

## Research Article

# Intelligent Resource Allocation Method for Wireless Communication Networks Based on Deep Learning Techniques

**Hancheng Hui** 

*China University of Geosciences (Wuhan), Wuhan, Hubei 430000, China*

Correspondence should be addressed to Hancheng Hui; 20191004352@cug.edu.cn

Received 14 September 2021; Revised 24 September 2021; Accepted 25 September 2021; Published 5 October 2021

Academic Editor: Guolong Shi

Copyright © 2021 Hancheng Hui. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, a deep learning approach is used to conduct an in-depth study and analysis of intelligent resource allocation in wireless communication networks. Firstly, the concepts related to CSCN architecture are discussed and the throughput of small base stations (SBS) in CSCN architecture is analyzed; then, the long short-term memory network (LSTM) model is used to predict the mobile location of users, and the transmission conditions of users are scored based on two conditions, namely, the mobile location of users and whether the small base stations to which users are connected have their desired cache states, and the small base stations select the transmission. The small base station selects several users with optimal transmission conditions based on the scores; then, the concept of game theory is introduced to model the problem of maximizing network throughput as a multi-intelligent noncooperative game problem; finally, a deep augmented learning-based wireless resource allocation algorithm is proposed to enable the small base station to learn autonomously and select channel resources based on the network environment to maximize the network throughput. Simulation results show that the algorithm proposed in this paper leads to a significant improvement in network throughput compared to the traditional random-access algorithm and the algorithm proposed in the literature. In this paper, we apply it to the fine-grained resource control problem of user traffic allocation and find that the resource control technique based on the AC framework can obtain a performance very close to the local optimal solution of a matching-based proportional fair user dual connection algorithm with polynomial-level computational complexity. The resource allocation and task unloading decision policy optimization is implemented, and at the end of the training process, each intelligent body independently performs resource allocation and task unloading according to the current system state and policy. Finally, the simulation results show that the algorithm can effectively improve the quality of user experience and reduce latency and energy consumption.

## 1. Introduction

Wireless communication technology has developed from the first generation of mobile communication technology, which emerged in the 1980s, to the fifth generation of mobile communication technology, from the beginning of satellite communication, radio transmission, and then developed into intelligent terminal devices, which makes wireless communication technology not only able to provide general voice communication or simple data services but also is fully integrated into people's daily life, becoming an indispensable part of today's society [1]. It has become an integral part of today's society, making people's lives more convenient and richer. The introduction of fourth-generation wireless tech-

nology provides a platform for moving towards higher data rates and reliable communication standards [2]. The growing demand for data services has led to the development of global interoperability for microwave access and the long-term evolution of wireless communication standards. However, with the increasing number of smart terminal devices, the exponential increase in wireless data demand and usage, and the introduction of emerging multimedia applications, it is very difficult to support the rapidly growing data rates and connected devices in the current 4GLTE cellular system. In the traditional radio access network architecture, the base station controller performs data transmission and reception by controlling the RF units, while the wireless network scales the system capacity mainly by expanding the number and

density of base station deployments; thus, the expansion of the number and density of base stations also poses many challenges and difficulties [3]. On the one hand, it is mentioned in the literature that in the traditional wireless access network architecture, the base stations mainly consume energy, and the increase in the number of base stations leads to a large increase in energy consumption; on the other hand, the traditional wireless access network architecture cannot be applied to the multiscenario and multiservice requirements of 5G wireless networks; therefore, a new network architecture that can achieve green energy saving and flexible deployment is urgently needed [4]. With the continuous increase of smart terminal equipment, the increase of wireless data demand and usage index, and the introduction of emerging multimedia applications, it is very difficult for the current 4GLTE cellular system to support the rapidly increasing data rate and connected devices. The highly dynamic nature of the network refers to the frequent changes in network topology and mobile data services; through the design of reasonable resource management techniques, users can access the network adaptively according to the current network topology and services, without low communication rate and quality of service due to frequent changes in network topology or frequent changes in mobile data services [5]. The variable density feature of the network helps to adapt the network to user flows at different densities, ensuring that the throughput of the network is maintained in the right range when the density of user flows increases sharply.

With the advent of the Internet era, marine wireless services are diversifying and the demand is growing exponentially, and ship users are putting forward higher requirements for communication service quality. Therefore, to meet the growing demand for services, the “intelligent” management of maritime information systems is realized [6]. The distribution of users and service equipment is not equal, and full coverage access is not possible. Different infrastructure networks have a different coverage in different sections of the sea, and a single network cannot meet all the communication needs, which greatly exceeds the supply of the network. In addition to deploying wireless networks at sea, it is difficult to arrange network base stations in a fixed location on the sea surface, except for using existing lighthouses and other fixed locations to deploy network nodes, so designing a network architecture adapted to the maritime environment is the primary challenge for the current deployment of wireless networks at sea [7]. To handle the requests of increasing data streams and break through the limitations of maritime communication, today’s maritime communication networks cannot meet the interoperability between networks as they deployed and interact through multiple layers and heterogeneous architectures, creating a maritime communication environment where many heterogeneous networks coexist to support a large and diverse user base. To meet the everincreasing business needs, the “intelligent” management of the maritime information system is realized. The distribution of users and service equipment is not equal, and full coverage access cannot be realized. The quality of service when communicating at sea cannot be met. Because ships are always in motion scenarios, the

network environment changes rapidly, and the user density is also unevenly distributed in the marine network, traditional communication networks can only meet the QoS requirements for real-time data, packet loss rate, and time delay in a single way, and cannot do comprehensive consideration. For example, real-time information-related service navigation channels have higher requirements for reliability and transmission delays, while the Internet and video services will be more concerned about bandwidth [8]. The issue of optimal allocation and scheduling that puts efficiency at the forefront of various network resources becomes a top priority. Therefore, the development of efficient solutions to deal with large-scale optimization problems becomes a new task for future wireless networks. The Internet and video services will pay more attention to bandwidth. Among various network resources, the optimization of allocation and scheduling, which puts efficiency in the first place, has become a top priority. Therefore, the development of efficient solutions for large-scale optimization problems has become a new task for future wireless networks.

To address the bandwidth bottleneck in centralized federation learning and to ensure the convergence performance of federation learning, at each synchronization time node, each node randomly selects a few nodes to transmit part of the model. Our goal is to maximize the bandwidth capacity between nodes, and to improve the convergence performance of the solution, “model replicas” are introduced to ensure that sufficient information is obtained from different nodes during the aggregation process. At the same time, we propose a synchronization mechanism at the granularity of model fragmentation. We “split” the model into a collection of slices that contain the same number of model parameters and do not overlap with each other. Nodes perform slice level updates by aggregating local slices with corresponding slices from other  $k$  nodes. To further speed up the convergence, we propose a bandwidth-aware node selection method based on a bandwidth-aware node selection method using the epsilon-greedy algorithm, where nodes always monitor and estimate the average bandwidth between peer nodes (Peer) and select the peer node with a high probability of fast transmission speed to transmit the model slice. Finally, we implemented Balcombe, a prototype system based on a model slicing-based gossip policy and a bandwidth-aware node selection policy, and experimental results show that the design significantly reduces the total training time by a factor of up to 18, while keeping the model accuracy constant.

