



Multi-Channel Scheduling for Short-Range Wireless Communication Networks Using a Q-Learning Feedback Mechanism

Li Yan^{1*}, Hongzhang Han²

¹ School of Information Engineering, Changzhou Vocational Institute of Industry Technology, 213164 Changzhou, China

² School of Computer Engineering, Jiangsu University of Technology, 213001 Changzhou, China

* Correspondence: Li Yan (yanli@ciit.edu.cn)

Received: 07-05-2024

Revised: 08-20-2024

Accepted: 09-03-2024

Citation: L. Yan and H. Z. Han, “Multi-channel scheduling for short-range wireless communication networks using a Q-learning feedback mechanism,” *Inf. Dyn. Appl.*, vol. 3, no. 3, pp. 171–183, 2024. <https://doi.org/10.56578/ida030303>.



© 2024 by the author(s). Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

Abstract: The traditional channel scheduling methods in short-range wireless communication networks are often constrained by fixed rules, resulting in inefficient channel resource utilization and unstable data communication. To address these limitations, a novel multi-channel scheduling approach, based on a Q-learning feedback mechanism, was proposed. The architecture of short-range wireless communication networks was analyzed, focusing on the core network system and wireless access network structures. The network channel nodes were optimized by deploying Dijkstra's algorithm in conjunction with an undirected graph representation of the communication nodes within the network. Multi-channel state characteristic parameters were computed, and a channel state prediction model was constructed to forecast the state of the network channels. The Q-learning feedback mechanism was employed to implement multi-channel scheduling, leveraging the algorithm's reinforcement learning capabilities and framing the scheduling process as a Markov decision-making problem. Experimental results demonstrate that this method achieved a maximum average packet loss rate of 0.03 and a network throughput of up to 4.5 Mbps, indicating high channel resource utilization efficiency. Moreover, in low-traffic conditions, communication delay remained below 0.4 s, and in high-traffic scenarios, it varied between 0.26 and 0.4 s. These outcomes suggest that the proposed approach enables efficient and stable transmission of communication data, maintaining both low packet loss and high throughput.

Keywords: Q-learning feedback mechanism; Short range; Wireless communication networks; Multi-channel scheduling; Channel state; Markov decision-making

1 Introduction

With the rapid development of wireless communication technology, short-range wireless communication networks have been widely used in the fields of smart home, industrial automation, and the Internet of Things (IoT). However, with the expansion of the network scale and the increase in the number of devices, the limited communication resources have become a key factor restricting the performance improvement of short-range wireless communication networks. The traditional single-channel communication method has difficulty meeting the demand for simultaneous communication of large-scale devices. Therefore, the multi-channel scheduling method has become an effective way to solve this problem [1, 2]. The method improves the overall throughput and communication efficiency of the network by reasonably allocating multiple communication channels so that devices can communicate simultaneously on different channels [3]. The multi-channel scheduling method can effectively alleviate the problem of tight communication resources, improve the communication capacity and efficiency of the network, and provide the possibility of concurrent communication for large-scale devices. A reasonable multi-channel scheduling strategy can reduce communication conflicts and interference, improve communication quality, and enhance the stability and reliability of the network [4]. In addition, research on multi-channel scheduling methods can also promote the integration of wireless communication technology with other technologies and facilitate the innovative development of wireless communication networks. Therefore, in-depth research on multi-channel scheduling methods for short-range wireless communication networks is of great significance for improving the performance of wireless communication networks and promoting the development of related industries.

Mohamed et al. [5] proposed a multi-channel scheduling method based on the Collaborative Software Review System (CSRS), which relies on effective evaluation metrics to estimate the quality of the available communication channels and exchanges blacklists of undesirable channels locally and statelessly, aiming to reduce the impact of data interference of undesirable channels during network communication. However, this method suffers from large communication delays. Based on an enhanced beetle antenna search algorithm for multi-channel scheduling, Kumar et al. [6] obtained the optimal solution to the static data flow and power measurement problems by reducing the overall regression of the network, realizing the effective allocation and scheduling of optimal channels. However, this method has the problems of failing to realize the dynamic allocation of channel resources and low utilization of channel resources. Hu et al. [7] proposed a channel scheduling and allocation method based on federated edge learning (FEEL), which sets workload constraints in cellular networks, transforms the channel scheduling and allocation problem into a problem of objective function integral variable summation and complex structural interference term solution, and obtains its suboptimal solution with lower computational complexity based on the learning advantage of the dual greedy strategy, effectively realizing the channel scheduling work. However, this method is easy to be affected by the network data quality in the problem transformation process and the problem of large breakfast communication delay. Luong et al. [8] proposed a network channel scheduling method based on blockchain technology, which takes the IoT device as a secondary transmitter, verifies, stores, and processes the network communication data collected from the IoT device in a decentralized but trustworthy way, effectively realizing the communication time scheduling between the IoT devices. However, the communication data verification process of this method is difficult to calculate and prone to the problem of low channel resource utilization.

