(IJACSA) International Journal of Advanced Computer Science and Applications,
Vol. 15, No. 9, 2024

Deep Reinforcement Learning-Based Carrier Tuning Algorithm for Mobile Communication Networks

Weimin Zhang*, Xinying Zhao
Zhengzhou Railway Vocational and Technical College, Zhengzhou 451460, Henan, China

***Abstract*—With the evolution of mobile communication networks towards 5G and beyond to 6G, managing network resources presents unprecedented challenges, particularly in scenarios demanding high data rates, low latency, and extensive connectivity. Traditional resource allocation methods struggle with network dynamics and complexity, including user mobility, varying network loads, and diverse Quality of Service (QoS) requirements. Deep Reinforcement Learning (DRL), an emerging AI technique, demonstrates significant potential due to its adaptive and learning capabilities. This paper integrates user mobility and network load prediction into a DRL framework and proposes a novel reward function to enhance resource utilization efficiency while meeting real-time QoS demands. We establish a system model involving base stations and receiving terminals to simulate communication services within coverage areas. Comparative experiments analyze the performance of the DRL approach versus traditional methods across metrics such as throughput, delay, and spectral efficiency. Results indicate DRL's superiority in handling dynamic environments and fulfilling QoS needs, especially under heavy loads. This study introduces innovative approaches and tools for future mobile network resource management, paving the way for practical DRL implementations and enhancing overall network performance.**

***Keywords*—DRL; mobile network; carrier tuning**

I. Introduction

With the rapid development of the mobile Internet and the Internet of Things (IoT), global mobile data traffic has shown exponential growth, which puts unprecedented pressure on modern mobile communication networks. Especially in 5G and even future 6G networks, the demand for high bandwidth, low latency and large-scale device connectivity puts higher requirements on network resource management and scheduling. However, the contradiction between the scarcity of spectrum resources, as a valuable asset for wireless communications, and the increasing user demand is becoming increasingly prominent, and has become one of the key factors constraining the quality of service (QoS) of mobile communications [1]. Moreover, in highly dynamic network environments, such as user mobility, signal interference, and variable network load, maintaining stable QoS becomes a complex and challenging task [2].

The popularization of 5G networks and the vision of 6G networks emphasize the need for communications with high speeds, low latency, and large numbers of connections, which puts more stringent requirements on the efficient management and scheduling of network resources. However, the increasingly sharp contradiction between the finiteness of spectrum resources and the infinite growth of user demand has become a major bottleneck constraining the performance improvement of modern communication networks, as shown in Fig. 1 for a specific mobile communication network model [3].

In dynamically changing network environments, such as the movement of user locations, fluctuations in signal strength, and ups and downs in network traffic, it becomes extremely difficult to maintain a stable QoS. Traditional carrier adjustment algorithms, although showing some effectiveness in dealing with simple and static network conditions, often appear to be inadequate in complex scenarios due to their reliance on preset rules and static models, and their lack of real-time response to dynamic changes in the network and intelligent decision-making capabilities, leading to inefficient utilization of spectrum resources and degradation of user experience [4].

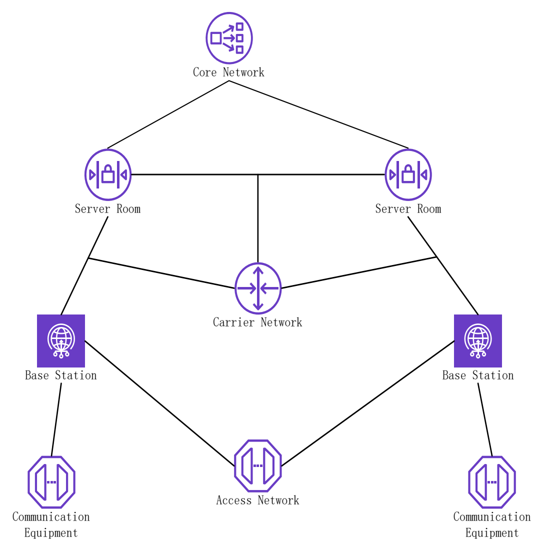

Fig. 1. Mobile communication network model.

Traditional carrier adjustment algorithms, although solving the resource allocation problem to a certain extent, they are often based on predefined rules and static models, making it difficult to adapt to rapidly changing network conditions. These algorithms usually rely on predefined thresholds and fixed policies, lack intelligent decision-making capabilities, and cannot respond to dynamic changes in the network in real time, leading to inefficient resource utilization, especially in complex multi-user scenarios, which is prone to spectrum wastage and degradation of user experience [5]. Facing the above challenges, in recent years, Deep Reinforcement Learning (DRL), as an emerging artificial intelligence technology, has shown great potential in the field of wireless communication due to its powerful learning ability and self-adaptability. DRL is able to learn the optimal policy through interaction with the environment and can handle complex, nonlinear decision problems without explicit programming. In mobile communication networks, DRL can be used to intelligently adjust carrier configurations to achieve dynamic resource optimization while maximizing network efficiency and user satisfaction. This approach overcomes the limitations of traditional algorithms and is able to make more flexible and efficient decisions in uncertain and dynamic environments, providing a new way to address spectrum scarcity and improve service quality [6].

This study aims to explore the application of deep reinforcement learning in the field of carrier tuning for mobile communication networks by proposing a novel DRL-based algorithm that aims to dynamically optimize the allocation of spectrum resources to meet the changing network demands and improve the spectrum utilization while guaranteeing the quality of service. Through experimental validation, we will demonstrate the superiority of the proposed algorithm in complex network environments and its significant value in addressing spectrum scarcity and improving QoS. This study focuses on the application of Deep Reinforcement Learning (DRL) to carrier adjustment in mobile communication networks, aiming to develop a new generation of algorithms that can intelligently adapt to dynamic changes in the network and optimize the allocation of spectral resources. With its powerful learning capability and adaptivity, DRL is able to learn the optimal strategy from the interaction with the environment and solve complex, nonlinear decision problems without human intervention. In the field of mobile communication, the application of DRL is expected to break through the limitations of traditional algorithms, achieve dynamic resource optimization, and improve network efficiency and user satisfaction. Specifically, this study will focus on the following key points: (1) Design and optimization of DRL algorithms: exploring how to combine the characteristics of DRL to design algorithms applicable to carrier adjustment in mobile communication networks, including the definition of state space and action space, and the design of reward functions, to ensure that the algorithms can efficiently learn the optimal resource allocation strategies. (2) Resource allocation in dynamic environments: study how to use DRL for real-time carrier allocation adjustment to improve spectrum utilization and network performance in complex dynamic environments with user mobility, signal interference and network load changes. (3) QoS guarantee mechanism: analyze the potential of DRL in ensuring quality of service, and explore how to achieve efficient utilization of spectrum resources and reduce resource wastage while meeting user demands.