## 2. Current Status of Research

With the development of artificial intelligence techniques, artificial intelligence algorithms have been applied to complex decision-making problems such as resource allocation [9]. Currently, AI techniques such as machine learning and deep learning can extract useful information from wireless systems, learn and make decisions from dynamic environments, and are considered potential solutions to complex and typical problems in future wireless networks that were previously difficult to solve [10]. Machine learning can not

only use data analysis to enhance the situational awareness and overall network operation of wireless networks but also effectively drive wireless network optimization. In addition, machine learning can play a key role in the physical layer of wireless networks. The application of machine learning in wireless networks has been studied, and research results and tutorials have been published [11]. The literature proposes some emerging learning frameworks suitable for IoT applications and summarizes the advantages, limitations, IoT applications, and key results of machine learning, sequential learning, and reinforcement learning. A new learning-based approach to wireless resource management was proposed in literature [12]. The core idea is to consider the input and output of the resource allocation algorithm as an unknown nonlinear mapping and to approximate it using deep neural networks. Literature [13] proposes a real-time multi-intelligent reinforcement learning method to manage the aggregated interference generated by multiple wireless area network systems. Literature [14] proposes an enhancement learning-based downlink power control scheme for nonorthogonal multiple access without considering interference and wireless channel parameters. The literature [15] proposes a Q-learning-based transmission scheduling mechanism that uses the idea of deep learning to solve the problem of how to transmit packets with different buffers over multiple channels to maximize the system throughput. Literature [16] uses deep learning to process wireless channels in an end-to-end manner and implicitly estimates the channel state information (CSI) and directly recovers the transmitted symbols based on the deep learning approach. Literature [17] proposes a design scheme for a cognitive engine and implements the learner in the cognitive engine using a neural network-based learning algorithm.

In addition to this, literature [18] proposes a cell selection-based aggregation method for multi-interface heterogeneous networks; for users at the edge of two LTE cells, the algorithm focuses on allowing the user to select the LTE cell where the LWA mode can be performed. Literature [19] proposes a self-optimizing algorithm that controls how data services are sliced by adjusting the aggregation mode between LTE, WLAN, and its results show that the proposed algorithm can derive the optimal control parameters for each cell under different load cases. The goal of the adaptive transmission mode selection system in the algorithm is to assign transmission modes to each user in the cell from the point of view of cell load and user throughput so that the resources of both technologies can be optimally utilized. For this purpose, the base station assigns LWA mode to a user if the user is in a good condition under both networks. When the user is in a poor LTE network, a full switch to WLAN transport mode results in a performance gain. The signal-to-noise interference ratio (SINR) threshold determines whether the WLAN transport mode is used, while the received signal strength indicator (RSSI) threshold determines the perceived range of the WLAN network.

If the current rate of a single connection cannot meet the smooth video playback requirements, then the user device can send a request to the core network manager, which selects a suitable access network for the user and establishes

a second connection with the user based on a built-in reinforcement learning policy to achieve the purpose of satisfying the user's quality of service. This paper presents a high dynamic density variable heterogeneous network resource management architecture, and the functions of the core network manager, access network reconfiguration manager, and user terminal manager in the network resource management architecture and the format and main workflow of the transmitted data packets in the network resource management architecture. And a simple user-adaptive access algorithm based on reinforcement learning is implemented based on the heterogeneous network resource management architecture to improve the utilization of network resources and user quality of service.

### 3. Deep Learning Techniques for Intelligent Resource Allocation in Wireless Communication Network Analysis

*3.1. Deep Learning Algorithms for Intelligent Resource Allocation Analysis.* DQN's experience playback uses a uniform distribution, and uniformly distributed sampling does not make efficient use of the data. Because the subject's experience is the experienced data, but this data is not equally important for training, the subject learns more efficiently in some states than in others. The idea of prioritized experience playback is to break the subject's original experience playback using uniform sampling and instead give greater sampling weights to the system states that are learned efficiently. An ideal criterion in the sampling of sampling weights problem is that the more efficient the subject learns at the system state, the larger the weight assigned to the samples corresponding to this system state [20]. Reinforcement learning also needs to explore the environment, learn from the effective behavior feedback that has been made, to generate rewards, and update the exploration actions, that is, trial and error search, to achieve the goal. The larger the difference between the Q value function at the system state and the TD objective function value, i.e., the larger the TD deviation, the more updates the subject must make when learning, and therefore, the more efficient the learning is at that location.

Reinforcement learning is the process by which an intelligence learns by "trial and error" and is guided to its next action by the rewards obtained through interaction with the environment, with the goal of maximizing the value of the rewards obtained by the intelligence. Reinforcement learning is mainly used to solve sequential multistep decision-making problems where the system (RLS) must rely on its own experience to learn due to the low information provided by the external environment, where the RLS gains experience in an action-evaluation environment and improves the action plan to fit the current environment. The characteristics of reinforcement learning delayed reward and trial-and-error search makes it an important branch of machine learning research. The intelligent body learning process needs to look at the whole picture and consider not only immediate rewards but also long-term cumulative

rewards, i.e., delayed rewards. In addition, reinforcement learning requires exploring the environment and learning from the feedback of valid behaviors made by itself to generate rewards and update exploratory actions, i.e., trial-and-error search, to achieve the final goal. In the iterative interaction between the decision-making behavior of the RLS and the state of the environment feedback as well as the evaluation, reinforcement learning continuously modifies the mapping strategy from the state to act in a learning manner to achieve the goal of optimizing the system performance. Reinforcement learning mapping from the state of the surrounding environment to the action behavior enables the strategy chosen by the intelligence to obtain the maximum reward value so that the external environment evaluates the RLS in some sense (or the operational performance of the whole system) as optimal.

$$R_t = \sum_{i=t}^T \gamma^{(i-t)} \cdot \gamma_i, \quad (1)$$

$$\begin{cases} Q_{k+1}(s_t, a_t) = Q_{k+1}(s_a, a_s) - s_a \cdot a_s, \\ \delta_k = r_{t+1} - \gamma \min Q_{k+1}(s_t, a_t). \end{cases} \quad (2)$$

Reinforcement learning is very effective in dealing with problems of high small-scale complexity in real-world applications and is used in many autonomous learning problems, such as autonomous driving, robot manipulation, and other humans. However, it is not as effective when facing real-world large-scale, complex problems, which limits the development of reinforcement learning, so the idea of combining deep learning with reinforcement learning is introduced, as shown in Figure 1.

