IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 28 June 2023, accepted 19 July 2023, date of publication 1 August 2023, date of current version 8 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3300870

RESEARCH ARTICLE

Cognitive Spectrum Scheduling Method for Internet of Vehicles Based on DNN and MCTS

XINYU CUI AND GUIFEN CHEN
School of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun 130022, China
Corresponding author: Guifen Chen (cclgdx123321@163.com)

ABSTRACT The application of Internet of Vehicles technology has led to an increase in the demand for vehicle scheduling, but the available computational resources and spectrum in current wireless networks are limited and scarce. Therefore, in order to improve the efficiency of vehicle scheduling, a cognitive spectrum scheduling method based on Deep Neural Network and Monte Carlo Tree Search for vehicle networking has been proposed. It evaluates the priority of MCTS and provides spectrum resource scheduling solutions, and uses DNN for offline training to obtain an environment model. In the simulation experiment for this method, the CUR of the proposed method increased by a maximum of 19.3% compared to other methods used for comparison. Its ALC is 20.4% higher than other methods. Its convergence time is lower than all methods used for comparison, with a maximum difference of 59.3%. The mean MAE of the proposed method is 0.793, and the mean RMSE is 0.628. The results of MAE and RMSE demonstrate that the proposed method in the experiment exhibits the lowest errors, both in training and testing processes. The proposed method provides a certain technical foundation for the cognitive spectrum scheduling of the Internet of Vehicles.

INDEX TERMS DNN, MCTS, vehicle networking, cognitive spectrum, scheduling methods.

I. INTRODUCTION

In recent years, Vehicle-to-everything related technologies and industries have received extensive attention and rapid development [1]. The huge data application market provides a strong foundation for the development of China's Vehicle-to-everything industry. However, with the rapid development of the Vehicle-to-everything, the data of the Vehicle-to-everything has also seen a blowout growth in recent years. In the face of huge and complex vehicle road environment, plus a large number of sensors, the Vehicle-to-everything has put forward high requirements for data task processing delay and network bandwidth resources. With the increasing maturity of emerging information and communication technologies, vehicles can achieve comprehensive network connectivity and resource sharing. Combined with artificial intelligence technology and mobile edge computing architecture, it can further realize a variety of computing intensive intelligent services, thus further improving the comprehensive information processing capacity of vehicles, so as to independently determine the driving scheme of vehicles in a complex road environment, and further ensure and improve road safety and travel efficiency [2]. As a wireless Internet of Things network, Vehicle-to-everything has the characteristics of network nodes moving and network topology changing from time to time. Obviously, when the number of intelligent connected vehicles shows a rapid growth trend, it will inevitably generate a large amount of resource demand, such as spectrum resources, computing resources, etc. At this time, a large number of network connections and differentiated data services generated by the Vehicle-to-everything put forward stricter requirements and challenges for reliable dynamic network resource scheduling schemes. In addition, in the dynamic Vehicle-to-everything environment, temporal data will gradually lose its usefulness over time. It is also very important to maintain high-quality temporal data. In order to improve the timeliness of network resource supply of intelligent networked vehicles, a more efficient and reliable Vehicle-to-everything wireless resource scheduling scheme is needed [3]. With the increase in the number of intelligent connected vehicles in cities, vehicles need to communicate frequently with ground cellular networks, resulting in a large

The associate editor coordinating the review of this manuscript and approving it for publication was Barbara Masini.

amount of data interaction and greatly exacerbating the serious shortage of spectrum resources. In order to cope with the huge challenge of spectrum resource shortage faced by the Vehicle-to-everything, cognitive Vehicle-to-everything came into being. Cognition of the Vehicle-to-everything allows vehicles to access idle authorized frequency bands opportunistically without affecting the use of authorized users, so as to alleviate the contradiction between the growing demand for spectrum use of the Vehicle-to-everything and the current situation of scarce spectrum resources. In order to improve the quality and efficiency of dynamic spectrum resource allocation, there have been a lot of research work based on Game theory, Swarm intelligence optimization algorithm, Deep reinforcement learning and other methods. However, existing spectrum scheduling techniques do not incorporate the driving status of vehicles themselves (i.e., the impact of individual information such as vehicle position and speed on the quality of spectrum scheduling) into the decision-making process of spectrum allocation. Secondly, the existing dynamic spectrum allocation methods cannot quickly adapt to the highly dynamic Vehicle-to-everything network environment. How to meet the fast and reliable demand of spectrum scheduling of vehicle users in cognitive Vehicle-to-everything has important research significance. Therefore, a frequency spectrum scheduling scheme based on vehicle driving status priority and scene simulation, namely Finder MCTS, is proposed for the urban Vehicle-to-everything. Subsequently, DNN was used for learning and training, resulting in a state predictor that can be applied even under offline conditions. A combined model of DNN and Finder MCTS was constructed to further improve the efficiency of vehicle scheduling in IoV. The main contribution of the research is to improve the spectrum resource utilization and the quality of user experience of the system. By adding interference constraints from cognitive radio resource scheduling to the tree search process, search efficiency can be improved. At the same time, the quality of radio resource management in vehicle-to-everything is ensured, and the efficiency of resource management is greatly improved.

II. RELATED WORK

Vehicle scheduling is an important factor in transport. A reasonable vehicle scheduling plan is related to the effective utilization of information. Many studies have conducted in-depth discussions on vehicle scheduling methods. Researchers have improved and optimized the vehicle scheduling method by using deep learning, swarm intelligence algorithm, etc. Cui H et al., in order to cope with the pressure of rescue service brought by multiple vehicle types in the rescue process, divided the vehicles into small, medium and large ones, modified the corresponding speed through six road conditions, and designed a multi-objective Decision model for rescue vehicle scheduling with two stages. At the same time, non dominated sorting genetic algorithm II with real number coding is designed. Compared with multi-objective gray wolf algorithm and traditional genetic algorithm, the proposed method has faster Rate of convergence and better scheduling effect [4]. To address the scheduling problem of public transportation vehicles and crew, Andrade Michel A's team designed an accurate constraint programming model that simulates driver absenteeism behavior using Monte Carlo methods and evaluates the obtained travel vehicle driver allocation. In the experimental results of randomly generated instances, it has been proven that this method has certain effectiveness and significant benefits in covering the travel distance [5]. Researchers have designed a scheduling scheme for electric vehicles using deep learning technology. This plan considers both the impact of risk preference on energy output and the impact of group charging on electric vehicle scheduling methods. The experimental results indicate that this method can improve energy utilization efficiency while reducing operational costs for enterprises. In practical scenarios, this scheduling scheme has high practical significance [6]. For electric vehicles, reasonable scheduling is also necessary to meet the vehicle's range and charging requirements [7]. The periodicity of public transportation has a certain impact on vehicle scheduling. In the study by Lieshout R et al., the periodicity of public transportation was included in the study of vehicle scheduling schemes. It improves related formulas such as constraint conditions and shrinkage techniques. The experimental results demonstrate that the improved formula has high efficiency and effectiveness in determining the number of vehicles and scheduling [8]. The vehicle scheduling method based on IoV can effectively solve the data processing problem caused by the amount of information. However, with the increasing amount of information, conventional methods can no longer meet the needs of data processing. Researchers consider factors such as resource allocation, energy optimization, and flow control into vehicle resource scheduling. They utilized 6G networks to build green IoV systems, thereby reducing energy consumption [9].

