data transmission and data feedback. After the access contention, the successful devices begin transmitting data, whereas the colliding devices are directed to the distributed queues for collision resolution. The simulation results show that this scheme has improved the throughput of M2M systems under the delay and resource constraints. Using virtual preambles is another proposed solution to reduce the RACH

collisions [22]. This scheme aims to virtually increase the number of available preambles by combining them with the RACH index used to transmit the preamble. This process allows the eNodeB to decode the preamble even if it is used by more than one device. The simulation results show that this scheme is efficient in reducing the collision rate and the access delay for M2M systems. Furthermore, a binary countdown scheme has been proposed to resolve the RACH collisions in 5G networks [23]. In this scheme, additional resources are allocated for the collided devices prior to the third step of the random access procedure. The amount of resources are determined based on the estimation of the backlog devices to achieve the minimum service time with minimum resources.

However, to the best of our knowledge, only a few studies have been conducted on the impact of the backoff procedure on the RACH performance during the massive access of M2M devices. For example, Gürsu *et al.* [24] have introduced a hybrid random access where they combine the congestion control techniques with the collision resolution techniques to increase the RACH throughput for synchronous M2M traffic. Specifically, they use the pre-backoff and the ACB to reduce the massive access requests by redistributing the arrivals with a predefined backoff period. After that, they use tree-based random access to resolve the RACH collisions by directing the colliding devices to predetermined resources to perform the next access attempt. It was found that the hybrid random access increased the RACH throughput and reduced the access delay since the tree-based random access optimises the usage of RACH resources. However, this scheme has a high complexity due to the additional communications overhead that could consume more energy and increase the processing delay. Furthermore, Vidal *et al.* [25] proposed another hybrid random access in which they combined the extended access barring (EAB) with the backoff scheme to enhance the access success probability. In this work, the backoff procedure is implemented by the unbarred devices before the access attempt to avoid collisions with other devices. The combination of EAB with the backoff scheme shows a great improvement in the access success probability and the mean number of preamble transmissions compared to EAB scheme. However, the combination of congestion control schemes such as EAB with the backoff scheme could increase the access delay which should be kept at the lowest possible value. This increment in the access delay results from barring time and the additional backoff time that the devices must wait before the access attempt. For this reason, we aim to address the issue of RACH collisions for synchronous M2M traffic by optimising the standard backoff procedure without any combination with other access control schemes to keep the access delay as low as possible while increasing the access success probability. We seek to reduce the communication overhead and the access delay using a dynamic backoff procedure where the backoff interval is dynamically adjusted based on the number of contending devices and the available resources.

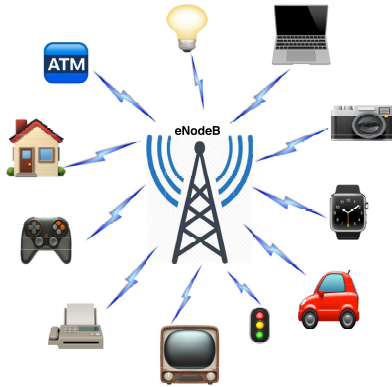


FIGURE 1. System model.

IV. SYSTEM MODEL

This work considers a single eNodeB serving  $N$  machine devices (UE access requests) that have previously registered with the eNodeB as shown in Fig. 1. We assume that each device is provided with a SIM which allows the device to connect directly to the base station without any gateway. The frame structure type 1 is considered for the RACH configuration where the subframe length is 1 ms. A maximum of one RAO is available in each subframe. The periodicity of RAO is determined by the RACH configuration index and it differs from 1 RAO every one subframe to 1 RAO every ten subframes (see Fig. 2). For the arrivals model, the burst M2M arrival model is considered. In this model, a massive number of UE requests arrive within a short activation time  $T_A$  due to occurrence of a specific event. Following the 3GPP standards and previous research, the beta distribution is used to model the burst arrivals of M2M traffic as shown in Fig. 3 [4] [24], [26] with the probability density function  $f(t)$  as follows:

$$f(t) = \frac{t^{\alpha-1}(T_A - t)^{\beta-1}}{T_A^{\alpha+\beta-1}B(\alpha, \beta)}, \quad 0 \leq t \leq T_A. \quad (2)$$

where  $B(\alpha, \beta)$  denotes the beta function  $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$ . According to [26],  $\alpha = 3$  and  $\beta = 4$ .

This work also considers the scenario of system start-up when all  $N$  devices try to access the RACH simultaneously in a single RAO. The Delta model is used to represent this scenario where  $T_A = 1$  RAO [24]. All the model notations are summarised in Table 1.

V. DBCR SCHEME

A DBCR scheme is proposed to optimise the success probability for the massive arrivals in the RACH. In this scheme, the value of  $BI$  is dynamic on the number of contending devices  $N$  and the available number of RACH resources (i.e., preambles  $M$ ).  $BI$  always indicates the maximum value of backoff time that the device must wait before it starts another access attempt. Most of the current settings of RACH consider the value of  $BI = 20$  ms. However, this value is not appropriate for the massive number of M2M devices since

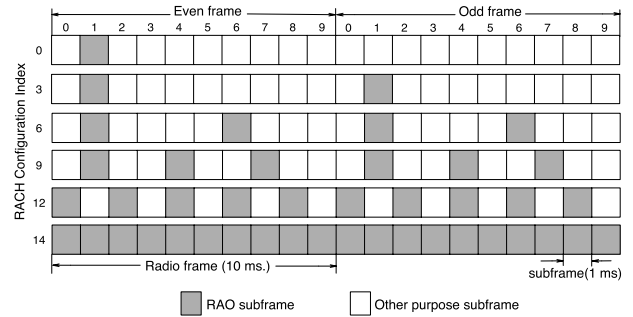


FIGURE 2. Frame structure type 1 with different RACH configurations [27].

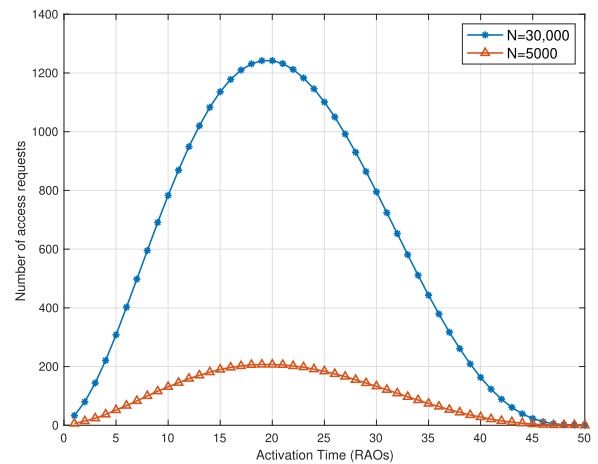


FIGURE 3. Arrivals activation model (Beta distribution).

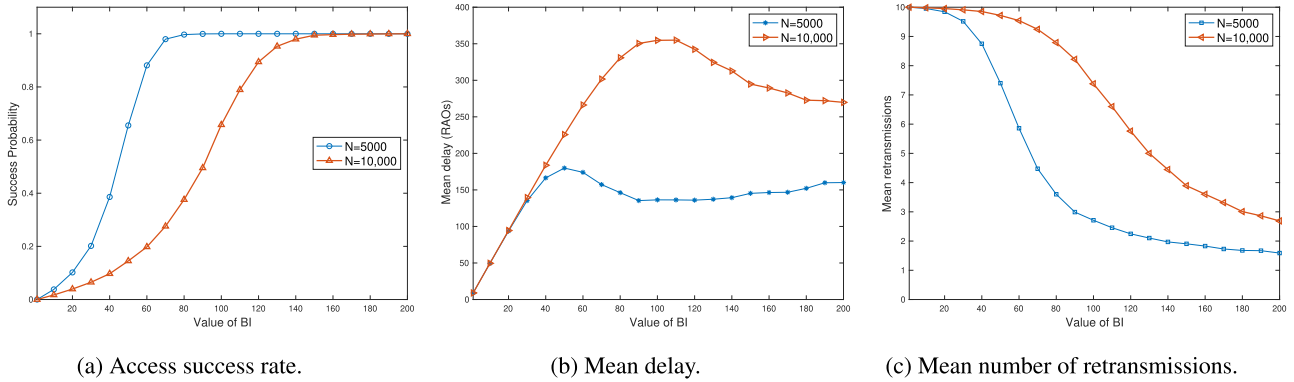
TABLE 1. Summary of model notations.