Cellular mobile communication is the use of cellular wireless networking to connect between terminals and network devices through wireless channels, which in turn enables users to communicate with each other during their activities. Its main feature is the mobility of the terminal with cross-area switching and automatic roaming across the local network [21]. In the small cell network, the coverage of small base stations is small, and the network throughput will be greatly affected by the movement of users. It is necessary to locate users in advance. Unlike the traditional one-step REL technology, it can store information for a long time. It has the ability to store the state of the previous hidden layer network. In a conventional cellular network consisting of macrobase stations, the movement of users has less impact on the network throughput due to the wide coverage area of macro base stations. In contrast, in small-cell networks, the small base stations cover a smaller area and the network throughput is affected more by user movement. From the previous section, it is known that each small base station has a certain caching capacity and the cached content of each small base station is not necessarily the same. Therefore, during a user's movement, there is no guarantee that all the small base stations connected to its mobile route have the requested content of that user on their caches. This section considers the impact on network throughput in terms of both the user's mobile location and whether the small base

station to which the user is connected has the cache state it needs, by which the small base station selects the few users with optimal transmission conditions.

$$Q^*(s_{\text{sys}}, a) = Q(s_{\text{sys}}, a) - a[R_{\text{once}} - \gamma \max Q_{k+1}(s_t, a_t)], \quad (3)$$

$$\begin{cases} L(\theta) = E[Q_{k+1}(s_a, a_s) - s_a \cdot a_s], \\ Q_k = R_{r+1} + \gamma \min Q_{k+1}(s_t, a_t). \end{cases} \quad (4)$$

As the speed of mobile communication is increasing, the variety of mobile communication services is growing, and data services have replaced traditional voice services as the mainstream services, the research on user mobility is not only limited to seamless switching and roaming but also experts are turning to the combination of user mobility and caching technologies. Therefore, in small-cell networks, it is particularly important to study the combination of user mobility and caching technology. In a traditional cellular network consisting of macrobase stations, user mobility has less impact on network throughput because the macrobase stations cover a wide area and have enough time to locate users in advance. Next, this paper allocates spectrum resources based on the established unweighted conflict graph. This method determines the user allocation order in a distance-aware manner and tries to allocate orthogonal subchannels for users that may cause strong interference with each other in the conflict graph. In the case of success, the user is allocated the subchannel with the largest net income to eliminate strong interference between users. In contrast, in small-cell networks, where the coverage area of small base stations is small, the network throughput is affected by the user's movement, and the user needs to be in advance. Unlike the traditional one-step RL technique, which can store information for a long time, where the memory can store previous hidden layer network states and can predict sequences of future network states, it is very powerful in predicting time series. The prediction for a given time step is affected by the network state of the previous time steps and the current input, so using LSTM to predict the user's location compounds the requirements of this paper, as shown in Figure 2.

The surge in demand for machine learning services has led to the investigation of adaptive solutions to reduce computational consumption while still providing satisfactory quality of service. The trade-off between accuracy and latency is exploited to allow consumers to choose the level of tolerance in machine learning as a service platform, thus allowing consumers to sacrifice the quality of the service's results by using a different version of the model, thereby improving other aspects of the service's quality, such as service response time and invocation cost. Interference management is an important tool to improve the spectrum efficiency of wireless networks in the resource allocation process, which is especially important for ultradense networks with severe interference. In this paper, a centralized user-centric resource allocation algorithm based on a coalition formation game is proposed. The coalition formation

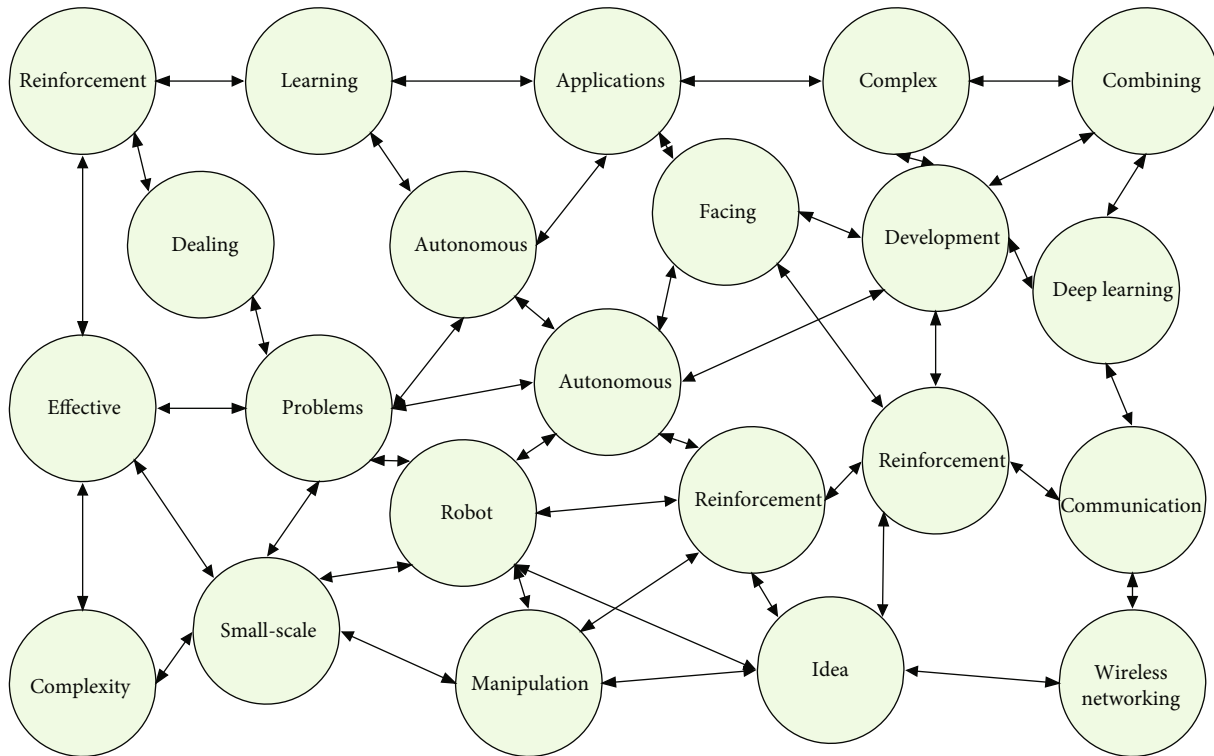


FIGURE 1: Deep learning resource intelligent allocation algorithm.