Spectrum resource allocation is a key research topic in IoV that affects the effectiveness of resource allocation. In recent years, deep learning methods such as DNN have yielded a series of research results in spectrum resource allocation. Sharif A et al. applied DNN to the optimal strategy learning of IoV technology by making use of its advantages of continuous learning and training, and designed an experience driven method based on the actor critic Deep reinforcement learning framework. The results show that the throughput of this method is increased by 35% and 14% respectively [10]. J. Elhachmi's team proposed an adaptive DNN technology that can update dynamic requests and achieve adaptation, and applied it to spectrum technology optimization. The results showed that this method has better superiority, stability, and robustness compared to traditional DNN [11]. To determine the optimal frequency of the spectrum, R Nandakumar et al. utilized DNN optimization spectrum technology and constructed a sequential user selection method. Compared to traditional scheduling schemes, the results verified that this method can effectively improve user

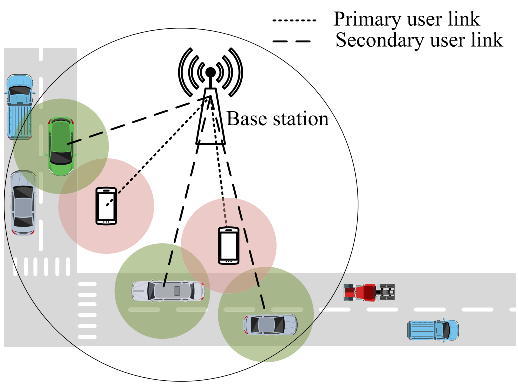


FIGURE 1. System scenario for spectrum scheduling in cvn.

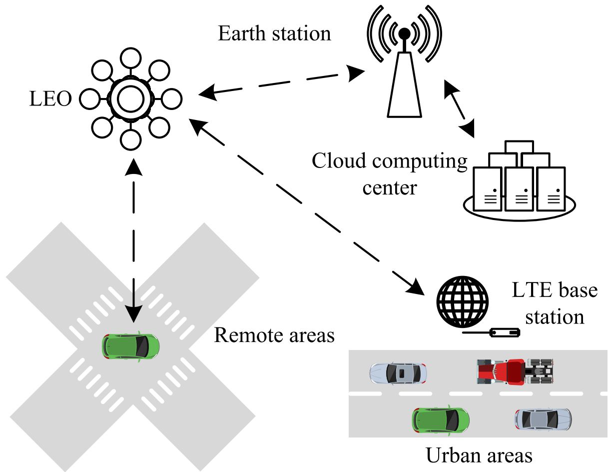


FIGURE 2. Three-dimensional structure of the integrated vehicle networking communication network.

satisfaction and maintain appropriate channel allocation status [12]. Nandakumar R et al. established a prediction model for spectral transmission using a neural network with time memory. The improved prediction method can reduce the waiting time of users and enhance the performance capacity of the network by predicting the spectral state [13].

In summary, spectrum resource allocation is a key issue in IoV. Researchers have optimized the application of technologies such as MCTS and DNN in IoV and its spectrum resource allocation, but issues such as unreasonable resource allocation still frequently arise in this field. Therefore, a frequency spectrum scheduling scheme based on vehicle driving status priority and scene simulation, namely Finder Monte Carlo Tree Search (MCTS), is proposed for the urban Vehicle-to-everything. Subsequently, DNN was used for continuous learning and training, resulting in a state predictor that can be applied even under offline conditions. A combined model of DNN and Finder MCTS, the DNN-F-MCTS model, was constructed to further improve the efficiency of vehicle scheduling in IoV.

III. DNN AND MCTS IN COGNITIVE SPECTRUM SCHEDULING METHODS FOR 2-CAR INTERNET

A. THE SYSTEM MODEL IN THE COGNITIVE SPECTRUM SCHEDULING METHOD FOR THE INTERNET OF VEHICLES

In the research related to IoV, as a newly proposed technical concept, CVN can solve the shortage of its spectrum resources. In CVN, spectrum sharing technology can be utilized to alleviate communication interference issues, thus meeting the differentiated communication business services in IoV. Figure 1 shows the spectrum allocation scenario in cognitive vehicle networking. The main user represents the user who authorizes the current network mobile phone; The second-level user represents the equipped vehicle. When the main user occupies the channel, a protection area will be formed around it (within the red circle range in Figure 1). When a secondary user occupies the channel, an interference area (within the green circle range in Figure 1) will be formed around it. The radiation of secondary users can cause interference to the primary user when it appears in their protected area.

When the above situation occurs, the established allocation algorithm will be applied to the base station in this study. The cognitive radio module can be used in the configuration of vehicle nodes. On this basis, the node can be used to sense whether the spectrum resources are idle and available. The vehicle is able to send access requests to the common control channel to the base station. After learning, the base station is able to allocate idle and available spectrum resources into the black circle range in Figure 1. Within this range is the approximate optimal strategy.

Based on CVN and other related technologies, the concept of Tiandi Integrated Vehicle Network (TIVN) has been derived. TIVN can be used to meet the needs of VN calculation in remote areas. Its characteristics mainly include the following aspects. The network topology of TIVN has the characteristics of dynamic changes, wide coverage effectiveness, and self-organization. In this network structure, Low Earth Orbit (LEO) can be used as a space-based node. LEO can be used for calculating vehicle demand and tasks in urban and remote areas. In urban areas, it includes ground base stations, drones, cloud computing centers, etc. The introduction of roadbed communication mode can make the service coverage of vehicle communication more targeted. The use of LEO nodes can reduce network workload and increase network capacity. Meanwhile, users can independently choose network servers, thereby improving the service experience. Figure 2 is a schematic diagram of the network structure of TIVN.

In CVN, it is necessary to evaluate the vehicle status, such as direction, speed, acceleration, and GPS coordinates, in order to obtain a priority level. The main evaluation indicators in this experiment include vehicle driving evaluation score, network utility score, and comprehensive priority evaluation score. Formula (1) is the specific calculation formula for the vehicle driving evaluation score.

$$Travelingscore_n = \frac{(1+\cos(\theta_n))}{4} \cdot (\frac{v_{\max} - v_n}{v_{\max} - v_{\min}} + \frac{1}{1+e^{a_n}}) \tag{1}$$

In formula (1), θ_n represents the angle between the current direction of travel and the line connecting the base station and vehicle coordinates; v_n and a_n are used to describe the speed and acceleration of vehicle n, respectively; $v_{\max}$ represents the maximum driving speed of the vehicle; $v_{\min}$ represents the minimum driving speed of the vehicle. The higher the value of the network utility score, the better the global communication capability, which means that the vehicle can receive a larger spectrum allocation weight. Formula (2) is the specific calculation formula for the network utility score.

$$Utility_n = \log_2(1+SNR_n) \cdot \frac{\sum_{1\le n,n'\le N, n'\ne n} Dispersion_{n,n'}}{N-1} \tag{2}$$

In formula (2), SNR_n represents the signal-to-noise ratio of user n when receiving base station signals; $\log_2(1+SNR_n)$ is used to describe the rate at which user n receives data; $\sum_{1\le n,n'\le N, n'\ne n} Dispersion_{n,n'}$ is used to describe the global user dispersion of user n. Formula (3) is the calculation formula for $Dispersion_{n,n'}$.