As a kind of killing and maiming learning algorithm, the Q-learning feedback mechanism can utilize its own reinforcement learning ability to interact with the application environment and learn how to output the optimal decision. In multi-channel scheduling of short-range wireless communication networks, the application of the Q-learning feedback mechanism can help network nodes learn how to select appropriate channels for transmission according to the current channel state and communication demand, and dynamically allocate spectrum resources to different users or devices to adapt to the changes in network load and different communication demands. Based on the above advantages, this research thoroughly studies the multi-channel scheduling method for short-range wireless communication networks based on the Q-learning feedback mechanism. By considering the reinforcement learning process as a Markov decision-making process, the innovative application of the Q-learning feedback mechanism can realize efficient and stable multi-channel scheduling for short-range wireless communication networks.

2 Multi-Channel Scheduling for Short-Range Wireless Communication Networks

2.1 Optimized Restructuring of the Short-Range Wireless Communication Network Topology

2.1.1 Analysis of the short-range wireless communication network architecture

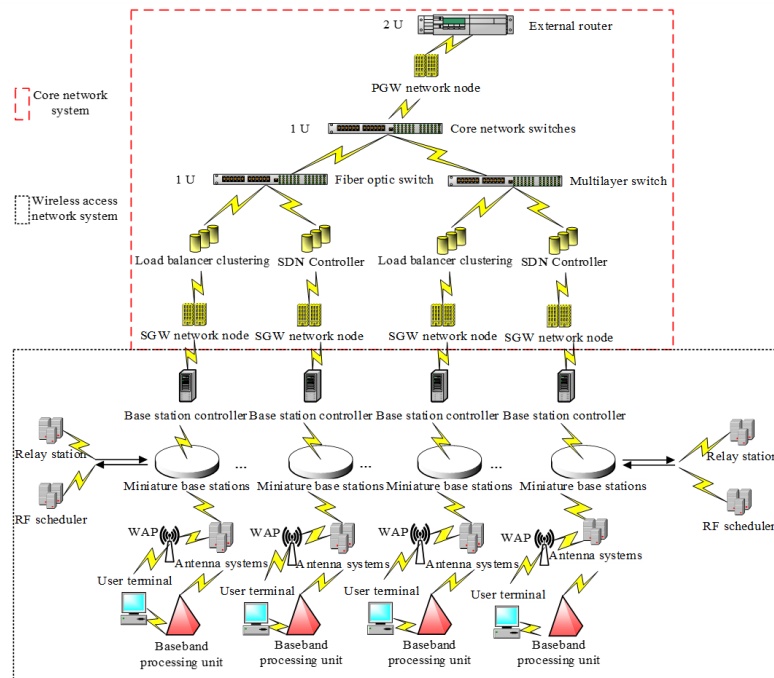


Figure 1. Short-range wireless communication network architecture

This study analyzes the architecture of short-range wireless communication networks by delving into the topology and underlying logic of the core network system architecture and the wireless access network system architecture. The short-range wireless communication network architecture is shown in Figure 1.

As shown in Figure 1, the architecture of short-range wireless communication networks is mainly composed of two parts: the core network system architecture and the wireless access network system architecture. Among them, the core network system architecture, as the backbone part of the network architecture, undertakes the important tasks of routing connection management and network communication data transmission processing. This system architecture mainly consists of external routers, packet network gateway (PGW) network nodes, core network switches, fiber optic switches, multilayer switches, load balancer clusters, software-defined networking (SDN) controllers, and serving gateway (SGW) network nodes.