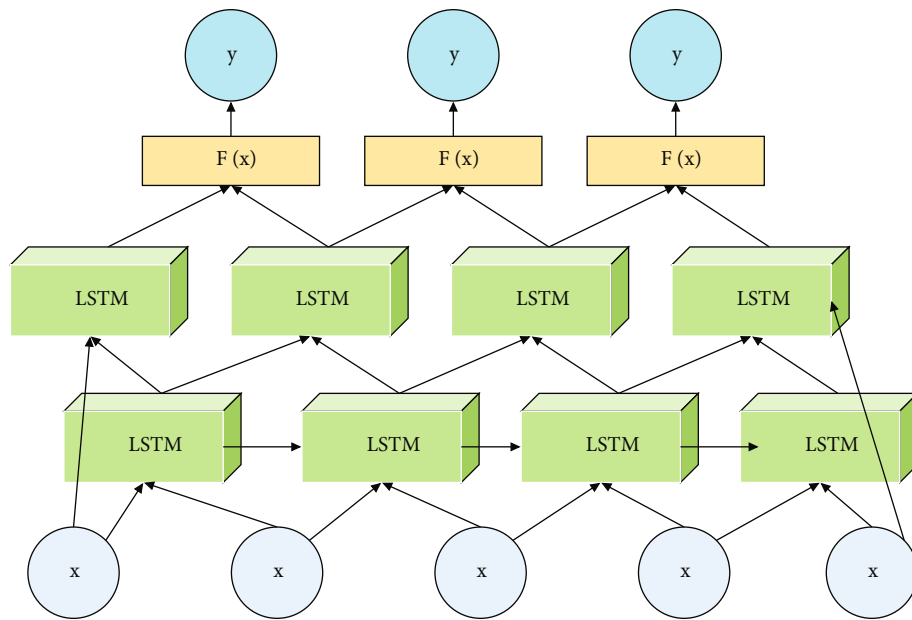


FIGURE 2: LSTM network.

game is used to find the optimal user coalition formation method under the defined resource allocation principle.

And the unprivileged conflict graph reflecting the interference relationship is constructed based on the network division (i.e., the result of coalition formation). Then, this paper performs spectrum resource allocation based on the established unweighted conflict graph, which determines

the user allocation order in a distance-aware manner and tries to allocate orthogonal subchannels for users in the conflict graph that may have strong interference with each other. The subchannel with the highest net gain is assigned to users in case of unavailability of free subchannels to eliminate strong interuser interference. In addition, to overcome the limitation of “only one subchannel per user” in the

previous coalition formation game, this paper proposes a low-complexity supplementary allocation algorithm to allocate the remaining subchannels to improve the spectrum efficiency of the system.

$$y_\tau = \sigma(W'h_\tau^2), \quad (5)$$

$$h_\tau = o_\tau \tanh(C_\tau^2). \quad (6)$$

To solve timing-related problems, such as predicting the next frame of a video or the next word in a text, traditional neural networks cannot work because they cannot “remember,” so recurrent neural networks are derived based on neural networks. In traditional neural networks, the output of the hidden layer at the next moment is only related to the input of the hidden layer at the next moment, while in recurrent neural networks, the output of the hidden layer at the next moment is also related to the output of the hidden layer at the previous moment, thus introducing the concept of temporal order. However, when the recurrent neural network expands too deeply in the temporal order, it tends to lead to the gradient dispersion problem, which makes the parameters of the neural network stop training and updating, so in practical applications, recurrent neural networks are rarely used directly, but a variant of recurrent neural networks. The difference between LSTM and recurrent neural networks is mainly the main difference between LSTM and recurrent neural network is that LSTM has an additional hidden parameter for recording the state of LSTM, which is generally called cell state. The LSTM also introduces a complex gating mechanism to automatically learn which information should be retained or forgotten by the cell state [22]. The gating mechanism of the LSTM contains three different types of gates, namely, input gates, forgetting gates, and output gates. The input gate determines which new information the cell state should retain, the forgetting gate determines which information the cell state should forget, and the output gate determines which information from the cell state is used as the output of the LSTM.

$$h_{gm}^m = \frac{1}{U} \sum_{u=1}^U h_{gmu}^n. \quad (7)$$

The value of the minimum correlation coefficient affects the performance of the user clustering algorithm. If the value of the minimum correlation coefficient is set too large, the number of users in the same cluster will be too small, which greatly reduces the network performance, and if the value of the minimum correlation coefficient is set too small, the beamwidth will be too large, which greatly increases the difficulty of beamforming. There is very little research on this aspect of the minimum correlation coefficient in domestic and international literature, and the minimum correlation coefficient is usually set as a constant.

*3.2. Experiment on Intelligent Resource Allocation for Wireless Communication Networks.* Consider a D2D-assisted cloud hybrid system for efficient offloading of ser-

vices and resource management. Each user finds the neighboring user that interferes most strongly with each other from the current user's interference list and tries to join the coalition in which that neighboring user is located, resulting in a new network partition. Subchannel assignment is performed under the new network division, and the potential gain (total system throughput) is calculated. When the packet arrival rate of data packets reaches 1.5 or above, the pressure of the buffer gradually increases. In order not to affect the data transmission, the system moves against the six-watt macro of the network between the right to increase the throughput. Attempts are made one by one to join the current user to the coalition where the neighboring strongly interfering user is located in the order of interference descending, and the above attempts are stopped when the total system throughput is improved, and the optimal network division and the maximum gain are updated to the current network division and gain. The merge or split operation occurs when and only when the system gain is improved. The game convergence reaches stability when no user can obtain higher system gain by the merge or split operation. Output the recursive kernel of the game at this point: the optimal network partition, the network again, and the result of subchannel assignment at this point. Each user joins the coalition of interfering users in its interference list in descending order of interference. Thus, strong interference in a wireless network can be suppressed by dividing users that are likely to interfere strongly with each other into a coalition and allocating orthogonal resources to users within the same coalition whenever possible. The process of such attempts continues in a 3-determined order until the game converges and reaches stability, at which point the most suitable network partition has been obtained for a given subchannel allocation strategy. It should be noted that this paper adopts the centralized architecture of C-RAN, where the control center “cloud” stores the network node location information collected from each base station. During the game, the stored network node location information is used to determine the proposed grouping of each user, and the road loss model and distance information are used to estimate the throughput and then determine whether the new coalition formation method can improve the network performance and whether the game updates the network division and its corresponding benefits, as shown in Figure 3.

Previous articles exploring how reinforcement learning can do user access or resource management in communication systems have been limited to the process of establishing connections to only one type of network. However, for future heterogeneous networks where multiple types of networks coexist, it has become a promising trend for users to aggregate data traffic from different networks to improve their communication rate, and the “fragmented” remaining resources in the communication system can be utilized by users in this way, and the utilization of the network will be greatly improved compared to the user. The utilization of the network will greatly improve compared with the single connection method. For the multiport resource aggregation technique, represented by the dual connection technique, many papers explore it in-depth and give algorithms for