$$Dispersion_{n,n'} = \begin{cases} 1, & D_{n,n'} > \varepsilon_n \\ 0, & D_{n,n'} \le \varepsilon_n \end{cases} \tag{3}$$

In formula (3), ε_n represents the threshold value of Dispersity; n, n' represents two secondary users; $D_{n,n'}$ is used to describe the average dispersion time between the two.From this, the formula for calculating the comprehensive priority evaluation score in formula (4) can be obtained.

$$Priorityscore_n = Travelingscore_n \cdot Utility_n \tag{4}$$

It is necessary to note that the vehicle driving evaluation score and network utility score are obtained by the base station through real-time collection and analysis of vehicle related feature information in the network. For vehicle users in the network who initiate service requests to the base station, the base station uses the collected vehicle information to calculate the priority score of the requested vehicle, and sorts it from top to bottom. Therefore, we can obtain a priority service score list for users in the cognitive network within the current allocation cycle_ List. This priority service order list will be used as a secondary user allocation order list, ensuring that vehicle users with different allocation weights in the network access based on differentiated priority, and improving the reliability of spectrum scheduling schemes.

B. COGNITIVE SPECTRUM SCHEDULING METHOD FOR VEHICLE NETWORKING BASED ON MCTS

The main problems with spectrum scheduling methods are insufficient real-time performance and insufficient solution quality [14], [15]. By utilizing MCTS, these issues can be addressed. Based on tree search, the classic MCTS (Basic MCTS, B-MCTS) exhibits strong learning ability. It can extend the tree search to nodes with high rewards and then simulate using the set strategy. This can update rewards and other information. MCTS is a heuristic search algorithm based on tree data structure that is still effective even when the search space is huge. The process of the Monte Carlo tree search algorithm is divided into four stages: selection, expansion, simulation and backpropagation. In the selection phase of the Monte Carlo tree search algorithm, the root node is taken as the starting point (i.e. the top black circle), a decision is made by comparing the value of the node to be selected, and the child node with the largest value is selected. In the expansion phase of the Monte Carlo tree search algorithm, after expanding the current node, a new node (i.e. the white circle at the bottom) is created in the decision tree as a new child node of the node, according to the confidence upper bound interval value of its child nodes. In the simulation phase of the Monte Carlo tree search algorithm, the basis of the current node movement selection is a random strategy: that is, a subnode is randomly selected from the subnodes to be selected and expanded to the last node. After the end of the Monte Carlo simulation phase, broadly speaking, the simulation results can be any value. In the backpropagation stage, any value propagates in the opposite direction along the decision tree to the root node, and the node state that initiates Monte Carlo simulation changes to visited.

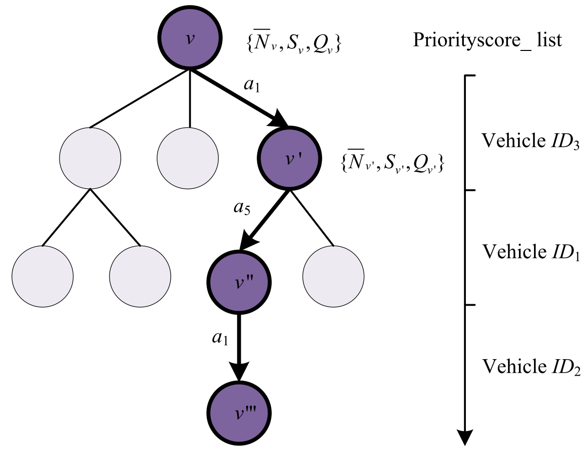


FIGURE 3. Specific search steps for f-mcts.

However, the search scale of B-MCTS expands with the expansion of search text, which reduces its search speed. Meanwhile, B-MCTS ignores the uncertainty brought about by environmental factors. The variance during simulation is large, which reduces its search performance. In response to these issues, a new scheduling scheme, F-MCTS, is proposed in this study. F-MCTS can use the vehicle's posture information and communication information to define the priority of spectrum allocation. Then, based on the prioritization results, MCTS is used to provide a real-time spectrum allocation scheme. In F-MCTS, the MDP state space is defined,

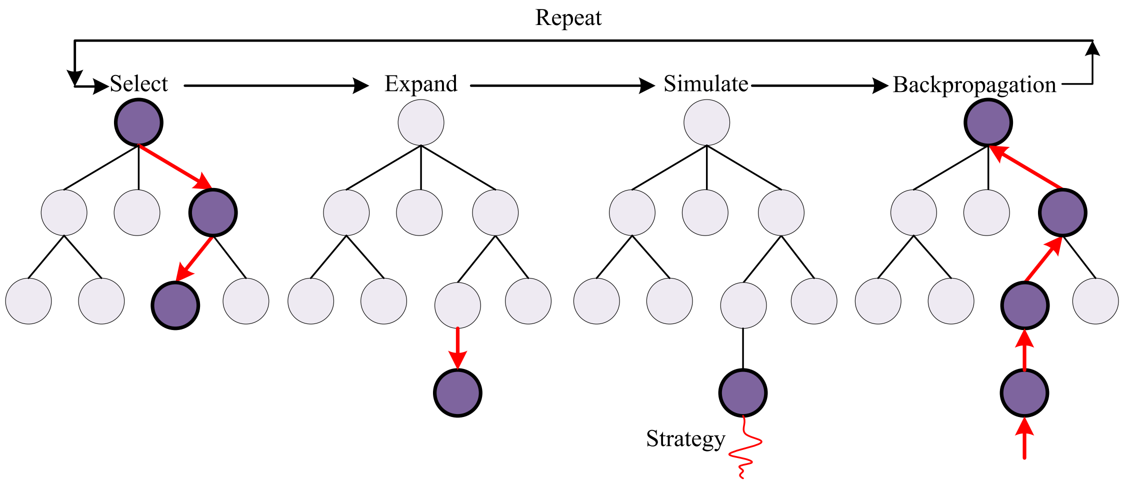


FIGURE 4. The iterative calculation process of dnn-f-mcts.

as shown in formula (5).

$$S = \{s_v \,|(\lambda_v, \varphi_v, \xi_v)\} \tag{5}$$

In formula (5), s_v represents the state; φ_v is the number of communication services that need to be allocated; ξ_v denotes the total broadband requests of φ_v users; λ_v denotes the vector of the remaining bandwidth on the base station side. It is shown in formula (6).

$$\lambda_v = \{\overline{\lambda_1}, \ldots, \overline{\lambda_m}, \ldots \overline{\lambda_M}\}_v \tag{6}$$

In formula (6), $\overline{\lambda_m}$ denotes the remaining bandwidth of channel m. The MDP action space is also defined, as shown in formula (7).

$$A = \{a_m \,|1 \le m \le M\} \tag{7}$$

In formula (7), a_m is used to describe the vehicles that can be scheduled and assigned when m is assigned by the agent; M represents the total number of channels; In MDP, node v and edges are the main components of a search tree; v represents the search tree node in the corresponding state s_v. In the search tree, the edges connecting the parent and child nodes are used to describe the actions that cause a state change to occur. Each node v needs to maintain its node status value in the search tree. The node status values mainly include s_v, v, the number of times the node has been accessed $\overline{N}_v$, and the accumulated reward value Q_v. Figure 3 depicts the specific search steps for F-MCTS.