This paper aims to solve the problem of how to efficiently manage and optimize wireless resources in dynamic network environment, especially in the face of high load and user mobility challenges. Existing resource management algorithms are often difficult to cope with the rapidly changing network conditions, resulting in degradation of quality of service. Therefore, this paper proposes a new method based on deep reinforcement learning to dynamically adjust network resource allocation, improve service quality and reduce energy consumption.

The simulation program based on deep reinforcement learning (DRL) developed in this study has many advantages. First, it can dynamically adjust network resource allocation to effectively cope with changes in user mobility and network load, thus improving overall network performance. Secondly, by introducing advanced reward function design, the program realizes strict guarantee of quality of service (QoS) and ensures high standard of service quality in various operation scenarios. In addition, its self-learning mechanism allows the system to continuously optimize policies over time, adapt to complex network environments, and ultimately achieve intelligent resource management effects.

The innovation of this paper lies in integrating user mobility and network load prediction into DRL framework systematically for the first time, thus achieving more accurate modeling of network state; and a novel reward function design is proposed, which can more effectively motivate algorithms to optimize resource utilization efficiency while satisfying real-time QoS constraints.

As mobile communication networks evolve towards 5G and future 6G, network resource management faces unprecedented challenges. Although existing literature explores the application of traditional resource allocation algorithms in high data rate, low latency and massive connectivity scenarios, there is still a lack of research on combining user mobility and dynamic changes in network load. This thesis aims to fill this research gap by introducing a Deep Reinforcement Learning (DRL) framework and proposing a new reward function design to optimize the resource utilization efficiency and satisfy the real-time QoS constraints, so as to provide a new way of thinking about resource management in dynamic network environments.

The structure of this paper is as follows: The second part introduces the theoretical basis of the study, including signal, channel and interference models; the third part describes the experimental environment and data set used in detail; the fourth part describes the experimental methods and evaluation indicators; the fifth part presents the experimental results and their analysis; the sixth part discusses the innovations and shortcomings of this paper. Finally, the seventh part summarizes the full text and looks forward to the future research direction.

II. Literature Review

Deep Reinforcement Learning (DRL) has demonstrated its unique advantages in resource allocation, network optimization, and spectrum management in the communication domain, providing a new perspective to address the limitations of traditional algorithms. Numerous cutting-edge researches have revealed the powerful capability of DRL in dealing with dynamic and complex network environments, opening the way for intelligent management of communication networks. An intelligent dynamic spectrum access scheme is designed by skillfully integrating DRL into the dynamic spectrum access strategy. This scheme is able to dynamically adjust the access strategy according to the real-time changes of the network, which significantly improves the spectrum utilization and system throughput, demonstrating the excellent performance of DRL in spectrum resource management [7]. A DRL-based network slice resource management algorithm is proposed in the literature, which is meticulously optimized for different quality of service (QoS) requirements, which not only confirms DRL's ability to deal with complex network environments, but also highlights its great potential in realizing fine-grained management of network resources [8]. In the literature, Proximal Policy Optimization (PPO) has been applied to optimize beamforming in multiuser MIMO systems. Compared with the traditional DQN, PPO shows better performance in continuous action space and achieves higher throughput and lower BER, and this result is a strong proof of the advantages of PPO in dealing with complex action space problems. It provides a new solution for resource allocation in multiuser MIMO systems [9]. Literature explores the application of DRL in beam selection and path switching in millimeter-wave communications, and they find that DRL can effectively cope with the signal fading and blocking problems in high-frequency communications, providing a strong guarantee for the reliability and stability of millimeter-wave communications. In the vast field of communication network optimization, different deep learning models have shown their strengths, providing a diverse toolbox for network resource management [10]. DQN, with its intuitive architecture and ability to deal with discrete action spaces, has dominated the early applications of DRL, and has especially excelled in dealing with simple decision problems. However, DQN has obvious limitations in continuous action space and in handling long temporal dependency problems. In contrast, Proximal Policy Optimization (PPO) and Trust Region Policy Optimization (TRPO) exhibit higher stability and convergence speed when dealing with continuous action spaces, and are especially suitable for complex environments like multi-user MIMO systems [11]. These algorithms are not only able to better adapt to the dynamic changes of the network, but also realize the fine tuning of resource allocation to ensure the efficiency and stability of network operation. In addition, the combination of recurrent neural networks (RNNs) and DRLs provides a powerful tool for processing time series data with long-term dependencies, such as in network traffic prediction and dynamic resource allocation. RNNs capture the intrinsic pattern of data evolution over time, and in combination with DRLs, they can realize accurate prediction of network state and prospective allocation of resources, further enhancing the intelligent management level of communication. The intelligent management level of the network is further enhanced [12].

Traditional carrier adjustment algorithms, such as rule-based and statistical modeling approaches, although perform well in static or relatively stable network environments, are severely challenged in terms of their flexibility and adaptability when facing dynamically changing and complex network conditions. These algorithms often rely on fixed thresholds and predefined rules, making it difficult to respond to dynamic changes in the network in real time and leading to inefficient resource allocation. Especially in multi-user scenarios, balancing system throughput and fairness becomes a difficult task, and the allocation of spectrum resources is often not reasonable, and user experience is significantly affected. In addition, when encountering uncertain or unforeseen network conditions, the model generalization ability of traditional algorithms is poor and difficult to respond effectively, limiting their application in complex network environments [13, 14].

III. THEORETICAL FOUNDATIONS

In mobile communication networks, we construct a system model consisting of base stations (transmitters) and receivers (terminals) that is designed to simulate how base stations provide communication services to terminals in their coverage areas. The network model integrates several key components, including a signal model, a channel model, and an interference model. The signal model describes in detail the physical process of signal propagation in the air, covering factors such as path loss, shadow fading, multipath effects and noise; the channel model reflects the characteristics of the channel over time, such as Rayleigh fading or Rice fading; and the interference model takes into account signal interference from other users, including co-channel and neighboring-frequency interference, in order to more accurately simulate the communication process in a real network environment [15].

