A NOMA-Enhanced 2-Step RACH Procedure for Low-Latency Access in 5G Networks

Dawei Nie, Wenjuan Yu, *Member, IEEE*, Chuan Heng Foh, *Senior Member, IEEE*, Qiang Ni, *Senior Member, IEEE*

***Abstract*—Random access channel (RACH) procedure is critical to support a multitude of devices transmitting small data payloads while ensuring low-latency access. In 3GPP Release 16, a two-step RACH is proposed to alleviate signaling overhead and access latency. While benefits are noticeable, collisions still persist. In this paper, we propose a novel non-orthogonal multiple access (NOMA)-enhanced 2-step RACH scheme (NOMA-RACH) that jointly leverages the benefits of access class barring (ACB), 2-step RACH, and NOMA random access (NOMA-RA) to further enhance the performance. We conduct a holistic study that accounts for entire access latency. The scheme optimizes NOMA access probabilities, utilizes an adjustable barring mechanism for delay-sensitive devices, and identifies the optimal barring rate for low latency. We develop a Markov chain model to analyze NOMA access and derive the optimal access probabilities and throughput of NOMA blocks. To cope with the practical scenarios with constantly changing user equipment (UE) traffic, we propose a deep contextual multi-armed bandit (DCMAB) model that optimizes the NOMA throughput and dynamically adjusts the barring rate based on the observable channel feedback. Our simulation results demonstrate that the DCMAB model performs better than benchmark schemes and remains close to the optimal latency confirming the effectiveness of our proposed scheme under changing UE traffic.**

***Index Terms*—Two-step random access, NOMA, user barring, Markov chain, multi-armed bandit.**

I. Introduction

In recent years, there has been an unprecedented surge in data traffic, with predictions indicating a continual exponential rise in both the number of devices and the demand for wireless connections [1], [2]. The advent of massive machine-type communication (mMTC) is considered indispensable, presenting a formidable challenge for forthcoming communication networks. This scenario envisions the deployment of millions of Internet of Things (IoT) devices per square kilometer to establish extensive connectivity [3]. While extensive research has explored wireless environments shaped by growing traffic, the efficient support for a massive volume of devices requiring low-latency access remains a complex and critical issue in cellular systems [1].

Within cellular networks, such as LTE-A, the initiation of an access request by user equipment (UE) triggers a contention-based procedure known as contention-based random access channel (or RACH in short) between the UE and the base station (BS) [4]. In response to the anticipated growing traffic and massive volume of devices in 5G communications, 3GPP Release 16 not only brought technical enhancements but also introduced novel features [5]. One such feature is the incorporation of the 2-step RACH which offers advantages in two scenarios [6]. Firstly, for burst transmissions of small packets, the simplicity of RACH proves appealing by mitigating the substantial overhead associated with Radio Resource Control (RRC) connection setup and resume procedures [7]. Secondly, for new radio unlicensed spectrum, streamlining random access (RA) contributes to a latency reduction in connecting UEs to the BS. As an enhanced version of the RA procedure for 5G, the 2-step RACH is expected to offer notable advantages, including reduced signaling overhead, diminished power consumption, and decreased latency.

The literature has studied various application scenarios and technologies for the 2-step RACH procedure [8]–[11]. In a pioneering work by Jones *et al.* in 2019 [9], a novel random access procedure named RAPID was introduced specifically tailored for delay-sensitive devices. A scheme for estimating the traffic characteristics of MTC devices was presented. By using a Markov chain model, the authors analyzed random access load and latency to determine the optimal number of preambles for RAPID. Building upon this, a subsequent contribution [10] expounded on the intricacies of the 2-step RACH defined in 3GPP Release 16. The work outlined challenges associated with various random access schemes and proposed a framework designed to assist UEs in estimating the

This work was supported in part by the Western O-RAN Deployment (ONE WORD) Project, in part by the EPSRC CHEDDAR TMF uplift SustainAIRA6G project, and in part by HORIZON-MSCA-2022-SE-01 under Grant Number 101131204. (Corresponding author: Wenjuan Yu.)

D. Nie, W. Yu, and Q. Ni, are with the School of Computing and Communications, InfoLab21, Lancaster University, LA1 4WA, U.K. C. Foh is with 5GIC & 6GIC, Institute for Communication Systems (ICS), University of Surrey, Guildford, Surrey, UK (Emails: {d.nie, w.yu8, q.ni}@lancaster.ac.uk, c.foh@surrey.ac.uk)

timing advance (TA) command. Another research work [11] focused on the channel structure design for the 2-step RACH procedure, specifically addressing both the preamble and data segments of MsgA, along with the associated receiver processing framework. The study found that employing smaller payload sizes or adopting distinct demodulation reference signal (DMRS) ports for UEs sharing the same physical resources can further enhance the performance of random access. However, concerns persist due to the purely random selection and limited availability of orthogonal preambles per cell, leading to increased collisions among MTC devices.

Furthermore, traffic bursts prevalent in the mMTC scenario exacerbate the volume of devices initiating RACH, leading to congestion and prolonged access latency [12]. To address these issues, effective solutions typically involve managing access traffic or enhancing the performance of RACH. For example, LTE-A specifications have incorporated access class barring (ACB), a mechanism that distributes UE accesses over time [13]. ACB categorizes users into distinct access classes. Periodically, the BS broadcasts the barring rate b for each access class, and each UE follows this setting to participate in the next RACH procedure. In scenarios with massive connectivity, ACB is crucial for stabilizing the RACH procedure. The barring rate can be set low to prevent excessive UEs from participating, avoiding preamble saturation, while a high barring rate allows more UEs but may cause collisions and increase latency. Therefore, finding the optimal barring rate is critical for balancing performance in massive connectivity systems.

Since 3GPP does not specify how to dynamically adjust ACB parameters, many studies have proposed their approaches. In a study by Li *et al.* [14], a quality of service (QoS)-based dynamic and adaptive mechanism was introduced, which prioritized preamble allocations for delay-sensitive devices while adaptively adjusting ACB parameters for both delay-sensitive and delay-tolerant devices. ACB mechanism can also integrate machine learning (ML) algorithms. For instance, studies such as [15] and [16] applied reinforcement learning (RL) algorithms, including Q-learning, to intelligently determine ACB parameters, effectively mitigating congestion and reducing access latency. In another study by Chen *et al.* [17], a two-sided learning approach based on multi-armed bandit (MAB) was proposed. This approach enabled devices to dynamically select resource blocks (RBs) for packet transmissions, optimizing throughput. Additionally, a contextual MAB scheme was proposed in [18] to jointly address user admission and channel access for the RACH procedure.

As mentioned earlier, there are mainly two approaches to mitigating collisions and reducing access latency. We have introduced the management of access traffic via ACB. Here we focus on addressing these issues by enhancing 2-step RACH performance [19], [20]. One effective method is the use of non-orthogonal multiple access (NOMA) [21], [22]. NOMA permits the simultaneous transmission of multiple non-orthogonal signals within the same time and frequency resources. Receivers can decode the superimposed signals in either the power or code domain [23], [24]. NOMA offers versatility in enhancing connectivity, security, and efficiency for future IoT networks. For example, Cao *et al.* propose a NOMA-assisted semi-grant-free transmission scheme that boosts massive connectivity and reduces access delays in IoT environments. Their security-optimized scheduling, including maximal user scheduling, mitigates eavesdropping risks while maintaining QoS for grant-based users [25]. Additionally, they address reliable and secure communications in wireless-powered NOMA systems, introducing a joint artificial noise and power allocation scheme that improves secrecy and throughput, particularly under energy constraints [26]. Specifically, power-domain NOMA-enabled RA, known as NOMA-RA, allows users to randomly select both a channel and a power level for transmission, utilizing a set of pre-determined power levels [3]. The BS with successive interference cancellation (SIC) then decodes all the received information. In contrast to the conventional RA schemes like pure ALOHA and slotted ALOHA, NOMA-RA is more spectrum-efficient, as it can leverage an additional dimension to accommodate more devices. Several research papers, including [27]–[30], have focused on evaluating the throughput performance of this novel NOMA-RA scheme. In our previous study [3], we analyzed the theoretical throughput performance of power-domain NOMA-RA and showed that NOMA-RA with four power levels has a potential to achieve a maximum throughput three times higher than that of an equivalent multi-channel slotted ALOHA.