In Figure 3, the Vehicle ID is used to represent the vehicle. In the specific search steps of F-MCTS, it is first necessary to create v and initialize it to obtain the state value. It is shown in formula (8).

$$v = \{\overline{N}_v, s_v, Q_v\} \tag{8}$$

In the second step, it is necessary to allocate the frequency spectrum of the vehicles in sequence according to the Priorityscore_list table in the comprehensive priority evaluation.

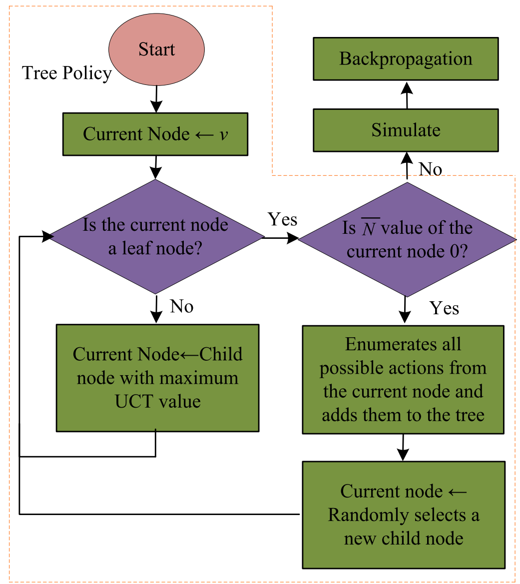


FIGURE 5. Flow chart of finder dnn-f-mcts.

Multiple iterative calculations are required during allocation. Taking v as an example, assume that the action when allocating the channel of Vehicle ID3 is a_1. At this point, the search tree expands downwards until the child node v'. During this process, continuous iterative calculations are carried out to continuously update the state values of nodes, as shown in formula (9).

$$s_{v'} = f_{ESP}(s_v, a_1) \tag{9}$$

In formula (9), f_{ESP} represents the state transition function.

In the third step, when the extension of the search tree meets the end condition of the iterative calculation, then the currently obtained set A^* needs to be returned. This set is used to describe the allocation actions of the optimal channel.

TABLE 1. Experimental environment.

Operating system	Windows 10	Processor	Intel Core CPU i99820X 3.50 GHz
Vehicle communication range	200m	RAM	64 GB
Vertical running off of the road	50m	Edge Server Communication Range	350m
V2V channel bandwidth	10MHZ	Vehicle transmission power	10P/dBm
V2I channel bandwidth	20MHZ	Path loss factor	2
Vehicle computing capacity	5×107 cycles/s	Attenuation coefficient	0.9
Edge Server Computing Power	7×108 cycles/s	Learning rate	0.001

C. SCHEDULING METHODS BASED ON DNN AND F-MCTS

Due to the uncertainty of spectrum occupation activities of primary users, the statistical state of the tree will be unstable when it expands from one node to the next. In other words, given a state and an action, the next state is uncertain. This uncertainty is caused by the unknown environment of Vehicle-to-everything. Therefore, in order to limit the scale of horizontal expansion of MCTS trees and accelerate search speed, it is necessary to gradually learn real environment models close to CVN during spectrum allocation. In the experiment, DNN was used for continuous learning and training, resulting in an Environmental State Predictor (ESP) that can be applied even under offline conditions. Note that in order to train for ESP, sufficient training data is required. So, first of all, during the cold start phase of F-MCTS, which is the initial running phase of the algorithm, ESP is not used. This does not affect the channel allocation solution of F-MCTS. Moreover, during the period after the cold start phase, our base station can calculate and obtain a considerable number of "state action transfer pairs" in real-time. Subsequently, these state action transition pairs are continuously input as training data into our ESP to obtain the state action transition function, which is an offline training process. After obtaining the transition function, the search of F-MCTS will accelerate convergence due to the reduction of branches (i.e., the reduction of uncertainty). Based on the accumulated state action transfer pairs mentioned above, we use the historical accumulated state action pairs as inputs and use DNN to train the transfer states. The network structure of DNN consists of one input layer, three hidden layers, and one output layer. Formula (10) showcases its Loss function.

$$Loss_{ESP} = \frac{1}{B}\sum_{B}(\left\| s_{v'} - \hat{s}_{v'} \right\|_2) \tag{10}$$

In formula (10), B is used to describe the size of the sample during the gradient descent process; $\|\cdot\|_2$ is used to describe the L2 norm. When the function reaches convergence state, the network parameter w_{ESP} in DNN will be updated. After learning and training, one can obtain ESP. For action a_m and state s_v, the state $\hat{s}_{v'}$ that exists in the extended node can be obtained, as shown in formula (11).

$$\hat{s}_{v'} = f_{ESP}(s_v, a_m | w_{ESP}) \tag{11}$$

Based on the above F-MCTS and DNN, the CVN vehicle scheduling method was further optimized in the experiment, namely DNN-F-MCTS. In the newly established vehicle scheduling method, there are still 4 steps of iterative operation to be calculated, as shown in Figure (4).

In the DNN-F-MCTS model, the first step is to select the optimal sub node, as shown in formula (12).

$$\underset{v' \in child(v)}{\arg\max}\,(\frac{Q_{v'}}{\overline{N}_{v'}} + c \cdot \sqrt{\frac{\ln(\overline{N}_v)}{\overline{N}_{v'}}}) \tag{12}$$

In formula (12), c is the coefficient, and its value is ≥ 0. $child(v)$ denotes the set of sub nodes in the search tree. During the iteration, $\overline{N}_v$ denotes the number of times the parent node v has been accessed; $\overline{N}_{v'}$ denotes the number of times sub node v' has been accessed; $Q_{v'}$ is the cumulative reward, and it is acquired by v'. Then there is the constraint oriented extension, where the optimal sub nodes obtained in the previous step need to be used for the subsequent extension. Constraint processing is necessary to prevent excessive expansion from increasing the computational complexity. Then there is a differentiated scenario simulation. Random returns are defined to adjust reward evaluation. Finally, there is backpropagation. The main purpose of backpropagation is to update the experience gained from previous exploration before proceeding to the next iteration. Therefore, the rewards in this process include the reward evaluation of all extension nodes. This can be used to reflect the performance of spectrum allocation. Meanwhile, the DNN-F-MCTS model achieves pre-processing before updating through formula (13).

$$\overline{N}_v \leftarrow \overline{N}_v + 1 \tag{13}$$

Then, it is combined with formula (14) to update the node state values.

$$Q_v \leftarrow Q_v + Q_{\tilde{v}} \tag{14}$$

After the above processing steps, DNN-F-MCTS can obtain a differentiated spectrum allocation method. Finally, after outputting the results, the optimal allocation scheme in CVN can be obtained. In DNN-F-MCTS, it is first necessary to input the current node and then determine whether it is a child node. If it is a child node, then in the third step, it is necessary to determine whether the $\overline{N}$ value of the current node is equal to 0. If its value is equal to 0, then the optimal allocation scheme is obtained through simulation and backpropagation. If in the second step, the current point is not a child node, then the child point with the maximum confidence upper limit (UCT) needs to be selected as the current node. It then determines whether it is a child node. In the third step, if the $\overline{N}$ value of the current point is not equal to 0,