The carrier tuning problem lies in how to optimally allocate network resources, such as frequency, power and time, to satisfy user demands and optimize the overall network performance. This problem can be abstracted as an optimization problem with the objective function and constraints shown below:

Objective function: maximize the network throughput or minimize the total power consumption, i.e., find $\mathbf{x}^*$ such that $f(\mathbf{x})$ is maximized or minimized, where $\mathbf{x}$ is a vector of decision variables containing frequency, power, and time allocations, which can be expressed as Eq. (1) and Eq. (2) [16].

$$\mathbf{x}^* = \arg\max_{\mathbf{x}} f(\mathbf{x}) \tag{1}$$

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} f(\mathbf{x}) \tag{2}$$

Constraints: These include QoS requirements, spectrum utilization rules, and physical layer constraints. QoS constraints ensure that each user's minimum data rate requirement is met. It is denoted as $r_i(\mathbf{x}) \geq R_i, \quad i = 1, 2, ..., N$ where $r_i(\mathbf{x})$ is the instantaneous throughput of user(i) and R_i is the minimum data rate requirement of user(i). The QoS constraint limits the total spectrum usage of each base station, denoted as wield. The spectrum usage constraint limits the total spectrum usage of each base station and is denoted wield Eq. (3), where x_i is the spectrum width used by base station i and F is the total available spectrum width. The power constraint limits the total power output of each base station and is denoted as Eq. (4), where p_i is the power output of base station i and P is the maximum allowed power [17].

$$\sum_{i=1}^{N} x_i \leq F \tag{3}$$

$$\sum_{i=1}^{N} p_i \leq P \tag{4}$$

The resource allocation problem can be described by the following mathematical model: assume that there are B base stations and U users in the network, and there exists a channel gain $(h_{b,u})$ between each base station b and user u. The power allocated by base station b to user u is $(p_{b,u})$, and the data rate of user u is determined by Shannon's formula, which is specified as Eq. (5) [18].

$$r_u = W \log_2 \left(1 + \frac{p_{b,u} h_{b,u}}{I_u + N_0 B} \right) \tag{5}$$

where W is the bandwidth allocated to the user, I_u is the interference to user u, N_0 is the noise power spectral density, and (B is the signal bandwidth. The objective is to minimize the total power consumption while satisfying the minimum data rate requirement for all users as shown in Eq. (6)-(8) [19].

$$\min_{\{p_{b,u}\}} \sum_{b=1}^{B} \sum_{u=1}^{U} p_{b,u} \tag{6}$$

$$\text{s.t. } r_u \geq R_{\min}, \quad u = 1, 2, ..., U \tag{7}$$

$$\sum_{u=1}^{U} p_{b,u} \leq P_b, \quad b = 1, 2, ..., B \tag{8}$$

Here, $R_{\min}$ is the minimum data rate requirement per user, and P_b is the maximum power limit of the base station (b

Deep Reinforcement Learning (DRL) can be applied to the above optimization problem by learning to dynamically adjust the carrier configuration in order to optimize the network performance. The DRL algorithm learns, by interacting with the environment, a policy π which selects the action a in a given state s in order to maximize the expected cumulative reward R [20].

IV. Methodology

The Deep Reinforcement Learning (DRL)-based resource management algorithm for mobile communication networks proposed in this paper is centered around three core parts: the design of state space, action space and reward function, resource allocation strategy in dynamic environment, and QoS guarantee mechanism. The state space S integrates key parameters such as channel state information, user data demand, and network load, while the action space A covers the dynamic adjustment of base station power and spectrum. The reward function R, on the other hand, combines several performance metrics such as throughput, delay, spectral efficiency and power consumption to guide the algorithm to learn the optimal resource allocation strategy [21].

In a dynamic network environment, the DRL algorithm dynamically adjusts resource allocation by observing the network state in real time and predicting changes in user mobility and network load to optimize network performance and user experience. The algorithm also introduces a dynamic QoS threshold adjustment mechanism to tune QoS parameters in real time according to the network state and user demand, ensuring that the quality of service meets high standards in various operational scenarios. Through iterative learning and policy updating, the DRL algorithm gradually approaches the optimal resource allocation policy, balances the multi-objective optimization problem, and realizes intelligent resource management in complex network environments [22].

A. Design and Optimization of DRL Algorithm

In designing deep reinforcement learning (DRL)-based carrier tuning algorithms for mobile communication networks, the key steps include defining the state space, action space, and designing the reward function to ensure that the algorithms are able to adapt to the dynamic characteristics and complex demands of mobile communication networks, and the specific deep reinforcement learning architecture is shown in Fig. 2 [23]. The state space S is the set of environmental states observed by the algorithm at each moment in time. In a mobile communication network, the states may include, but are not limited to, channel state information (CSI), user data demand, network load, spectrum resource allocation, and base station power state. A typical state vector s_t may contain, as specified in Eq. (9) [24].

$$s_t = \left[\mathbf{h}_t, \mathbf{d}_t, \mathbf{p}_t, \mathbf{f}_t \right] \tag{9}$$

where $\mathbf{h}_t$ is the channel gain vector indicating the channel conditions between each base station and user. $\mathbf{d}_t$ is the data demand vector, denoting the data transmission demand of each user at time (t). $\mathbf{p}_t$ is the base station power allocation vector, denoting the power allocated to each user by each base station. $\mathbf{f}_t$ is the spectrum allocation vector, denoting the spectrum resources allocated to each user by each base station [25].

The action space (A) defines all possible actions, i.e., resource allocation strategies, that the algorithm can take. In carrier tuning, an action can be to change the power and spectrum allocation of the base station. A typical action a_t can be Eq. (10).

$$a_t = \left[p_{1,t}, p_{2,t}, ..., p_{B,t}; f_{1,t}, f_{2,t}, ..., f_{B,t} \right] \tag{10}$$

where $(p_{i,t})$ and $(f_{i,t})$ represent the power and spectrum allocation of the base station (i at time (t, respectively.

The reward function (R quantifies the immediate effect of performing an action (a_t) in a particular state (s_t). The design of the reward function is crucial as it guides the learning direction of the DRL algorithm. In mobile communication networks, the reward function may be based on network throughput, QoS satisfaction level and power consumption. A basic reward function can be defined as Eq. (11).

$$R(s_t, a_t) = \alpha T(s_t, a_t) - \beta P(s_t, a_t) \tag{11}$$

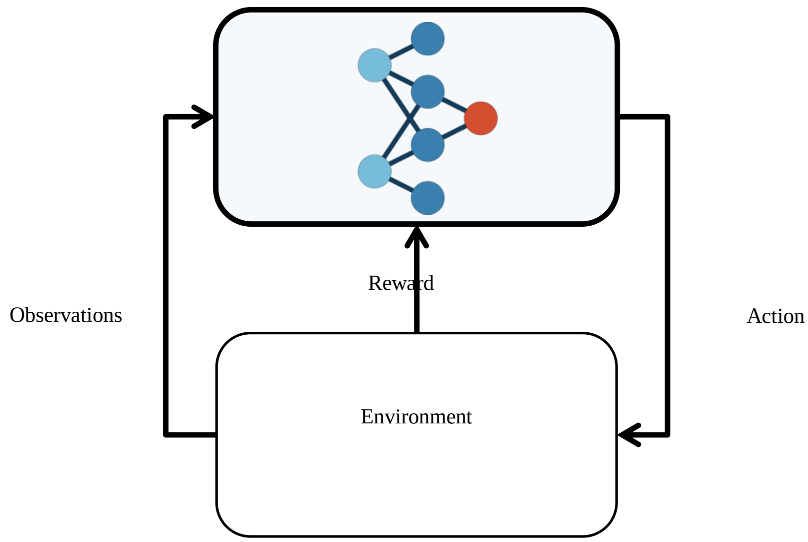

Fig. 2. Deep reinforcement learning architecture.