Specifically, the main operation principle of the architecture is as follows: the core router located in the center of the core network system connects the PGW core network node, which handles IP address allocation, quality of service (QoS) policy control, and packet filtration. Upon establishing a connection between the user equipment and the external IP network, the core router facilitates data exchange and directs the forwarding of data packets from the edge equipment and the external network based on destination address information. This process ensures accurate packet delivery to the appropriate next-hop network for forwarding decisions. When the packet forwarding operation starts, the fiber optic switches and multilayer switches located in the core network system centralize the packet exchange and forwarding operation by connecting different network devices, servers, data centers and other network nodes. At the same time, during the data transmission process, after receiving the packets forwarded by the upper-layer network switches, the load balancer cluster and the SDN controller jointly perform load and traffic balancing distribution and management operations among multiple servers. The operations are facilitated by SGW network nodes, which function as portable gateways, while also managing mobility and ensuring the security of the short-range wireless network communication.

The main operation principle of the wireless access network system architecture is as follows: when multiple servers perform traffic load balancing distribution and management in parallel through the SGW network node, the short-distance network communication enters into the wireless access network system, and the base station controller, which is responsible for the wireless resource allocation management and call control of multiple base stations, coordinates the wireless network communication between base stations.

Since the network topology has short-range characteristics, the base station attributes in this structure are miniature base stations capable of providing localized wireless coverage, high-capacity and high-data-rate wireless connectivity within a limited coverage area as well as good indoor coverage. In order to achieve short-distance wireless communication, the radio frequency (RF) scheduler and the relay station schedule and manage the wireless resources between the miniature base stations, and through the mutual cooperation of the two, expand the coverage range and strength of the wireless communication signals, extend the signal transmission distance, and optimize the signal quality of the wireless coverage area.

Specifically, the RF scheduler mainly converts and transmits communication signals through the RF transceiver module and the RF front-end module. When the relay station makes up for the signal attenuation wirelessly, the signal continues to be transmitted to the RF transceiver module, and the communication signal is converted into a digital signal. Then the digital signal is converted into a RF signal to be sent to the antenna system, and the RF front-end module performs signal amplification, filtering, adjustment and other enhancement processing operations in the sending process. When the signal passes through the antenna system, it reaches the baseband processing unit, which digitizes the signal and performs coding and decoding operations simultaneously. When the data processing operation is completed, the signal is fed back to multiple user terminals through the wireless access point, finally realizing short-distance wireless communication.

2.1.2 Optimized deployment of channel nodes

The operation of short-range wireless communication networks involves the adjustment and processing of signal coverage and quality between base stations. Therefore, in order to effectively expand the signal coverage, reduce multi-channel interference during data transmission, and avoid signal blind zones or coverage overlap and other problems, it is necessary to optimize the deployment of network nodes to improve the stability of data transmission.

This study combines the Dijkstra algorithm with the undirected graph of short-distance wireless communication network nodes to optimize the deployment of network channel nodes. The Dijkstra algorithm is a classic shortest path algorithm, with the basic idea of gradually expanding the shortest path from the starting point to all other nodes until the shortest path of the target node is found. This algorithm adopts a greedy strategy, selecting the node closest to the starting point from unlabeled nodes at each step for labeling, and updating the distance value of its neighboring nodes. This gradually determines the shortest distance from each node to the starting point. When applying the Dijkstra algorithm to the deployment problem of network channel nodes on undirected graphs, each node can be considered as a candidate position for channel nodes, with edges representing the connection relationships between nodes and edge

weights representing the distance or cost between nodes. By solving the shortest path, the optimal layout between channel nodes can be determined to achieve more effective communication. The specific steps are as follows:

Step 1: Creation of an undirected graph of routing nodes for short-range wireless communication networks

Based on the topology of Figure 1 and the connectivity between network devices, the design of the undirected graph structure is shown in Figure 2.

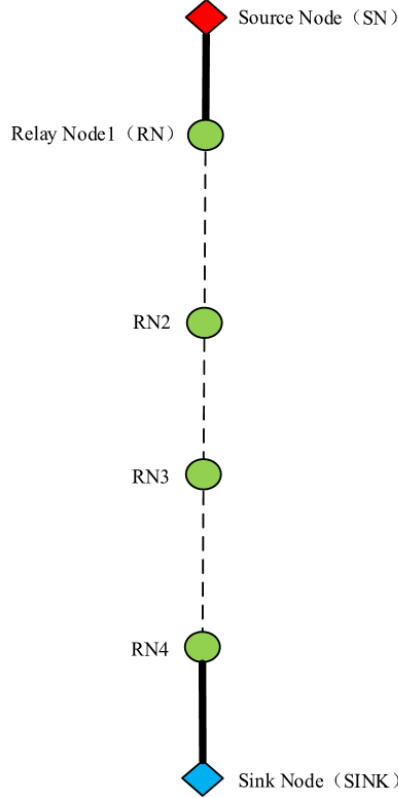


Figure 2. Undirected graph of routing nodes for short-range wireless communication networks