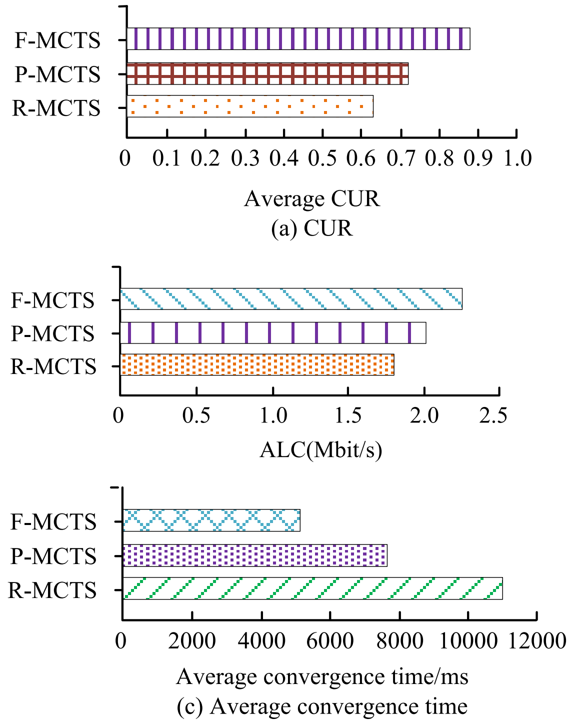


FIGURE 6. Comparison of different indicators of three mcts algorithms.

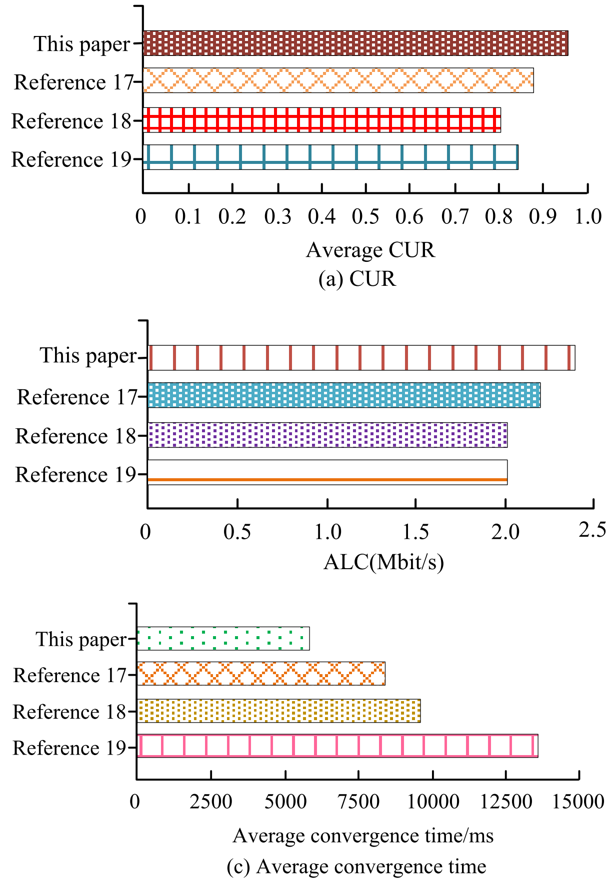


FIGURE 7. Comparison of indicators of different algorithms.

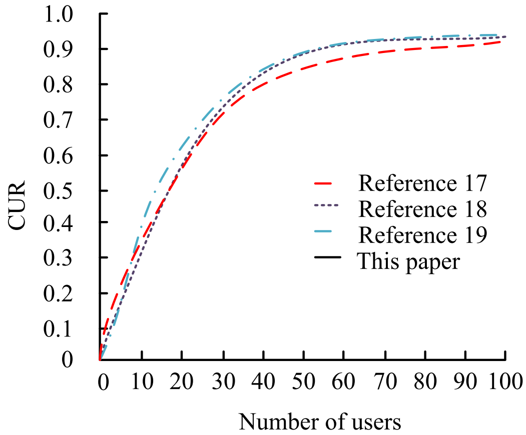
it is necessary to enumerate the possible actions that may exist in the point and append them to the search tree. Then it randomly selects a new child point as the current node and returns to the second step. Figure 5 specifically illustrates the DNN-F-MCTS method.

IV. VERIFICATION OF COGNITIVE SPECTRUM SCHEDULING METHOD FOR 3-VEHICLE NETWORKING

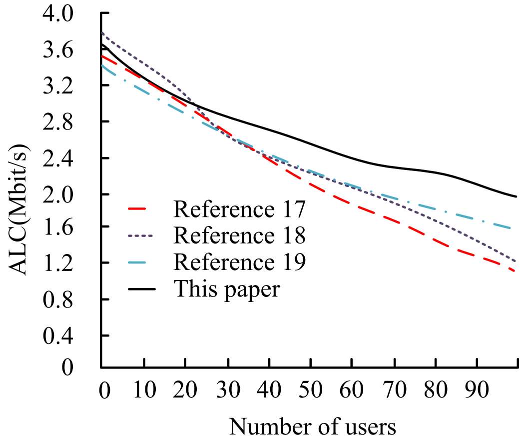
In the method validation section, Table 1 provides the specific experimental environment. All validation experiments were conducted on the Windows 10 operating system. The simulators are used to set the vehicle and channel attributes, and are random. Simulate the experiment in MATLAB.The processor is Intel Core CPU i99820X 3.50 GHz, with 64 GB of RAM. Edge servers have a computing power of 7 × 108 cycles/s, vehicle computing power is 5 × 107 cycles/s, V2V channel bandwidth is 10MHZ, V2I channel bandwidth is 20MHZ, and vehicle communication range is 150m. The vertical distance of the road is 50m, the communication range of the edge server is 300m, the vehicle transmission power is 10P/dBm, the attenuation coefficient is 0.9, the learning rate in the strategy network is 0.001, and the path loss factor is 2. The parameter settings are shown in Table 1.

To verify the advantages of the F-MCTS proposed in the experiment, this method was compared with the R-MCTS and P-MCTS spectrum scheduling methods based on random order and priority. At the same time, the number of users selected is 1862. Figure 6 showcases the comparison results of Channel Utilization Ratio (CUR), Average Link Capacity (ALC), and convergence time for three methods. The CUR of F-MCTS increased by 23.9% and 41.9% respectively compared to P-MCTS and R-MCTS. The ALC of F-MCTS increased by 12.5% and 25.0% respectively compared to P-MCTS and R-MCTS. The average convergence time of F-MCTS decreased by 32.9% and 52.9% respectively compared to P-MCTS and R-MCTS. The CUR and ALC values of F-MCTS are lower than those of P-MCTS and R-MCTS. The average convergence time of F-MCTS is lower than that of P-MCTS and R-MCTS. This indicates that F-MCTS can learn quickly in a time-varying environment, to find the optimal solution. This significantly improves the availability of spectrum resources in the system.