B. Resource Allocation in a Dynamic Environment

In the dynamic environment of mobile communication networks, real-time adjustment and adaptation of resource allocation strategies are crucial to ensure optimization of network performance and enhancement of user experience. Deep Reinforcement Learning (DRL) provides a powerful solution that dynamically adjusts resource allocation strategies by predicting user mobility and network load changes in real-time to achieve intelligent management of network resources. The core idea of the DRL algorithm consists of defining the state space (S), which contains key environment characteristics such as channel state information (CSI), user data demand, network load, spectrum resource allocation, and base station power state, and other key environmental features; defining the action space (A), which defines all possible actions that the algorithm can take, i.e., resource allocation strategies, including operations such as adjusting the power, channel allocation, and frequency selection of the BTS; designing the reward function (R), which is used to quantify the instantaneous effect of executing an action in a specific state, such as an increase in throughput, a decrease in latency, or an improvement in the spectral efficiency; and policy learning (π), the DRL algorithm dynamically adjusts the resource allocation strategy by observing the network state and predicting user mobility and network load changes, with the goal of maximizing the long-term cumulative rewards. Dynamic resource adjustment is one of the key features of the DRL algorithm, which adjusts the resource allocation of base station c in real time based on the prediction of user location changes, e.g., a user moves from the coverage area of base station b to the coverage area of base station c, to ensure the quality of service. In addition, the resource allocation strategy involves solving a multi-objective optimization problem, including maximizing throughput, minimizing delay, and maximizing spectral efficiency, etc. The DRL algorithm gradually approaches the optimal strategy through iterative learning while balancing the different objectives through a composite reward function. The iterative learning process starts from the initialization state, the algorithm selects an action based on the current state at each time step t, updates the environment state after executing the action, computes the rewards, stores the experience to the experience playback pool and periodically samples from the pool, and updates the parameters of the DRL model using gradient descent until the termination condition is reached. Through these steps, the DRL algorithm can dynamically adapt to the changing network environment and user demands, optimize network performance, and guarantee service quality [26, 27].

In dynamic network environments, real-time and adaptive resource allocation is key. Deep Reinforcement Learning (DRL) provides a mechanism that enables algorithms to dynamically adapt resource allocation strategies based on the real-time state of the network to optimize network performance and user experience. In the DRL framework, an intelligent body (i.e., the DRL algorithm) decides the action a_t, i.e., the resource allocation strategy for the next moment, by observing the current state s_t of the network, including the channel state information (CSI), the user's data demand, the network load, and the allocation of spectrum resources. The mapping between the state s_t and the action a_t is defined by the policy π, while the goal of the policy π is to maximize the long-term cumulative reward. In mobile communication networks, intelligence need to adjust the power and spectrum allocation of base stations in real time to cope with changes in user locations and fluctuations in network load. This process can be described by the following Eq. (12) [28].

$$a_t = \pi(s_t) \tag{12}$$

where a_t is the resource allocation action at time t, s_t is the current state of the network, and π is the policy function obtained through DRL learning.

Dynamic adjustment of resource allocation is realized by observing the changes in network state and evaluating the effects of different resource allocation schemes. In terms of user mobility, the DRL algorithm predicts changes in the user's location and adjusts the resource allocation accordingly to maintain service continuity and quality. For example, when a user u is detected to move from the coverage area of base station b to that of base station c, the algorithm predicts this change and adjusts the resource allocation of base station c accordingly to ensure that the quality of service of user u is not affected [29].

The adjustment of resource allocation can be quantized as $\Delta p_c = \alpha \cdot \Delta d_{uc}$, $\Delta f_c = \beta \cdot \Delta d_{uc}$ by the following equations. Where Δp_c and Δf_c are the incremental adjustments of power and spectrum resources of base station c, respectively, Δd_{uc} is the amount of change in the distance between user u and base station c, and α and β are the adjustment coefficients, which can be optimized according to the actual situation of the network [30].

In order to integrate the user mobility model into the DRL algorithm, we can define the state space S, the action space A, and the reward function R(s,a,s'). The state space S can contain information such as the user's current position, velocity, and target location; the action space A can be operations such as adjusting the base station's transmit power, channel assignment, frequency selection, etc. and the reward function R measures the change in the network's performance after executing a certain action, such as an increase or decrease in the throughput, delay, or spectral efficiency [31].

The implementation of resource allocation policies usually involves solving multi-objective optimization problems, where the objectives may include maximizing throughput, minimizing delay, and maximizing spectral efficiency, etc. The DRL algorithm learns iteratively to gradually approximate the optimal policy. In a multi-objective optimization scenario, the DRL algorithm can balance multiple objectives by defining a composite reward function, which is defined in this paper as $R(s_t,a_t)=\alpha T(s_t,a_t)-\beta D(s_t,a_t)+\gamma E(s_t,a_t)$. Where, $D(s_t,a_t)$ is the average delay when the network executes the action a_t in the state s_t . $E(s_t,a_t)$ is the spectral efficiency when the network executes the action in state s_t a_t . α , β, γ , δ is a weighting factor used to balance throughput, delay and spectral efficiency.

The implementation of a resource allocation policy usually involves solving a multi-objective optimization problem, where the objectives may include maximizing throughput, minimizing delay, and maximizing spectral efficiency, etc. The DRL algorithm learns iteratively to gradually approximate the optimal policy.

The specific steps and pseudo-code are shown in Table I.

TABLE I. RESOURCE ALLOCATION STRATEGY PSEUDO-CODES

1. initialization state(s (including user location, network load, etc.). 2. For each time step (t: The smart body selects an action a_t based on the current state s_t . Execute the action a_t to update the environment status to s_{t+1} . Calculate rewards $r_t = R(s_t, a_t, s_{t+1})$. Store experience (s_t, a_t, r_t, s_{t+1}) to the experience playback pool. Sample a batch of experiences from the experience playback pool and update the parameters of the DRL model using gradient descent. Repeat step 2 until the termination condition is reached.