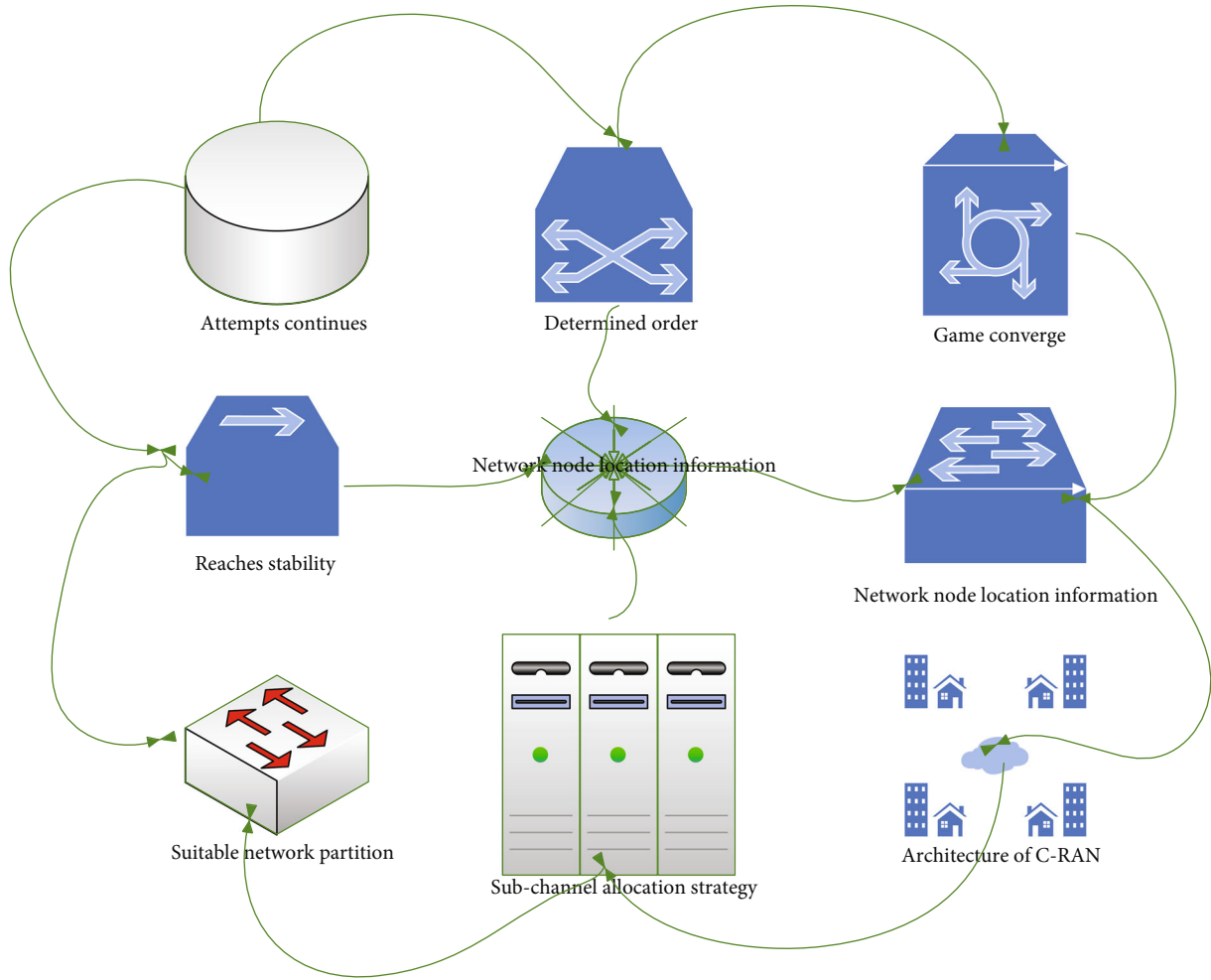


FIGURE 3: System scenario.

the maximum utilization that can be achieved when the user supports the dual connection transmission mode in the system. However, some algorithms, although reaching the theoretical upper bound on the system utilization, are required to search for the optimal solution of the system utilizing optimized traversal or multiple iterations and therefore have high complexity, and both the user access process and the process of resource control in the actual system require short time delays, so these algorithms are not of practical significance. To solve the above problems, the DQN-based subscriber access multipoint technique proposed in this section will be analyzed by simulation results to illustrate the gain of the algorithm for the communication system and its feasibility in the real system.

Assuming the cell simulation scenario, the WLAN AP and microbase station locations are obtained by uniformly randomly scattering points within the cell area, the corresponding system utilization can be obtained by simulating different user number scenarios using the DQN-based user multipoint access algorithm, and the simulation parameters are shown in Table 1.

To represent the gap between the system utilization of dual connection transmission and single connection trans-

TABLE 1: Network simulation parameters.

Cellular network parameters	Values
Channel bandwidth	11.34
Noise power	2.76
Macrocell base station transmit power	2.03
Microbase station transmit power	2.04
Path loss between macrobase station and user	9.57
Path loss between microbase station and user	9.64
Lognormal distribution shadow fading	11.69

mission, this paper simulates the relationship between the number of users and the utilization of the system when all the users are single connection; to represent the gap between the DQN-based user access algorithm and the traditional user access algorithm, this paper simulates the user nearest neighbor access algorithm and the signal-to-noise ratio (SNR) maximum access algorithm, respectively. To represent the gap between the DQN-based user access algorithm and the traditional user access algorithm, we simulate the user nearest neighbor access algorithm and the signal-to-

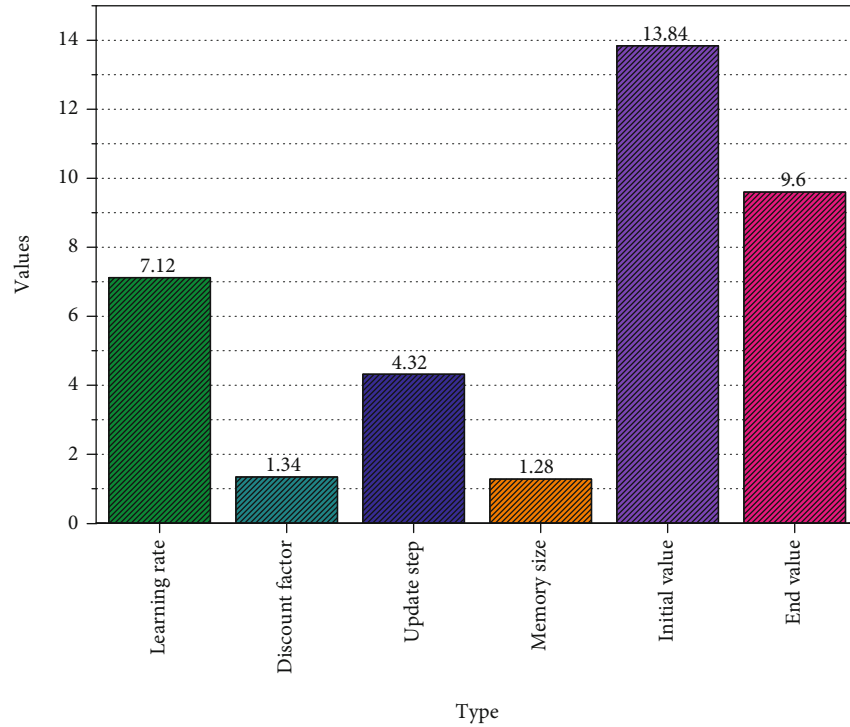


FIGURE 4: DQN model parameters.

noise ratio (SNR) maximum access algorithm, where the user nearest neighbor access algorithm indicates that each user selects the nearest microbase station to establish the second connection, and the SNR maximum access algorithm indicates that the user selects the microbase station with the largest SNR to establish the second connection. To represent the difference between the DQN-based user access algorithm and the theoretical upper bound, the matching-based proportional fair user dual connection access algorithm is also simulated in this paper, as shown in Figure 4.