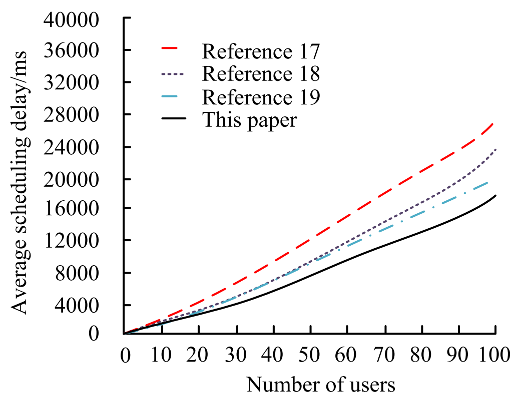
In order to verify the advantages of DNN-F-MCTS proposed in the experiment, this method was compared in the experiment with the methods in the literature. Reference [17] is the task offload of vehicle-mounted edge computing network realized by deep reinforcement learning. Reference [18] is an adaptive attitude determination method for biomimetic polarization integrated navigation systems based on reinforcement learning strategies. Reference [18] is a resource allocation method based on Markov decision-making. The comparison results of CUR, ALC, and convergence time for the four methods are shown in Figure 7. The



(a) CUR

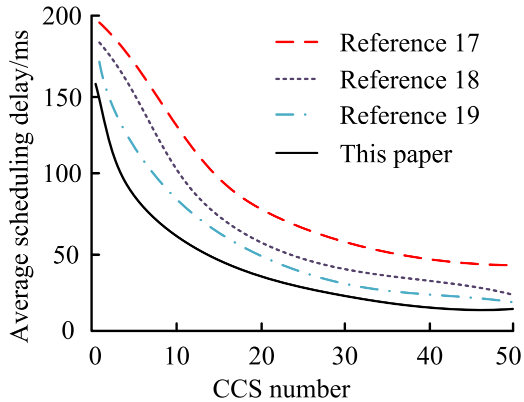


(b) ALC

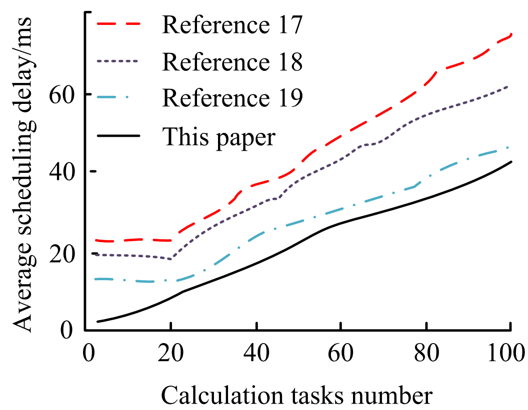


(c) Average scheduling delay

FIGURE 8. Performance comparison under different cognitive vehicle user numbers.



(a) CCS number and average scheduling delay



(b) Calculation tasks number and average scheduling delay

FIGURE 9. Relationship between ccs number and computing tasks number and the average scheduling delay.

CUR of DNN-F-MCTS was increased by 9.1%, 19.3%, and 15.7% compared to the methods in references [17] and [18], respectively. The ALC of DNN-F-MCTS was increased by 7.1%, 20.4%, and 20.4% compared to the methods in references [17] and [18], respectively. The average convergence time of DNN-F-MCTS was reduced by 30.7%, 32.3%, and 59.3% compared to the methods in references [17] and [18], respectively. The CUR and ALC values of DNN-F-MCTS are lower than those of the methods in references [17] and [18]. The average convergence time of DNN-F-MCTS is lower than that of the methods in references [17] and [18].

The number of users is an important factor affecting the performance of different vehicle scheduling methods. Figure 8 compares the CUR, ALC and convergence times of different methods for different numbers of users. Figure 8 indicates that the CUR values of each method increase as the number of users increases until a convergence state is reached. This may be due to the accumulation of experience due to the increase in the number of users, thereby improving the quality of vehicle scheduling solutions. The ALC values of each method decrease as the number of users increases. It may be because the increase in the number of users makes it difficult for various methods to find the optimal solution, thus achieving a global convergence state. The average convergence time of each method increases with

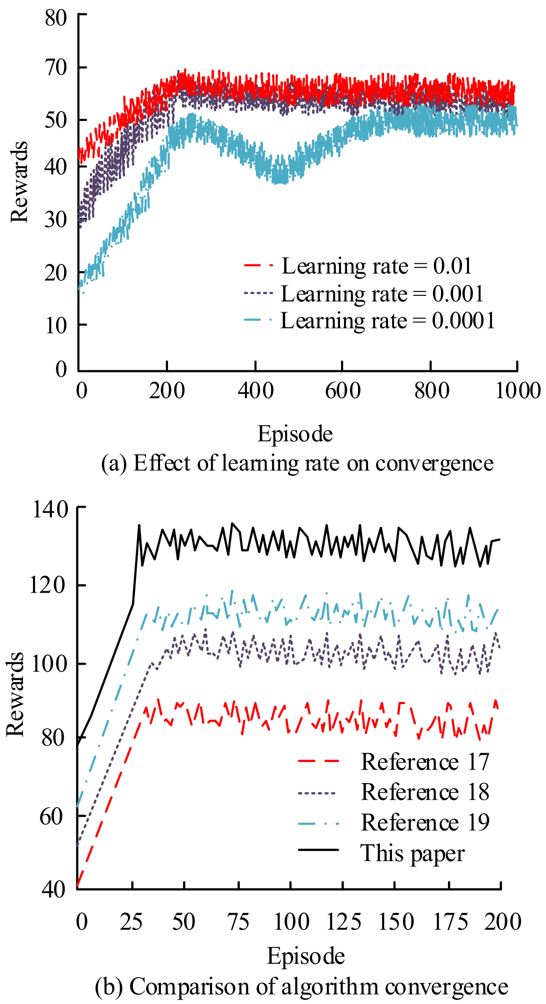


FIGURE 10. Comparison of algorithm convergence and relationship between vehicle number and average scheduling delay.

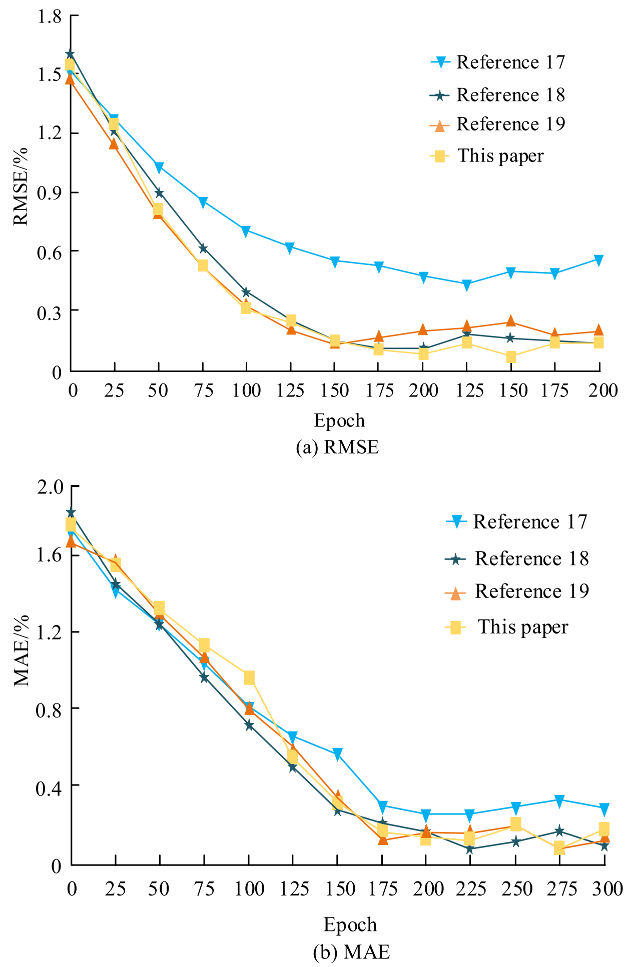


FIGURE 11. Comparison of rmse and mae.

the increase of the number of users, while the convergence time of DNN-F-MCTS changes the slowest.