Despite the current research efforts to improve the performance of RACH procedure, lengthy RA delays and intensive collisions persist. We notice that existing literature often focuses on a single aspect of improvement. Besides, most studies focused on the conventional 4-step RACH, their solutions cannot be directly applied to 2-step RACH due to incompatibility. This paper aims to address the research gap by enhancing the newly introduced 2-step RACH in the 5G standard to meet the growing demands of future networks. To achieve better performance over any single approach, we jointly

leverage the benefits of ACB and NOMA within the 2-step RACH for the first time. The NOMA-enhanced 2-step RACH, named 2-step NOMA-RACH, is proposed and analyzed for the first time in this work. The scheme optimizes the access probabilities of the NOMA blocks inside RACH, employs an adjustable barring mechanism for massive delay-sensitive devices, and determines the optimal barring rate to minimize latency. Furthermore, when concurrently applying the external adjustment of the ACB barring rate and the internal optimization of the 2-step NOMA RACH procedure, neither of them can be considered in isolation due to their inherent interdependence, emphasizing the need for our holistic approach. The following describes the main contributions of this paper:

- Based on the newly introduced 2-step RACH in the 3GPP 5G standard, a novel 2-step NOMA-RACH scheme is proposed to address the increasingly stringent latency and mMTC requirements of future networks. This scheme integrates intelligent ACB and NOMA into the 2-step RACH process. Its latency performance is analyzed using a Markov chain model for the first time, as well as the optimal barring rate. Simulation results validate our analysis and demonstrate a significant reduction in latency compared to other benchmark models.
- To fully leverage the potential of applying NOMA within the 2-step RACH for the first time, our proposed scheme allows collided UEs to access remaining NOMA blocks after prioritizing primary UEs[1] in the top levels of NOMA blocks. We derive analytical expressions for optimal access probabilities and NOMA throughput and validate them in simulation.
- Finally, focusing on a more practical scenario with dynamic UE traffic, a newly proposed deep contextual multi-armed bandit (DCMAB) agent is designed to jointly optimize NOMA access probabilities and barring rate under practice scenarios. Simulations show that the designed DCMAB model outperforms the benchmark schemes and maintains latency close to the theoretical minimum.

The remainder of this paper is organized as follows. We first present the system model in Section II and introduce the proposed 2-step NOMA-RACH scheme. Then, we develop a Markov chain model and conduct the theoretical analysis on latency and NOMA access probabilities in Section III. In Section IV, the DCMAB agent is designed Simulation results are given in Section V, followed by conclusions in Section VI.

[1]A primary UE refers to a UE with a fallback RAR whose preamble has been successfully detected, as will be introduced later.

II. System Model

Collision is a critical issue in the RACH procedure, leading to an increase in access latency. As previously discussed, managing UEs traffic with ACB can alleviate collisions. Additionally, integrating NOMA within the RACH procedure can further enhance UEs access success rates, effectively resolving collisions. As mentioned earlier, our system model incorporates the ACB mechanism and the 2-step RACH assisted with NOMA to mitigate collisions. In this section, we first revisit the ACB mechanism and the 2-step RACH procedure, followed by presenting our proposed NOMA-RACH.

A. ACB Mechanism

ACB is a UE traffic control mechanism. It uses the following steps to regulate UE traffic. The BS first disseminates a barring rate, denoted as b, as part of the system information before preamble transmission. Each UE then generates a random number between 0 and 1, and only participates in RACH if the generated number is smaller than b, otherwise, it is barred until the next round. With the permission to participate, the UE shall transmit a preamble to the BS requesting for access.

The selection of the barring rate b is thus pivotal in managing the challenge of RACH overload. However, this choice involves a delicate trade-off between resource utilization and collision management. A smaller b restricts the number of participating UEs, reducing collisions but risking underutilization of resources. Conversely, a larger b increases the number of participating UEs, potentially leading to excessive preamble collisions. For delay-sensitive devices, the impact of b is significant. A smaller barring rate may extend the time spent waiting to pass barring, while a larger b may result in transmission failures and backoff, potentially leading to increased latency. Hence, finding the optimal value for b becomes a critical consideration in system design. By regulating traffic loads, the ACB can control the load such that RACH will operate in its optimal point for latency and throughput performance. In addition to traffic regulation, there is an opportunity to enhance the RACH process by incorporating NOMA. This integration will be introduced in subsection II-C.

B. 2-step RACH Procedure

Let us proceed by providing an overview of the 2-step RACH protocol. This protocol is initiated when a UE requests for an access to the network. The 2-step RACH can be conceptualized as a condensed version of the 4-step RACH, as illustrated in Fig. 1. Despite the reduced steps, it involves more complex logic. In the initial

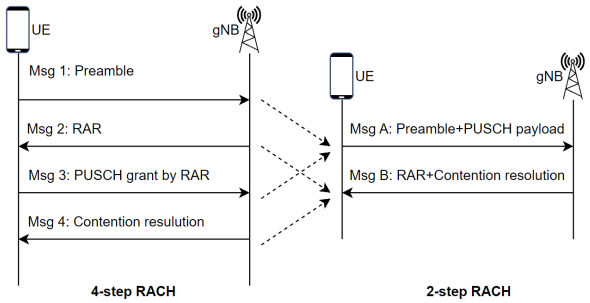


Fig. 1: Procedures for 4-step and 2-step RACH.

step, MsgA of UE encompasses the Msg1 preamble and Msg3 payload, akin to the 4-step RACH procedure. UEs transmit their preamble and payload over physical random access channel (PRACH) and physical uplink shared channel (PUSCH) respectively. In 2-step RACH, the PUSCH resource used for payload transmission is determined by a mapping sequence set by the BS, based on the selected preamble ID.

Upon receiving MsgA, the BS proceeds with detecting the preamble from PRACH and decoding the payload from PUSCH. Meanwhile, the UE awaits a response from the network within a configured window W_{RAR}. This response, known as MsgB, varies depending on the detection of the preamble and the decoding outcomes of the payload at the BS. Different cases may arise, guiding the actions to be taken in the next step, as detailed below.

- **Case 1**: With successful detection and decoding of the PRACH preamble and PUSCH payload, the BS replies MsgB containing a successful random access response (RAR) and a Timing Advance (TA) command. Upon receiving MsgB by the UE, the 2-step RACH handshake is completed successfully.
- **Case 2**: It is possible that a PRACH preamble is successfully detected, allowing the BS to determine the preamble's reception time, but the decoding of the associated PUSCH payload fails. This case is typically caused by channel errors in the PUSCH or by multiple UEs transmitting their payloads on the same PUSCH resource[2]. In this case, the preamble is successfully detected, and the BS will utilize the preamble reception time to transmit a fallback indication (fallback RAR) in MsgB. The handshake then proceeds by falling back to the third and fourth steps of the 4-step RACH.
- **Case 3**: It is also possible that a PRACH preamble is detected, but the decoding of the corresponding PUSCH payload fails, and the BS is unable to determine the preamble reception time. This case is typically caused by multiple identical preambles being received at the BS[3]. In this case, the preamble is said to be collided, and the BS will transmit a backoff indication to the collided UEs, instructing them to back off and attempt random access again.
- **Case 4**: If the UE does not receive a response by the end of the RAR window W_{RAR}, it assumes that the BS has failed to detect MsgA. Consequently, there will be no RAR transmission between the BS and the UE. The UE will perform a backoff procedure.

The flowchart delineating the standard 2-step RACH procedure is presented in Fig. 2. Different cases arise based on the outcomes of preamble detection (PBD), preamble reception time configuration (RTC), and payload decoding (PLD). The UE that successfully receives the contention resolution (CR) completes the random access procedure. In cases where the random access procedure is deemed unsuccessful, UEs will enter the backoff (BO) state and initiate the re-transmissions of MsgA. The number of times of MsgA re-transmission is monitored, and when the count reaches a predetermined threshold m, the 2-step RACH procedure is considered unsuccessful, prompting the UE to revert to the 4-step RACH procedure.

C. The Proposed 2-step NOMA-RACH

Having introduced the 2-step RACH and ACB mechanisms, we now provide a detailed description of our proposed 2-step NOMA-RACH procedure. Unlike the standard 2-step RACH, our proposed scheme incorporates NOMA into the RACH procedure, thereby enhancing its effectiveness by mitigating collisions. This is made possible because power-domain NOMA permits multiple UEs to transmit on the same NOMA resource. To manage the massive UE traffic in the proposed scheme, when a UE initiates the access, it first participates in the ACB mechanism, as described in Section II-A. If the UE successfully passes the ACB, it will then participate in the 2-step NOMA-RACH procedure by transmitting MsgA. Upon receiving MsgA, the BS attempts to detect the preamble and decode the payload, resulting in the outcome cases discussed in Section II-B. If the BS cannot determine the RTC of the preambles (Case 3), the preamble is considered collided. Conversely, the preamble is deemed successful if the RTC is successful, regardless of whether the PUSCH payload in MsgA is decoded successfully (Cases 1 and 2). Different from

[2]In many-to-one mapping, different preamble IDs can be mapped to the same PUSCH resource for payload transmission, which can cause PUSCH payload collisions [10].

[3]Collisions on PRACH preambles can result in payload decoding failures. For further details, please refer to [10], [31].

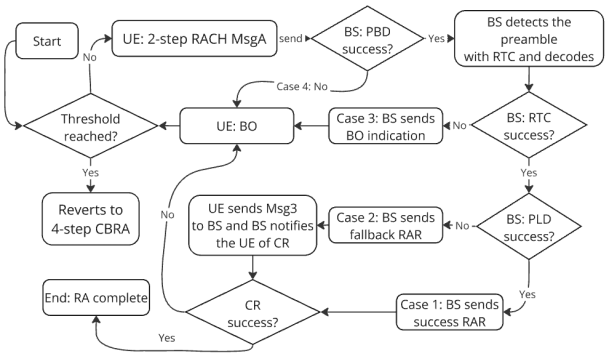


Fig. 2: 2-step RACH procedure.

the original 2-step RACH illustrated in Section. II-B, in our proposed 2-step NOMA-RACH procedure, NOMA is selectively applied in Cases 2 and 3.

