WILEY | Hindawi

Research Article

Resource Scheduling in URLLC and eMBB Coexistence Based on Dynamic Selection Numerology

Lei Wang ,[1,2] Sijie Tao,[1] Lindong Zhao,[1,2] Dengyou Zhou,[1] Zhe Liu,[1] and Yanbing Sun [3]

[1]School of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]Key Laboratory of Broadband Wireless Communication and Sensor Network Technology Ministry of Education, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[3]School of Foreign Studies, Jiangsu Open University, Nanjing 210036, China

Correspondence should be addressed to Yanbing Sun; sunyb@jsou.edu.cn

This paper focuses on the resource allocation problem of multiplexing two different service scenarios, enhanced mobile broadband (eMBB) and ultrareliable low latency (URLLC) in 5G New Radio, based on dynamic numerology structure, mini-time slot scheduling, and puncturing to achieve optimal resource allocation. To obtain the optimal channel resource allocation under URLLC user constraints, this paper establishes a relevant channel model divided into two convex optimization problems: (a) eMBB resource allocation and (b) URLLC scheduling. We also determine the numerology values at the beginning of each time slot with the help of deep reinforcement learning to achieve flexible resource scheduling. The proposed algorithm is verified in simulation software, and the simulation results show that the dynamic selection of numerologies proposed in this paper can better improve the data transmission rate of eMBB users and reduce the latency of URLLC services compared with the fixed numerology scheme for the same URLLC packet arrival, while the reasonable resource allocation ensures the reliability of URLLC and eMBB communication.

1. Introduction

The exponential growth in quantity and quality of mobile services has led to a surge in demand for high data rates, high-reliability, power effectiveness, low latency, etc. To meet the rapidly growing demand for future mobile Internet services and bring new business experiences to users, 5G is divided into the following three main application scenarios based on different quality of service (QoS) requirements: enhanced mobile broadband (eMBB), ultrareliable low-latency communications (URLLC), and massive machine type communications [1].

The required QoS will vary depending on the different scenarios. eMBB is mainly used for services with larger data volumes, such as 4K video, virtual reality, and augmented reality, which require higher data transmission rates, while meeting certain reliability requirements. The specific application scenarios of URLLC, such as autonomous driving and telemedicine, require low latency and high-reliability, which

means transmission latency should be short. The former is the main service of 5G because of the large number of user accesses and large traffic volume; the latter service has less data volume with sporadic and burst characteristics, requiring timely responses. Therefore, when the above scenarios exist simultaneously in a system, conflicts will inevitably arise, and it is necessary to study how to meet the different QoS requirements at the same time, aiming to achieve optimal system performance.

To solve the above problems, the easiest way is to reserve resources to ensure that the URLLC can be transmitted at the first time whenever it arrives to meet the low latency and high-reliability requirements of URLLC transmission. However, it will cause a great waste of resources when there is no URLLC transmission, so overlay multiplexing and puncture punching are proposed later. Overlay multiplexing enables eMBB and URLLC users to share the same spectrum resources, improving the spectrum efficiency of the system, ensuring the
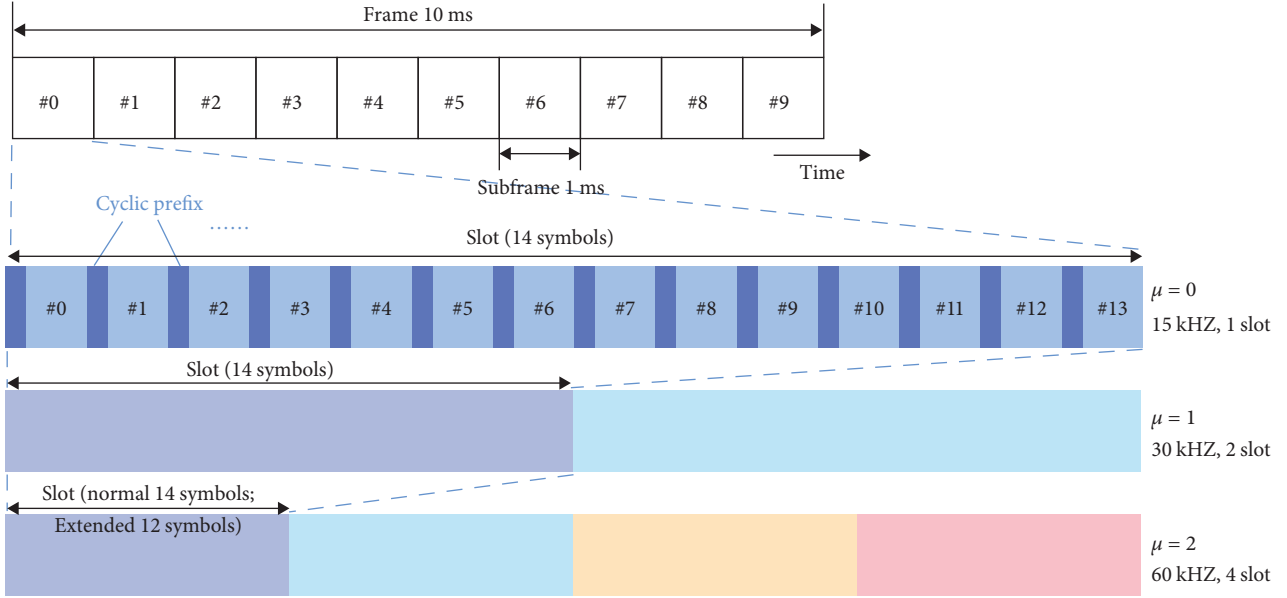
FIGURE 1: 5G NR basic structure.

throughput of eMBB users, and reducing the latency of URLLC users. However, interference occurs when the two are overlapped and multiplexed, which can be decoded by continuous interference cancelation technology, but it still has an impact on the reliability of URLLC users, so attention should be paid to the balance of ensuring throughput and meeting the requirements of high-reliability of URLLC. As for puncture punching, the base station (BS) first allocates all the free resources to the eMBB service and then perforates the resources already allocated to the eMBB service when the next time slot comes, transmitting the arriving URLLC service so as to ensure the delay of transmission and scheduling. The punctures will satisfy the reliability and low latency of URLLC users both but will also lead to the loss of throughput of eMBB.

Therefore, to cope with various deployment scenarios, 5G New Radio (NR) defines techniques such as numerology architecture, mini-time slot transmission, and punctured scheduling. Specifically, NR supports a wide range of carrier frequencies (within the two possible ranges, FR1 and FR2) and channel bandwidths. To achieve this flexibility, 3GPP proposes a flexible framework structure, multinumerology, where different values of the numerology have different subcarrier spacing (SCS) and orthogonal frequency division multiplexing (OFDM) symbol lengths [2]. Therefore, resource block (RB) structures are different but always include 12 subcarriers and 14 OFDM symbols, as shown in Figure 1. Based on multiple numerologies, users can flexibly schedule the resource distribution by choosing the SCS and OFDM symbol lengths that satisfy different QoS. Therefore, different transmission time intervals can be achieved with the help of mini-time slot transmission and multiple numerologies, while puncture punching scheduling ensures orthogonality of resource reuse and reduces interference.

3GPP introduces a new time interval in 5G NR, while inheriting the frame length (10 ms) and subframe length (1 ms) from the original 4G LTE. Considering the high-reliability and low

delay characteristics of URLLC services, the traditional Shannon formula is no longer applicable to URLLC traffic, so URLLC traffic needs to be sent in the form of finite code length transmission.

In this paper, we consider the resource multiplexing problem of puncturing based on eMBB and URLLC services in the FR1 band. First, we establish the basic channel model to maximize the system throughput, which is defined as an NP-hard problem. Then, we decompose this problem and use DRL to dynamically select the optimal numerology value at the beginning of each time slot. Later, the simulation analysis finds that the dynamic selection of numerology value can improve the performance of eMBB with guaranteed URLLC delay and reasonably allocate the system resources and power.