Symbol	Description
$N$	Total number of UE requests (devices)
$M$	Total number of preambles
$T_A$	Activation time period
$T_{RA}$	Periodicity of RAOs
$BI$	Backoff indicator or backoff interval length
$T_{BO}$	Backoff time
$k_{max}$	Maximum number of preamble retransmissions
$k$	number of preamble retransmissions
$p$	Access barring rate
$P_1$	Success probability for event 1
$P_2$	Total success probability for event 2
$P_k$	Success probability of the $k^{th}$ transmission
$P_s$	Total access success probability for event 1+ event 2
$N_{s_k}$	Number of successful devices for the $k^{th}$ transmission
$N_{c_k}$	Number of collided devices for the $k^{th}$ transmission
$N_s$	Total number of successful devices in event 2

it causes excessive collisions in the RACH. This is due to the burst arrivals that share the same time interval with the backoff time  $T_{BO}$ . Thus, the backoff devices compete with the new burst arrivals which results in new congestion and collisions.

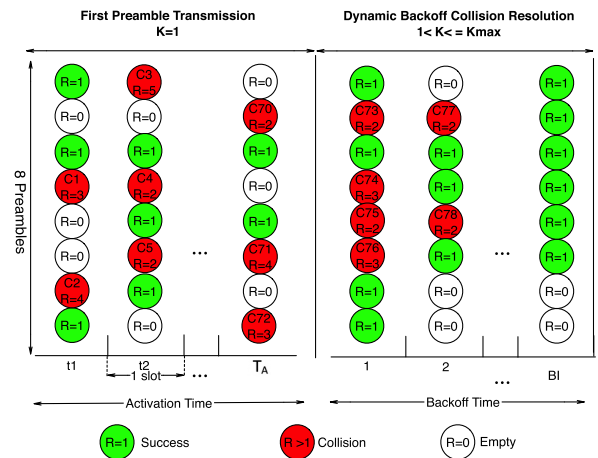
A. IMPACT OF MODIFYING THE VALUE OF BI

To clarify the impact of modifying the value of  $BI$  on the RACH performance, let us examine the access success



**FIGURE 4.** Impact of different values of  $BI$  on different performance indices of SRA scheme for different number of devices  $N$ , with  $M = 54$  preambles and  $k_{max} = 10$ .

probability for SRA with different values of  $BI$  within its standard range from 0 to 960 ms. The researchers in [3], [24], [28] analyse the SRA and present the simulation results considering a fixed value for  $BI$  (equal to 20 ms). However, we re-simulated the SRA for beta traffic arrivals with different values of  $BI$  according to the standard range determined by the 3GPP to find the best value of  $BI$  that could increase the access success probability to the highest possible value. Fig. 4 shows the relation between the value of  $BI$  and different performance indices for SRA scheme. It is clearly seen from Fig. 4a that the highest success probability for each number of devices is achieved with a different value of  $BI$ . The increased number of devices  $N$  requires a higher value of  $BI$  to increase the backoff time and in turn increase the number of RAO for the backoff interval. The higher number of RAO increases the access success probability for the backlog devices in their next access attempt. Moreover, the higher  $BI$  value allowed the collided devices to be distributed among longer backoff intervals which helps reduce the congestion in the RACH. However, it is observed from Fig. 4a that for each number of devices, there is an optimal value of  $BI$  that achieves the highest success probability after which the success probability becomes stable. For example, when  $N = 5000$ , the highest success probability is achieved when  $BI \approx 90$  ms. Whereas this value (90 ms) is not the best when  $N = 10,000$ . Furthermore, the value of  $BI$  has a great impact on the access delay because  $BI$  represents the maximum time that a UE must wait before trying another access attempt. Therefore, as the value of  $BI$  increases, the access delay could be increased as well. However, it is seen from Fig. 4b that the access delay keeps increasing with the increased number of  $BI$  until some point after which the delay starts to drop. The access delay here refers to the mean number of RAO required for all devices until they successfully transmit unique preambles. It is noticeable that the value of  $BI$  achieves the optimal success rate ( $BI \approx 90$  for  $N = 5000$ ), is the same value that achieves the lowest access delay for the same number of devices  $N$  whereas the lower value of  $BI$  leads to lower access success rate and less delay. This is because the lower value of  $BI$  results in many access attempts within short time which results in many collisions and therefore the access requests



**FIGURE 5.** Example of dynamic backoff collision resolution RA.

are dropped in short time. When the value of  $BI$  is increased, it allows the devices to try another access attempt after longer time (when the massive new arrivals are reduced) which means higher success probability with a slight increment in the access delay. Moreover, the value of  $BI$  influences the mean number of retransmissions as shown in Fig. 4c. It is clearly observed that as the value of  $BI$  increased, the mean number of retransmissions decreased because the probability of collision reduced with a longer backoff interval. Therefore, it is possible to enhance the performance of the RACH by correctly adjusting the value of  $BI$  based on the number of devices  $N$  and the number of available preambles. It is observed that adopting a small value of  $BI = 20$  ms is not efficient for the massive arrivals of M2M traffic. For this reason the value of  $BI$  must be optimised to achieve the highest access success rate with the lowest possible delay and reasonable number of mean retransmissions even for the massive number of arrivals.

**B. THE PROPOSED DBCR SCHEME**

To overcome the limitations of the static backoff scheme, a DBCR scheme is proposed. In DBCR, the value of  $BI$  is

updated according to the number of contending devices and the available number of preambles. To illustrate the process of dynamic backoff collision resolution, there is an example in Fig. 5. The diagram contains two main parts. The first one is the activation time where the burst access requests arrive and transmit their first preamble based on SRA, ACB or DAB. Note that the activation time for delta arrivals is only a single RAO which represents the system start up. In this stage, some devices successfully access the network in the first preamble transmission attempt (the green preambles where the number of requests  $R$  is equal to 1). However, most requests fail in the first access attempt due to the massive number of requests which results in several collisions such as  $C1$  which includes 3 collided requests, (i.e.  $R = 3$ ). These devices must go through the dynamic backoff collision resolution which is the second part of the diagram. All collided devices must wait for a backoff time  $T_{BO}$  which is generated randomly by the devices in the range of 0 to  $BI$ , where  $BI$  is dynamically determined by the base station according to the number of requests and the available preambles. The device retries the second preamble transmission immediately after the backoff time. In case it fails in the second access attempt, the same process is repeated until all requests are successfully served. By the end of the backoff interval, all devices are supposed to access the network successfully with a maximum of  $BI$  backoff time. Note that the activation time of late arrivals may overlap with the backoff time of early arrivals for beta activation model. However, they are separated in the diagram only to clarify the process of the DBCR. The key point in DBCR is that the value of  $BI$  is not static. It is updated based on the number of collided devices  $N_c$  that have failed in the previous transmission attempt, and the number of available preambles  $M$  with awareness of maximising the access success probability of the RACH. We seek to derive the optimal value of  $BI$  that leads the access success probability to achieve the optimal value even for a massive number of devices. For more understanding, Algorithm 1 and 2 explains the details of the DBCR scheme for SRA and ACB/DAB respectively. The following subsection shows the access success probability analysis and the derivations of the optimal value of  $BI$  for SRA, ACB and DAB.