In Case 2, the BS can identify the UEs and send fallback RARs. A unique NOMA resource can be assigned to each identified UE. The UEs that are assigned unique NOMA blocks are designated as primary UEs, while collided UEs refer to UEs whose preamble has collided. In our design, the BS first schedules primary UEs onto sufficient NOMA blocks in a sequential manner by sending advanced fallback RAR[4], following a highest-power-level-first principle. Subsequently, the BS broadcasts a NOMA block availability map, inviting collided UEs to opportunistically access the unoccupied NOMA blocks. Following the assignment, and similar to the approach in [18], all remaining NOMA resource blocks are made available for access by the collided UEs in Case 3. These collided UEs wait for the NOMA block availability map and receive the advanced fallback RAR from the BS. Thus, instead of enforcing BO, we allow these collided UEs to re-transmit the payload and opportunistically utilize the remaining NOMA resource blocks to resolve preamble collisions, thereby enhancing overall performance. UEs that successfully pass the SIC decoding of NOMA will receive the contention resolution message from the BS and complete the RA procedure. Similar to the original 2-step RACH, if the random access procedure is unsuccessful, UEs will enter the BO state and initiate the re-transmission of MsgA. The flowchart of the proposed 2-step NOMA-RACH procedure is depicted in Fig. 3, with the adoption of NOMA highlighted in the grey area.

III. Theoretical Analysis

In the proposed 2-step NOMA-RACH scheme, the introduction of NOMA further complicates the setting of ACB parameter values, as various parameters influence each other within the transmission and BO loops outlined in the flowchart. The choice of the ACB barring rate not only impacts the optimal access probabilities of NOMA blocks but also influences overall system latency. Finding the optimal access probabilities and analyzing access latency involves complex calculations. Here, we model our proposed 2-step NOMA-RACH scheme using a Markov chain to analyze its performance.

A. The Markov Chain Model

We begin by making several assumptions. Firstly, we assume no process error within the RACH procedure and no transmission failures due to channel errors. With these assumptions, Case 4 will not occur. The UE operation can be modeled by the Markov chain depicted in Fig. 4, where $S_{i,j}$ is the state of each status. Here, $i \in \mathcal{M}$ represents the re-attempt count after BO where $\mathcal{M} = \{0, 1, \ldots, m\}$, and $j \in \{1, 2, 3, \text{B,T,CR}\}$ is the current UE status. Let $\mathcal{N} = \{1, 2, 3\}$. Since UEs revert to the 4-step RACH procedure after $m+1$ failures, and the impact of the 4-step components in the proposed scheme can be disregarded if m is appropriately set, we focus solely on the 2-step RACH components to effectively model the proposed scheme. The following defines state elements and state transition probabilities of the model.

- **Barring State** ($S_{0,\text{B}}$): In this state, the UE has a probability P_b of passing the barring state, meaning that the random number generated by the UE in ACB is less than the barring rate b.
- **MsgA Transmission States** ($S_{i,\text{T}}$): The UE transmits MsgA to the BS, which includes a randomly selected preamble. Upon receiving MsgA, the BS begins to detect the preamble and decode the payload, resulting in three possible cases: Cases 1, 2, and 3.
- **Successful UE States** ($S_{i,1}$): The BS has a probability P_1 of successfully detecting a preamble with RTC and decoding the payload.
- **Primary UE States** ($S_{i,2}$): The BS has a probability P_2 of successfully detecting a preamble with RTC but failing to decode the payload. Each primary UE will be allocated a unique NOMA block for transmitting Msg3, and its SIC is assumed to be successful.
- **Collided UE States** ($S_{i,3}$): The BS has a probability P_3 of receiving multiple identical preambles and being unable to specify the RTC. In such states, the collided UEs will opportunistically access the remaining NOMA blocks according to pre-calculated access probabilities. Then, there is a probability P_4 that these collided UEs successfully access the

[4]In proposed scheme, fallback RARs include NOMA allocation information, referred to as 'advanced fallback RARs' in this paper.

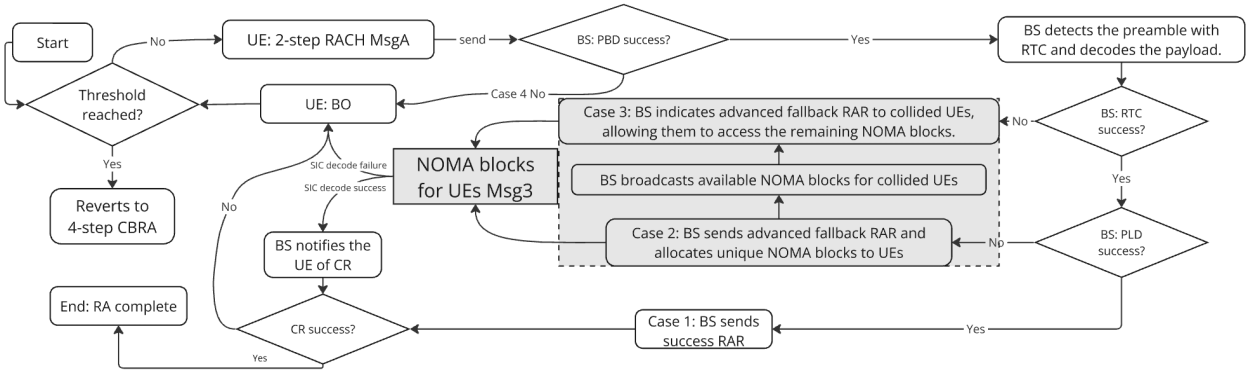


Fig. 3: 2-step NOMA-RACH procedure.

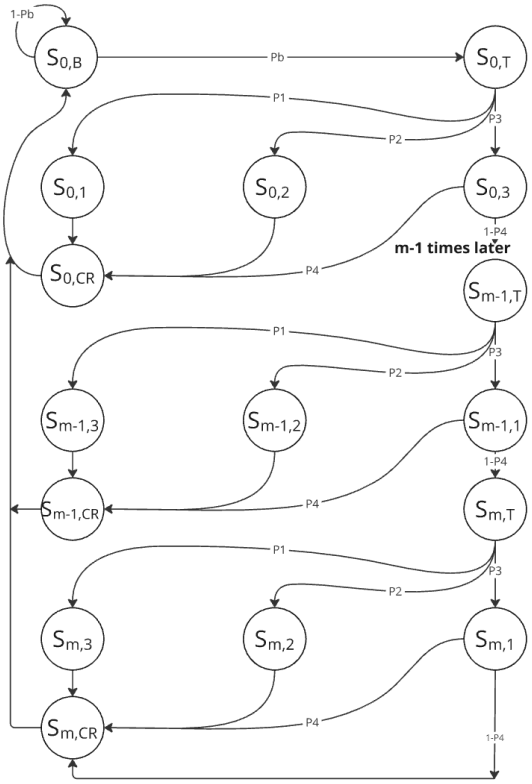


Fig. 4: Markov chain model for 2-step NOMA RACH

NOMA blocks.

- **Content Resolution States** ($S_{i,\mathrm{CR}}$): BS sends a contention resolution message to the UE.

B. Stationary Distribution Derivation

We denote $\pi_{i,j}$ as the stationary probability of state $S_{i,j}$. These probabilities, derived from the stationary distribution, are expressed in (1).

$$\pi_{i,j} = \begin{cases} \pi_{0,\mathrm{B}}, & i = 0, j = \mathrm{B} \\ P_b \pi_{0,\mathrm{B}} \prod_{l=0}^{i} (P_3(1-P_4))^l, & i \in \mathcal{M}, j = \mathrm{T} \\ P_j P_b \pi_{0,\mathrm{B}} \prod_{l=0}^{i} (P_3(1-P_4))^l, & i \in \mathcal{M}, j \in \mathcal{N} \\ P_b \pi_{0,\mathrm{B}} (P_1 + P_2 + P_3 P_4) \prod_{l=0}^{i} (P_3(1-P_4))^l, & i \in \mathcal{M}, j = \mathrm{CR}. \end{cases} \tag{1}$$

For states $S_{i,j}$, j depends on the current status, while i denotes the re-attempt count for that state. The stationary distribution of a particular state can be represented by the summation of all possible i values, denoted as $\pi_j = \sum_{i=0}^{L} \pi_{i,j}$. The expression for π_j is provided in (2).

$$\pi_j = \begin{cases} \pi_{0,\mathrm{B}}, & j = \mathrm{B} \\ \frac{P_b \pi_{0,\mathrm{B}} (1-(P_3(1-P_4))^{L+1})}{1-P_3(1-P_4)}, & j = \mathrm{T} \\ \frac{P_j P_b \pi_{0,\mathrm{B}} (1-(P_3(1-P_4))^{L+1})}{1-P_3(1-P_4)}, & j \in \mathcal{N} \\ \frac{P_b \pi_{0,\mathrm{B}} (P_1+P_2+P_3P_4)(1-(P_3(1-P_4))^{L+1})}{1-P_3(1-P_4)}, & j = \mathrm{CR}. \end{cases} \tag{2}$$

The value of each particular state can be determined by summing the stationary probabilities. We can solve the stationary probability by leveraging the fact that the sum of all stationary probabilities equals one.