In summary, the combination of user mobility model, network load prediction and DRL algorithm can realize a smarter and more efficient resource allocation strategy to adapt to the changing network environment and user demands.

C. QoS Assurance Mechanism

Deep Reinforcement Learning (DRL) algorithm maintains or enhances Quality of Service (QoS) in mobile communication networks through dynamic resource allocation policies. It maximizes the overall performance of the network by designing a comprehensive reward function which considers both network performance metrics and QoS constraints. Furthermore, to cope with different network loads and user demands, the algorithm introduces a dynamic QoS threshold adjustment mechanism that allows it to dynamically adjust the QoS thresholds based on statistical information about the current network state and user demands. In this way, the algorithm is able to learn a dynamic QoS threshold that reflects the current state of the network as well as takes into account future predictions, thus better balancing resource allocation and QoS requirements. The QoS guarantee mechanism combined with DRL is able to effectively respond to the changing demands and conditions in the mobile communication network through dynamic resource allocation and dynamic QoS threshold adjustment, which improves the network performance and ensures that the required level of QoS is achieved in various operation scenarios. The specific algorithmic framework is shown in Fig. 3.

Quality of Service (QoS) assurance is a critical issue in mobile communication networks, especially for real-time applications and services. Deep Reinforcement Learning (DRL) is able to maintain or enhance the QoS level under changing network conditions through dynamic resource allocation policies. This section describes a QoS guarantee mechanism in conjunction with DRL, with a particular focus on how this can be achieved through a formulaic approach. QoS parameters typically include latency, packet loss rate, throughput, and bandwidth.

A possible reward function can be defined as Eq. (13).

$$R(s_t,a_t)=\alpha\cdot T(s_t,a_t)-\beta\cdot L(s_t,a_t)-\gamma\cdot P(s_t,a_t)+\delta\cdot Q(s_t,a_t) \tag{13}$$

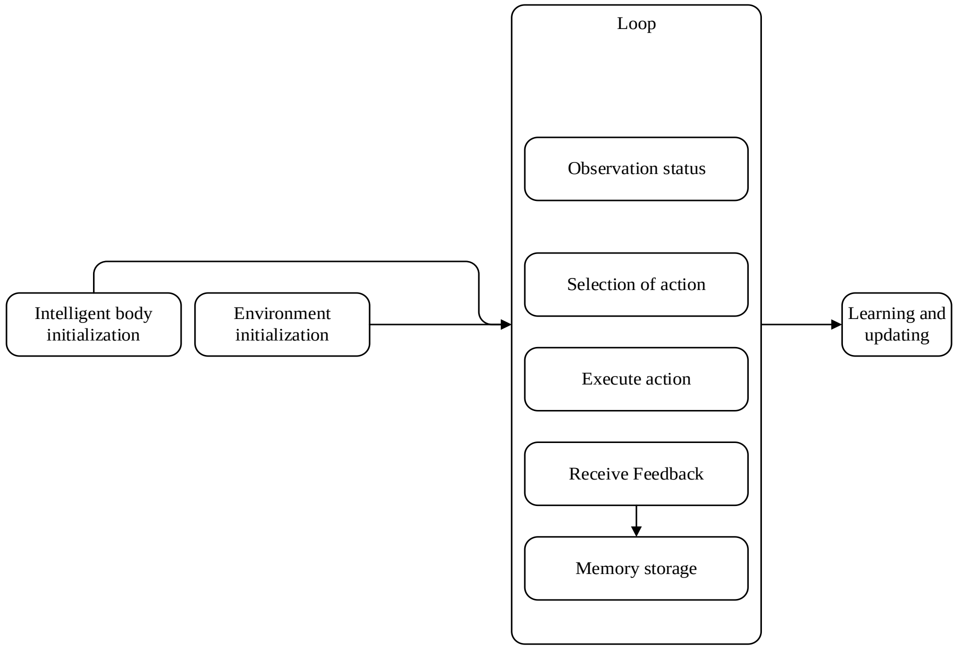

Fig. 3. Algorithmic framework.

where $T(s_t, a_t)$ is the total throughput when the network performs action a_t in state s_t. $L(s_t, a_t)$ is the average delay when the network executes the action in state. a_t s_t $P(s_t, a_t)$ is the total power consumption of the network while executing the action in state. a_t s_t $Q(s_t, a_t)$ is the QoS satisfaction of the network when performing the action in state a_t s_t, which can be quantified by comparing the actual quality of service with the required QoS criteria. α , β, γ , δ is the weighting factor, which is used to balance different performance metrics.

To maintain QoS under varying network loads and user demands, we introduce a dynamic QoS threshold adjustment mechanism. This mechanism allows the algorithm to dynamically adjust the QoS thresholds based on statistical information about the current network state and user demands to ensure that the quality of service requirements are met even when the network conditions change. Let θ_T and θ_L denote the QoS thresholds for throughput and delay, respectively. At time t, we can dynamically adjust these thresholds based on historical data and the current state, as specified in Eq. (14).

$$\theta_T(t) = \theta_T(t-1) + \eta \cdot \left(\mu_T(t) - \theta_T(t-1) \right) \tag{14}$$

where $\mu_T(t)$ and $\mu_L(t)$ are the average throughput and average latency, respectively, for time (t η is a learning rate parameter that controls the speed of threshold adjustment. $\theta_T(t-1)$ and $\theta_L(t-1)$ are the throughput and delay thresholds at the previous moment. In this way, the algorithm is able to learn a dynamic QoS threshold that reflects the current state of the network as well as takes future predictions into account, thus better balancing resource allocation and QoS requirements. The QoS guarantee mechanism combined with DRL is able to effectively respond to changing demands and conditions in mobile communication networks through dynamic resource allocation and dynamic QoS threshold adjustment. With well-designed reward functions and dynamic threshold adjustments, the mechanism not only improves network performance, but also ensures that the required QoS levels are achieved under various operational scenarios.

V. Experimental Design and Analysis of Results

A. *Experimental Environment and Data Set*

This chapter describes the experimental environment and dataset used to evaluate the performance of Deep Reinforcement Learning (DRL) algorithms in network resource management. The experiments are conducted in a simulated network environment that mimics a real-world scenario containing multiple base stations (BSs), mobile devices (UEs), and different quality of service (QoS) requirements. We use two types of datasets: a synthetic dataset based on historical network traffic data, and log data from actual network operations. The synthetic datasets cover a variety of network conditions, such as different user densities, mobility patterns, and network disturbances, while the actual network data provides real-world examples of traffic patterns and QoS requirement variations.