### C. ACCESS SUCCESS PROBABILITY ANALYSIS AND OPTIMAL BI DERIVATIONS

This section analyses the access success probability for SRA, ACB and DAB considering the length of backoff interval  $BI$ . After that, the optimal value of  $BI$  that leads the access success probability to be equal to 1 is derived. We consider the delta model arrivals in the analysis for more simplicity since there is no overlap between the new arrivals and the backoff requests. However, this model represents the scenario of the most congestion in which all the  $N$  requests are activated in a single RAO. This scenario is considered realistic in M2M systems because it represents the system start up where all devices are contending to access the base station in the same time instant. Generally, the access success probability

#### Algorithm 1 DBCR Scheme for SRA

**Input:**  $N$  is the total number of devices;  $M$  is the number of preambles;  $k_{max}$  is the maximum number of retransmissions;  $T_{RA}$  is the periodicity of RAO.

**Initialisation:**

```

1:  $BI_0 \leftarrow 0$ ;
2:  $N_{s_0} \leftarrow 0$ ;
3:  $N_{c_0} \leftarrow 0$ ;
4: for  $k = 1$  to  $k_{max}$  do
5:   if ( $k = 1$ ) then
6:     Compute:  $P_{(k)}$  based on Equation (6);
7:     Update:  $N_{s_{(k)}} \leftarrow M * P_{(k)}$ ;
8:     Update:  $N_{c_{(k)}} \leftarrow N - N_{s_{(k)}}$ ;
9:   else
10:    Update:  $BI_{(k)} \leftarrow \left\lceil \frac{N_{c_{(k-1)}}}{0.31813M} \right\rceil \times T_{RA}$ ;
11:    Compute:  $P_{(k)}$  based on Equation (11);
12:    Update:  $N_{s_{(k)}} \leftarrow M * BI * P_{(k)}$ ;
13:    Update:  $N_{c_{(k)}} \leftarrow N_{c_{(k-1)}} - N_{s_{(k)}}$ ;
14:   end if
15:   if ( $N_{c_{(k)}} = 0$ ) then
16:     break;
17:   end if
18: end for

```

of DBCR random access depends on the occurrence of two disjoint events that can be defined as follows:

- Event 1: The UE access request is successful in the first access attempt during the activation time  $T_A$ , which considered to be done based on SRA, ACB or DAB, with a probability  $P_1$ .
- Event 2: The UE request is successful in the  $k^{th}$  access attempt within a backoff interval  $BI$  (more than one preamble transmission [i.e.  $2 \leq k \leq k_{max}$ ] with a probability  $P_2$ ).

The total success probability  $P_s$  can be defined as the total number of successful devices for both events which can be expressed as the union of the probability of the two events, therefore:

$$P_s(P_1 \cup P_2) = P_1 + P_2. \quad (3)$$

Now, we can analyse the success probability for each scheme as follows:

#### 1) SRA WITH DBCR

- Event 1: Consider  $N$  UE requests arriving in a single RAO based on the Delta model. According to SRA scheme, each UE randomly selects one of the available preambles  $M$  with equal probability  $1/M$ , the probability that a given preamble contains  $n$  requests, where  $n = 0, 1, 2, \dots, N$ , can be given as follows:

$$P_1(R = n) = \binom{N}{n} \left(\frac{1}{M}\right)^n \left(1 - \frac{1}{M}\right)^{(N-n)}. \quad (4)$$

**Algorithm 2** DBCR Scheme for ACB/DAB

**Input:**  $N$  is the total number of devices;  $M$  is the number of preambles;  $k_{max}$  is the maximum number of retransmissions;  $T_{RA}$  is the periodicity of RAO;  $p$  is the barring rate.

**Initialisation:**

```

1:  $BI_0 \leftarrow 0$ ;
2:  $N_{s_0} \leftarrow 0$ ;
3:  $N_{c_0} \leftarrow 0$ ;
4: for  $k = 1$  to  $k_{max}$  do
5:   if ( $k = 1$ ) then
6:     Compute:  $P_{(k)}$  based on Equation (23);
7:     Update:  $N_{s(k)} \leftarrow M * P_{(k)}$ ;
8:     Update:  $N_{c(k)} \leftarrow (N * p) - N_{s(k)}$ ;
9:   else
10:    if ( $p=M/N$ ) then
11:      Update:  $BI_{(k)} \leftarrow \left\lceil \frac{1}{0.31813} \right\rceil \times T_{RA}$ ;
12:    else
13:      Update:  $BI_{(k)} \leftarrow \left\lceil \frac{p*N_{c(k-1)}}{0.31813M} \right\rceil \times T_{RA}$ ;
14:    end if
15:    Compute:  $P_{(k)}$  based on Equation (29);
16:    Update:  $N_{s(k)} \leftarrow M * BI * P_{(k)}$ ;
17:    Update:  $N_{c(k)} \leftarrow (N_{c(k-1)} * p) - N_{s(k)}$ ;
18:  end if
19:  if ( $N_{c(k)} = 0$ ) then
20:    break;
21:  end if
22: end for

```

The transmission is considered successful if the number of requests  $R$  in a given preamble  $M$  is equal to 1, therefore, the access success probability  $P_1$  of SRA for a given preamble in a given RAO can be written as:

$$P_1(R = 1) = \binom{N}{1} \left(\frac{1}{M}\right) \left(1 - \frac{1}{M}\right)^{(N-1)}, \quad (5)$$

which can be approximated as:

$$P_1 \approx \left(\frac{N}{M}\right) e^{(-N/M)}. \quad (6)$$

The number of successful preambles/devices  $N_{s_1}$  that have succeeded in the first access attempt can be calculated as:

$$\begin{aligned} N_{s_1} &= M \times \left(\frac{N}{M}\right) e^{(-N/M)}. \\ &= N \times e^{(-N/M)}. \end{aligned} \quad (7)$$

Since the total number of devices  $N$  is known, the number of collided (backlog) requests  $N_{c_1}$  that have failed in the first access attempt can be written as:

$$N_{c_1} = N - N_{s_1}. \quad (8)$$

Note that  $N_{c_1}$  represents the number of backlog devices that have failed on their first transmission and will be

used in the following calculations to derive the optimal value of  $BI_{(k)}$ .

- Event 2: Each UE that failed in the first access attempt uniformly randomly selects one of the available preambles  $M$  in the backoff interval  $BI_{(k)}$  with equal probability  $1/(M \times BI_{(k)})$ . Note that  $BI_{(k)}$  is the length of the backoff interval for the  $k^{th}$  transmission and is equal to the value of the backoff indicator  $BI$  which represents the maximum value of backoff time. Considering the worst case, we assume that all the  $N$  requests have failed in the first access attempt (i.e.,  $N_{c_1} = N$ ). Therefore, distributing  $N$  requests over  $(M \times BI_{(k)})$  preambles covered by the backoff interval  $BI_{(k)}$  results in the following probability mass function (pmf):

$$P_2 = \binom{N_{c_1}}{n} \left(\frac{1}{M \times BI}\right)^n \left(1 - \frac{1}{M \times BI}\right)^{(N_{c_1}-n)}. \quad (9)$$

The probability that a given preamble contains only a single request (successful transmission) in the backoff interval  $BI_{(k)}$  of SRA for the  $k^{th}$  transmission can be given as follows:

$$P_k = \binom{N_{c(k-1)}}{M \times BI_{(k)}} \left(1 - \frac{1}{M \times BI_{(k)}}\right)^{(N_{c(k-1)}-1)}, \quad (10)$$

where  $k = 2, 3, \dots, k_{max}$ .