To solve the existing problems, this paper combines the advantages of the above methods to study the optimal resource allocation problem based on mixed numerologies, mini-time slot transmission, and puncture punching scheduling in multiplexing scenarios. In this paper, we propose to achieve low latency communication by mixed numerologies and mini-time slots, combined with DRL to select numerology fetching values at the beginning of the time slot to meet the different requirements of different application scenarios. The main contributions of this paper are shown below:

(1) Transforming the multiplexed resource allocation problem into an optimization problem, which is solved as a mixed integer linear problem, and the original problem is decomposed into two convex optimization subproblems, eMBB resource allocation, and URLLC scheduling, reducing the calculation complexity and maximizing the data transmission rate of eMBB users under the constraints of URLLC.

(2) Mixed numerologies selection method. The numerology values are determined at the beginning of the time slot, and a DQN-based dynamic selection numerology

algorithm is proposed to achieve a flexible time slot scheduling strategy, meeting the QoS requirements in eMBB and URLLC coexistence scenarios. The optimal time–frequency resource allocation in multiuser scenarios using mixed numerologies and mini-time slot methods is investigated. The numerology selection and resource allocation optimization methods under the mini-time slot system are proposed, and all the necessary constraints concerning packet size and delay are analyzed in detail.

(3) Compared to traditional fixed timeslot scheduling, flexible timeslot selection offers significant advantages in terms of URLLC service performance. It can better reduce interference to eMBB users and significantly decrease latency for URLLC users.

2. Related Works

Constructing a framework for efficient resource allocation with limited resources is the optimal solution to the resource multiplexing problem for eMBB and URLLC in a model-based approach, making many unknown quantities as known conditions, solving the resource slicing problem based on optimization ideas to maximize system performance under the constraints of URLLC (performance targets can be data transfer rates, throughput, etc. for eMBB). Manzoor et al. [3] investigated the possibility of scheduling URLLC in existing eMBB services and applied one-to-one matching game theory to find eMBB-URLLC pairs that can be stabilized in the same spectrum by representing the optimization problem of coexistence through dynamic overlay/punching, and also designed a contract theoretical framework for overlay reuse. In [4], eMBB transmissions are allocated on time slots. The authors studied the joint scheduler under such systems to reduce the delay on mini-time slots preemptively overlap/puncture URLLC and derived the optimal joint scheduler for the linear rate loss model and convex loss model, respectively. Alsenwi et al. [5] proposed a risk-sensitive formulation based on introducing the conditional value at risk as a risk measurement for eMBB propagation, expressing the reliability constraint of URLLC as an opportunity constraint, and relaxing it based on Markov inequalities to decompose the problem into two convex form subproblems solved by eMBB user scheduling and URLLC user scheduling.

Mathematical model-based methods are generally difficult to obtain more accurate optimal solutions, while with the advancement of machine learning, many studies focus on the application of DRL in communication to assist or even lead in making resource allocation decisions. In order to simplify the computational complexity and obtain the exact optimal solution, this paper turns to deep reinforcement learning (DRL) that does not require priori information on dynamic changes in URLLC data traffic and channel gain changes, etc. Li et al. [6] jointly optimized the offloading decision and computational resource allocation with the goal of minimizing the total cost of the considered mobile edge

computing (MEC) system and proposed Q-learning-based and DRL-based optimization frameworks to solve the resource allocation problem in wireless MEC, respectively. Considering the proportional fairness among multiple users, Yin et al. [7] proposed a new URLLC downlink scheduler to formulate the resource allocation problem in each small slot as an integer programing problem, analyzed the performance of two algorithms, concave-convex process and greedy algorithm in practice, and gave a preference on the algorithm selection. Yin et al. [7] decomposed the problem into two subproblems, eMBB and URLLC, and proposed the penalized successive upper bound minimization algorithm and the optimal transmission model to solve the resource scheduling problem for eMBB and URLLC users on time slots and mini slots, respectively. Bairagi et al. [8] proposed to solve the eMBB and URLLC resource allocation problems with the help of a heuristic algorithm and the one-sided matching game, respectively. Bairagi et al. [9] proposed a DRL-based optimization-assisted approach for resource slicing in eMBB and URLLC intelligent punching scheduling. Alsenwi et al. [10] and Luong et al. [11] explored communication problems, including but not limited to DRL, to solve dynamic network access, data rate control, etc. It is found that solving communication resource allocation problems with DRL can arbitrarily and effectively reduce the average time spent and improve the effectiveness and agility of slicing, and the latter is further proposed to have great advantages compared to traditional computing [12].

In addition, most of the above studies consider only the case of fixed numerology in the slicing scheme, which means numerologies do not change with time. In [13], an analytical system model was developed for internumerology interference (InterNI) analysis in OFDM systems considering mixed numerologies, and a power allocation study was performed using the derived analytical method to give a resolution of the optimization problem of sum-rate maximization. Mao et al. [14] developed metrics for quantifying the INI level using a continuous-time approach and directly revealed the main influencing factors of INI through the INI analytical expressions in terms of the mean square error and error vector magnitude. Mcwade et al. [15] proposed a two-stage suboptimization approach to deal with the problem of optimizing subcarriers and power allocation to maximize the spectral efficiency of multinumerology-nonorthogonal multiple access. Thus, we start to investigate resource allocation between different numerologies [16–21].

The former solves the multiplexing problem with the help of optimization ideas, mostly under the limitations of URLLC (e.g., delay and reliability) to solve how to maximize a certain performance of eMBB, such as maximizing eMBB user throughput, data transmission rate, etc., and obtains the approximate optimal solution through range expansion and contraction; the latter is represented by DRL, which gives a more accurate optimal solution for allocation. The eMBB-URLLC coexistence problem of existing work mostly uses independent optimization, which has certain limitations and cannot capture well to deal with dynamic changes.

TABLE 1: 5G NR numerologies.

Numerology	Cyclic prefix	SCS (kHZ)	Slot time (ms)	Number of slots per subframe	Number of OFDM symbols per slot
0	Normal	15	1	1	14
1	Normal	30	0.5	2	14
2	Normal	60	0.25	4	14

3. Preliminaries

To address the coexistence of URLLC and eMBB services on the same radio resource, this paper integrates the techniques of numerology, time slot-based transmission, and punctured scheduling to study the resource multiplexing slicing problem in OFDM-based downlink eMBB and URLLC. In this section, we introduce the basics used in this paper.

3.1. Basic Structure of 5G NR. Specifically, numerology is to select different values of SCS, time slot duration, symbol duration, etc., by means of parameter sets. In the time domain, each subframe can be divided into 2^{μ} time slots, each time slot has 14 OFDM symbols, the corresponding SCS is $15 \times 2^{\mu}$ kHZ. In other words, as the value increases, each subframe contains 1/2/4 time slots, with the duration of each time slot being 1/0.5/0.25 ms, respectively, during which SCS is set to 15/30/60 kHz correspondingly. When the SCS is 60 kHZ, there is an extended CP, which is about four times the normal CP time and is suitable for networks with long delays, which is not considered in this paper. Meanwhile, NR can dynamically adjust the corresponding time slot ratio in the frame structure according to the ratio of upstream and downstream services. Some of the 5G NR numerology structures are shown in Table 1.

In addition, 5G NR supports mini-time slot transmission, which can include only 1 OFDM symbol according to Release 15, but mini-time slots are generally limited to 2, 4, and 7 OFDM symbols. Due to the high requirements of URLLC transmission delay, it is not possible to defer URLLC traffic to the start of the next timeslot, so with the help of mini time-slots, it can achieve lower delay and reduce waiting time.

3.2. Machine Learning. Machine learning can be classified as supervised learning, unsupervised learning, and reinforcement learning (RL). RL is mainly used by the agent to find the corresponding training samples by interacting with the environment, and train the machine learning model in the nonstop interaction and exploration with the environment. RL algorithms can be used in scenarios such as resource allocation, user connectivity, and channel selection. In this paper, we try to solve the eMBB and URLLC user resource reuse problem with the help of DRL and consider the dynamic selection numerology algorithm based on DNQ.