As shown in Figure 2, the undirected graph mainly consists of one source node, one aggregation node, four relay nodes, and five edges. The network communication nodes in the undirected graph mainly represent the device entities participating in the wireless network communication and packet forwarding; the edges represent the wireless communication connections between the devices; and the independent weights corresponding to each edge denote the actual distance between the two device entities and the communication cost [9–11]. In this study, $G = (V, E, W)$ is the vectorless graph of routing nodes for short-range wireless communication networks, where V denotes the set of nodes, E denotes the set of edges, W is the independent mapping weight coefficients of the edges, s_1 is the source node, s_2, s_3, s_4, s_5 are all relay nodes, and s_6 is the aggregation node.

The source node represents the originating node of the communication data or control information, i.e., the initial location where the starting data is generated or the request is initiated. In a short-range wireless communication network, the source node usually represents the sender of the data packet or the issuer of the information, i.e., the sender of the data to the other nodes in the network, which triggers the important conditions for the data to be transmitted and communicated in the channel.

The relay node represents the node that assumes the forwarding function in the data transmission process. In the process of data transmission from the source node to the aggregation node, the relay node demonstrates the important function of connecting the transmission link and extending the coverage through data forwarding, relaying, and routing as well as other operations [12–14], that is, to help the data packets to be transmitted to the different nodes and paths in the network through the channel to reach the aggregation node.

The aggregation node denotes the final receiving node of the data or information, i.e., the end point or destination of the data flow. In a wireless communication network, the aggregation node represents the final destination of the communication packet, where the communication node receives and processes the channel transmission data from the source or relay node to complete the entire communication process.

Step 2: Initialization of the distance array and the precursor array

On the basis of obtaining the node layout structure of short-range wireless communication networks, the distance

array and the antecedent period array in Dijkstra's algorithm were introduced to represent the updating distance of node paths in the undirected graph [15]. For the distance array $d[\cdot]$, its initialization setting process can be specifically expressed as follows:

$$d[s_1] = 0 \quad (1)$$

$$d[s_n] = \infty, n \neq 1 \quad (2)$$

where, s_n denotes any node, n denotes the node ordinal identity and $d[s_n]$ denotes the shortest distance value from the source node to any node s_n .

In the initial phase of node optimization deployment, the distance initialization of the s was set to 0, indicating that the distance from the source node to itself is 0. The other node distances were initialized to ∞ , which is a tagged value and does not involve the direction of infinitely varying intervals, indicating that it is unreachable at the initial time.

For the antecedent array $p[\cdot]$, it is mainly used to record the previous node (antecedent node) of each node on the path in the shortest path from the source node to the other nodes. Through the antecedent array, it is possible to backtrack the specific path, i.e., from the target node all the way back to the source node [16–18], to get all intermediate nodes on the shortest path. In the initialization process, the predecessor nodes of all nodes were set to \emptyset (null or non-existent value).

$$p[s_n] = \emptyset, n = R \quad (3)$$

where, $p[s_n]$ denotes the antecedent array of all nodes. The initialization state is that the shortest path from the source node to the node has not yet been determined. When the algorithm updates the shortest path of a node, it simultaneously updates the information of the predecessor node of the corresponding node.

Step 3: Creation of an update store collection

During the actual execution of the algorithm, the set S was used to store the nodes for which the shortest distance was determined during the execution of the algorithm, with $S = [\emptyset]$ initially.

Step 4: Construction of the shortest path tree

As the algorithm was executed, the nodes in S were continuously updated during the loop processing until the shortest path from the source node to all other nodes was found. Specifically, the loop processing session first involves determining the shortest path candidate nodes by traversing all nodes that have not joined S and selecting a node s_d with the shortest distance to the source node among the nodes other than S . After joining s_d to S , each of the neighboring nodes of s_d was defined as \bar{s}_d , and the shortest path candidate nodes were determined by calculating the distance d' from which s_d arrives to \bar{s}_d .

$$d' = d(s_d) + W(s_d, \bar{s}_d) \quad (4)$$

where, $d(s_d)$ denotes the current shortest distance of s_d , and $W(s_d, \bar{s}_d)$ denotes the independent mapping weight coefficients of the edge s_d, \bar{s}_d .