$P_k$  can be approximated as:

$$P_k \approx \binom{N_{c(k-1)}}{M \times BI_{(k)}} e^{(-N_{c(k-1)}/M \times BI_{(k)})}. \quad (11)$$

The total success probability  $P_2$  for event 2 can be written as:

$$P_2 = \sum_{k=2}^{k_{max}} P_k. \quad (12)$$

The number of successful devices of the  $k^{th}$  transmission  $N_{s_k}$  which is equal to the number of preambles that have been chosen by only one device can be computed as:

$$N_{s(k)} = M \times BI_{(k)} \binom{N_{c(k-1)}}{M \times BI_{(k)}} e^{(-N_{c(k-1)}/M \times BI_{(k)})}. \quad (13)$$

The number of collided devices that have failed in the  $k^{th}$  transmission  $N_{c(k)}$  can be computed as:

$$N_{c(k)} = N_{c(k-1)} - N_{s(k)}. \quad (14)$$

The total number of successful devices  $N_s$  that have succeeded in the backoff interval  $BI$  can be written as:

$$N_s = \sum_{k=2}^{k_{max}} N_{s(k)}. \quad (15)$$

We stated earlier that the total success probability  $P_s$  of the RACH is the summation of the probabilities  $P_1$  and



$P_2$ . To achieve the optimal access success probability,  $P_1 + P_2$  must be equal to 1 (i.e.,  $P_1 + P_2 = 1$ ). Therefore,  $P_2$  can be written as:

$$P_2 = 1 - P_1, \quad (16)$$

since we consider the massive access scenario where  $N \rightarrow \infty$ , which makes  $P_1 = 0$ . Therefore, the optimal success probability can be achieved if  $P_2 = 1$ . Since the maximum total number of devices  $N$  and the number of preambles  $M$  are known, it is possible to derive the value of  $BI_{(k)}$  for the  $k^{th}$  transmission that makes  $P_k$  equal to 1 as follows:

$$\left(\frac{N_{c_{(k-1)}}}{M \times BI_{(k)}}\right) e^{(-N_{c_{(k-1)}}/M \times BI_{(k)})} = 1. \quad (17)$$

The optimal value of  $BI_{(k)}$  for the  $k^{th}$  transmission can be written as:

$$BI_{(k)} = -\frac{N_{c_{(k-1)}}}{M \times W(-1)}$$

$$BI_{(k)} \times W(-1) = -\frac{N_{c_{(k-1)}}}{M}, \quad (18)$$

where  $W$  is the product logarithm function. Since  $W(-1) \approx -0.31813 + 1.33723i$ , Eq. (18) can be written as:

$$BI_{(k)} \times (-0.31813 + 1.33723i) = -\frac{N_{c_{(k-1)}}}{M},$$

$$-0.31813BI_{(k)} + i(1.33723BI_{(k)}) = -\frac{N_{c_{(k-1)}}}{M},$$

$$BI_{(k)} = \frac{N_{c_{(k-1)}}}{0.31813M} - i(1.33723BI_{(k)}). \quad (19)$$

Since the value of  $BI_{(k)}$  must be a positive integer number, we discard the imaginary number  $i$ . Therefore, the value of  $BI_{(k)}$  for the  $k^{th}$  transmission can be approximated as:

$$BI_{(k)} \approx \left\lceil \frac{N_{c_{(k-1)}}}{0.31813M} \right\rceil, \quad (20)$$

where  $\lceil \cdot \rceil$  is the ceiling function to ensure that  $BI$  is an integer. The value of  $BI$  in Eq. (20) is typical when the RAO is available in each subframe (i.e., *RACH Configuration Index* = 14). However, considering the other configurations index when the RAO is not repeated in each subframe, the value of  $BI$  can be written as:

$$BI_{(k)} \approx \left\lceil \frac{N_{c_{(k-1)}}}{0.31813M} \right\rceil \times T_{RA}, \quad (21)$$

which indicates the value of  $BI_{(k)}$  for the  $k^{th}$  duration is updated based on the number of collided devices  $N_{c_{(k-1)}}$  that have collided in the  $k^{th-1}$  transmission. The new value of  $BI_{(k)}$  is computed and updated at the end of the previous duration of  $BI_{(k-1)}$  when all devices have finished their  $k^{th-1}$  transmission. This means that Algorithm 1 and 2 are executed with a maximum of  $k_{max}$  times. The updated value of  $BI_{(k)}$  is broadcasted by the base station to all devices through the next system information block (SIB) message.

## 2) ACB AND DAB WITH DBCR

With the same flow, it is possible to derive the access success probability for the ACB and DAB under the DBCR as follows:

- Event 1: Consider  $N$  access requests/UEs arriving in a single RAO according to the Delta model. After a UE has passed the ACB check with a probability of  $p$  (barring rate or  $P_{ACB}$ ), it selects one of the available preambles  $M$  in a given time slot  $t_i$  with equal probability  $1/M$ . The conditional probability that a given preamble contains  $n$  requests can be written as:

$$P_1(R = n) = \binom{N}{n} \left(\frac{p}{M}\right)^n \left(1 - \frac{p}{M}\right)^{(N-n)}, \quad (22)$$

where  $n = 0, 1, 2, \dots, N$ . A transmission is considered successful if  $n = 1$ , therefore, the success probability for ACB/DAB can be given as follows:

$$P_1(R = 1) = \binom{N}{1} \left(\frac{p}{M}\right) \left(1 - \frac{p}{M}\right)^{(N-1)},$$

$$= \left(\frac{N \times p}{M}\right) \left(1 - \frac{p}{M}\right)^{(N-1)}. \quad (23)$$

Note that the value of  $p$  in ACB scheme is a random number between 0 and 1. However, according to Duan et al. [4], the optimal value of  $p$  for DAB scheme is:

$$p = \frac{M}{N}. \quad (24)$$

Inserting  $p$  to Eq. (23) gives:

$$P_{1DAB} = \left(1 - \frac{1}{N}\right)^{(N-1)}. \quad (25)$$

The number of successful preambles/devices  $N_{s_1}$  for ACB/DAB can be written as:

$$N_{s_1} = M \times P_1,$$

$$= M \times \left(\frac{N \times p}{M}\right) \left(1 - \frac{p}{M}\right)^{(N-1)},$$

$$= (N \times p) \left(1 - \frac{p}{M}\right)^{(N-1)}. \quad (26)$$

The number of collided devices  $N_{c_1}$  that have failed in the first transmission for ACB/DAB can be calculated as:

$$N_{c_1} = (N \times p) - N_{s_1}. \quad (27)$$

- Event 2: Each UE that failed in the first access attempt selects one of the available preambles  $M$  in the back-off interval  $BI_{(k)}$  with equal probability  $p/(M \times BI_{(k)})$ . Considering  $N_{c_1}$  devices failed in the first transmission, the conditional probability that a given preamble in the backoff interval contains  $n$  requests can be written as:

$$P_2 = \binom{N_{c_1}}{n} \left(\frac{p}{M \times BI_{(k)}}\right)^n \left(1 - \frac{p}{M \times BI_{(k)}}\right)^{(N_{c_1}-n)}. \quad (28)$$

The success probability of ACB/DAB in the backoff interval  $BI_{(k)}$  for the  $k^{th}$  transmission can be written as:

$$\begin{aligned}
 P_k &= \left( \frac{N_{c(k-1)} \times p}{M \times BI_{(k)}} \right) \left( 1 - \frac{p}{M \times BI_{(k)}} \right)^{(N_{c(k-1)}-1)} \\
 &= \left( \frac{N_{c(k-1)} \times p}{M \times BI_{(k)}} \right) \left( 1 - \frac{p}{M \times BI_{(k)}} \right)^{(N_{c(k-1)}-1)}.
 \end{aligned} \tag{29}$$

The number of successful preambles/devices  $N_{s_k}$  for the  $k^{th}$  transmission can be written as:

$$N_{s(k)} = M \times BI_{(k)} \times P_k. \tag{30}$$

The number of collided devices  $N_{c_k}$  for the  $k^{th}$  transmission can be calculated as:

$$N_{c(k)} = (N_{c(k-1)} \times p) - N_{s(k)}. \tag{31}$$

Considering that  $N$ ,  $M$  and  $p$  are known for the ACB scheme, the value of  $BI_{(k)}$  that makes  $P_k$  equal to 1 can be derived as:

$$\begin{aligned}
 BI_{(k)} &= \frac{p \times N_{c(k-1)}}{M \times W(-1)} \\
 &= \frac{p \times N_{c(k-1)}}{0.31813M} - i(1.33723BI_{(k)}).
 \end{aligned} \tag{32}$$