In TIVN, the number of Central Cloud Server (CCS) on the ground and the quantity of computing tasks will have an impact on the efficiency of vehicle scheduling schemes. Figure 9 shows the average scheduling delay variation under different CCS and computational task quantities. Figure 9 (a) shows the average scheduling delay when the number of CCS grows from 1 to 50 when the quantity of computing tasks is set to 20. As the number of CCS increases, the average scheduling delay of each method gradually decreases. This may be because increasing the number of CCS provides more vehicle scheduling options for scheduling schemes. The DNN-F-MCTS vehicle scheduling method proposed in this experiment has the highest reduction in average scheduling delay. In Figure 9 (a), when the quantity of CCS is set to 20, the average scheduling delay is calculated when the quantity of tasks increases from 1 to 100. It can be concluded that as the quantity of computing tasks grows, the average scheduling delay of each method gradually increases. When the quantity of computing tasks is less than 20, the variation in the average scheduling delay of the methods in references [17] and [18] is not significant. When the quantity of computing tasks exceeds 20, the average scheduling delay of the methods in references [17] and [18] increases with the number of computations. When the number of tasks is 1-100, the average scheduling delay of DNN-F-MCTS increases as the number of tasks increases, showing a certain linear relationship. This may be because the algorithms in the literature are significantly influenced by constraints. Therefore, the efficiency of problem solving is affected, resulting in a higher average scheduling delay. The method proposed in this experiment is less affected by constraint conditions, so the efficiency of problem solving is higher.Overall, the efficiency and convergence speed of the DNN-F-MCTS model proposed in the study are better than those in references [17] and [18] because the proposed method incorporates the state of the vehicle itself into the process of IoV spectrum allocation, while adapting to the characteristics of dynamic IoV networks. Moreover, this method uses the vehicle's posture information and communication information to define the

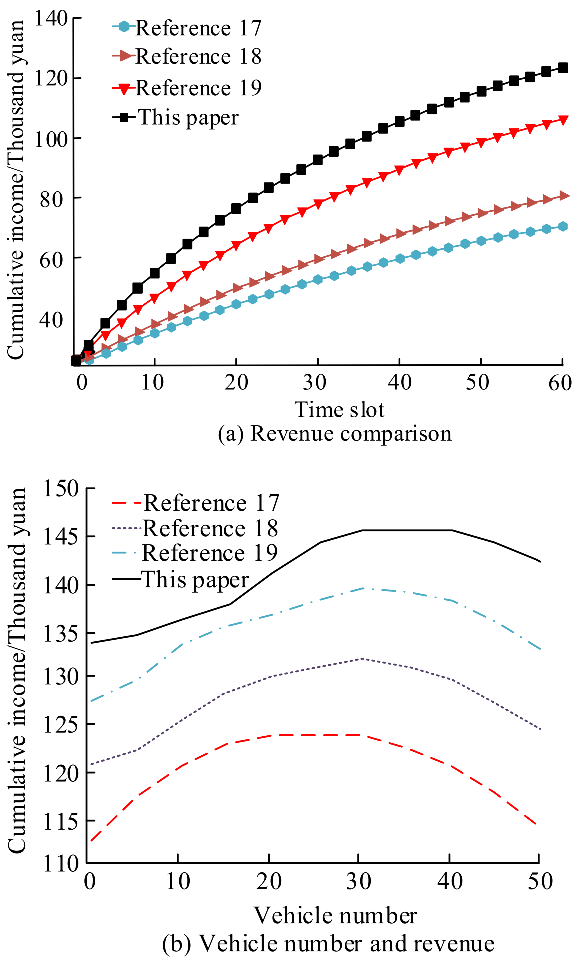


FIGURE 12. Revenue comparison and relationship between vehicle number and revenue.

priority of spectrum allocation, and provides a real-time spectrum allocation scheme through MCTS, So the effect is better.

In the iterative calculation of the algorithm, the learning rate also affects the convergence of the method. The impact of different learning rates on the vehicle scheduling method proposed in the experiment is shown in Figure 10 (a). It demonstrates that when Episode is below 200, different learning rates increase with the increase of Episode. When Episode is above 200, the method reaches convergence at learning rates of 0.01 and 0.001. When the learning rate is 0.00001, the convergence speed of the proposed method in the experiment changes slowly and requires longer training time. When the learning rates are 0.01 and 0.001, the difference in convergence speed between the two is small. The learning rate used in this experiment is 0.001. Figure 10 (b) shows the comparison of convergence rates among different methods. It indicates that the method proposed in this experiment reached a convergence state when the episode was 23, with high convergence efficiency.

Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are indicators to measure the accuracy of algorithms. These two indicators present the error situation of the algorithm from different perspectives. The smaller the MAE and RMSE values of the algorithm, the stronger its performance. Figure 11 (a) showcases the MAE indicator test results, and Figure 11 (b) indicates the RMSE indicator test results. The experiment in the test set indicates that the MAE and RMSE of the proposed algorithm in this experiment show an overall decrease with the increase of the epoch and finally reach convergence. In the comparison of MAE and RMSE indicators of different methods, the error curves of the proposed algorithm are all at the bottom. This indicates that the overall error of the proposed method in the experiment is the lowest. Figure 11 illustrates that the MAE mean and RMSE mean of the proposed algorithm in the experiment are 0.793 and 0.628, respectively. The results of MAE and RMSE demonstrate that the proposed method in the experiment exhibits the lowest errors in both training and testing processes.

For testing the practicality of the method in this study, Figure 12 demonstrates the calculation of benefits under different time slots and vehicle numbers. Figure 12 (a) shows that the cumulative benefits of the method gradually increase as the time slot increases. This indicates that different vehicle scheduling methods can achieve better vehicle scheduling and thus achieve better returns. Relatively speaking, the proposed method in this experiment achieved the highest cumulative returns. This proves that this method exhibits better scheduling performance in the actual vehicle scheduling. Figure 12 (a) indicates that as the quantity of vehicles grows, the cumulative benefits of the different vehicle scheduling methods tend to increase and then decrease. This may be because as the number of vehicles increases, it leads to an increase in computational tasks, thereby reducing the scheduling effectiveness of vehicle scheduling methods.