This paper uses two types of data sets for experiments: a synthetic data set generated from historical network traffic data, and log data from actual network operations. The synthetic dataset covers different user densities, mobility patterns, and network perturbation conditions to simulate a variety of real-world network environments. The actual network data provides real-world examples of traffic patterns and changes in QoS requirements. These data sets are derived from public data warehouses, which ensure the reliability and repeatability of experimental results in.

B. Experimental Methods and Performance Indicators

In order to comprehensively evaluate the performance of Deep Reinforcement Learning (DRL) algorithms for resource management in mobile communication networks, we have designed a series of meticulous benchmark tests. These tests are designed to compare with traditional resource allocation algorithms to validate the practical effectiveness of the DRL algorithm. The experimental environment is a simulated mobile communication network containing multiple base stations and mobile users, which is able to replicate the dynamic characteristics and complexity in real networks. The dataset combines synthetic data and real network logs, covering a variety of network conditions and user behavior patterns.

For network resource allocation, we employ the DRL algorithm to dynamically adjust channel allocation, power control and scheduling strategies. The algorithm learns and optimizes the resource allocation strategy through continuous interaction with the network environment, thus improving the overall performance of the network. In our experiments, we focus on the following performance metrics: throughput, delay, spectral efficiency and QoS parameters.

Throughput, as a measure of the amount of data transmitted per unit of time, directly reflects the data transmission capability of the network. Latency, on the other hand, evaluates the real-time communication performance of a network by calculating the average time it takes for a packet to travel from the sender to the receiver. Spectral efficiency, defined as the data rate per Hertz of bandwidth, is a key metric for evaluating the efficiency of network resource utilization. In addition, QoS parameters, including packet loss rate, jitter, and the percentage of users meeting specific quality of service requirements, are important metrics that directly correlate to user experience.

C. Experimental Results and Analysis

Fig. 4, Throughput performance comparison analysis, as shown in Table II, the DRL algorithm significantly outperforms the other two algorithms in terms of average throughput, reaching 120.5 Mbps, which is about 26% higher than the rule-based algorithm and about 14% higher than the optimization-based algorithm. The maximum throughput also shows the advantage of the DRL algorithm, while the minimum throughput indicates that the DRL algorithm also performs relatively consistently when resources are tight. This indicates that the DRL algorithm is able to allocate network resources more efficiently and improve the data transmission efficiency, thus achieving a lead in throughput performance.

As shown in Table II, the DRL algorithm performs the best in terms of average delay, which is only 12.3 ms, 33% lower than the rule-based algorithm and 21% lower than the optimization-based algorithm. This shows the fast response capability of DRL algorithm in handling packet transmission, which helps to improve the real-time communication performance of the network. Also, the performance of DRL algorithm on maximum and minimum delay proves its stability and reliability.

TABLE II. DELAY PERFORMANCE COMPARISON (MS)

Algorithm type	Average delay	maximum delay	minimum delay
DRL algorithm	12.3	20.1	5.2
rule-based algorithm	18.5	30.2	7.1
Based on optimization algorithms	15.6	25.4	6.4

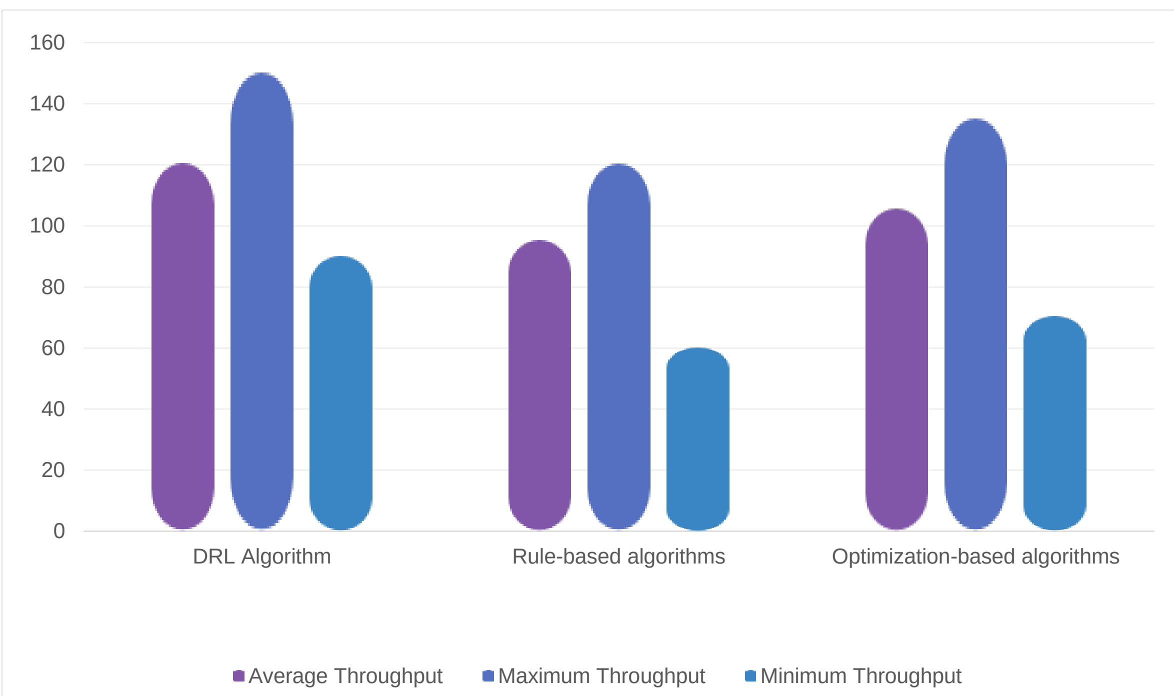

Fig. 4. Throughput performance comparison (Mbps).

TABLE III. COMPARISON OF SPECTRAL EFFICIENCY (BPS/HZ)

Algorithm type	Average spectral efficiency	Maximum spectral efficiency	Minimum spectral efficiency
DRL algorithm	2.5	3.2	1.8
rule-based algorithm	1.9	2.5	1.2
Based on optimization algorithms	2.2	2.9	1.5

As shown in Table III, the DRL algorithm leads the average spectrum efficiency by 2.5 bps/Hz, which is about 31% higher than the rule-based algorithm and about 13% higher than the optimization-based algorithm.