Discarding the imaginary number and considering the periodicity of RAO, the value of  $BI_{(k)}$  for ACB/DAB can be approximated as:

$$BI_{(k)} \approx \left\lceil \frac{p \times N_{c(k-1)}}{0.31813M} \right\rceil \times T_{RA}, \tag{33}$$

where  $0 < p < 1$  for ACB scheme. However, considering the optimal value of  $p$  for DAB as per Eq.(24),  $BI_{(k)}$  of DAB can written as:

$$BI_{(k)} \approx \left\lceil \frac{\frac{M}{N_{c(k-1)}} \times N_{c(k-1)}}{0.31813M} \right\rceil \times T_{RA}, \tag{34}$$

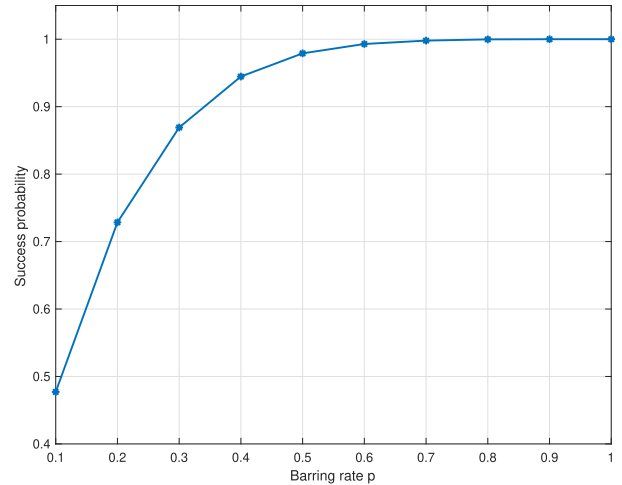
which can be simplified as:

$$BI_{(k)} \approx \left\lceil \frac{1}{0.31813} \right\rceil \times T_{RA}. \tag{35}$$

To summarise, three different values of  $BI$  have been derived for three different random access schemes. The optimal value of  $BI$  differs according to the scheme that is adopted for the random access. Therefore, the derived values of  $BI_{(k)}$  for the three schemes SRA, ACB and DAB can be written as:

$$BI_{(k)} \approx \begin{cases} \left\lceil \frac{N_{c(k-1)}}{0.31813M} \right\rceil \times T_{RA} & \text{for SRA} \\ \left\lceil \frac{p \times N_{c(k-1)}}{0.31813M} \right\rceil \times T_{RA} & \text{for ACB, } 0 < p < 1 \\ \left\lceil \frac{1}{0.31813} \right\rceil \times T_{RA} & \text{for DAB, } p = \frac{M}{N_{c(k-1)}}. \end{cases} \tag{36}$$

The numerical results of the DBCR analysis for the success rate are shown in Fig. 4a for the Delta traffic model.



**FIGURE 6.** Relation between barring rate  $p$  and the access success probability  $P_k$  for ACB when  $BI$  is computed based on Equation (33),  $N = 30,000$ ,  $M = 54$  and  $k_{max} = 10$ .

It is observed from Fig. 4a that the derived values of  $BI$  have achieved 100% access success rate for SRA and DAB. However, the ACB achieved 97% access success rate using the derived value of  $BI$ . In fact, this result is affected by the settings of the barring rate  $p$  which is set to 0.5 in this study. Nevertheless, when  $p = 0.9$ , the derived value of  $BI$  allows the access success rate of ACB to reach 100%. The relation between the barring rate  $p$  and the success probability is shown in Fig 6. It is clearly seen that the success probability naturally increased with the barring rate which indicates that the derived value of  $BI$  is efficient for the massive access. In detail, when  $p = 0.9$  almost 90% of the activated devices will pass the ACB check and proceed to the first preamble transmission after which the DBCR is implemented for the collided devices according to the derived value of  $BI$  which achieve 99.9% access success rate.

## VI. PERFORMANCE EVALUATION

This section will evaluate the performance of SRA, ACB and DAB under dynamic backoff collision resolution through extensive simulations, and will compare their performance according to three benchmarks that use static backoff for collision resolution.

### A. SIMULATION SETUP

The evaluations were conducted through medium access control (MAC) layer simulations for the standard LTE-A RACH system with a MATLAB discrete event simulator. The total number of devices  $N$  from 5000 to 30,000 and  $M$  is equal to 54 preambles. A total of 10 retransmission attempts are considered for preambles transmission. Table 2 shows the detailed simulation settings which follow the standard settings suggested by the 3GPP [1]. For each combination of traffic model and number of users  $N$ , 100 independent simulations were performed which resulted in a 95% confidence interval.

**TABLE 2. Simulation settings.**

Parameter	Setting
Total number of UE devices $N$	30,000
Total number of preambles $M$	54
Activation time period $T_A$	50 ms for beta and 1 ms for delta
RACH Configuration Index	14
Subframe length	1 ms
Periodicity of RAOs $T_{RA}$	1 ms
Maximum number of preamble retransmissions $k_{max}$	10
Static backoff indicator $BI$	20 ms
Dynamic backoff indicator $BI$	see Eq. (36).
Access barring rate $P_{ACB}$	0.5 for ACB and $M/N$ for DAB

## B. BENCHMARKS

Three benchmarks that implement static backoff collision resolution are considered to compare their performance with dynamic backoff collision resolution. The first benchmark is the SRA with static backoff interval [29]. The second benchmark is the ACB with static backoff collision resolution [3]. The third benchmark is the DAB with static backoff collision resolution [4]. All three benchmarks consider a static uniform  $BI = 20$  ms. We also compared our proposed scheme to the recent state-of-the-art random access techniques which include the optimal ACB scheme with timing advance information [10] and the exponential backoff scheme in which the value of  $BI$  is computed as follows [3]:

$$T_{BO} = \mathcal{U}(0, 10 \times 2^{(k-1)}), \quad (37)$$

where  $k$  is the number of transmission attempts performed by a UE and its value ranging from 1 to  $k_{max}$ .

## C. PERFORMANCE METRICS

The performance of the dynamic backoff collision resolution can be evaluated based on four metrics as follows [1]:

- 1) Access success rate: defined as the ratio between the number of devices  $N_s$  that have successfully accessed the network to the total number of devices  $N$ .
- 2) Mean delay: defined as the mean time required for a UE to successfully complete the random access procedure.
- 3) Mean number of preambles retransmissions  $k$ : this metric reflects the mean number of retries performed by the successful UEs after the first access attempt failed.
- 4) Collision rate  $C_r$  defined as

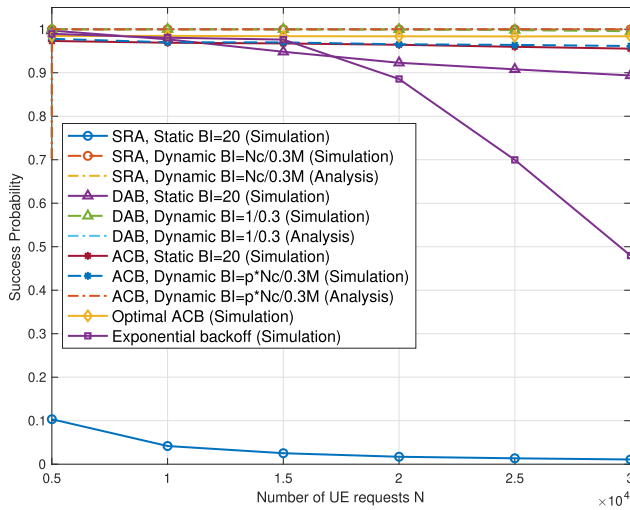
$$C_r = \frac{\text{NumberOfCollidedPreambles}}{M} \quad (38)$$

## VII. RESULTS EVALUATION AND DISCUSSION

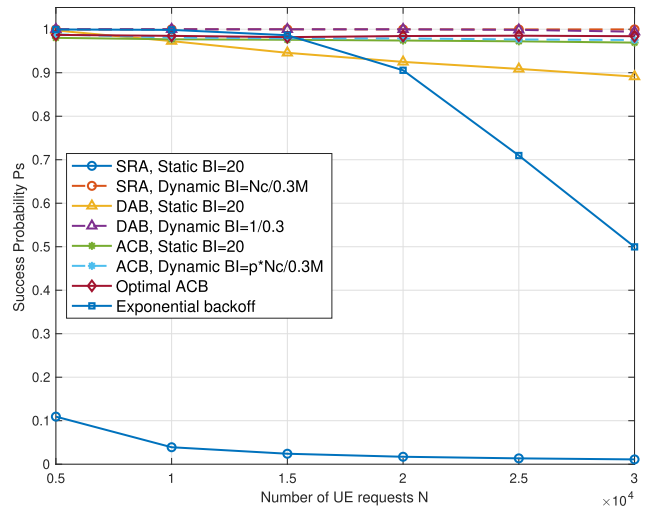
This section will present the simulation results for DBCR scheme based on the aforementioned performance metrics. We plot the results obtained from simulations as a function of the total number of devices  $N$ . We also plot the numerical results of the mathematical analysis of the access success rate for SRA, ACB and DAB based on Eq. (11) and Eq. (29), respectively. Fig. 7a and 7b shows the access success rate for SRA, ACB and DAB with the dynamic backoff collision resolution for delta and beta distributions respectively.