Formula (4) was repeated. $d(\bar{s}_d)$ was updated to d' , which is less than $d(s_d)$, and $p[\bar{s}_d]$ was set to s_d . If \bar{s}_d was not considered as a candidate node for the shortest path, it should be added to the candidate set.

At the end of the loop, the shortest path tree was constructed [19]. $p[\cdot]$ at each node records the predecessor nodes on the shortest path from the source node to that node.

Step 5: Determination of the optimal deployment of network routing nodes

After iterating the shortest path tree and analyzing the frequency of each node appearing on the shortest path, the optimal deployment of routing nodes was determined to achieve the optimal reorganization of the short-range wireless communication network topology [20–22]. The determination process pseudo-code for the final optimized deployment of the routing node s_r is as follows:

```
# Assume shortest_paths is a dictionary where the keys are the destination nodes and the values are the list of
shortest path nodes from the source node to that destination node (excluding the source node)
# For example: shortest_paths[j] = [node1, node2, ... , nodek]
# Initialize the node frequency counter
frequency = node: 0 for node in G.nodes() # G is an undirected graph
# Traverse the shortest path tree and update node frequencies
for target, path in shortest_paths.items(): for node in path: for node in G.nodes()
    for node in path.
        frequency[node] += 1
```



```

# Determine the routing nodes
# For example, select the top N nodes with the highest frequency as routing nodes
N = 5 # Assume the first 5 nodes are selected as routing nodes
sorted_nodes = sorted(frequency, key=frequency.get, reverse=True)
routing_nodes = sorted_nodes[:N]
print("Selected routing nodes:", routing_nodes)

```

2.2 Multi-Channel State Prediction for Short-Range Wireless Communication Networks

The core of the Dijkstra algorithm is to calculate the shortest path from the source node to all other nodes in the network. Therefore, during the acquisition process (the routing node optimization deployment process), the algorithm considers all possible communication paths and selects the communication path that offers the minimum delay, maximum bandwidth, or optimal overall performance. When evaluating the application of the Dijkstra algorithm to the deployment problem of network channel nodes on undirected graphs, it is first necessary to understand the characteristics and principles of the Dijkstra algorithm. The algorithm is a classic shortest path algorithm, with the basic idea of gradually expanding the shortest path from the starting point to all other nodes until the shortest path to the target node is found. This algorithm adopts a greedy strategy, selecting the node closest to the starting point from unlabeled nodes at each step for labeling, and updating the distance value of its neighboring nodes. This gradually determines the shortest distance from each node to the starting point. When applying the Dijkstra algorithm to the deployment problem of network channel nodes on undirected graphs, each node can be considered as a candidate position for channel nodes, with edges representing the connection relationships between nodes and edge weights representing the distance or cost between nodes. By solving the shortest path, the optimal layout between channel nodes can be determined to achieve more effective communication. The specific steps for constructing a channel state prediction model usually include data collection and preprocessing, feature extraction and selection, model selection and training, model evaluation and optimization, and model application and deployment. The advantages include helping network managers optimize network resource allocation, improving network performance, and enhancing the user experience. The disadvantages involve high accuracy dependence, real-time and robustness challenges, data privacy and security issues. In summary, building a channel state prediction model is an important means to improve network management efficiency and performance, but it needs to be continuously optimized and improved in practice to overcome the challenges and limitations.

The core of Dijkstra's algorithm is to calculate the shortest path from the source node to all other nodes in the network. Therefore, in the process of obtaining s_r (routing node optimization deployment process), the algorithm considers all possible communication paths and selects the one consisting of s_r and having the smallest latency, the largest bandwidth, or the best overall performance. At the same time, since the routing node can support multi-channel communication and dynamically adjust the use of channels according to the network state, the multi-channel structure under the influence of s_r exhibits the comprehensive attributes of reasonable allocation, high utilization of network resources, and high communication efficiency.