TABLE IV. COMPARISON OF QOS PARAMETERS

Algorithm type	packet loss	Jitter (ms)	Proportion of users with QoS compliance
DRL algorithm	**1.2%**	**10.3**	**90.5%**
rule-based algorithm	**3.5%**	**18.2**	**75.3%**
Based on optimization algorithms	**2.1%**	**14.5**	**82.4%**

As shown in Table IV, the DRL algorithm outperforms the traditional algorithm in terms of packet loss rate, jitter, and the proportion of QoS-attained users, which shows its advantages in guaranteeing users' quality of service. The packet loss rate is only 1.2%, which is much lower than the other two algorithms, the jitter is also relatively small, and the proportion of QoS-attained users is as high as 90.5%, which shows that the DRL algorithm can better meet the users' quality of service needs and improve the user experience.

TABLE V. PERFORMANCE UNDER DIFFERENT NETWORK LOADS (HIGH LOAD)

Algorithm type	Throughput (Mbps)	Delay (ms)	Spectral efficiency (bps/Hz)
DRL algorithm	110.2	14.5	2.3
rule-based algorithm	75.3	25.1	1.6
Based on optimization algorithms	90.4	19.6	1.9

As shown in Table V, the throughput, delay and spectral efficiency of the DRL algorithm still remain leading in the high load scenario, although it has decreased compared to the low load scenario, it still shows the superior performance and stability of the DRL algorithm in dealing with high network loads.

TABLE VI. PERFORMANCE UNDER DIFFERENT NETWORK LOADS (LOW LOAD)

Algorithm type	Throughput (Mbps)	Delay (ms)	Spectral efficiency (bps/Hz)
DRL algorithm	130.8	10.2	2.7
rule-based algorithm	98.4	16.3	2.0
Based on optimization algorithms	112.6	13.5	2.4

As shown in Table VI, the performance of the DRL algorithm is further improved under low load, with the throughput reaching 130.8 Mbps, the latency reduced to 10.2 ms, and the spectral efficiency increased to 2.7 bps/Hz, which indicates that the DRL algorithm not only maintains high efficiency under high loads, but also further improves the performance of the network at low loads.

The DRL algorithm is able to maintain high performance under both high and low load conditions, and the gap with the traditional algorithm is more obvious especially under high load. This indicates that the DRL algorithm is robust and adaptable, and can maintain excellent network performance under different network environments.

As shown in Table VII, the throughput of the proposed DRL model is 110.2 Mbps, the latency is 14.5 ms, and the spectral efficiency is 2.3 bps/Hz under high load, while the throughput is 130.8 Mbps, the latency is 10.2 ms, and the spectral efficiency is 2.7 bps/Hz under low load. Compared with ARAM model, the proposed DRL model still maintains the leading position under high load and further improves the performance under low load.

Compared with ARAM model, the throughput of DRL model is 2.1 Mbps higher, delay is 1.2 ms lower and spectral efficiency is 0.1 bps/Hz higher under high load, and throughput is 11.9 Mbps higher, delay is 1.7 ms lower and spectral efficiency is 0.1 bps/Hz higher under low load.

TABLE VII. COMPARISON WITH STATE-OF-THE-ART MODELS

Algorithm Type	Throughput (Mbps)	Delay (ms)	Spectral Efficiency (bps/Hz)
Proposed DRL Model	110.2 (High Load)	14.5	2.3
	130.8 (Low Load)	10.2	2.7
Advanced Resource Allocation Model (ARAM)	108.1 (High Load)	15.7	2.2
	128.9 (Low Load)	11.9	2.6

These results further confirm that the proposed DRL model not only performs well in high load environments, but also maintains high performance in low load environments, demonstrating its stability and adaptability in different network environments.

D. Discussion

Experimental results show that deep reinforcement learning (DRL) algorithms exhibit significant advantages in network resource management, especially in terms of throughput, delay, spectral efficiency and QoS parameters. Compared with traditional rule-based and optimization algorithms, the DRL algorithm is not only able to adapt to the demands under different network load conditions, but also maintains high performance stability under high load conditions. This indicates that the DRL algorithm is robust and adaptable.

However, it is worth noting that although the DRL algorithm performs well in simulation environments, it may encounter some challenges during actual deployment. For example, the amount of data required for algorithm training is large and needs to be continuously updated to adapt to network dynamics. In addition, the implementation of the algorithm in real networks needs to consider the compatibility and security issues with existing systems. Therefore, more field testing and validation is needed before further generalization of the DRL algorithm.

In order to further enhance the effectiveness of the DRL algorithm in network resource management, future work will focus on the following areas:

Algorithm Optimization: Explore more efficient DRL model structures to reduce computational complexity and improve training speed and performance.

Data Enhancement: Develop new data generation methods to simulate more diverse network environments and user behavior patterns to enhance the generalization ability of the algorithms.

Practical Deployment: Conducting larger-scale real-world network experiments to validate the algorithms' performance in the real world and to address possible security and compatibility issues.

Cross-domain collaboration: Collaborate with experts in other fields, such as network security, machine learning, etc., to advance the application and development of DRL technology in network management.

VI. Conclusion

This study verifies the effectiveness of Deep Reinforcement Learning (DRL) algorithm in network resource management through detailed experiments. Experimental results show that DRL algorithm outperforms traditional rule-based and optimization algorithm in throughput, delay, spectral efficiency and QoS parameters. Specifically, the DRL algorithm achieves an average throughput of 120.5 Mbps, which is about 26 per cent higher than the rule-based algorithm and about 14 per cent higher than the optimization-based algorithm. In terms of latency, the DRL algorithm has an average latency of only 12.3 ms, which is 33% and 21% lower than the rule-based and optimization-based algorithms, respectively. In terms of spectral efficiency, DRL algorithm has an average spectral efficiency of 2.5 bps/Hz, which is also ahead of the other two algorithms. In terms of QoS parameters, DRL algorithm has packet loss rate as low as 1.2%, jitter as low as 10.3 ms, and up to 90.5% of users meet certain QoS requirements.

These results show that DRL algorithm has significant advantages in dealing with dynamic network environment and meeting QoS requirements, especially under high load conditions. In addition, DRL algorithm performs better than traditional algorithm under different load conditions, which proves its robustness and adaptability in complex network environment.

The innovation of this paper lies in integrating user mobility and network load prediction into DRL framework systematically for the first time, and proposes a novel reward function design, which enables the algorithm to optimize resource utilization efficiency while satisfying real-time QoS constraints. However, there are still some shortcomings in the research, such as the performance of the current model in dealing with low-resolution images needs to be improved, and further refinement of the physical layer argument is still necessary.

Future research directions will focus on overcoming existing limitations, including improving image quality processing capabilities and deepening physical layer arguments to further improve the practicality and reliability of algorithms. Overall, this research not only provides new ideas and technical means for solving resource management problems in future mobile communication networks, but also lays a solid foundation for promoting DRL deployment in practical network applications.