These results are compared to the SRA, ACB and DAB with static backoff collision resolution and with the exponential backoff scheme and the optimal ACB scheme. The dynamic backoff achieves 99.9% success probability for SRA and DAB in both scenarios which exceeds the performance of the optimal ACB. The simulation results of the dynamic backoff scheme match the results of mathematical analysis in which the access success achieves 99.9% which verifies the correctness of the proposed scheme. The results show that the dynamic backoff has increased the success probability for SRA by more than 90% and up to 10% for DAB. This increment results from the extended backoff interval which offers more RAOs for the backoff devices and in turn increases their opportunity for success. Even though the dynamic backoff for ACB does not achieve the optimal result, it has increased the success probability with approximately 1%. In fact, the results of success probability for ACB are affected by the chosen value for ACB rate. For example, choosing low value for  $p$  prevents more devices from starting the random access procedure. Thus, we expect a lower number of collisions and in turn a smaller backoff interval is needed. For this reason, we chose the value of 0.5 for  $P_{ACB}$  to reduce this effect as much as possible by choosing the middle value between 0 and 1. However, the results of access success probability for ACB with dynamic backoff achieved the highest value, equal to 1 when  $P_{ACB} = 0.9$ . Moreover, the DBCR scheme improved the access success probability for the massive arrivals ( $N = 30,000$ ) by approximately 50% compared to the exponential backoff scheme which indicates that the DBCR scheme is efficient even for large number of devices.

Furthermore, the dynamic backoff for ACB has reduced the mean delay by approximately 16% compared to the static backoff as shown in Fig. 8. Note that the results in Fig. 8 are presented in terms of the number of RAOs where each RAO is equal to one millisecond. The time delay here is equal to the number of RAOs because the periodicity of RAOs considered in this work is equal to one millisecond. However, when considering different RACH configuration index, the results must multiplied by the periodicity of the RAOs to get the real time delay. It is observed in Fig. 8a that the mean delay for the delta arrivals scenario is higher than the mean delay for beta arrivals in Fig. 8b. This is because all the  $N$  arrivals in delta distribution are activated in a single RAO which causes intensive congestion that requires a longer barring time which in turn increases the total delay. However, the dynamic backoff causes a slight increment around 10% in the mean delay for DAB. This is due to the longer backoff interval in the dynamic backoff. However, this increment is insignificant for delay-tolerant devices where the most important objective is to access the network successfully. Turning to SRA, it is observed that the mean delay with the static backoff is much lower than SRA with dynamic backoff. The low delay for SRA with static backoff results from the short backoff time which allows the UEs to retransmit after short period of time. This process

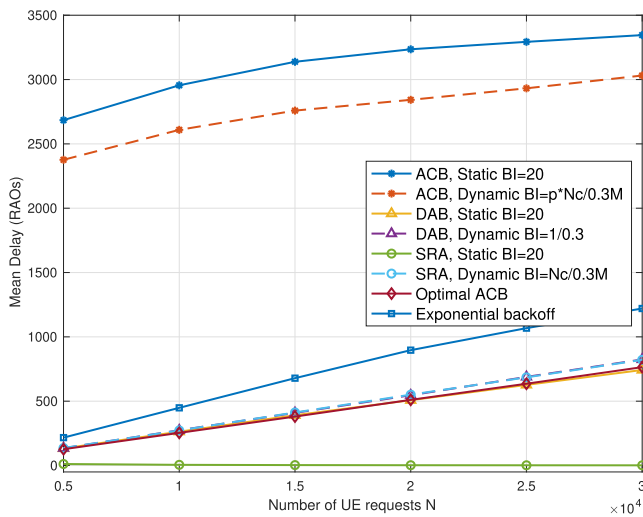


(a) Delta distribution.

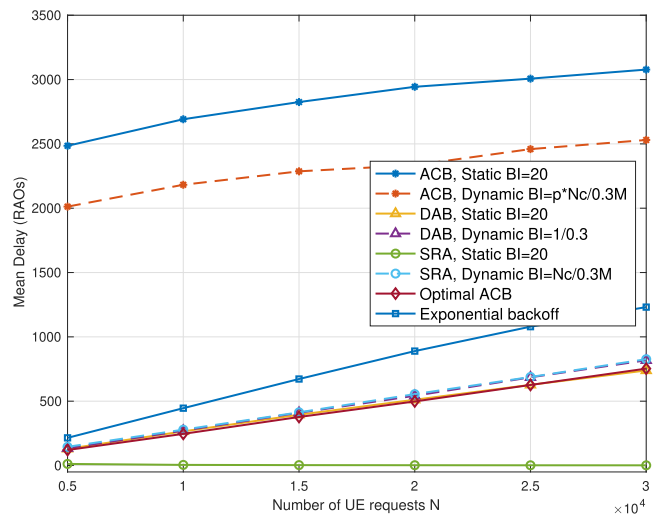


(b) Beta distribution.

FIGURE 7. Access success rate.



(a) Delta distribution.

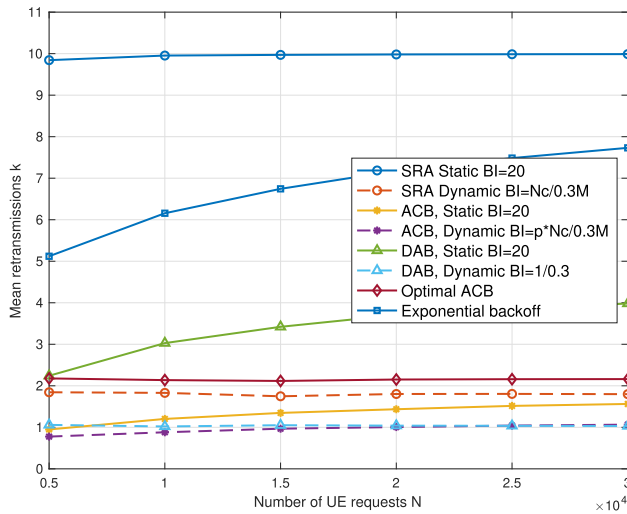


(b) Beta distribution.

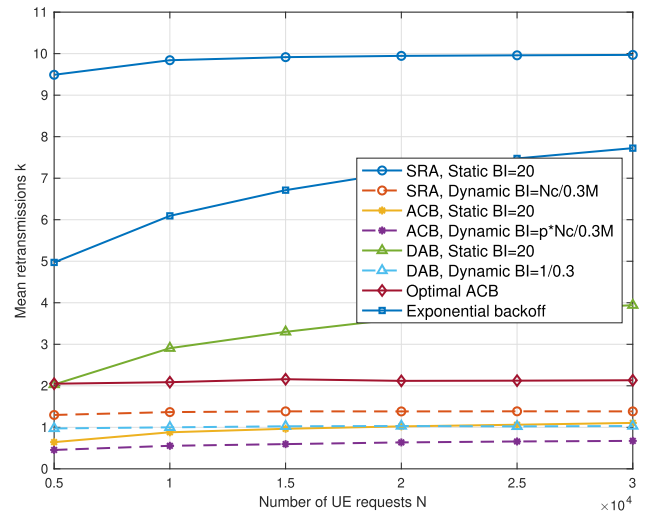
FIGURE 8. Mean delay.

causes the UEs to collide again due to the massive requests that share the short backoff interval and in turn most of requests exceed the maximum number of retransmissions. As a result, most of the requests are dropped within short period of time. However, the dynamic backoff reduced the access delay for the SRA around 30% compared to the exponential backoff. The reason is that the exponential backoff indicator increased as the number of retransmission attempts increased which in turn increases the access delay. In contrast, the dynamic backoff indicator reduced as the number of transmissions increased since this depends on the number of collided devices which reduced after each transmission attempt.