The basic model of reinforcement is shown in Figure 2, where the agent first observes its current state and then takes action, the environment responds to it, and the agent is rewarded with its new state. The observed information is used to adjust the agent's policy, and the process is repeated until the agent's policy is close to the optimal one. Also, RL can only be performed based on each experience, mainly because the external
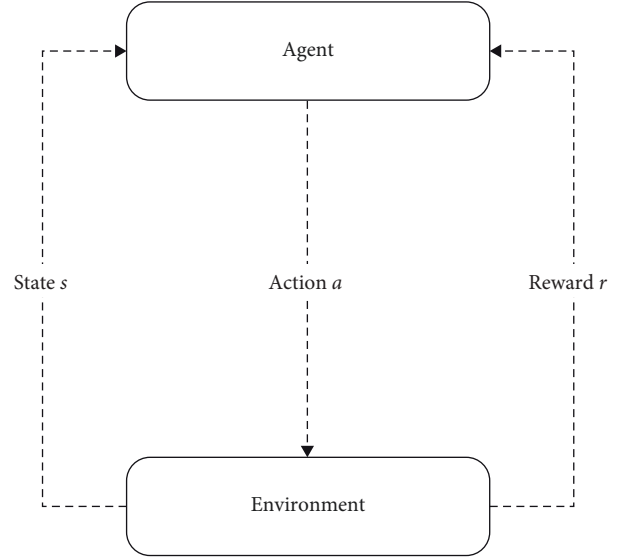

FIGURE 2: Basic model of reinforcement.

environment generates less information, so the only way to adapt the strategy to the environment is by adjusting it through its own experience.

The RL model focuses on the environmental state S, the action A and the reward R. γ is the discount factor. The agent takes action a_t, gets reward r, updates action a_{t+1}, and continuously optimizes its strategy until it reaches the optimal strategy. The value function is shown as follows:

$$V^{\pi}(s) = E[R(s,a)] + \gamma \sum_{s' \in S} P_{ss'}(a) V^{\pi}(s'), \quad (1)$$

where $P_{ss'}(a)$ denotes the state transfer probability from state s to s' under action a. The objective is to find the optimal strategy π^* to make the value function reach the maximum; thus, the expected feedback Q value is defined as follows:

$$Q^{\pi}(s, a_t) = E[R(s,a)] + \gamma \sum_{s' \in S} P_{ss'}(a) Q^{\pi}(s'). \quad (2)$$

Let $V^*(s)$ and $Q^*(s,a)$ under the optimal strategy be $V^*(s) = \max_a Q^*(s,a)$ and $Q^*(s,a) = Q^{\pi*}(s,a) \forall s, a$, respectively, thus the update criterion of Q-value is as follows:

$$Q_{t+1}(s,a) = (1-\alpha)Q_t(s,a) + \alpha \left[R_t(s,a) + \gamma \max_{a \in A} Q_t(s_{t+1}, a_{t+1}) \right], \quad (3)$$

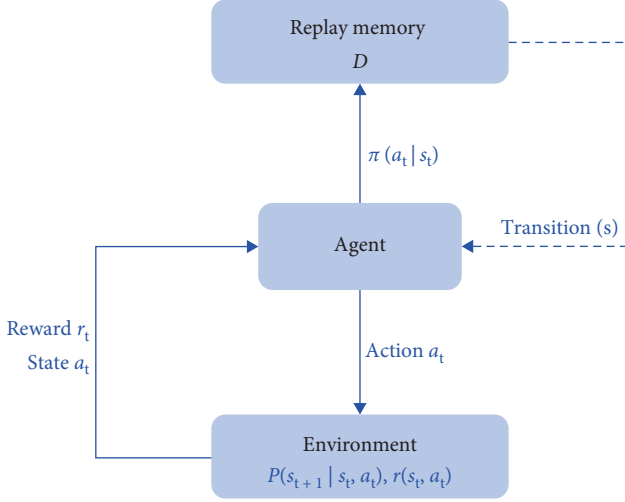where α denotes the learning rate, which measures the speed of convergence of the Q-learning algorithm.

FIGURE 3: DNQ flowchart.

The Q-learning algorithm is to directly select the optimal Q-function, such that the updated Q-function is about the strategy π, thus leading to problems such as overdependence of the objective function on parameters and overcorrelation of sample data. To address this drawback, DQN is proposed to resolve two problems of Q-learning. DQN first freezes the target network so that the parameters of the objective function can be fixed at a time, and thus, the target can be determined. The second is to use experience replay to remove the correlation of sample data by the construction of an experience pool, and this experience pool is composed of the experience of the intelligence, while the samples are randomly selected from the experience pool to replace the current samples for training during the training process, breaking the correlation between samples and also avoiding the local optimum. The specific flow chart is shown in Figure 3.

4. Problem Equation Establishment

Consider two kinds of service downlink requests, eMBB and URLLC, with several eMBB and URLLC users under a BS (gNB), and assume that the downlink is always carrying out eMBB data transmission so that there are U_e eMBB users and U_u URLLC users under a gNB, and the total number of RBs is B, an RB $b \in \{1, 2, \ldots, B\}$. The gNB selects the numerology $\mu \in \{0, 1, 2\}$ by learning the comprehensive performance of the system under different numerologies, with the purpose of selecting the appropriate SCS for each user and updating the data scheduling eMBB users and URLLC users at the beginning of each time slot ($t = 1/2^\mu$ ms), shown in Figure 4.

In the frequency domain, multipath fading radio communication channels are often divided into multiple subchannels, which are modulated using multiple carriers. The number of RBs used for eMBB transmission in each time slot depends on the resource scheduling of the previous time slot. When a URLLC service arrives at the BS, it is first necessary to determine the packet size and choose different numerologies according to the size, giving the value of numerology μ to determine the way the URLLC user punctures the currently transmitted eMBB resource.

The puncturing of eMBB transmission by URLLC traffic has an impact on the data rate of eMBB users. Assuming that the arrival of URLLC data follows a Poisson distribution, the data rate of eMBB user k on RB b can be approximated by the Shannon formula. As for a continuously propagating eMBB user, the data rates that can achieve in time slot t are as follows:

$$r_{e,b}^t = B\log_2\left(1 + \delta_{e,b}^t\right)d,\qquad(4)$$

where $B = S \times SCS$ refers to the bandwidth of an RB, $\delta_{e,k}^t = S_e/N_e = (|h_e|^2 P)/N_e$ represents the signal-to-noise ratio SNR of the signal, and S_e, h_e, P_e, N_e is the signal received power, transmit power, channel gain, and channel noise, respectively. We refer to the Nakagami-m channel model [22, 23] to provide the channel gain.

Thus, the data transmission rate of eMBB user k in the case of multiple RB usage can be expressed as follows:

$$r_e^t = \sum_{b \in B} x_{e,b}^t r_{e,b}^t,\qquad(5)$$

where $x_{e,b}^t$ refers to whether the current RB b is allocated to an eMBB user at time slot t, as shown in the following expression:

$$x_{e,b}^t = \begin{cases} 1 \text{ RB } b \text{ is allocated for eMBB user} \\ \quad \text{at time slot } t \\ 0 \text{ otherwise.} \end{cases},\qquad(6)$$

where δ_u^t represents the SNR of the time slot $t \in T$ for URLLC users, $D_u = 1 - 1/(1 + \delta_u^t)^2$ indicates the channel dispersion, and $Q^{-1}(x)d$ is the inverse of the Gaussian function $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$, $x_{u,b}^t$ refers to whether the current RB b is allocated to an URLLC user at time slot t, as follows:

$$x_{u,b}^t = \begin{cases} 1 \text{ RB } b \text{ is allocated for URLLC user} \\ \quad \text{at time slot } t \\ 0 \text{ otherwise.} \end{cases}\qquad(7)$$