C. State Sojourn Time and Latency Analysis

To determine the stationary distribution and calculate the transition probabilities, it is essential to express the average number of UEs performing preamble transmission within the Markov chain framework. From the perspective of a single UE, the average time spent on preamble transmission can be derived through stationary distribution analysis. This requires determining the time proportion that a UE spends on preamble transmission, denoted by τ. With this time proportion, we can statistically estimate the average number of UEs performing

preamble transmission in the system's steady state. To calculate the time proportion τ, we must first determine the sojourn time, which is the time a UE spends in each state before moving to the next. Let $T_{i,j}$ represent the sojourn time for state $S_{i,j}$. If the processing time remains consistent across all reattempts, then $T_j = T_{i,j}$ is independent of i. The average sojourn time for each state is detailed in (3) below.

$$T_{i,j} = \begin{cases} t_w, & j = \text{B} \\ t_T, & j = \text{T} \\ W_{\text{RAR}}, & j = 1 \\ t_{\text{NOMA}} + W_{\text{RAR}} + t_3, & j = 2 \\ t_{\text{NOMA}} + W_{\text{RAR}} + t_3 + (1 - P_4)\overline{t_{\text{BO}}} +, j = 3 \\ W_{\text{CR}}, & j = \text{CR}. \end{cases} \tag{3}$$

In the above, t_w, t_T and $\overline{t_{BO}}$ denote the barring waiting time, preamble transmission time and the average BO time window respectively. Moreover W_{RAR}, t_{NOMA}, t_3 and W_{CR} represent the RAR time window, the overall NOMA allocation time at the BS, the UE processing delay plus RRC message transmission time and the contention resolution time window respectively. We define the overall average sojourn time for all states as the average step latency of the proposed scheme which is denoted as T_{ave}. The overall average sojourn time can be calculated using (4) shown below.

$$\begin{aligned} T_{\text{ave}} &= \sum_{i=0}^{L} \sum_{j} \pi_{i,j} T_{i,j} = \pi_{0,B} t_w + \pi_T t_T \\ &\quad + \sum_{j=1}^{3} \pi_j T_j + \pi_{CR} W_{RES}. \end{aligned} \tag{4}$$

Note that the time proportion τ can be expressed as the ratio of the preamble transmission time to the overall average sojourn time T_{ave}, which includes UEs that have successfully passed the barring stage and completed the BO process. Specifically, at the steady state, τ is given by $\tau = \dfrac{\pi_T t_T}{T_{\text{ave}}}$. Consequently, the average number of UEs performing preamble transmission at the steady state, denoted by $\bar{N}$, is

$$\bar{N} = \tau U_T \approx \left\lceil U_T \frac{\pi_T t_T}{T_{\text{ave}}} \right\rceil,$$

where U_T represents the total incoming UE traffic and $\lceil \cdot \rceil$ represents the ceiling operation. The transition probabilities of the states as well as the access probabilities for the NOMA blocks, can be derived using $\bar{N}$.

D. Derivation of Optimal NOMA Access Probabilities

1) Sequential NOMA Access Analysis

To obtain the state transition probabilities related to NOMA, we need to derive P_1, P_2, P_3 and P_4 through the 2-step NOMA-RACH procedure. Since NOMA accommodates both collided UEs and primary UEs, it is essential to analyze the outcomes of preamble selection and transmission. As described earlier, the average number of UEs performing preamble transmission at steady state is denoted by $\bar{N}$. After preamble selection and transmission, some preambles will be successfully detected, while others will collide. Let N_s denote the number of UEs with successfully detected preambles (successful UEs), and N_c the total number of UEs with collided preambles (collided UEs). Thus, we have $\bar{N} = N_s + N_c$.

Since each UE initiating access randomly selects one preamble from M available preambles, the probability that only one UE selects a specific preamble can be given as $\binom{\bar{N}}{1} \frac{1}{M} (1 - \frac{1}{M})^{\bar{N}-1}$. Hence, the expected number of the successful UEs is given by

$$N_s = \bar{N} \left(1 - \frac{1}{M}\right)^{\bar{N}-1}. \tag{5}$$

Among the UEs with successful preambles, some will also have successful payload decoding, while others will not. Let N_f represent the expected number of UEs with successful preamble detection but failed payload decoding, i.e., primary UEs, and let N_{s1} denote those with successful payload decoding. Thus, $N_s = N_f + N_{s_1}$. To calculate N_f, assuming p_f is the probability that the BS fails to decode the payload, we have $N_f = \lceil N_s p_f \rceil$.

By utilizing the aforementioned parameters, the success rate of NOMA access can be determined. Consider a $C \cdot L$ power-domain NOMA with C subchannels[5] and L power levels. We allocate N_f primary UEs to the upper levels of NOMA, while allowing N_c collided UEs to opportunistically access the remaining resources. We assume that there are sufficient radio resources to accommodate all primary UEs, i.e., $N_f \leq C \cdot L$. We illustrate this process with an example shown in Fig. 5. Primary UEs are sequentially assigned to the upper levels of the NOMA resources. After this assignment, the power levels may be fully or partially occupied. In scenarios where a power level is only partially occupied, we denote the two different numbers of occupied power levels as L_1 and L_2, where $L_1 - L_2 = 1$. The corresponding numbers of subchannels for L_1 and L_2 are represented by C_1 and C_2, respectively. Let L_s be the number of power levels where all NOMA blocks are occupied. For any given N_f, L_s is calculated as

$$L_s = \frac{N_f - N_f \mod C}{C},$$

where $\mod$ is the modulus operator. Subsequently, the quantities L_1, L_2, C_1 and C_2 can be expressed by:

[5]In this paper, a subchannel of NOMA refers to a time-frequency resource block, while a NOMA block refers to a specific power level within a subchannel.

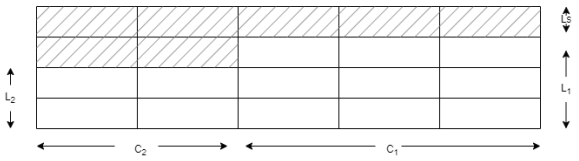


Fig. 5: An example of sequential NOMA block allocation.

$$\begin{aligned} L_1 &= L - L_s, \\ C_1 &= C - N_f \mod C, \\ L_2 &= L - L_s - 1, \\ C_2 &= N_f \mod C. \end{aligned} \tag{6}$$

After completing the sequential allocation of NOMA blocks, the remaining available subchannels and levels can be treated as two distinct regions, denoted as $R_1 = C_1 \cdot L_1$ and $R_2 = C_2 \cdot L_2$. The total number of remaining NOMA blocks is given by $K = C \cdot L - N_f = C_1 \cdot L_1 + C_2 \cdot L_2$. This scenario can be viewed as collided users accessing the two remaining regions, R_1 and R_2. If the collided users access the remaining NOMA blocks randomly, and the probability of selecting any power level within a subchannel is $\frac{1}{K}$, where $K = C \cdot L - N_f$, then the corresponding success probabilities for each region, P_{s_1} and P_{s_2}, as well as the overall success probability P_s, can be expressed as:

$$\begin{aligned} P_{s_1} &= C_1 \sum_{a=1}^{L_1} \tfrac{1}{K} \sum_{b=0}^{L_1-a} \left[\tbinom{N_c-1}{b} \tbinom{L_1-a}{b} b! \right. \\ &\quad \left. \times (\tfrac{1}{K})^b (1 - \tfrac{L_1-a+1}{K})^{N_c-b-1} \right], \\ P_{s_2} &= C_2 \sum_{a=1}^{L_2} \tfrac{1}{K} \sum_{b=0}^{L_2-a} \left[\tbinom{N_c-1}{b} \tbinom{L_2-a}{b} b! \right. \\ &\quad \left. \times (\tfrac{1}{K})^b (1 - \tfrac{L_2-a+1}{K})^{N_c-b-1} \right], \\ P_s &= P_{s_1} + P_{s_2}. \end{aligned} \tag{7}$$

Additional detailed explanations of the derivation steps are included in Appendix A. Consequently, the throughput of the collided UEs accessing the NOMA blocks can be expressed as $P_s N_c$.

2) Optimal NOMA Access Probabilities Analysis

In Subsection III-D1, we assume that collided UEs randomly access the remaining NOMA blocks. However, that is not the optimal approach. The reason is that the decoding process in power-domain NOMA will fail if multiple UEs select the same NOMA block. Moreover, due to the higher-power dominance issue [3], this failure affects all lower power levels within the same block. Therefore, an optimal approach is to use different access probabilities for different levels instead of using a common access probability across all levels. Let a_1 and a_2 represent the levels that can be chosen by UEs when selecting subchannels in C_1 and C_2, respectively. The probability of a UE selecting a level within a subchannel is denoted by l_{a_1} and l_{a_2}, where $\sum_{a_1=1}^{L_1} l_{a_1} = \sum_{a_2=1}^{L_2} l_{a_2} = 1$. The probability of a UE selecting a subchannel from C_1 or C_2 is $\frac{L_1}{K}$ or $\frac{L_2}{K}$, respectively. Consequently, the success probabilities P_{s_1}, P_{s_2}, and the overall success probability P_s are given by

$$\begin{aligned} P_{s_1} &= C_1 \sum_{a_1=1}^{L_1} l_{a_1} \tfrac{L_1}{K} \sum_{b=0}^{L_1-a_1} \left[\tbinom{N_c-1}{b} \tbinom{L_1-a_1}{b} b! \right. \\ &\quad \left. \times (1 - \tfrac{L_1}{K} + \sum_{j=1}^{a_1-1} \tfrac{l_{a_1} L_1}{K})^{N_c-b-1} \prod_{i=1}^{b} \tfrac{l_b L_1}{K} \right], \\ P_{s_2} &= C_2 \sum_{a_2=1}^{L_2} l_{a_2} \tfrac{L_2}{K} \sum_{b=0}^{L_2-a_2} \left[\tbinom{N_c-1}{b} \tbinom{L_2-a_2}{b} b! \right. \\ &\quad \left. \times (1 - \tfrac{L_2}{K} + \sum_{j=1}^{a_2-1} \tfrac{l_{a_2} L_2}{K})^{N_c-b-1} \prod_{i=1}^{b} \tfrac{l_b L_2}{K} \right], \\ P_s &= P_{s_1} + P_{s_2}. \end{aligned} \tag{8}$$

When level a_1 or a_2 in a subchannel is chosen, b represents the maximum number of UEs that can be successfully transmitted above the chosen level, where l_b is the probability of the UE choosing its corresponding level above the chosen level. The optimal access probabilities across power levels for each subchannel, l_{a_1} and l_{a_2}, can be found using an exhaustive search. Additional detailed explanations of the derivation steps for (8) are included in Appendix A.