The mean number of preamble retransmissions for delta and beta distributions is shown in Fig. 9a and 9b respectively. This metric shows the number of retrials performed by a UE until it successfully accesses the network. The range of retransmissions in this study is between 0 to 10 which is the maximum number of preamble retransmissions based on the simulation settings. The number zero implies that a UE has succeeded in the first access attempt and performed no retransmission. We aimed to reduce the mean number of retransmissions to save the devices energy which are mostly battery operated. It is clearly seen from Fig. 9a and 9b that the dynamic backoff reduced the mean number of preamble retransmissions for SRA, ACB and DAB compared to the

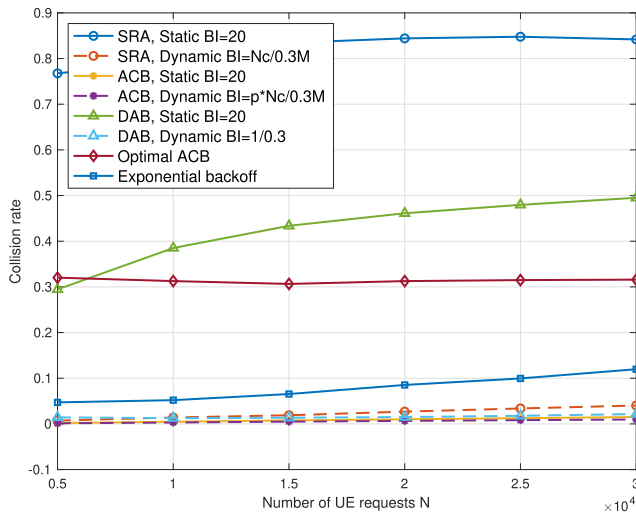


(a) Delta distribution.

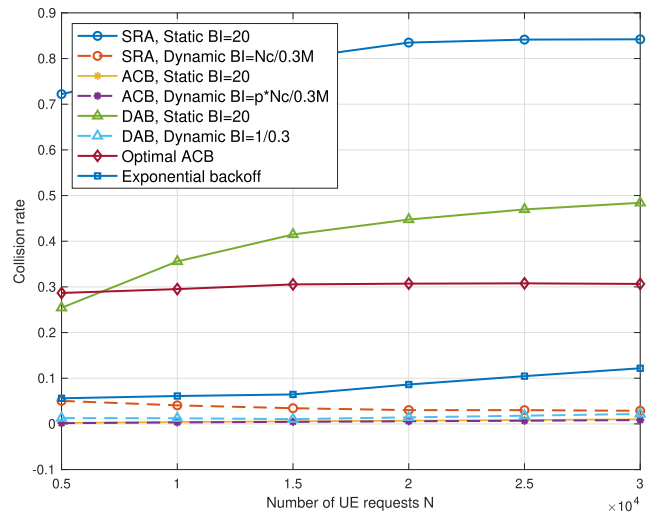


(b) Beta distribution.

FIGURE 9. Mean number of preamble retransmissions.



(a) Delta distribution.



(b) Beta distribution.

FIGURE 10. Collision rate.

static backoff. By observing the SRA with static backoff, it can be noticed that it reaches the maximum number of retransmissions which greatly consumes the energy of M2M devices. However, the dynamic backoff has decreased the mean number of retransmissions for SRA by around 70% which greatly improves the energy saving for M2M devices. Moreover, the dynamic backoff of SRA reduced the number of retransmissions around 75% compared to the exponential backoff. Furthermore, the mean number of retransmissions for ACB and DAB with dynamic backoff ranges from 0.5 to 1 which indicates that most devices have succeeded in their first or second access attempt. The dynamic backoff of ACB and DAB reduced the mean number of retransmissions up

to 50% compared to the optimal ACB. However, ACB with dynamic backoff is the most efficient in energy saving compared to SRA and DAB. The number of retransmissions is directly influenced by the collision rate, as shown in Fig. 10. It is observed that the collision rate for SRA with static backoff is very high, and this causes the devices to make more access attempts. However, dynamic backoff reduced the collision rate for SRA by approximately 70% compared to the exponential backoff which causes the same reduction in the mean number of retransmissions. The dynamic backoff has also reduced the collision rate for ACB and DAB with around 90% compared to the optimal ACB scheme which interprets the decrease in the mean number of preamble

retransmissions. Moreover, the ACB has the lowest collision rate which implies that ACB is effective in reducing the RACH congestion and in turn reducing the collision rate. Finally, we notice that there is a trade-off between the performance metrics. For example, the ACB with dynamic backoff achieved a high success probability and a low collision rate with a higher delay which makes it a good choice for the delay-tolerant battery-operated devices. However, the SRA and DAB with dynamic backoff have an optimal success probability with a higher energy consumption (collision rate) and a reasonable delay which makes them suitable for delay-sensitive applications. Therefore, the choice of the appropriate scheme depends on the QoS requirements imposed by the system administrator.

### VIII. CONCLUSION AND FUTURE WORK

This work investigated the issue of the RACH collisions with the aim of reducing the collision rate and increasing the access success probability. The paper explored the impact of modifying the BI value and the relation between the number of devices and the BI value. After that, a dynamic backoff collision resolution scheme was proposed with an optimal backoff indicator. The proposed scheme was implemented for three well-known random access schemes, SRA, ACB and DAB. The access success probability has been analysed based on the proposed dynamic backoff scheme. The performance of the dynamic backoff scheme was evaluated through extensive simulations, the results of which match the results of the analysis. The proposed scheme has been compared with the static backoff for SRA, ACB and DAB. The results show that the dynamic backoff achieved the optimal access success probability for SRA and DAB with a slight increase in the access delay. Moreover, the dynamic backoff increased the access success probability for ACB and reduced the access delay as well. Furthermore, the dynamic backoff reduced the collision rate and the number of retransmissions for SRA, ACB and DAB which in turn improved the energy efficiency for M2M devices.

There are several aspects considered for future work. For example, future work should study the issue of delay increments for SRA and DAB under the dynamic backoff scheme. Furthermore, future work should investigate the efficiency of the dynamic backoff collision resolution with different scenarios and different arrival distributions. Developing a collision avoidance scheme for M2M random access is another aspect that should be considered for future work.

### ACKNOWLEDGMENT

The authors would like to thank everyone who provided support to improve the content of this article.