From Figure 4, it can be seen that the system utility tends to decrease with the increase of the number of users, and it can be observed that the system utility is the worst when all users take a single connection; the system utility is the highest when all users adopt matching-based dual connection access algorithm; the performance of reinforcement learning-based user access algorithm and nearest neighbor entry algorithm and SNR maximum access algorithm is in between the user single connection access algorithm and matching-based user access algorithm, the performance of the nearest neighbor access algorithm and SNR maximum access algorithm are basically similar, and the DQN-based user access algorithm is better than the nearest neighbor access algorithm and SNR maximum access algorithm, but there is still some gap compared to the optimal performance curve [23]. As can be seen from Figure 4, the DQN-based user access algorithm curve continues to lie between the matching-based proportional fair user access algorithm and the traditional user access algorithm, except for the case when all users use a single connection, where the system utility increases with the number of small cells. When all users use a single connection, as the number of small cells

increases, increased users will leave the macro base station with high probability and switch to a single connection with small cells, and since the transmit power of small cells is much smaller than that of microcells, the throughput of users accessing small cells decreases; thus, the system utility tends to decrease with the increase in the number of small cells.

The matching-based proportional fair user access algorithm achieves a locally optimal solution to the system model but at the cost of increased computational complexity, while the DQN-based user access algorithm is well balanced compared to the nearest neighbor access algorithm and the SNR maximum access algorithm as it improves the system utility at the expense of only a small fraction of the latency.

## 4. Analysis of Results

*4.1. Deep Learning Algorithm Performance Results.* Overall, although the reinforcement learning model requires a lot of learning time in the early exploration of the system, when the parameters of the neural network in the AC framework slowly converge to a stable value as the learning rate decreases, for any kind of user distribution, the reinforcement learning model based on the AC framework can quickly give a reasonable user access and resource matching scheme, with a performance close to that of the matching-based local optimal access algorithm, but with much delay savings, which means that AI techniques can replace traditional engineering optimization methods to bring gains in throughput as well as long-term downlink utilization to the communication system, as shown in Figure 5.



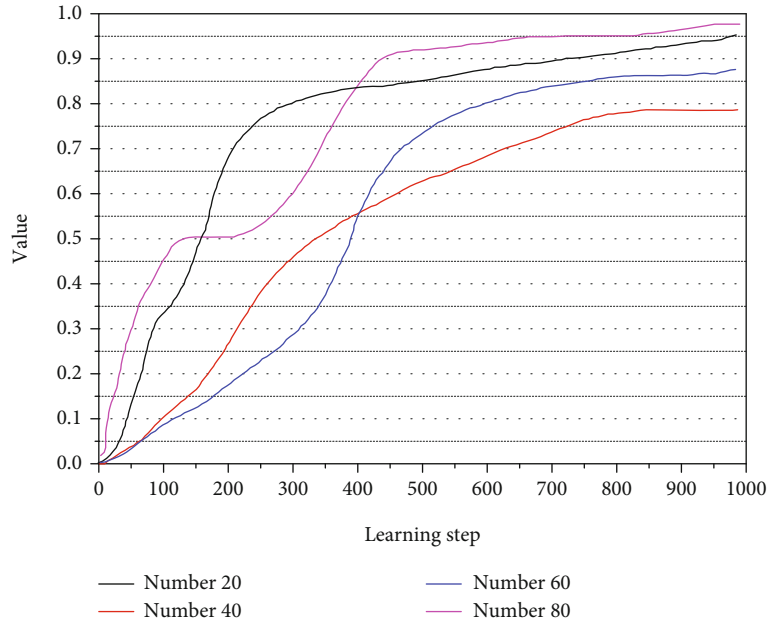


FIGURE 5: Convergence plot of the reward function.

Figure 5 compares the average throughput and average utility values of the three algorithms. The utility value indicates the link utilization of the system during data transmission. When the packet arrival rate increases, the system helps it to reduce the cache pressure in the form of increasing the outgoing volume. However, this practice is relatively energy-intensive. The proposed algorithm in this paper normalizes the Q value so that a slight increase in Q value produces a large fluctuation and thus transmits more packets, so the S-DQN algorithm transmits more packets and has higher throughput. As the data volume increases, the cache pressure on each network node increases. For routing, the traditional routing algorithm does not consider whether other channels are in congestion and just makes the decision of the shortest path to the source node with self as the core, which will lead to a significant decrease in the link utilization of the system. The pressure on the cache gradually increases when the packet arrival rate reaches 1.5 or more. In order not to affect the data transmission, the system day-activates the inverse companion right between the net of six Walden macros, thus thinking of high spit volume. At the same time, as the modulation method changes, the agent increases the frequency of interaction with the environment, thus increasing the link utilization. Moreover, in Figure 6, the DQN algorithm does not compare in the same plot as the data packet arrival rate increases due to its high computational complexity, resulting in relatively low average utility values.

Since Dijkstra's algorithm sometimes falls into local optimum, which results in the network resource allocation scheme given by the agent not being the best policy and the average energy consumption is too high. Dijkstra's algorithm consumes more energy compared to the other two AI algorithms. Due to the limited cache capacity, the energy consumption curve generated by the S-DQN algorithm

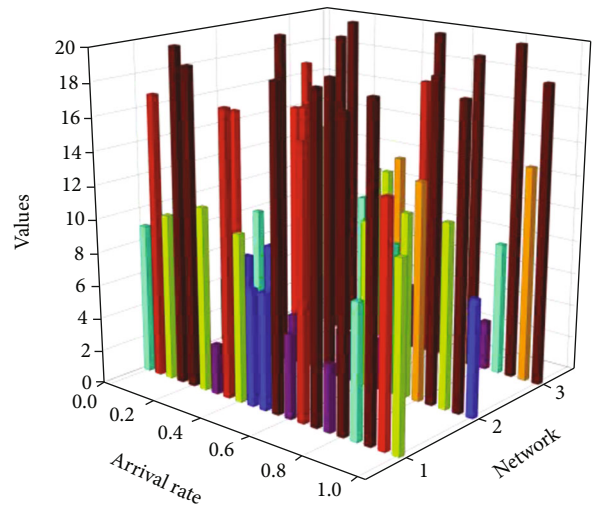


FIGURE 6: Throughput (grouping) comparison.