URLLC requests cannot be postponed to the next time slot before puncturing, and considering the reliability and delay requirements of URLLC users, let $L_u^t d$ be the size of the URLLC packet that follows the Poisson distribution and Φ_u^t to represent the size of the resources allocated to URLLC users at time slot t. Therefore, there is an inequality as follows:

$$L_u^t \le \Phi_u^t.\qquad(8)$$

In the multiplexing scenario, eMBB users are punctured, so the throughput loss of eMBB users can be expressed as follows, indicating that the throughput loss occurs only when RB b is occupied by both eMBB and URLLC users.
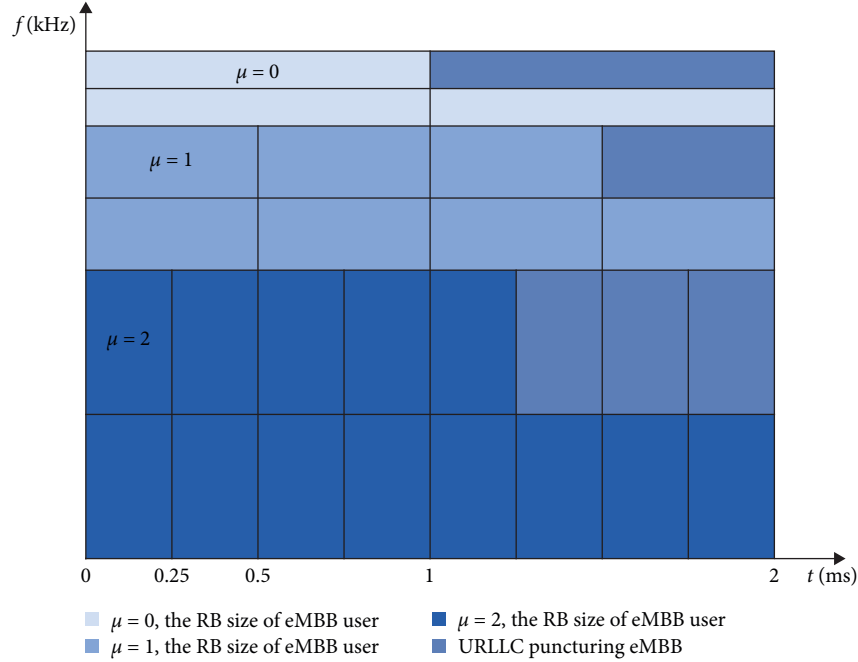
FIGURE 4: RB sizes in different numerology.

$$r_{e,\text{loss}}^t = \sum_{b \in B} r_{e,b}^t \left| \left| x_{e,b}^t - x_{u,b}^t \right| - 1 \right|. \tag{9}$$

Therefore, the actual achievable data transfer rates for eMBB users are as follows:

$$R_e^t = r_e^t - r_{e,\text{loss}}^t . \tag{10}$$

Above all, we establish a data expression that aims to maximize eMBB throughput while satisfying the low latency and high-reliability requirements of URLLC, expressed as follows:

$$P_0 : \max_{x_{u,b}^t, x_{e,b}^t} E\left(\sum_{t=1}^{|T|} R_e^t \right), \tag{11}$$

subject to the following:

$$\begin{aligned}
&P\left(\sum_{u \in U} \Phi_u^t < U \right) \leq \varepsilon \\
&\Phi_u^t L_u^t \leq r_u^t \ \forall b, t \\
&\sum_{u \in U} x_{u,b}^t \leq 1 \ \forall b, t \\
&\sum_{b \in B} x_{e,b}^t \geq 1 \ \forall e, t \\
&\sum_{b \in B} \Phi_u^t x_{u,b}^t \geq 1 \ \forall u, t \\
&\sum_{b \in B} x_{e,b}^t + \sum_{u \in U} \sum_{b \in B} \Phi_u^t x_{u,b}^t \leq |B| \ \forall t \\
&x_{e,b}^t, \ x_{u,b}^t \in \{0, \ 1\} \ \forall e, u, b, t \\
&\Phi_u^t \in \{0, \ 1, \ \ldots, \ B\} \ \forall u, \ b, \ t
\end{aligned} \tag{12}$$

5. The Optimal Algorithm

To obtain the optimal solution of the above problem, it is necessary to consider all numerology values possibilities, as well as the corresponding data transmission rates for eMBB users and punching time slots for URLLC users simultaneously, providing the highest possible data transmission rates for eMBB users while satisfying the requirements of URLLC users for low-latency and high-reliability. Therefore, we cannot use the traditional optimization algorithm to deal with this nonconvex problem. To simplify the original equations and reduce the computational complexity, this paper first decomposes the entire optimization problem P_0 into two convex optimization subproblems, P_1 eMBB resource allocation and P_2 URLLC scheduling problem. A decomposition- and relaxation-based strategy is applied to the eMBB and URLLC resource allocation problems, and the binary variables $x_{e,b}^t$, $x_{u,b}^t$ are relaxed to continuous variables of [0,1]. The dynamic selection numerology algorithm is then proposed based on DQN to get the optimal solutions. The specific implementation process is as follows.

5.1. P_1 eMBB Resource Allocation.

Considering the eMBB resource allocation problem, at this time, we fixed $x_{u,b}^t$, which means the URLLC user puncture punching situation is known. The original integer nonlinear programing problem is relaxed to find the optimal solution of the approximate range, the noninteger solution is rounded to get the integer solution, and at this time, the resource allocation subproblem can be approximated as follows:

$$\widetilde{P}_1 : \max_{x_{e,b}^t} E\left(\sum_{t=1}^{|T|} R_e^t \right), \tag{13}$$

subject to the following:

$$\sum_{b \in B} \widetilde{x}_{e,b}^t \geq 1 \; \forall e, t$$
$$\sum_{b \in B} \Phi_e^t \widetilde{x}_{e,b}^t \leq |B| + \theta \; \forall t \; . \qquad (14)$$
$$0 \leq \widetilde{x}_{e,b}^t \leq 1 \; \forall e, b, t$$

Finally, after solving the above problem, the continuous variables are converted to the original binary variables by threshold rounding, assuming that $\tau \in [0, 1]$ is the threshold, which is then expressed as follows:

$$x_{e,b}^t = \begin{cases} 1, & \widetilde{x}_{e,b}^t \geq \tau \\ 0, & \text{otherwise} \end{cases} . \qquad (15)$$

The above expression utilizes a threshold rounding technique to enforce the variables to remain integers, where τ is a linear relaxation threshold. The value of τ is provided in [24].

The error in the relaxed integer solution obtained by $\tilde{P}_1$ may potentially violate the RB resource allocation constraints of eMBB resources. To address this issue, the integer linear programing problem $\tilde{P}_1$ can be modified as follows:

$$\widetilde{P}_1 : \max_{x_{e,b}^t} E\left(\sum_{t=1}^{|T|} R_e^t \right) + \alpha \Delta, \qquad (16)$$

subject to the following:

$$\sum_{b \in B} \widetilde{x}_{e,b}^t + \Delta \geq 1 \; \forall e, t$$
$$\sum_{b \in B} \Phi_e^t \widetilde{x}_{e,b}^t \leq |B| + \theta \; \forall t \; . \qquad (17)$$
$$0 \leq \widetilde{x}_{e,b}^t \leq 1 \; \forall e, b, t$$

In this context, α represents the weight of Δ, which is a negative value. The objective of the $\tilde{P}_1$ problem is to maximize the data transmission rate of eMBB data while minimizing the rounding error Δ, under the condition of a specified value for $\tilde{x}_{e,b}^t$. Consequently, feasible solutions are obtained in the presence of relaxed errors.