E. Markov Chain Solution

After determining the optimal NOMA access probabilities, we can obtain P_1, P_2, P_3, P_4 by

$$\begin{aligned} P_1 &= \tfrac{N_{s_1}}{N}, P_2 = \tfrac{N_f}{N}, \\ P_3 &= \tfrac{N_c}{N}, P_4 = P_s. \end{aligned} \tag{9}$$

where P_4 represents the success probability of collided UEs accessing NOMA blocks, as determined by either (7) or (8), depending on whether uniform or optimal access probabilities are used.

Upon substituting the above probabilities into (2), the expressions in (2) then solely dependent on the steady-state distributions π. By leveraging the fact that the sum of all stationary probabilities equals one, all stationary distributions can be solved. Subsequently, the theoretical values for average step latency, NOMA access probabilities and success rate can be obtained. Furthermore, with all variables within the Markov Chain solved, the expected access time for a UE from barring to CR can be calculated through random walk. This access time is referred to as access latency.

IV. Algorithm Design

To address dynamic traffic load in real-world scenarios, an intelligent agent is designed to dynamically adjust the barring rate in response to fluctuating load

conditions. Additionally, it optimizes NOMA access probabilities to enhance RACH performance with aim to reduce collisions and user access latency. This agent, known as the DCMAB agent, fine-tunes the ACB rate and optimizes NOMA access probabilities, based on environment inputs and insights from our theoretical analysis. The entire process is illustrated in Fig. 6.

A. DCMAB Configuration

Before introducing our proposed DCMAB agent, we first describe contextual and non-contextual bandits. Both types involve making decisions from an action space denoted as $\mathcal{A}$, with a subsequent stochastic reward r revealed only for the chosen action. The actions in $\mathcal{A}$ are called arms, and the objective is to maximize the received rewards by identifying the optimal arm. This framework aligns with the need to adjust the barring and NOMA access rates in practical scenarios, where each barring rate in ACB represents an arm for exploring the optimal rate. Extensions of MAB, such as Contextual MAB, incorporate additional context to improve decision-making in complex scenarios. Generating contexts can be challenging in certain scenarios. To address this, we draw inspiration from neural networks (NN) used in Multi-Armed Bandit (MAB) schemes, as demonstrated in [32]. Specifically, we integrate an NN prediction model within the agent to forecast the context and broadcast NOMA access probabilities, which are then applied in the RA procedures.

To capture relevant information about the current environment, the main part of the context $\mathcal{C}$ includes the current load U_T and barring rate b as contextual information, with the flexibility to incorporate additional information as needed. The primary goal of NN prediction model is to estimate U_T using observable data. By leveraging the Markov chain model, we can derive the correlation between observable contents and U_T, facilitating the generation of targeted training data set 1. Since the BS does not know the exact number of collided UEs, the estimation of U_T is key to determining the optimal NOMA access probabilities. Using the estimated U_T, we can obtain N_c and N_f values to calculate the optimal NOMA access probabilities.

Finding the optimal NOMA access probabilities α_L[6] is complex, especially with an exhaustive search as outlined in (8). To utilize the NN model to broadcast the optimal NOMA access probabilities without real-time search, the training data set 2 needs to be pre-

[6]To simplify matters, let α_L denote the set comprising the probability of a collided user selecting a NOMA region, determined according to the region's weight, alongside the normalized optimal access probabilities l_{a_1} and l_{a_2}, which are determined based on the specific NOMA region being selected.

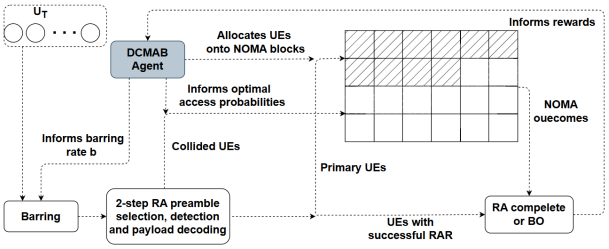


Fig. 6: The design of DCMAB agent.

searched for different U_T values. As a result, two sets of training data are employed to refine the NN prediction model. To mitigate the burden of excessive training data, we postulate a ceiling on the maximum number of serviceable incoming UEs, denoted as U_T^{max}. In the steady state, N_c and N_f are typically not excessively large, which permits us to impose a maximum limit on the number of serviceable collided UEs, thereby further reducing the scope of the training data. The range of N_f is set between 0 and the number of NOMA blocks, $C \times L$, while the range of collided UEs is determined based on the estimated U_T. This NN prediction model considers factors such as the size of NOMA ($C \times L$), the number of preambles (M), the current barring rate (b), the number of UEs completing the 2-step NOMA-RACH procedure (U_s), and the corresponding average step latency ($T = T_{\text{ave}}$). It predicts U_T as part of the context and indicates the optimal NOMA access probabilities, represented by $P(U_T, \alpha_L | b, T, U_s, M, C, L)$. In the rare event that incoming UEs exceed these limitations, the model can be retrained to ensure optimal performance.

B. Agent Design

We now elucidate the design of the DCMAB agent. The action space of the agent, denoted as $\mathcal{B}$, comprises a list of barring rates, represented by $\mathcal{B} = \{b_1, b_2, ..., b_n\}$. The number of users admitted to the RACH procedure and NOMA, specifically U_T, N_c, and N_f, is unknown to the agent and thus estimated using the aforementioned NN prediction model. Building upon the NN prediction model, Algorithm 1 is formulated to dynamically adjust the barring rate. It outlines the process for each barring round i of the DCMAB agent changing barring rate. In each round, the agent selects a barring rate b and executes a two-step handshake over a specific number of time slots. Since changes in the barring rate b do not immediately affect the reward, there is a delay of at least δt before the impact of the new b setting is reflected in the reward. Consequently, the agent's decision to change the barring rate is made over a period of δt or longer, with δt determined by simulation.

Once the reward for the previously selected barring

rate b is obtained, the agent estimates the current number of incoming UEs U_T^i as part of the context associated with that barring rate b^i in this round. Then, the agent broadcasts the optimal NOMA access probabilities α_L for the next round. The broadcast and estimation are achieved using the NN prediction model $P(U_T, \alpha_L | b, T, U_s, M, C, L)$. The data is stored in the agent's context matrix $\mathcal{C}$, with the context led by the estimated U_T^i. If the barring rates of an estimated U_T in context have not been fully explored, there is a probability ϵ of randomly selecting an unexplored arm. Otherwise, the agent selects the arm with the lowest average step latency[7] among the explored arms. If all barring rates of an estimated U_T in context are fully explored, the agent chooses the arm with the lowest average step latency. When all arms within a context are explored and the system estimates this context again, it is considered *seen*. For a seen context, the DCMAB employs the learned optimal arm for the specific number of UEs. Conversely, if an unseen context is estimated, the system initiates exploration. Hence, after sufficient exploration, when new traffic corresponds to a seen context, the agent can make the optimal choice directly. The computational complexity remains low, as there are no internal loops and the neural network model is pre-trained.

V. Simulation Results

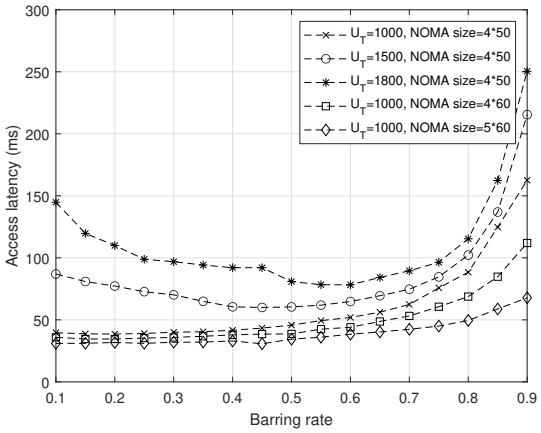


Fig. 7: Access latency of the proposed scheme.