V. CONCLUSION

A cognitive spectrum scheduling method for vehicle networking based on DNN and MCTS is proposed to address the issue of low data availability and utilization efficiency in the current vehicle networking industry. According to the simulation test results, the CUR of F-MCTS increased by 23.9% and 41.9% respectively compared to P-MCTS and R-MCTS. The ALC of F-MCTS increased by 12.5% and 25.0% respectively compared to P-MCTS and R-MCTS. The average convergence time of F-MCTS reduced by 32.9% and 52.9% respectively compared to P-MCTS and R-MCTS. The CUR and ALC of this method are both improved by more than 7% compared to similar algorithms used for comparison. The highest increase reached 20.4%.The mean MAE of this algorithm is 0.793, and the mean RMSE is 0.628. The results of the error indicator show that the algorithm has the lowest error among the centralized algorithms. In terms of practicality, the cumulative benefits of the proposed algorithm are the highest under different time slots. The method proposed in this study still has certain shortcomings. The proposed algorithm lacks sufficient granularity in spectrum resource requirements. Therefore, it may not be able to maintain performance when faced with multi-objective optimization scenarios. Testing and optimizing its performance in multi-objective scenarios is the next research direction.

REFERENCES

[1] B. Gomathi, S. T. Suganthi, T. N. Prabhu, and A. Kovalenko, ''Multi-objective immune algorithm for Internet of Vehicles for data offloading,'' *Intell. Autom. Soft Comput.*, vol. 34, no. 3, pp. 1851–1860, 2022.

[2] X. Song, K. Wang, L. Xu, Y. Tan, and J. Miao, ''Efficient joint resource allocation for cognitive Internet of Vehicles networks based on asymmetric relay transmission,'' *IET Commun.*, vol. 15, no. 13, pp. 1655–1664, Aug. 2021.

[3] T.-D. Le and G. Kaddoum, ''LSTM-based channel access scheme for vehicles in cognitive vehicular networks with multi-agent settings,'' *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 9132–9143, Sep. 2021.

[4] H. Cui, L. Liu, Y. Yang, and M. Zhu, ''A two-stage scheduling model for the tunnel collapse under construction: Rescue and reconstruction,'' *Energies*, vol. 15, no. 3, pp. 743–759, 2022.

[5] A. Andrade-Michel, Y. A. Ríos-Solís, and V. Boyer, ''Vehicle and reliable driver scheduling for public bus transportation systems,'' *Transp. Res. B, Methodol.*, vol. 145, pp. 290–301, Mar. 2021.

[6] S. S. G. Perumal, R. M. Lusby, and J. Larsen, ''Electric bus planning & scheduling: A review of related problems and methodologies,'' *Eur. J. Oper. Res.*, vol. 301, no. 2, pp. 395–413, Sep. 2022.

[7] J. Wang, K. Zhu, and E. Hossain, ''Green Internet of Vehicles (IoV) in the 6G era: Toward sustainable vehicular communications and networking,'' *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 391–423, Mar. 2022.

[8] R. Lieshout, ''Integrated periodic timetabling and vehicle circulation scheduling,'' *Transp. Sci.*, vol. 55, no. 3, pp. 768–790, 2021.

[9] J. Gao, Y. Yang, F. Gao, and P. Liang, ''Optimization of electric vehicles based on Frank-Copula-*GlueCVaR* combined wind and photovoltaic output scheduling research,'' *Energies*, vol. 14, no. 19, pp. 6080–6094, 2021.

[10] A. Sharif, J. P. Li, M. A. Saleem, G. Manogran, S. Kadry, A. Basit, and M. A. Khan, ''A dynamic clustering technique based on deep reinforcement learning for Internet of Vehicles,'' *J. Intell. Manuf.*, vol. 32, no. 3, pp. 757–768, Mar. 2021.

[11] J. Elhachmi, ''Distributed reinforcement learning for dynamic spectrum allocation in cognitive radio-based Internet of Things,'' *IET Netw.*, vol. 11, no. 6, pp. 207–220, Nov. 2022.

[12] R. Nandakumar, V. Ponnusamy, and A. K. Mishra, ''LSTM based spectrum prediction for real-time spectrum access for IoT applications,'' *Intell. Autom. Soft Comput.*, vol. 35, no. 3, pp. 2805–2819, 2023.

[13] C. Yin, Z. Su, and A. Kortun, ''Performance analysis of cognitive vehicular networks under unreliable backhaul,'' *EAI Endorsed Trans. Ind. Netw. Intell. Syst.*, vol. 8, no. 26, Apr. 2021, Art. no. 169178.

[14] X. Liu, B. Lin, M. Zhou, and M. Jia, ''NOMA-based cognitive spectrum access for 5G-enabled Internet of Things,'' *IEEE Netw.*, vol. 35, no. 5, pp. 290–297, Sep. 2021.

[15] X. Li, Y. Zheng, W. U. Khan, M. Zeng, D. Li, G. K. Ragesh, and L. Li, ''Physical layer security of cognitive ambient backscatter communications for green Internet-of-Things,'' *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 3, pp. 1066–1076, Sep. 2021.

[16] E. Karimi, Y. Chen, and B. Akbari, ''Task offloading in vehicular edge computing networks via deep reinforcement learning,'' *Comput. Commun.*, vol. 189, pp. 193–204, May 2022.

[17] H. Bao, T. Du, and L. Sun, ''Adaptive attitude determination of bionic polarization integrated navigation system based on reinforcement learning strategy,'' *Math. Found. Comput.*, vol. 6, no. 2, pp. 161–177, 2023.

[18] T. Leng, P. Duan, D. Hu, G. Cui, and W. Wang, ''Cooperative user association and resource allocation for task offloading in hybrid GEO-LEO satellite networks,'' *Int. J. Satell. Commun. Netw.*, vol. 40, no. 3, pp. 230–243, May 2022, doi: 10.1002/sat.1436.

XINYU CUI was born in Changchun, Jilin, China, in 1988. She received the bachelor's degree in communication engineering from the Changchun University of Science and Technology, China, in 2011, and the master's degree in signal and information processing from Jilin University, in 2014. She is currently pursuing the Ph.D. degree with the Changchun University of Science and Technology. She has published one SCI paper and one EI paper as the first author: ''Research on Distributed In-Vehicle Wireless Self-Organized Routing Protocol Distribution Mechanism'' and ''Application of Improved Ant Colony Optimization in Vehicular Ad-Hoc Network Routing.'' Her research interests include cooperative communication in the Internet of Vehicles and sensor networks.

GUIFEN CHEN was born in Yushu, Jilin, China, in 1964. She received the B.S. and M.S. degrees in information and communication engineering from the Jilin University of Technology, China, in 1986 and 1991, respectively, and the Ph.D. degree in optical engineering from the Changchun University of Science and Technology, China, in 2009. She is currently a Professor of information and communication engineering with the Changchun University of Science and Technology. The projects she is responsible for include ''National Thirteen-Five Year Plan Equipment Pre-Research Joint Fund: Unmanned Platform XX Communication Technology Research'' and ''National Thirteenth-Five-Year Arms Industry Joint Fund: XX Equipment XX Research.'' The patents she invented include ''An optimization method for wireless ad hoc network collaboration time synchronization'' and ''Heterogeneous cognitive wireless sensor network clustering routing method.'' She has published two SCI papers as the corresponding author: ''Research on Multi-Agent D2D Communication Resource Allocation Algorithm Based on A2C'' and ''A Method of Relay Node Selection for UAV Cluster Networks Based on Distance and Energy Constraints.'' Her research interests include optical information and wireless communication technology, the Internet of Things, and sensor networks.