5.2. P_2 URLLC Scheduling.

With respect to the URLLC scheduling problem, fix $x_{e,b}^t$ while expressing the chance constraint as a linear constraint with the help of Markov inequalities as follows:

$$P\left(\sum_{u \in U} \Phi_u^t < U \right) \leq \varepsilon \Rightarrow P\left(\sum_{u \in U} \Phi_u^t < U \right) \leq \frac{E[U]}{\sum\limits_{u \in U} \Phi_u^t}. \qquad (18)$$

Therefore, similar to eMBB resource allocation, the URLLC resource scheduling problem can be reformulated as follows:

$$\widetilde{P}_2 : \max_{x_{u,b}^t} E\left(\sum_{t=1}^{|T|} R_e^t \right), \qquad (19)$$

subject to the following:

$$\varepsilon \sum_{u \in U} \Phi_u^t \geq E[U]$$
$$0 \leq \widetilde{x}_{u,b}^t \leq 1 \; \forall u, b, t \qquad . \qquad (20)$$
$$\Phi_u^t \in \{0, \; 1, \; \dots, \; B\} \; \forall u, \; b, \; t$$

5.3. Dynamic Selection of Numerology.

To address the above problems, considering that URLLC traffic is carried out sporadically and randomly, it is very efficient to allocate resources by interactively and intelligently selecting numerology values in the environment, where numerology values can be updated at the beginning of each time slot to obtain the target optimal solution.

State S: the system state consists of URLLC and eMBB user basic information, including eMBB allocated packet size, URLLC packet size, channel throughput, etc. Therefore, we define $s(t) = \{\mu, L_u^t, R_e^t\}$ as the state space of the whole system at timeslot t.

Action A: the action space is the number of punctured at timeslot t. Therefore $a(t) = \{x_{e,b}^t, x_{e,b}^t\}$, when both are 1 represents the timeslot t is punctured.

To address the above problem, this paper designs the policy $\pi(a_t, s_t)$ that the next state depends only on the current state and the action taken on the current time slot. Therefore, the eMBB and URLLC resource allocation problem can be considered as a Markov decision process (MDP), and this paper designs the corresponding reward function based on the DQN framework to obtain the optimal policy.

Reward R: the reward function is related to the objective function. Considering that our objective is to obtain the maximum eMBB data transfer rate under the URLLC constraints, the value of the reward should be positively related to the magnitude of the eMBB data transfer rate while considering all actions of the intelligence. Therefore, the reward function is defined as the sum of the weights of the system throughput and satisfying the URLLC constraints.

$$R(s(t), a(t)) = \gamma E(\Phi_u^t - L_u^t) + \varphi R_e^t d. \qquad (21)$$

Specifically, the DQN-based dynamic numerology selection algorithm roughly follows Algorithm 1. At the beginning of each time slot, the current numerology and URLLC packet size, etc., are input as the current state; in the current state, the predicted Q-value of each numerology value is output using the deep neural network model, and the highest Q-value is output according to the policy selection corresponding numerology value; calculate the reward value based on the currently selected numerology value and the resource allocation result, and the goal of this reward function is to maximize the data transmission rate of eMBB users under the limitation of URLLC users; update the weights based on the reward value and the selected action so that it can select

```
1      Initialize replay memory D, capacity is N
2      Initialize action-value function Q with random weights γ and random target Q with γ′ = γ
3   For episode = 1, J do
4      repeat
5         With probability ε select a random action a_k, update numerology value μ, with probability 1 − ε select a random action
   a_k = arg max Q(s_k, a, θ)
6         Execute action a_k, observe reward R and new station s′
7         Store (s_k, a_k, R_k, s_{k+1}) in replay memory D
8         Collect sample data(s_j, a_j, R_j, s_{j+1}) from replay memory D randomly
9         Update action-value function Q with limit γ
10          Every C steps reset γ′ = γ
11          until S termination stat
12   End For
13   Return optimal strategy π∗
```

ALGORITHM 1: DQN-based resource allocation algorithm.

the optimal numerology value based on the current state and this update process uses techniques such as experience replay, gradient descent, etc.; repeat the above process until the system completes the resource allocation process and the DRL algorithm iteratively updates the weights of the neural network model to adjust the selection strategy of the optimal numerology value based on the state and reward value of each time slot.

6. Performance Evaluation

This section implements and evaluates the performance of the proposed model with the help of MATLAB. We consider a wireless network where the BS is located at the center of the coverage area, and many eMBB and URLLC users randomly appear in the area. The former is always in transmission, while the latter has a sudden sporadic character. The duration of a frame is 1 ms, and each frame is further divided into equidistant time slots depending on the numerology taken. In the considered mixed numerology network, each RB consists of 12 subcarriers with 14 OFDM symbols, each mini-time slot consists of 2 OFDM symbols, and the modulation and demodulation modes are QPSK and 16QAM. The total resources used in this paper are four frames with a duration of 40 ms. The RB size varies depending on the value of μ, but the total resources remain fixed. Each subframe has 1 URLLC user data packet, where the arrival packet size for URLLC users follows a Poisson distribution traffic model with $\lambda = 1$, and the URLLC data packet sizes are not uniform. The numerology takes the values $\mu = \{0, 1, 2\}$ and the corresponding subcarrier frequency $\Delta f \in \{15, 30, 60\}$, kHZ. The channel noise is set to 0 dB in the presence of puncturing, representing no eMBB data transmission. In the absence of puncturing, the signal-to-noise ratio is 10 dB, indicating smooth eMBB data transmission.

First, the performance of the system with a fixed numerology, i.e., μ is taken to be constant and simulated by MATLAB. The throughput, as well as the data transmission rate in the normal eMBB scheduling process, remain basically
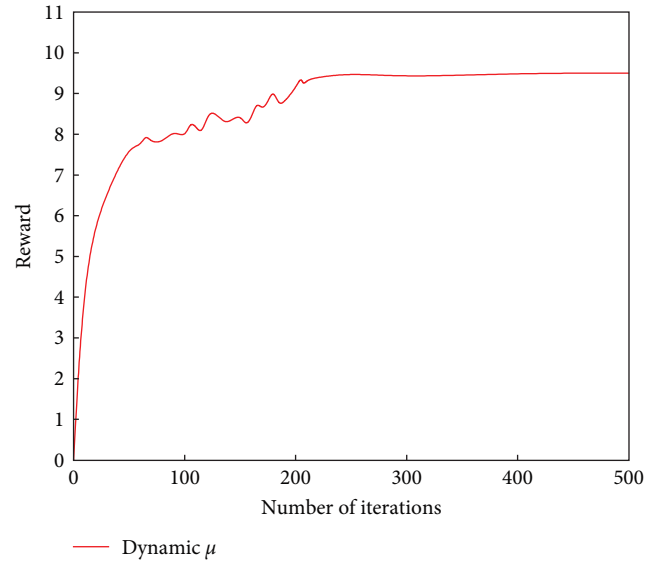

FIGURE 5: The reward function value of the algorithm.

fixed for the two modulation modes. Then, the performance corresponding to different numerologies is further investigated by varying the SNR within a certain range and observing the throughput in the transmission process. For different values of numerology, a reduction in numerology is inevitably accompanied by a reduction in throughput. Therefore, in this paper, we hope to choose suitable numerologies to reduce the impact of puncturing on channel throughput while increasing the eMBB data transmission rate, given a fixed signal power of the transmitter.

In Figure 5, the convergence of the reward function for the proposed algorithm is illustrated over training iterations. It can be observed that throughout the training process, the convergence curves of the total reward function for all agents emerge. Moreover, with the increase in training iterations, the resource allocation scheme within the system undergoes continuous changes, and the system's reward exhibits
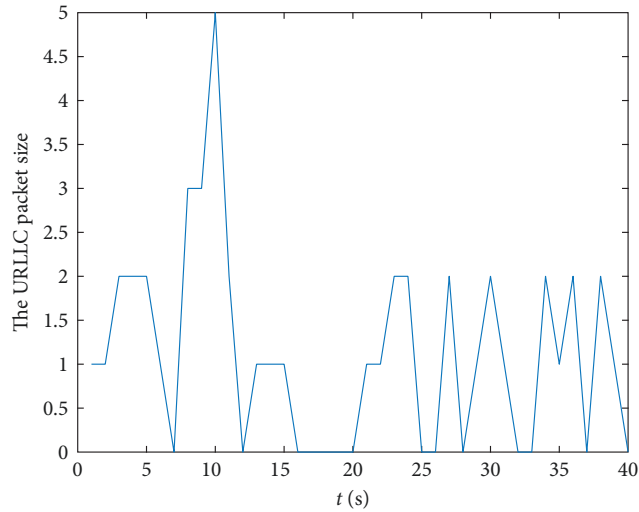
FIGURE 6: URLLC packet size under QPSK.