In this section, we present simulation results to show the proposed 2-step NOMA-RACH performance and its performance advantages against other benchmark schemes. The simulation results are also used to validate our theoretical analysis. Furthermore, we evaluate the performance of the designed DCMAB algorithm, discussing its convergence and correctness. For the latency simulations, the following assumptions are made: preamble transmission time $t_T = 1$ ms, barring waiting time $t_w = 4$ ms. Moreover, the average BO time window is $\overline{t_{\text{BO}}} = 25$ ms, the BO time threshold is set to be 15, RAR window size $W_{\text{RAR}} = 4$ ms, and contention resolution time $W_{\text{RES}} = 5$ ms, unless otherwise indicated. The size of the NOMA blocks is $C \cdot L = 50 \cdot 4$ with overall process time $t_{\text{NOMA}} = 1$ ms. Additionally, we adopt a one-to-one

[7]Minimizing the average step latency, T_{ave}, and minimizing access latency, T_{a}, yield the same result, as demonstrated in Section V. Therefore, in this algorithm, we focus solely on T_{ave} to avoid duplication.

Algorithm 1 Barring Algorithm with DCMAB Agent

Input: M, C, L, $\mathcal{B} = \{b_1, b_2, ..., b_n\}$, δt, $\mathcal{C} = \emptyset$.

$i = 0$,$t_i = 0$, $b^i \leftarrow$ Random element from the set, $\alpha_L^i \leftarrow$ Random access probabilities.
while RACH is on **do**
  Barring round i begins, set barring rate $\leftarrow b^i$
  Set NOMA access probabilities $\leftarrow \alpha_L^i$
  Obtain U_s and T_{ave} at time $t_i + \delta t$.
  $U_T^i \leftarrow P(U_T | b^i, T_{\text{ave}}, U_s, M, C, L)$
  Set reward $R_i = -T_{\text{ave}}$
  if $[U_T^i, b^i] \in \{[U_T, b, R, \textit{seen}] | \mathcal{C}\}$ **then**
    Set $i \leftarrow i + 1$, Set $t_i \leftarrow$ Current time
    Set $b^i \leftarrow \underset{[U_T, b, R] \in \mathcal{C}, \forall U_T = U_T^i}{\arg\max} R$
    $\alpha_L^i \leftarrow P(\alpha_L | b^i, T_{\text{ave}}, U_s, M, C, L)$
    continue
  else
    Set $\mathcal{C} \leftarrow \mathcal{C} \bigcup \{[U_T^i, b^i, R]\}$,
    Set $\mathcal{B}' = \{b | [U_T, b, R] \in \mathcal{C}, \forall U_T = U_T^i\}$ to represent current explored arms of U_T^i.
  end if
  if $\mathcal{B} \subset \mathcal{B}'$ **then**
    Set $U = U_T^i$, Mark the context of U_T^i as *seen*:
    $\{[U_T, b, R] \leftarrow [U_T, b, R, \textit{seen}] | [U_T, b, R] \in \mathcal{C}, \forall U_T = U\}$
    Set $b \leftarrow \underset{[U_T, b, R] \in \mathcal{C}, \forall U_T = U_T^i}{\arg\max} R$
  else if $\mathcal{B} \not\subset \mathcal{B}'$ **then**
    if rand$(0, 1) \leq \epsilon$ **then**
      Set $b \leftarrow$ a random arm $\in \mathcal{B}$ and $\notin \mathcal{B}'$
    else
      Set $b \leftarrow \underset{[U_T, b, R] \in \mathcal{C}, \forall U_T = U_T^i}{\arg\max} R$
    end if
  end if
  set $i \leftarrow i + 1$, set $t_i \leftarrow$ Current time, set $b^i \leftarrow b$
  $\alpha_L^i \leftarrow P(\alpha_L | b^i, T_{\text{ave}}, U_s, M, C, L)$
end while

mapping strategy with low p_f for MsgA transmissions, tailored for NOMA and ACB configurations catering to a substantial number of UEs.

Recall that access latency is defined as the expected time for a UE to successfully complete the RA procedure, including the barring process and the handshake. Fig. 7 illustrates the access latency of the proposed 2-step NOMA-RACH scheme under varying barring rates and different values of U_T and NOMA sizes. It is observed that access latency increases as the number of incoming UEs grows or as the reduction of NOMA size occurs for a given barring rate. It can be confirmed from the figure that there exists an optimal barring rate that minimizes access latency. Additionally, the plot shows that the access latency in the proposed scheme decreases with the barring rate at first, reaches a minimum, and then increases within a certain range. The rationale behind this behavior is that a low barring rate increases the likelihood of UEs repeating the barring procedure, leading to longer waiting times. Conversely, a high barring rate results in collisions on most preambles, increasing the probability of transmission failure or backoff, ultimately increasing access latency. In comparison, Fig. 8 presents the access latency and the number of successful UEs of all cases (i.e., access throughput) for the traditional 2-step RACH without NOMA and the proposed scheme, with U_T set to 1000. The figure demonstrates that the proposed scheme consistently achieves lower access latency and higher successful UEs across all barring rates compared to the traditional approach. Notably, the maximum access throughput and minimum latency are achieved at the same optimal point. This improved performance is attributed to the incorporation of NOMA, which enhances the access success probability of UEs.

To validate the analysis presented in Section III, Fig. 9 depicts the average step latency for the proposed scheme, comparing numerical results obtained from (4) with random walk simulations for $U_T = 1500$. The plot demonstrates a close match between the theoretical and the simulation results. Furthermore, when compared to the access latency for $U_T = 1500$ in Fig. 7, it is evident that the average step latency across all states follows a similar trend, initially decreasing with the barring rate and then increasing after reaching its minimum value at the same optimal barring rate of $b = 0.45$.

To validate the analysis of NOMA access probabilities presented in Section III-D1, we compare the analytical results derived from expression (7) and Monte Carlo simulations. Fig. 10 illustrates the success rate for NOMA access in relation to the number of collided UEs. We present the results for a scenario involving a 4×10

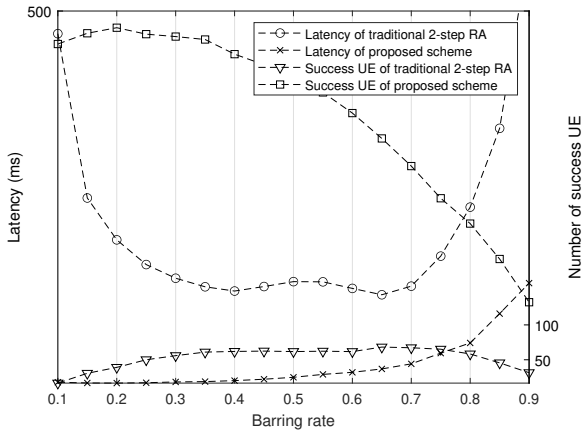


Fig. 8: Comparison of access latency between traditional 2-step RACH and the proposed scheme.

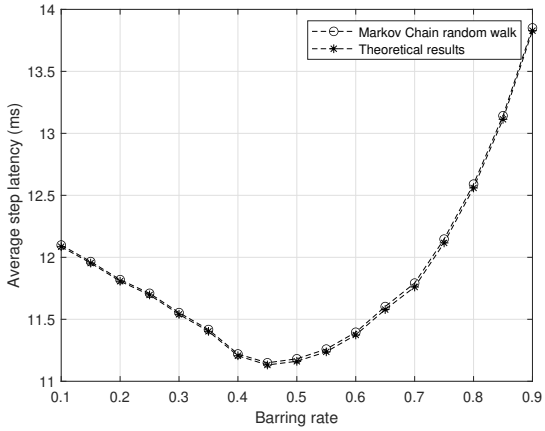


Fig. 9: Comparison of average step latency between analytical results and random walk simulations

NOMA block with 14 primary users occupying 2.5 levels, leaving 1.5 levels for the collided UEs to compete for access. The analytical results align perfectly with the Monte Carlo simulations in Fig. 10, thereby validating the accuracy of the derived expression (7). The plot also reveals that the success rate diminishes significantly as the number of collided UEs (N_c) increases, highlighting the adverse impact of a large number of collided UEs on the success rate, i.e., an excessive number of UEs will significantly reduce the success rate of NOMA access.

In Fig. 11, we study the throughput of collided UEs versus the load N_c when accessing the remaining NOMA blocks under different conditions with different values of primary users, N_f. The throughput quantifies the number of collided UEs that successfully complete NOMA-RA, and it is calculated using (7), with random

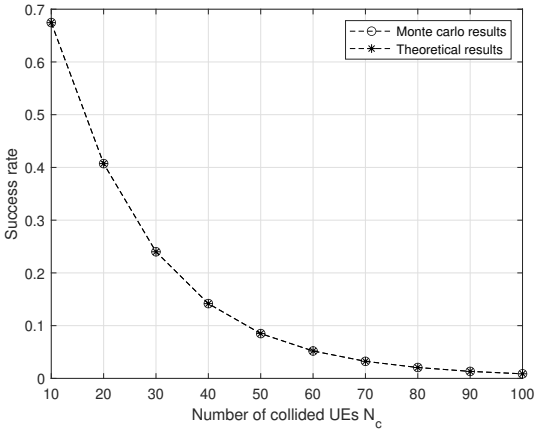


Fig. 10: Comparison of access success rate: analytical results vs. Monte-Carlo simulations.