tends to grow slowly and then stabilize. When the amount of data reaches the load limit of the cache, the cache pressure does not increase and eventually stabilizes. Compared with the performance of the coloring method 1/2/3, the performance of this method is improved by 43.88%, 62.00%, and 88.86%, respectively. In addition, the corresponding algorithms are also simulated when the network load is 50%, 70%, and 90%. Dijkstra's algorithm also leads to a decrease in transmission efficiency when the data packet arrival rate increases, and packets are not transmitted in the time leading to packet loss. The S-DQN algorithm produces much lower latency than Dijkstra's algorithm. This is because the S-DQN algorithm makes the system throughput higher when the data packet arrival rate increases, putting more

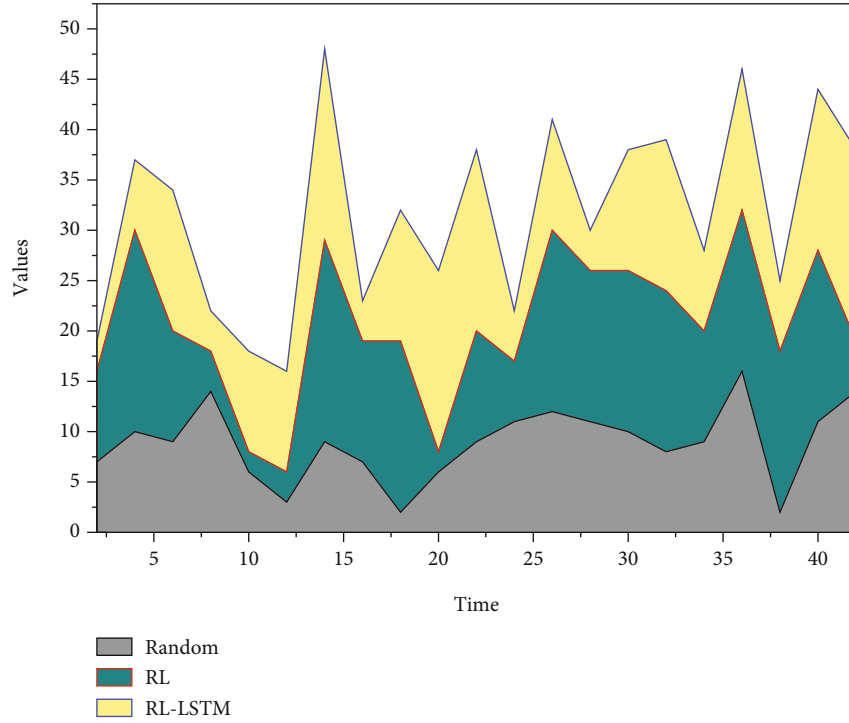


FIGURE 7: Plot of energy efficiency versus time.

pressure on the cache and forcing the relay network nodes to choose higher-order modulation to transmit data whenever possible, thus guaranteeing the real-time performance of the system. Since the s-DQN algorithm outputs the probability of the Q value, it is equivalent to quantizing the Q value, reducing the range of fluctuations in each state transfer, and thus reducing the delay jitter.

**4.2. Experimental Results of Intelligent Resource Allocation for Wireless Communication Networks.** Figure 7 reflects the relationship between the energy efficiency of the system and time. The horizontal axis is time, and the vertical axis is the energy efficiency of the system. Overall, it appears that the energy efficiency of the RL-LSTM-based resource allocation algorithm is significantly higher than that of the RL algorithm and the random access algorithm. It is worth noting that the energy efficiency of both the RL-LSTM algorithm and the RL algorithm gradually decreases as time increases. This is because as the RL-LSTM algorithm serves increased users, the energy efficiency of the system decreases for all users connected to the small base station within the coverage area of the small base station for which the transmission condition is best has been completed, and the small base station may have to serve users with longer transmission distances.

As the number of small base stations increases, the proportion of the load volume served by each small base station decreases. In addition, reducing the number of subchannels leads to a decrease in the proportion of the load volume served. Although the number of subchannels is not a player in the game, they affect the choice of spectrum allocation action for each small base station. As the number of sub-

channels increases, the action space of the channel selection vector increases, thereby increasing the amount of load served by the small base station.

When the number of subchannels is less than 10, the SAR value decreases faster; when the number of subchannels keeps increasing, the SAR value decreases gradually slower. In other words, combined with Figure 7, because the network has already reached the optimal subchannel allocation rate, continuing to divide the system bandwidth into more subchannels does not continue to increase the total throughput. Since too much interference reduces the total throughput, no more spectrum resources should be allocated to users after the optimal subchannel allocation rate is reached. Also, the subchannels of the allocated users when the optimal subchannel allocation rate is reached should not be removed because it will lead to underutilization of the subchannels and thus reduce the total network throughput. Moreover, since subchannel allocation is independent of network size and network density, the variation curve of the subchannel allocation rate with the number of subchannels holds for different network densities. Therefore, in a sparse network scenario, a suitable number of subchannels can be found at a low computational cost and applied to a high-density network, as shown in Table 2.

The curves of the system spectral efficiency with the number of FAPs in the network for a certain network load are given in Table 2. The proposed load-aware resource allocation method is compared with three recently published representative coloring methods in Table 2. When we increase the number of FAPs in the network (scale up the network), the proposed load-aware FAP resource allocation method has better spectral efficiency than all the benchmark

TABLE 2: Statistics of results.

Data set	Node	Parameter	Sampling/node	
			Mean	Standard deviation
Feminist	4.2	1.2	1.14	12.3
Synthetic-C5-W10	10.94	7.06	3.2	2.54
Synthetic-C5-W40	10.55	2.98	6.14	8.61
Synthetic-C5-W80	10.72	7.87	10.88	2.12
Synthetic-C10-W50	14.15	1.37	13.21	7.1

algorithms. When the network load is 30% and there are 128 FAPs in the network, the performance improvement of the method compared to the coloring method 1/2/3 is 43.88%, 62.00%, and 88.86%, respectively. In addition, the corresponding algorithm is also simulated at 50%, 70%, and 90% of the network load. The method outperforms all the benchmark methods for most network sizes and network loads.

In this section, the problem of interference modeling and resource allocation for downlink communication in the case of unavailability of geographic location information of network nodes is investigated. The relative interference intensity of the uplink is modeled based on the association rule algorithm using the large amount of data generated in the network. In addition, this section proposes a load-aware resource allocation method, which calculates the bounds for multiplexing the same spectrum resources and allocating orthogonal spectrum resources for each user with its interference source based on the modeled relative interference intensity and the network load in each TTI. The set of orthogonal interference sources for each user is generated based on the time-varying multiplexing/orthogonal boundaries for each user, and then, the spectrum resources are allocated based on the set of orthogonal interference sources for each user. In the simulation results section, the accuracy of the relative interference intensity modeling scheme based on the association rule algorithm proposed in this section is evaluated, and the simulation results show that the method can achieve high modeling accuracy with fewer samples. In addition, the simulation analysis section also analyzes the performance of the load-aware resource allocation algorithm proposed in this section under different network loads and different network sizes, and the results show that the method achieves good performance in most network density and network load cases.