$\mu = 0$
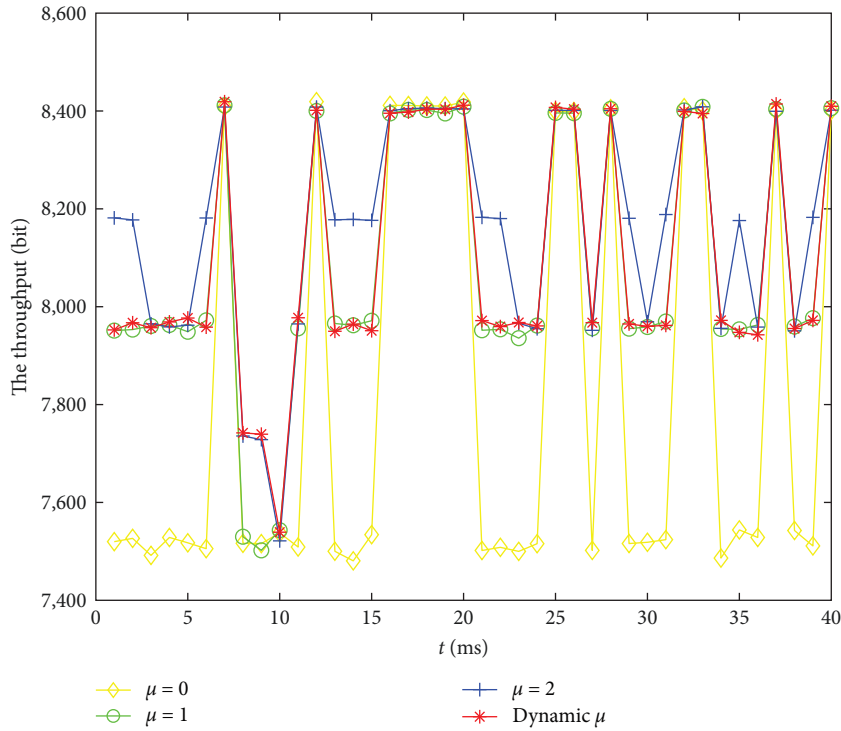$\mu = 1$
$\mu = 2$
Dynamic μ

FIGURE 7: eMBB throughput for different numerology values under QPSK.

a gradual upward trend, ultimately converging as training progresses.

It is known that URLLC user packets follow a Poisson distribution traffic model, and the performance of the system is first simulated and analyzed for different arrival cases of URLLC users with QPSK modulation. The packet size of URLLC following Poisson distribution with time is shown in Figure 6.

Figure 7 shows the performance of eMBB throughput for different numerology values in the case of URLLC packet illustration. From the comparison of simulation results, it

is obvious that regardless of the chosen numerology, the whole throughput of the system must drop when there is URLLC user puncture punching, and the larger the URLLC user packet is at the same numerology value, the throughput loss due to puncture rises, but the level of degradation varies for different numerologies. Therefore, in this paper, we formulate the multiplexing process of eMBB and URLLC as an MDP problem, propose a dynamic adjustment of the numerology-taking algorithm based on DQN, introduce a value function, and evaluate all the trained policies through continuous iterations to obtain the optimal policy, which
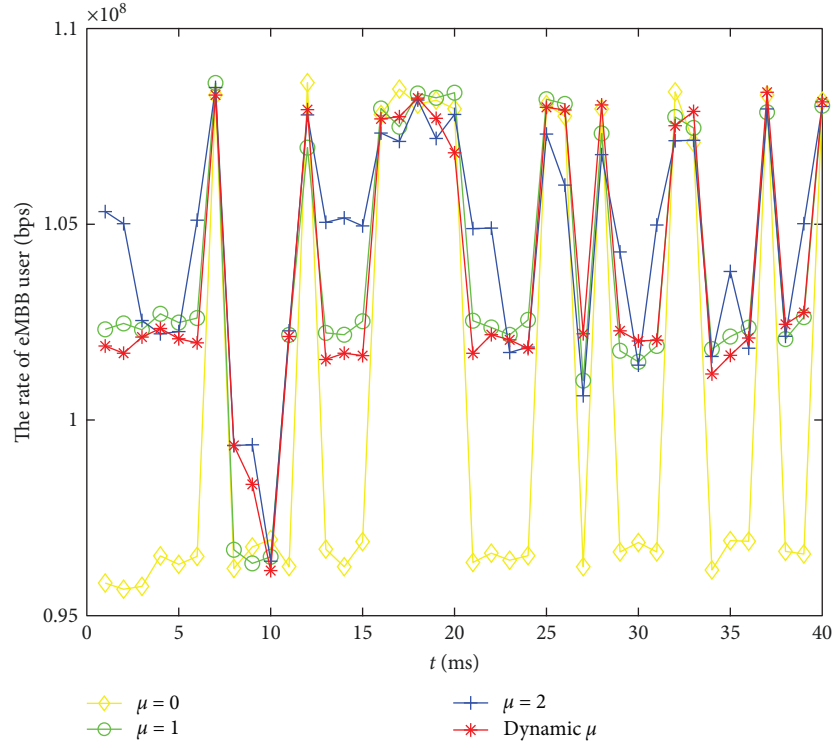
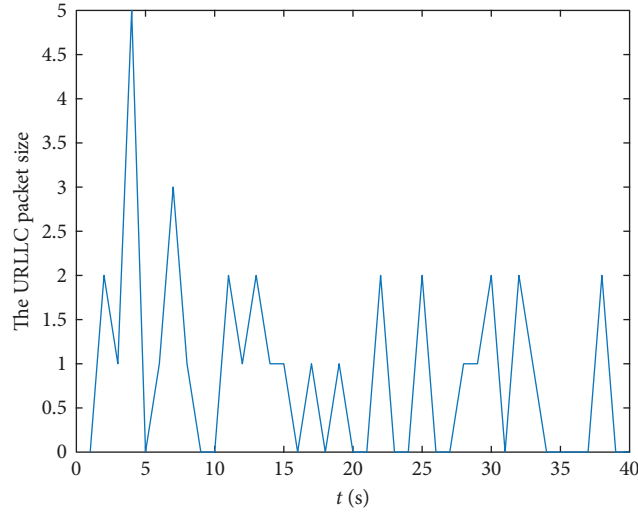FIGURE 8: eMBB data transmission rate for different numerology values under QPSK.

FIGURE 9: URLLC packet size under 16QAM.

moderates the decreasing trend of data throughput and shows that the use of mixed numerologies is meaningful.