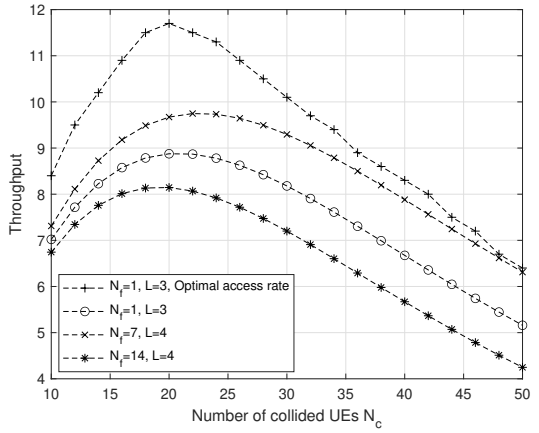
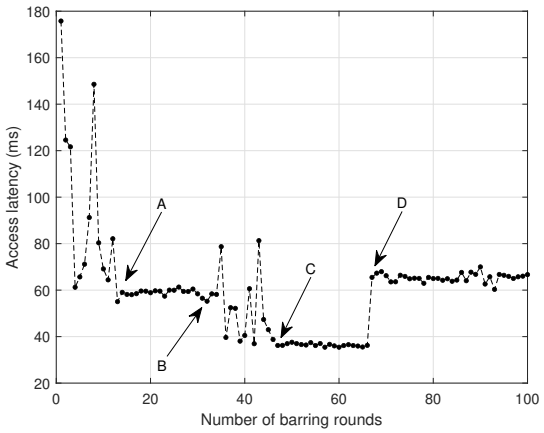


Fig. 11: Comparison of throughput between different scenarios of sequential NOMA.

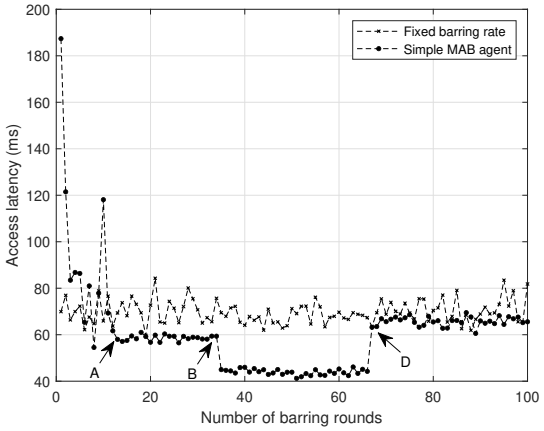
NOMA access probabilities applied, except for the case of optimal access rate. The figure shows that throughput initially increases with load but gradually decreases after reaching a peak value. This observation confirms the existence of a unique maximum throughput and optimal load. Furthermore, the plot highlights a reduction in throughput with an increase in the number of primary UEs N_f as throughput for $N_f = 14, L = 4$ is lower than that of $N_f = 7, L = 4$. This is because more NOMA blocks will be occupied by primary UEs, leaving fewer blocks available for collided UEs to access. More remaining NOMA blocks lead to higher throughput for collided UEs. This relationship between throughput, the number of primary UEs, and remaining NOMA blocks underscores the optimization potential to enhance NOMA performance. Furthermore, for the case where $N_f = 1, L = 3$, we see that the optimal throughput obtained by (8) significantly outperforms that without optimization. This observation signifies the importance of optimal access probabilities, which outperform random access probabilities and contribute to improved NOMA performance.

To assess the effectiveness of the DCMAB algorithm, Fig. 12(a) presents the access latency over a series of barring rounds, accounting for continuously changing numbers of UEs. In this evaluation, the number of UEs U_T decreases from 1500 to 1000 at point B and then recovers to 1500 at point D. The plot shows that when the agent encounters a new context, signifying a different number of UEs, the agent achieves optimal latency for that context after exploration. The exploration process, depicted in the region before point A and between points B and C in Fig. 12(a), demonstrates that the designed DCMAB algorithm converges within 10 to 20 barring rounds. After the exploration, as shown in the regions after point C and between points A and B, the agent successfully converges to a near-optimal access latency in a limited number of barring rounds. In the region after point D, we observe that since the context of $U_T = 1500$ has been fully explored at point A and recognized by the algorithm, no further exploration is required. Instead, the algorithm employs the optimal arm learned during previous barring rounds. It can be observed that the minimum access latency achieved by the algorithm at $U_T = 1000$ and $U_T = 1500$ matches the theoretical optimal values presented in Fig. 7, which is approximately 40 ms and 60 ms, respectively. The key advantage of the DCMAB agent lies in its ability to dynamically adjust the barring rate based on the current traffic conditions. In our simulations, the execution step mentioned in Section IV-B is set to $\delta t = 0.5s$. Reducing the barring round step sizes would increase precision but slow down the convergence speed.

To further evaluate the performance of the 2-step NOMA RACH with the DCMAB agent, Fig. 12(b) plots its access latency over 100 barring rounds and compares it with two benchmark schemes for 2-step NOMA RACH: 1) a simple MAB agent without load estimation, i.e., Algorithm 1 without the NN prediction model and context, relying solely on the MAB scheme to explore optimal barring rate, and 2) a fixed barring rate scheme without any agent where collided UEs randomly access NOMA. For the MAB agent scheme, it is assumed that the agent does not possess a load estimation function and only has a memory for $U_T = 1500$ once converged. For the fixed barring rate scheme, collided UEs access NOMA at a random rate, with b set to be the optimal



(a) Access latency performance for DCMAB agent



(b) Access latency performance for benchmark schemes.

Fig. 12: Comparison between DCMAB agent and benchmark schemes.

barring rate for $U_T = 1500$. The primary difference between these schemes becomes apparent when U_T changes. Comparing the results between points B and D in Figs. 12(b) and 12(a), it is evident that the DCMAB agent achieves lower latency than the MAB scheme after exploration when the UE traffic changes. This is because, unlike the DCMAB approach, the MAB scheme does not dynamically adjust barring rate. Since the MAB scheme also implements optimal NOMA access probabilities, it performs similarly to DCMAB before the change in traffic conditions. However, the fixed barring rate scheme did not have the ability to adjust NOMA access probabilities, and therefore its performance drifted below that of the MAB scheme. In conclusion, as DCMAB is able to detect changes in traffic and adjust both the barring rate and NOMA access probabilities, it is able to maintain optimal performance once converged.

VI. Conclusions

This research focused on exploring the potential and advantages of integrating the innovative 2-step NOMA-RACH scheme into future communication networks, particularly in the scenario of mMTC. The proposed scheme introduced a power dimension, enabling multiple UEs to utilize the resource, resulting in improved NOMA access success rates and reduced latency. ACB mechanism was also proposed to address the inherent challenge of traffic burstiness in mMTC. The investigation involved theoretical probability derivation, Markov chain analysis, and extensive comparative studies, revealing substantial improvements in RACH performance. To cope with the practical scenarios with dynamically changing traffic load, a DCMAB agent with load estimation is designed to jointly optimize NOMA access probabilities and barring rates. When the DCMAB agent was used, not only the superiority of the agent over other benchmark schemes was demonstrated, its dynamic ability to tune the barring rate and facilitate optimal user participation in the RACH procedure amidst ever-changing environmental conditions was also confirmed. The proposed scheme, emerges as a promising solution for mMTC scenarios in future communication networks.

Appendix A

To derive the optimal power level access probabilities for each subchannel, we first consider NOMA blocks of size $C \times L$ with C subchannels and L power levels. Let N_c users randomly access this block. The goal is to derive the success rate for NOMA access. Assuming that for any subchannel in NOMA blocks, the probability of a UE choosing a certain block is $\frac{1}{K}$. For sequential NOMA with random access, the success rate depends on two conditions: only the current UE chooses this block, or the users in the upper levels of the block are successfully decoded.

If level a is selected, the remaining upper level $L - a$ should be successfully decoded, where $L = L_1$ or L_2. Assuming that there are b UEs in the upper levels that have been successfully decoded, the probability of condition 2) can be expressed as:

$$\sum_{b=0}^{L-a} \binom{L-a}{b} b! \frac{1}{K^b} \tag{10}$$

In the sequential NOMA system depicted in Fig. 5, the available remaining blocks are given by $K = C \times L - N_f = C_1 \times L_1 + C_2 \times L_2$. If a UE selects level a, and the b UEs choose upper levels intending to be successfully

transmitted, the remaining $N_c - 1 - b$ UEs must select other subchannels or levels below level a in the current subchannel. The corresponding probability for this scenario can be expressed as $\frac{1}{K^b}\left(1 - \frac{L-a+1}{K}\right)^{N_c-b-1}$.

In this expression, b needs to be selected from the pool of $N_c - 1$ users and arranged in the $L - a$ remaining levels. Since a denotes the specific level to choose, ranging from 1 to L, the comprehensive success rate is expressed as:

$$C\sum_{a=1}^{L}\frac{1}{K}\sum_{b=0}^{L-a}\left[\binom{N_c-1}{b}\binom{L-a}{b}b!\left(\frac{1}{K}\right)^b \times\left(1-\frac{L-a+1}{K}\right)^{N_c-b-1}\right] \tag{11}$$

Considering the fact that there are different parts of one NOMA block with sizes of $C_1 \times L_1$ and $C_2 \times L_2$, the final success rate is represented in (7).