### REFERENCES

- [1] *Study on RAN Improvements for Machine Type Communications*, document TR 37.868, 3GPP, 2011.
- [2] *Medium Access Control (MAC) Protocol Specification (Release 8)*, Access, Evolved Universal Terrestrial Radio, document TS 36.321, V8. 0.0, 3GPP, 2008.
- [3] L. Tello-Oquendo, I. Leyva-Mayorga, V. Pla, J. Martinez-Bauset, J.-R. Vidal, V. Casares-Giner, and L. Guijarro, "Performance analysis and optimal access class barring parameter configuration in LTE—A networks with massive M2M traffic," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3505–3520, Apr. 2018.
- [4] S. Duan, V. Shah-Mansouri, Z. Wang, and V. W. S. Wong, "D-ACB: Adaptive congestion control algorithm for bursty M2M traffic in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9847–9861, Dec. 2016.
- [5] *5G; NR; Medium Access Control (MAC) Protocol Specification (Release 15)*, document TS 38.321, V15.6.0, 3GPP, 2019.
- [6] *Technical Specification Group Services and System Aspects, Service accessibility (Release 9)*, document TS 22.011, V9.4.0, 3GPP, 2010.
- [7] *Further Performance Evaluation of EAB Information Update Mechanisms*, document R2-120270, 3GPP, 2012.
- [8] A. Sanchis-Cano, L. Guijarro, V. Pla, and J. R. Vidal, "Economic viability of HTC and MTC service provision on a common network infrastructure," in *Proc. 14th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2017, pp. 1044–1050.
- [9] S. Gharbi and N. Zangar, "Adaptive multiple access class barring factors for M2M communications in LTE—A networks," in *Proc. 12th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2015, pp. 605–606.
- [10] Z. Wang and V. W. S. Wong, "Optimal access class barring for stationary machine type communication devices with timing advance information," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5374–5387, Oct. 2015.
- [11] H. Althumali and M. Othman, "A survey of random access control techniques for machine-to-machine communications in LTE/LTE—A networks," *IEEE Access*, vol. 6, pp. 74961–74983, 2018.
- [12] I. Leyva-Mayorga, L. Tello-Oquendo, V. Pla, J. Martinez-Bauset, and V. Casares-Giner, "On the accurate performance evaluation of the LTE—A random access procedure and the access class barring scheme," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 7785–7799, Dec. 2017.
- [13] R. Harwahyu, X. Wang, R. F. Sari, and R.-G. Cheng, "Analysis of group paging with pre-backoff," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, pp. 1–9, Dec. 2015.
- [14] M. El Tanab and W. Hamouda, "Machine-to-machine communications with massive access: Congestion control," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3545–3557, Apr. 2019.
- [15] A. S. A. El-Hameed and K. M. F. Elsayed, "A Q-learning approach for machine-type communication random access in LTE-advanced," *Telecommun. Syst.*, vol. 71, no. 3, pp. 397–413, Jul. 2019.
- [16] L. Tello-Oquendo, D. Pacheco-Paramo, V. Pla, and J. Martinez-Bauset, "Reinforcement learning-based ACB in LTE—A networks for handling massive M2M and H2H communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–7.
- [17] H. Han, Y. Li, W. Zhai, and L. Qian, "A grant-free random access scheme for M2M communication in massive MIMO systems," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3602–3613, Apr. 2020.
- [18] Q. Zhang, S. Jin, and H. Zhu, "A hybrid-grant random access scheme in massive MIMO systems for IoT," *IEEE Access*, vol. 8, pp. 88487–88497, 2020.
- [19] G. C. Madueno, S. Stefanovic, and P. Popovski, "Efficient LTE access with collision resolution for massive M2M communications," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 1433–1438.
- [20] C. Yoon, "Distributed queuing with preamble grouping for massive IoT devices in LTE random access," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2016, pp. 103–105.
- [21] J. Yuan, H. Shan, A. Huang, T. Q. S. Quek, and Y.-D. Yao, "Massive Machine-to-Machine communications in cellular network: Distributed queuing random access meets MIMO," *IEEE Access*, vol. 5, pp. 2981–2993, 2017.
- [22] J. S. Kim, S. Lee, and M. Y. Chung, "Efficient random-access scheme for massive connectivity in 3GPP low-cost machine-type communications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6280–6290, Jul. 2017.
- [23] M. Vilgelm, S. Rueda Linares, and W. Kellerer, "Dynamic binary countdown for massive IoT random access in dense 5G networks," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6896–6908, Aug. 2019.
- [24] H. M. Gursu, M. Vilgelm, W. Kellerer, and M. Reisslein, "Hybrid collision avoidance-tree resolution for M2M random access," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 4, pp. 1974–1987, Aug. 2017.
- [25] J.-R. Vidal, L. Tello-Oquendo, V. Pla, and L. Guijarro, "Performance study and enhancement of access barring for massive machine-type communications," *IEEE Access*, vol. 7, pp. 63745–63759, 2019.

- [26] *MTC LTE Simulations*, document RAN WG2 no. 71, R2-104663, 3GPP, 2010.
- [27] *Physical Channels and Modulation*, document TS 36.211 V15.7.0, 3GPP, 2019.
- [28] O. Arouk, A. Ksentini, and T. Taleb, "How accurate is the RACH procedure model in LTE and LTE—A?" in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Sep. 2016, pp. 61–66.
- [29] M. Ouaiassa, M. Benmoussa, A. Rhattoy, M. Lahmer, and I. Chana, "Performance analysis of random access mechanisms for machine type communications in LTE networks," in *Proc. Int. Conf. Adv. Commun. Syst. Inf. Secur. (ACOSIS)*, Oct. 2016, pp. 1–6.

**HUDA DAKHILALLAH ALTHUMALI** (Graduate Student Member, IEEE) received the bachelor's degree in computer science from Taif University, Saudi Arabia, in 2009, and the master's degree in computer science from Universiti Putra Malaysia, in 2017, with a specialization in distributed computing, where she is currently pursuing the Ph.D. degree in computer networks. She is currently a Lecturer with Imam Abdulrahman Bin Faisal University, Saudi Arabia. Her research interests include wireless networks, the IoT, cloud computing, and network security.



**MOHAMED OTHMAN** (Senior Member, IEEE) received the Ph.D. degree (Hons.) from the National University of Malaysia. He was the Deputy Director of the Information Development and Communication Centre, where he was in charge of the UMPNet network campus, uSport Wireless Communication Project, and the UPM Data Centre. He is currently a Professor in computer science with the Department of Communication Technology and Networks, Universiti Putra Malaysia (UPM). He is also an Associate Researcher and a Coordinator of high-speed machines with the Laboratory of Computational Science and Informatics, Institute of Mathematical Science, UPM. He is also a Visiting Professor with South Kazakhstan State University, Shymkent, Kazakhstan, and L. N. Gumilyov Eurasian National University, Astana, Kazakhstan. He has filed six Malaysian, one Japanese, one South Korean, and three U.S. patents. He has published more than 300 International journal articles and 330 proceeding articles. His main research interests include computer networks, parallel and distributed computing, high-speed interconnection networks, network design and management (network security, and wireless and traffic monitoring), consensus in the IoT, and mathematical models in scientific computing. He is a Life Member of the Malaysian National Computer Confederation and the Malaysian Mathematical Society. In 2017, he received the Honorary Professorship from SILKWAY International University (formerly known as South Kazakhstan Pedagogical University), Shymkent, and the Best Ph.D. Thesis by Sime Darby Malaysia and the Malaysian Mathematical Science Society, in 2000.

**NOR KAMARIAH NOORDIN** (Senior Member, IEEE) received the B.Sc. degree in electrical engineering with a major in telecommunications from The University of Alabama, Tuscaloosa, AL, USA, in 1987, the master's degree from University Teknologi Malaysia, and the Ph.D. degree from Universiti Putra Malaysia (UPM). She has been with UPM, since 1988. She was appointed as an Associate Professor, in 2006, and a Professor, in 2012. During her tenure, she has been the Head of the Department, the Deputy Dean of academics, and the Director of corporate strategy and communication of the university. She is currently the Dean of the Faculty of Engineering, UPM. She has been with UPM for more than 25 years. During that time, she has published more than 200 journal articles and conference papers and has secured more than 30 research and consultancy projects.



**ZURINA MOHD HANAPI** (Member, IEEE) received the B.Sc. degree in computer and electronic systems from the University of Strathclyde, Glasgow, U.K., in 1999, the M.Sc. degree in computer and communication system engineering from Universiti Putra Malaysia (UPM), Selangor, Malaysia, in 2004, and the Ph.D. degree from Universiti Kebangsaan Malaysia, in 2011. She is currently an Associate Professor with the Department of Communication Technology and Networks, Faculty of Computer Science and Information Technology, UPM. She has been a Lecturer with UPM, since 2004. She has authored over 70 papers in cited journals and conferences in the area of security and wireless sensor networks. Her current research interests include security, routing, wireless sensor networks, wireless networks, distributed computing, and cyber-physical-systems. She is a member of the Malaysian Security Research Committee.

...