Figure 8 shows the performance of the eMBB data transmission rate in the case of the URLLC packet illustration above. From the comparison of simulation results, it is obvious that regardless of the chosen numerology, once the URLLC users puncture, eMBB data transmission rate must drop, and the larger the URLLC user packet at the same numerology value, the greater the impact of perforation on eMBB data transmission rate, but the degradation is different for different numerologies. Therefore, by dynamically adjusting the numerology values, it can be seen in Figure 8 that the dynamic selection of the numerology algorithm proposed in this paper maintains the eMBB data transfer rate at a high level with better performance. According to the simulation results, when the eMBB transmission rate is relatively low, the throughput difference shown in Figure 7 is not significant. This indicates that the dynamic μ algorithm proposed in this paper achieves comparable performance with fewer resources, demonstrating a certain degree of superiority.
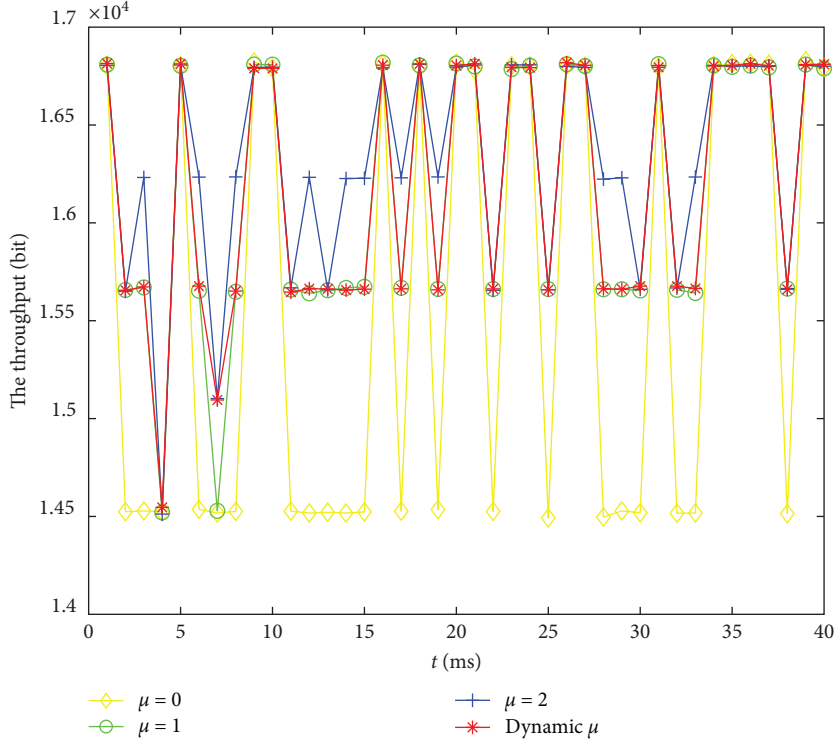
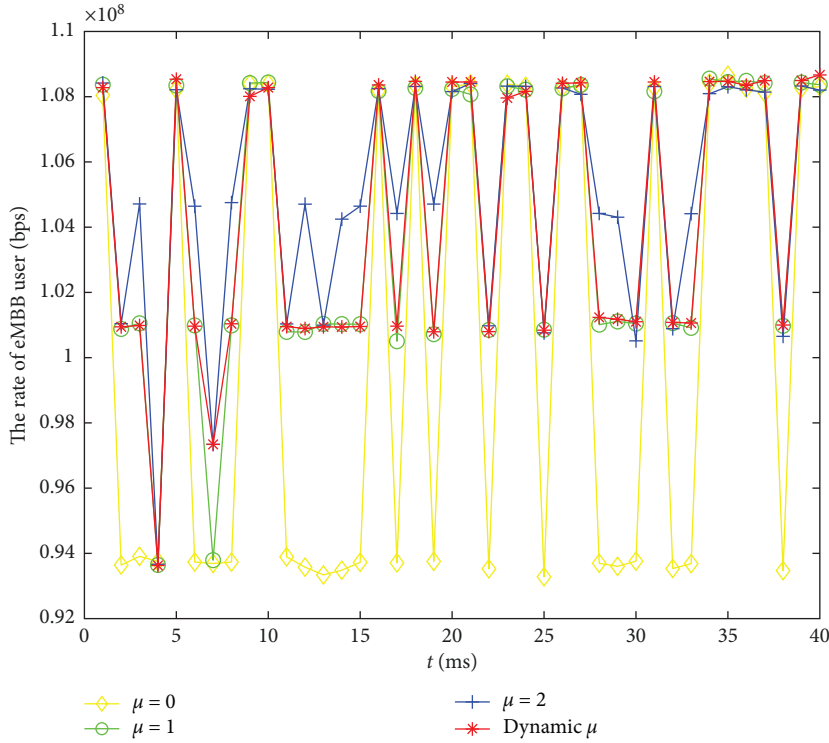FIGURE 10: eMBB throughput for different numerology values under 16QAM.

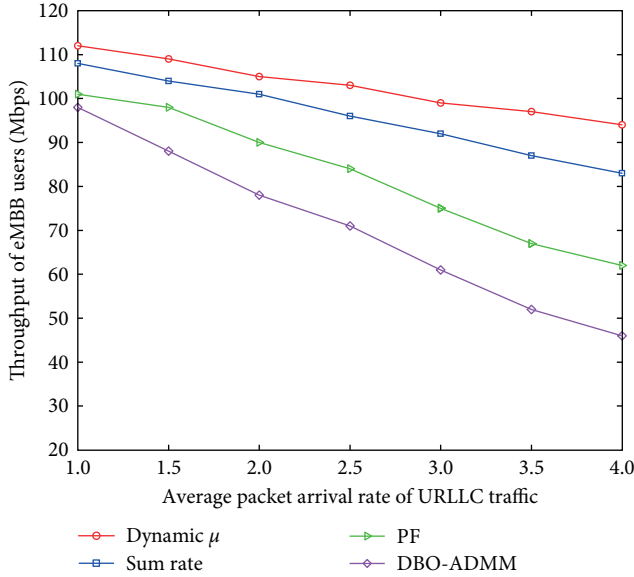FIGURE 11: eMBB data transmission rate for different numerology values under 16QAM.

FIGURE 12: Relationship between eMBB data rate and URLLC traffic packet arrival rate.

FIGURE 13: URLLC user data reliability changes with data packets under different methods.

Keeping the rest of the conditions unchanged and changing the modulation to 16QAM, the URLLC packet size is shown in Figure 9.

Figure 10 depicts the overall throughput performance of 16QAM modulation with different numerology values. It is obvious from the comparison of the simulation results that the overall throughput of 16QAM is increasing compared to QPSK. However, no matter what numerology is chosen, the overall throughput decreasing trend is still unavoidable once punctured, and the throughput loss caused by puncture and URLLC user packet size are positively correlated under the same numerology value, with different levels of decrease for different numerologies. At this time, the dynamic selection of the numerology algorithm can provide the system with the optimal numerology to smooth out the throughput change curve.

Figure 11 shows the performance of the eMBB data transmission rate in the case of the URLLC packet illustration. From the comparison of simulation results, it can be found that different modulation methods reduce the peak difference of eMBB data transmission rate, which is stable at about 100 Mbps, and the low peak of eMBB data transmission rate occurs when the numerology is fixed, which greatly reduces the transmission reliability of the whole system. Therefore, by dynamically adjusting the numerology values, the dynamic selection numerology algorithm proposed in this paper maintains the eMBB data transmission rate at a stable value and performs better in terms of reliability.

In the comparative experiments, the dynamic association scheme proposed in this paper is compared with the proportional fairness (PF) algorithm [7], the algorithm for maximizing eMBB data rate using puncturing strategy (Sum-Rate), and the distributed bandwidth optimization based on ADMM (DBO-ADMM) algorithm [25].
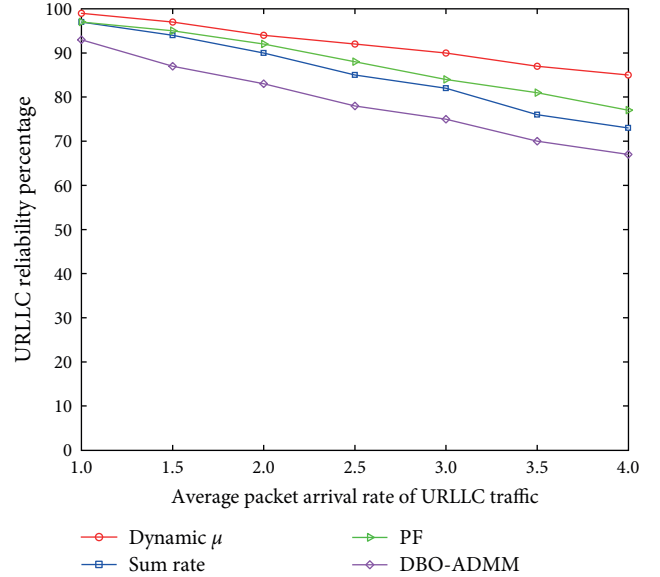
The relationship between the data transmission rate of eMBB users and the average packet arrival rate of URLLC is illustrated in Figure 12. It can be seen that the eMBB average data transmission rate of the dynamic numerology selection scheme proposed in this paper is 6.12% higher than the Sum-Rate algorithm, 19.54% higher than the PF algorithm, and 68.33% higher than the DBO-ADMM algorithm. As the average packet size of URLLC increases, the data transmission rate of eMBB users decreases. However, due to the dynamic allocation of digital resources proposed in this study, the impact on the data transmission rate of eMBB users is significantly less compared to other algorithms.