## 5. Conclusion

With the rapid development of wireless communication technology, users can enjoy increasingly high-speed data services. This has further stimulated the use of mobile terminals (smartphones, tablets, smartwatches, etc.) by users, and as a result, the volume of data services in cellular networks has grown dramatically. Mobile network architectures need to be continuously innovated to adapt to the network environment where the volume of data services is increasing dramatically. Small-cell networks and caching technologies are two of the most promising technologies currently being studied by experts. The current resource allocation algo-

rithms applied to traditional cellular networks are no longer applicable to small-cell network architectures with caching, so new resource allocation algorithms need to be proposed. It is in this context that the study of resource allocation for small-cell network architectures with cache is of great importance in this paper. The basic concepts of wireless access networks, prospective technologies for network development are briefly explained, and a small-cell network architecture with caching are introduced. After that, the throughput of small cells in CSCN architecture is calculated by two transmission links, and the conditions for selecting users in small cells are optimized in terms of both the user's mobile location and the cache state of the small cells to which the user is connected. The problem is then modeled using the knowledge of game theory and an RL-LSTM-based resource allocation algorithm is proposed based on the knowledge of deep augmented learning. Based on the small base stations selecting good users according to their transmission conditions, the algorithm considers each small base station as an intelligent learning body. Using the encoder-decoder model, the semantic vector is obtained by encoding the input historical traffic sequence, after which the decoder decodes the semantic vector to obtain the action sequence of the small base station, and the optimal weight matrix parameters of the model are obtained by continuous iterative training. Then, the small base station action sequences are learned based on the optimal weight parameters so that the objective function is maximized.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] H. Y. Le Liang, G. Yu, and G. Y. Li, "Deep-learning-based wireless resource allocation with application to vehicular networks," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 341–356, 2020.
- [2] A. Zappone, M. Di Renzo, M. Debbah, T. T. Lam, and X. Qian, "Model-aided wireless artificial intelligence: embedding expert knowledge in deep neural networks for wireless system

- optimization,” *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 60–69, 2019.
- [3] K.-S. Kong, W. Lee, Y.-H. Han, M.-K. Shin, and H. R. You, “Mobility management for all-IP mobile networks: mobile IPv6 vs. proxy mobile IPv6,” *IEEE Wireless Communications*, vol. 15, no. 2, pp. 36–45, 2008.
- [4] Y. Zhou, Z. M. Fadlullah, B. Mao, and N. Kato, “A deep-learning-based radio resource assignment technique for 5G ultra dense networks,” *IEEE Network*, vol. 32, no. 6, pp. 28–34, 2018.
- [5] A. Zappone, M. Di Renzo, and M. Debbah, “Wireless networks design in the era of deep learning: model-based, AI-based, or both?,” *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7331–7376, 2019.
- [6] D. Wang, B. Song, D. Chen, and X. Du, “Intelligent cognitive radio in 5G: AI-based hierarchical cognitive cellular networks,” *IEEE Wireless Communications*, vol. 26, no. 3, pp. 54–61, 2019.
- [7] W. Ahsan, W. Yi, Z. Qin, Y. Liu, and A. Nallanathan, “Resource allocation in uplink NOMA-IoT networks: a reinforcement-learning approach,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 8, pp. 5083–5098, 2021.
- [8] Q. Mao, F. Hu, and Q. Hao, “Deep learning for intelligent wireless networks: a comprehensive survey,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2595–2621, 2018.
- [9] G. Zhu, D. Liu, Y. Du, C. You, J. Zhang, and K. Huang, “Toward an intelligent edge: wireless communication meets machine learning,” *IEEE Communications Magazine*, vol. 58, no. 1, pp. 19–25, 2020.
- [10] E. E. Tsiropoulou, G. K. Katsinis, and A. Filios, “On the problem of optimal cell selection and uplink power control in open access multi-service two-tier femtocell networks,” in *International Conference on Ad-Hoc Networks and Wireless*, pp. 114–127, Springer, Cham, 2014.
- [11] H. Huang, S. Guo, G. Gui et al., “Deep learning for physical-layer 5G wireless techniques: opportunities, challenges and solutions,” *IEEE Wireless Communications*, vol. 27, no. 1, pp. 214–222, 2020.
- [12] F. Tang, Y. Kawamoto, N. Kato, and J. Liu, “Future intelligent and secure vehicular network toward 6G: machine-learning approaches,” *Proceedings of the IEEE*, vol. 108, no. 2, pp. 292–307, 2020.
- [13] D. Wang, D. Chen, B. Song, N. Guizani, X. Yu, and X. Du, “From IoT to 5G I-IoT: the next generation IoT-based intelligent algorithms and 5G technologies,” *IEEE Communications Magazine*, vol. 56, no. 10, pp. 114–120, 2018.
- [14] M. Yan, G. Feng, J. Zhou, Y. Sun, and Y.-C. Liang, “Intelligent resource scheduling for 5G radio access network slicing,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7691–7703, 2019.
- [15] S. Yu, X. Chen, L. Yang, M. B. Di Wu, and J. Zhang, “Intelligent edge: leveraging deep imitation learning for mobile edge computation offloading,” *IEEE Wireless Communications*, vol. 27, no. 1, pp. 92–99, 2020.
- [16] N. Kato, Z. M. Fadlullah, F. Tang et al., “Optimizing space-air-ground integrated networks by artificial intelligence,” *IEEE Wireless Communications*, vol. 26, no. 4, pp. 140–147, 2019.
- [17] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, “Artificial neural networks-based machine learning for wireless networks: a tutorial,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3039–3071, 2019.
- [18] H. Ye, G. Y. L. Le Liang, J. B. Kim, L. Lu, and M. Wu, “Machine learning for vehicular networks: recent advances and application examples,” *IEEE Vehicular Technology Magazine*, vol. 13, no. 2, pp. 94–101, 2018.
- [19] X. L. Huang, X. Ma, and F. Hu, “Editorial: Machine learning and intelligent communications,” *Mobile Networks and Applications*, vol. 23, no. 1, pp. 68–70, 2018.
- [20] Y. Liu, S. Bi, Z. Shi, and L. Hanzo, “When machine learning meets big data: a wireless communication perspective,” *IEEE Vehicular Technology Magazine*, vol. 15, no. 1, pp. 63–72, 2020.
- [21] J. Chen, S. Chen, Q. Wang, B. Cao, G. Feng, and J. Hu, “iRAF: a deep reinforcement learning approach for collaborative mobile edge computing IoT networks,” *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 7011–7024, 2019.
- [22] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, “Ten challenges in advancing machine learning technologies toward 6G,” *IEEE Wireless Communications*, vol. 27, no. 3, pp. 96–103, 2020.
- [23] F. Tang, B. Mao, Z. M. Fadlullah, and N. Kato, “On a novel deep-learning-based intelligent partially overlapping channel assignment in SDN-IoT,” *IEEE Communications Magazine*, vol. 56, no. 9, pp. 80–86, 2018.