For (8), the probability of a UE choosing level $a = a_1$ or a_2 in a subchannel and b UEs choosing $L - a$ levels is not $\frac{1}{K}$ because the access probabilities of each level is not random, where $L = L_1$ or L_2. The probability of b UEs choosing the remaining levels $L - a$ and the remaining $N_c - b - 1$ UEs choosing other levels or subchannels becomes:

$$\left(1-\frac{L}{K}+\sum_{j=1}^{a-1}\frac{l_a L}{K}\right)^{N_c-b-1}\prod_{i=1}^{b}\frac{l_b L}{K} \tag{12}$$

where l_a represents the probability of the current UE choosing level $a = a_1$ or a_2 in this subchannel, and l_b represents the access probability when b users choose the corresponding remaining level in this subchannel. Finally, the overall success rate is

$$C\sum_{a=1}^{L}l_a\frac{L}{K}\sum_{b=0}^{L-a}\left[\binom{N_c-1}{b}\binom{L-a}{b}b! \times(1-\frac{L}{K}+\sum_{j=1}^{a-1}\frac{l_a L}{K})^{N_c-b-1}\prod_{i=1}^{b}\frac{l_b L}{K}\right] \tag{13}$$

and the success rate of different parts of one NOMA block with sizes of $C_1 \times L_1$ and $C_2 \times L_2$ is represented by (8). The optimal access probabilities can then be determined by performing numerical search over l_a.

REFERENCES

[1] S. Park, S. Lee, and W. Choi, "Markov chain analysis for compressed sensing based random access in cellular systems," in *2019 International Conference on Computing, Networking and Communications (ICNC)*, 2019, pp. 34–38.

[2] T. Miyazawa, K. Ishizu, H. Asaeda, H. Tsuji, and H. Harai, "Energy-efficient power management for o-ran base stations utilizing pedestrian flow analytics and non-terrestrial networks," *IEICE Transactions on Communications*, pp. 1–16, 2024.

[3] W. Yu, C. H. Foh, A. U. Quddus, Y. Liu, and R. Tafazolli, "Throughput analysis and user barring design for uplink noma-enabled random access," *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6298–6314, 2021.

[4] 3GPP TS 36.211, "Evolved universal terrestrial radio access (E-UTRA) physical channels and modulation," 3GPP, Tech. Rep., Release 13, 2016.

[5] 3GPP, "Nr; NR and NG-RAN overall description; stage 2," 3GPP, Technical Specification 3GPP TS 38.300, March 2021.

[6] 3GPP RP-190711, "3GPP work item description, 2-step rach for NR," 3GPP, Tech. Rep., Sep. 2019.

[7] 3GPP TS 38.331, "NR; radio resource control (RRC); protocol specifi- cation," 3GPP, Tech. Rep., ver. 15.8.0, Jan. 2020.

[8] Z. Li, L. Tian, Y. Yin, and W. Cao, "On contention-based 2-step random access procedure," in *2020 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2020, pp. 771–776.

[9] J. Kim, S. Kim, T. Taleb, and S. Choi, "Rapid: Contention resolution based random access using context id for iot," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 7, pp. 7121–7135, 2019.

[10] J. Kim, G. Lee, S. Kim, T. Taleb, S. Choi, and S. Bahk, "Two-step random access for 5g system: Latest trends and challenges," *IEEE Network*, vol. 35, no. 1, pp. 273–279, 2021.

[11] E. Peralta, T. Levanen, F. Frederiksen, and M. Valkama, "Two-step random access in 5g new radio: Channel structure design and performance," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 2021, pp. 1–7.

[12] M. Grau, C. H. Foh, A. u. Quddus, and R. Tafazolli, "Preamble barring: A novel random access scheme for machine type communications with unpredictable traffic bursts," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, 2019, pp. 1–7.

[13] 3GPP TS 22.011, "Technical specification group services and system aspects," 3GPP, Tech. Rep., Release 13, 2016.

[14] L. Zhao, X. Xu, K. Zhu, S. Han, and X. Tao, "Qos-based dynamic allocation and adaptive acb mechanism for ran overload avoidance in mtc," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[15] 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 *IEEE Int. Conf. Commun (ICC)*, 2018, pp. 1–7.

[16] M. Jihun and L. Yujin, "A reinforcement learning approach to access management in wireless cellular networks," *Wireless Communications & Mobile Computing*, vol. 2017, pp. 1–7, 2017.

[17] J. Choi, "Two-sided learning for noma-based random access in iot networks," *IEEE Access*, vol. 9, pp. 66 208–66 217, 2021.

[18] W. Wang, W. Yu, C. H. Foh, D. Gao, and Q. Ni, "User scheduling in noma random access using contextual multi-armed bandits," in *IEEE Globecom Workshops*, 2022, pp. 112–117.

[19] Y. Piao and T.-J. Lee, "Integrated 2–4 step random access for heterogeneous and massive iot devices," *IEEE Transactions on Green Communications and Networking*, vol. 8, no. 1, pp. 441–452, 2024.

[20] S. Song, H. Xie, and H. Jin, "Optimizing random access procedure to support massive and delay-critical mtcs," in *2023 12th International Conference on Awareness Science and Technology (iCAST)*, 2023, pp. 169–172.

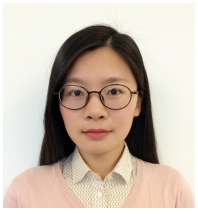
[21] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6g: Vision, enabling technologies, and applications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 1, pp. 5–36, 2022.

[22] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5g and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 615–637, 2021.

[23] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Non-orthogonal multiple access for 5G and beyond,"

Proceedings of the IEEE, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.
[24] W. Yu, L. Musavian, and Q. Ni, "Link-layer capacity of NOMA under statistical delay QoS guarantees," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4907–4922, Oct. 2018.
[25] K. Cao, H. Ding, B. Wang, L. Lv, J. Tian, Q. Wei, and F. Gong, "Enhancing physical-layer security for iot with nonorthogonal multiple access assisted semi-grant-free transmission," *IEEE Internet of Things Journal*, vol. 9, no. 24, pp. 24 669–24 681, 2022.
[26] K. Cao, B. Wang, H. Ding, L. Lv, J. Tian, H. Hu, and F. Gong, "Achieving reliable and secure communications in wireless-powered noma systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1978–1983, 2021.
[27] E. Balevi, F. T. A. Rabee, and R. D. Gitlin, "ALOHA-NOMA for massive machine-to-machine IoT communication," in *IEEE Int. Conf. Commun. (ICC)*, May 2018.
[28] S. A. Tegos, P. D. Diamantoulakis, A. S. Lioumpas, P. G. Sarigannidis, and G. K. Karagiannidis, "Slotted ALOHA with NOMA for the next generation IoT," *IEEE Trans. Commun.*, Jul. in press, 2020.
[29] Z. Chen, Y. Liu, S. Khairy, L. X. Cai, Y. Cheng, and R. Zhang, "Optimizing non-orthogonal multiple access in random access networks," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Antwerp, Belgium, Belgium, May 2020.
[30] Y. Wang, T. Wang, Z. Yang, D. Wang, and J. Cheng, "Throughput-oriented non-orthogonal random access scheme for massive MTC networks," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1777–1793, Mar. 2020.
[31] 3rd Generation Partnership Project, "3GPP TS 38.213 V16.6.0 (2021-06) Technical Specification," 3rd Generation Partnership Project; Technical Specification Group Radio Access Network, Technical Specification 38.213.
[32] M. Collier and H. U. Llorens, "Deep contextual multi-armed bandits," 2018.

Dawei Nie is a Research Associate at the School of Computing and Communications, Lancaster University, UK. He earned his Ph.D. in Communication Systems from Lancaster University in 2024, an M.Sc. in Wireless and Optical Communication from University College London (UCL) in 2018, and a B.Sc. in Electrical and Electronic Engineering from the University of Liverpool in 2017. His research focuses on cognitive radio networks, random access, edge-fog computing, and resource allocation in wireless communications.



Wenjuan Yu (Member, IEEE) received her Ph.D. degree in Communication Systems from Lancaster University, U.K. She is currently a Lecturer at the School of Computing and Communications, Lancaster University. Previously, she was a Research Fellow at the 5G Innovation Centre (5GIC), Institute for Communication Systems, University of Surrey, U.K., and a part-time Research Officer at the School of Computer Science and Electronic Engineering, University of Essex, U.K. Her research interests include radio resource management, low-latency communications, and machine learning for communications.

Qiang Ni (Senior Member, IEEE) is a Professor and the Head of the Communication Systems Group, School of Computing and Communications, Lancaster University, Lancaster, UK. He received the Ph.D. degrees in engineering from Huazhong University of Science and Technology, China, in 1999. His research interests include green communications and networking, millimeter-wave wireless, cognitive radio networks, 5G and 6G, SDN, IoT, cyber physical systems, AI and machine learning. He has authored or co-authored 300+ papers in these areas.

Chuan Heng Foh (Senior Member, IEEE) received the M.Sc. degree from Monash University, Melbourne, VIC, Australia, in 1999, and the Ph.D. degree from the University of Melbourne, Melbourne, in 2002. After the Ph.D., he spent six months as a Lecturer with Monash University. In December 2002, he joined Nanyang Technological University, Singapore, as an Assistant Professor until 2012, and he is currently with University of Surrey, U.K. as an Associate Professor. He has authored or coauthored more than 200 refereed papers in international journals and conferences. His research interests include protocol design, machine learning and AI applications, and performance analysis of advanced networks. He served as a Vice Chair (Europe/Africa) for IEEE TCGCC from 2015 to 2017. He is currently the Vice Chair of the IEEE VTS Ad Hoc Committee on Mission Critical Communications. He is on the editorial boards of several international journals and a Senior Editor for IEEE Access.