As shown in Figure 13, when the URLLC average packet arrival rate increases, the reliability of URLLC user data transmission decreases for all four algorithms. When the number of URLLC packets is small, the data reliability of the proposed dynamic numerology selection method is comparable to PF and Sum-Rate algorithms. However, with the increase in URLLC user data volume, the URLLC data reliability of the dynamic numerology selection method is significantly superior to other solutions. When the URLLC user data volume is 4, the reliability of the dynamic numerology selection method is 11.67% higher than that of the PF algorithm, 17.8% higher than the Sum-Rate algorithm, and 28.36% higher than the DBO-ADMM optimization algorithm.

7. Conclusions

In this paper, we propose a dynamic selection numerology scheme for dynamic resource slicing in CP-OFDM-based downlink with the aim of satisfying URLLC constraints while improving the reliability of eMBB. To facilitate joint scheduling of eMBB and URLLC traffic, a combination of numerology, mini-time slot-based transmission, and orthogonal and

punctured scheduling methods is used to solve the optimal resource allocation problem. In particular, this paper analyses the selection of numerology values, decomposing the optimization problem into two convex optimization subproblems to reduce the computational difficulty and applying the DRL algorithm to determine the optimal numerology configuration parameters at the beginning of the time slot. Through simulation, it is found that the dynamic numerology selection scheme proposed in this paper can adapt to the network dynamics and maximize the data transmission rate of eMBB in each time slot while satisfying the high-reliability and low latency of URLLC and the utilization of computational resources increases as the number of URLLC holes increases, which helps maintain the reliability of data transmission and achieve lower latency for URLLC slices. The superiority of the scheme in this paper in the scenario of the coexistence of two services is verified.

Data Availability

Data sets are not used in this article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] M. Series, "Minimum requirements related to technical performance for IMT-2020 radio interface (s)," Report 2410: 2410-2017, 2017.

[2] Part, Transaction Capabilities Application, "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; 3G security; Network Domain Security (NDS)," 2011.

[3] A. Manzoor, S. M. A. Kazmi, S. R. Pandey, and C. S. Hong, "Contract-based scheduling of URLLC packets in incumbent EMBB traffic," *IEEE Access*, vol. 8, pp. 167516–167526, 2020.

[4] A. Anand, G. de Veciana, and S. Shakkottai, "Joint scheduling of URLLC and eMBB traffic in 5G wireless networks," *IEEE/ACM Transactions on Networking*, vol. 28, no. 2, pp. 477–490, 2020.

[5] M. Alsenwi, N. H. Tran, M. Bennis, A. K. Bairagi, and C. S. Hong, "eMBB-URLLC resource slicing: a risk-sensitive approach," *IEEE Communications Letters*, vol. 23, no. 4, pp. 740–743, 2019.

[6] J. Li, H. Gao, T. Lv, and Y. Lu, "Deep reinforcement learning based computation offloading and resource allocation for MEC," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, Barcelona, Spain, 2018.

[7] H. Yin, L. Zhang, and S. Roy, "Multiplexing URLLC traffic within eMBB services in 5G NR: fair scheduling," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 1080–1093, 2021.

[8] A. K. Bairagi, M. S. Munir, M. Alsenwi et al., "Coexistence mechanism between eMBB and uRLLC in 5G wireless networks," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1736–1749, 2021.

[9] A. K. Bairagi, M. S. Munir, M. Alsenwi, N. H. Tran, and C. S. Hong, "A matching based coexistence mechanism between eMBB and uRLLC in 5G wireless networks," in *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, pp. 2377–2384, Association for Computing Machinery, 2019.

[10] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for eMBB and URLLC coexistence in 5G and beyond: a deep reinforcement learning based approach," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4585–4600, 2021.

[11] N. C. Luong, D. T. Hoang, S. Gong et al., "Applications of deep reinforcement learning in communications and networking: a survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3133–3174, 2019.

[12] R. Li, Z. Zhao, Qi Sun et al., "Deep reinforcement learning for resource management in network slicing," *IEEE Access*, vol. 6, pp. 74429–74441, 2018.

[13] J. Mao, L. Zhang, P. Xiao, and K. Nikitopoulos, "Interference analysis and power allocation in the presence of mixed numerologies," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5188–5203, 2020.

[14] J. Mao, A. Farhang, L. Zhang, Z. Chu, P. Xiao, and S. Gu, "Interference analysis in multi-numerology OFDM systems: a continuous-time approach," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, IEEE, Montreal, QC, Canada, 2021.

[15] S. Mcwade, M. F. Flanagan, J. Mao, L. Zhang, and A. Farhang, "Resource allocation for mixed numerology NOMA," *IEEE Wireless Communications Letters*, vol. 10, no. 10, pp. 2240–2244, 2021.

[16] C.-Y. Su, C.-H. Fang, L.-H. Shen, and K.-T. Feng, "Effective capacity maximization for multi-numerology based 5G NR networks," in *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*, pp. 1–5, IEEE, Victoria, BC, Canada, 2020.

[17] M. Zambianco and G. Verticale, "Mixed-numerology interference-aware spectrum allocation for eMBB and URLLC network slices," in *2021 19th Mediterranean Communication and Computer Networking Conference (MedComNet)*, pp. 1–8, IEEE, Ibiza, Spain, 2021.

[18] M. Zambianco and G. Verticale, "A reinforcement learning agent for mixed numerology interference-aware slice spectrum allocation with non-deterministic and deterministic traffic," *Computer Communications*, vol. 189, pp. 100–109, 2022.

[19] A. Esmaeily, H. V. K. Mendis, T. Mahmoodi, and K. Kralevska, "Beyond 5G resource slicing with mixed-numerologies for mission critical URLLC and eMBB coexistence," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 727–747, 2023.

[20] N. Ferdosian, S. Berri, and A. Chorti, "5G new radio resource allocation optimization for heterogeneous services," in *2022 International Symposium ELMAR*, pp. 1–6, IEEE, Zadar, Croatia, 2022.

[21] A. F. Demir and H. Arslan, "Inter-numerology interference management with adaptive guards: a cross-layer approach," *IEEE Access*, vol. 8, pp. 30378–30386, 2020.

[22] C. Tellambura and A. D. S. Jayalath, "Generation of bivariate Rayleigh and Nakagami-m fading envelopes," *IEEE Communications Letters*, vol. 4, no. 5, pp. 170–172, 2000.

[23] S. A. Dianat, "SNR estimation in Nakagami fading channels with arbitrary constellation," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP '07*, pp. II–325-II-328, IEEE, Honolulu, HI, USA, 2007.

[24] K.-S. Goetzmann, S. Stiller, and C. Telha, "Optimization over integers with robustness in cost and few constraints," in *Approximation and Online Algorithms. WAOA 2011*, R. Solis-Oba and G. Persiano, Eds., vol. 7164 of *Lecture Notes in Computer Science*, pp. 89–101, Springer, Berlin, Heidelberg, 2012.

[25] P. Yang, X. Xi, T. Q. S. Quek, J. Chen, X. Cao, and D. Wu, "How should i orchestrate resources of my slices for bursty URLLC service provision?" *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 1134–1146, 2021.