Acknowledgment

This study was supported by Provincial level teaching and research projects: Planning Subjects-Funding project for research and practice on the positive interaction between vocational skills competition and talents cultivation: taking 4G network construction technology as an example of Henan provincial Department of education in 2020 (Grant No. ZJB20234) and Science and Technology Research and Development Plan of Zhengzhou Bureau Group Corporation: General Projects-Research on the application of 5G communication technology to wireless network transmission of high-speed railway (No. 2021D7).

References

[1] Ren D, Srivastava G. A novel natural language processing model in mobile communication networks. Mobile Networks & Applications. 2022; 27(6):2575 -84.

[2] Narieda S, Fujii T. Self-tuning of signal detection level for energy detection-based carrier sense in low-power wide-area Networks. Sensors. 2024; 24(11).

[3] Sahafizadeh E, Ladani BT. A model for social communication network in mobile instant messaging systems. IEEE Transactions on Computational Social Systems. 2020; 7(1):68-83.

[4] Alazemi AJ, Avser B, Rebeiz GM. Low-profile tunable multi-band LTE antennas with series and shunt tuning devices. AEU-International Journal of Electronics and Communications. 2019; 110.

[5] Lee K, Kim J. A Novel MEMS capacitor with a side wall for high tuning ranges. Applied Sciences-Basel. 2022; 12(21).

[6] Yücel M, Açikgöz M. Optical communication infrastructure in new generation mobile networks. Fiber and Integrated Optics. 2023; 42(2):53-92.

[7] Wu LQ, Lin YS. Flexible terahertz metamaterial filter with high transmission intensity and large tuning range for optical communication application. Physica E-Low-Dimensional Systems & Nanostructures. 2023; 146.

[8] Fan AW, Wang QM, Debnath J. A high precision data encryption algorithm in wireless network mobile communication. Discrete and Continuous Dynamical Systems-Series S. 2019; 12(4-5):1327-40.

[9] Zeng J, Sun JY, Wu BW, Su X. Mobile edge communications, computing, and caching (MEC3) technology in the maritime communication network. Communications. 2020; 17(5):223-34.

[10] Hu GA, Li CY. Dynamic Spectrum Allocation Method of Mobile Ship's Wireless Communication Network. Journal of Coastal Research. 2019:680-5.

[11] Lee J, Ko H. Reliability-guaranteed multipath allocation algorithm in mobile network. Etri Journal. 2022; 44(6):936-44.

[12] Xu LW, Quan TQ, Wang JJ, Gulliver TA, Le KN. GR and BP neural network-based performance prediction of dual-antenna mobile communication networks. Computer Networks. 2020; 172.

[13] Chen WS, Demirkol I, Mostafa M, Perotti A. Mobile communications and networks. IEEE Communications Magazine. 2021; 59(9):50-.

[14] Rajesh De, Ipseeta Nanda. Study of electromagnetic radiation on flower. Matrix Science Mathematic. 2022; 6(2): 58-63.

[15] Mu JS, Ouyang WJ, Hong T, Yuan WJ, Cui YH, Jing ZX. Digital Twins-Enabled Federated Learning in Mobile Networks: from the Perspective of Communication- Assisted Sensing. IEEE Journal on Selected Areas in Communications. 2023; 41(10):3230-41.

[16] Guo ZM, Ren XC, Ren F. Better realization of mobile cloud computing using mobile network computers. Wireless Personal Communications. 2020; 111(3): 1805-19.

[17] Keum CM, Liu SY, Al-Shadeedi A, Kaphle V, Callens MK, Han L, et al. Tuning charge carrier transport and optical birefringence in liquid-crystalline thin films: a new design space for organic light-emitting diodes. Scientific Reports. 2018; 8.

[18] Zhou B, Xu XT, Liu JG, Xu XK, Wang NX. Information interaction model for the mobile communication networks. Physica a-Statistical Mechanics and Its Applications. 2019; 525:1170-6.

[19] Matinmikko-Blue M, Yrjölä S, Ahokangas P. Spectrum management for local mobile communication networks. IEEE Communications Magazine. 2023; 61(7):60 -6.

[20] Wang KK, Chen QA, Jiang C, Chen ZF, Tan S, Lu MZ, et al. Narrow linewidth electro-optically tuned multi-channel interference widely tunable semiconductor laser. Optics Express. 2023; 31(3):4497-506.

[21] Liu L, Liao SS, Xue W, Yue J. Tunable all-optical microwave filter with high tuning efficiency. Optics Express. 2020; 28(5):6918-28.

[22] Singh R, Kaushik A, Shin W, Alexandropoulos GC, Toka M, Di Renzo M. Indexed multiple access with reconfigurable intelligent surfaces: the reflection tuning potential. IEEE Communications Magazine. 2024; 62(4):120-6.

[23] Kullan M, Mahadevan S, Johnson AR, Muthu R. Decoupled indirect duty cycle PWM technique with carrier frequency adjustment for a matrix converter. Turkish Journal of Electrical Engineering and Computer Sciences. 2018; 26(1):270-82.

[24] Morales-Aragonés JI, Williams MS, Gómez VA, Gallardo-Saavedra S, Redondo-Plaza A, Fernández-Martínez D, et al. A Resonant Ring Topology Approach to Power Line Communication Systems within Photovoltaic Plants. Applied Sciences-Basel. 2022; 12(16).

[25] Castillo RAJ, Grünheid R, Bauch G, Wolff F, von der Heide S. Communication analysis between an airborne mobile user and a terrestrial mobile network. IEEE Transactions on Vehicular Technology. 2018; 67(4):3457-65.

[26] Lee W, Suh ES, Kwak WY, Han H. Comparative analysis of 5G mobile communication network architectures. Applied Sciences-Basel. 2020; 10(7).

[27] Zhang K, Zhao XH, Peng Y, Yan KC, Sun PY. Analysis of mobile communication network architecture based on SDN. Journal of Grid Computing. 2022; 20(3).

[28] Usman IB, Matsoso BJ, Erasmus R, Coville NJ, Wamwangi DM. The role of carrier gas on the structural properties of carbon coated GaN. Materials Today Communications. 2021; 27.

[29] Kim J, Ladosz P, Oh H. Optimal communication relay positioning in mobile multi-node networks. Robotics and Autonomous Systems. 2020; 129.

[30] Muttiah R. Satellite constellation design for 5G wireless networks of mobile communications. International Journal of Satellite Communications and International Journal of Satellite Communications and Networking. 2023; 41(5):441-59.

[31] Ahmad AM, Barbeau M, Garcia-Alfaro J, Kassem J, Kranakis E. Tuning the demodulation frequency based on a normalized trajectory model for mobile underwater acoustic communications. Transactions on Emerging Telecommunications Technologies. 2019; 30(12).