In order to understand the availability, loading, and potential interference factors of each channel, the load differences between different channels were analyzed so that data transmission tasks can be evenly distributed among the channels to avoid overloading some channels while others are idle. Multi-channel state prediction is needed based on the multi-channel structure under the influence of s_r [23]. The characterization parameters of the multi-channel state are shown below.

The objective of the channel time-domain feature estimation is to determine the impulse response of the channel during the time-domain transmission of the signal from the transmitter to the receiver.

Specifically, for each moment l_1 , the formula for the channel time domain feature estimation is expressed as follows:

$$\eta_1 = \left(\iota_1 \frac{\hat{x}}{\iota_2 \sqrt{y}} \right)^{s_r d'} - \iota_1 R(\hat{x}) + \frac{\rho(h | y)}{\iota_2^2 y} \quad (5)$$

where, η_1 denotes the multi-channel time-domain characteristic parameter, which represents the impulse response of the channel during the time-domain transmission of the time-domain signal from the transmitter to the receiver; l_1 denotes the discrete time index, i.e., the sampling point of the signal in the time domain space; l_2 denotes the noise interference factor in the received signal; y denotes the received time-domain signal, i.e., the result of the transmitted signal after multiple channels and noise effects; \hat{x} denotes the transmitted guide frequency sequence; $R(\cdot)$ denotes the autocorrelation function, which is used to describe the correlation of variables at different points in time; $\rho(h | y)$ denotes the posterior probability density function of the channel value h when y is known; and l_2^2 denotes the mean square error of the noise.

In the process of η_1 calculation, it is necessary to focus on the influence of multi-channel noise factors. Therefore, the multi-channel signal-to-noise ratio is calculated as a multi-channel state characteristic.

$$\eta_2 = \frac{P(l_2) - P(y)}{\sqrt[l_1]{\hat{x} - |l_2|^{s_r d'}}} \quad (6)$$

where, η_2 denotes the multi-channel signal-to-noise ratio; $P(l_2)$ denotes the noise power; and $P(y)$ denotes the signal power.

Using the Fourier transform to convert y to a frequency domain signal [24], the formula for the channel frequency domain feature estimation is expressed as follows:

$$\eta_3 = \delta \left(y + \sqrt[l_1]{\hat{x} - |l_2|^{s_r d'}} \right) + \frac{\hat{x}}{y'''} \quad (7)$$

where, η_3 denotes the feature parameter of the multi-channel frequency domain, which represents the frequency response characteristics of the multi-channel at different frequencies; $\delta(\cdot)$ denotes the Fourier transform factor [25]; \hat{x} denotes the transmitter-side frequency-domain derivative sequence; and y''' denotes the signal spectrum variation coefficient at the receiving end.

Since the signals on multiple channels propagate over different distances, superimposed interference occurs when they arrive at the receiving end. Therefore, if the signal phases of different channels are in the same phase, the signal amplitude can be enhanced under superimposed interference, i.e., the multipath gain effect. On the contrary, if the signal phases of multiple channels are in anti-phase, the superimposed interference can weaken the signal amplitude, forming the multipath attenuation effect. Affected by this environmental condition, the multiple channels receive signals in the wireless communication network environment, which produces a large change in the signal amplitude, resulting in multi-channel energy attenuation. The shift characteristic of the multi-channel signal phase is expressed as follows:

$$\eta_4 = \partial_1 + 2\partial_2 \frac{\beta(\hat{x}) - y'''}{\beta(y''') + \hat{x}} \quad (8)$$

where, η_4 denotes the offset characteristic parameter of the multi-channel signal phase; ∂_1 denotes the signal gain amplitude (no imaginary part units involved); ∂_2 denotes the complex gain offset angle; $\beta(\hat{x})$ denotes the transmitter signal offset; and $\beta(y''')$ denotes the signal offset at the receiving end.

Similarly, on the basis of analyzing the equivalent value of the phase shift of the multi-channel signal, the Doppler shift generated by the relative motion between the transmitter and receiver was further analyzed as one of the multi-channel state characteristics.

$$\eta_5 = \frac{2 \cos \theta}{v - f \tan \theta} \quad (9)$$

where, η_5 denotes the Doppler shift coefficient of the multi-channel signal; θ denotes the angle between the motion direction of the communication antenna system and the direction of signal propagation; v denotes the relative velocity between the signal sender and the signal receiver; and f denotes the center frequency used in wireless communications.

Associating Eqs. (5)-(9), the channel state prediction model was constructed based on the characteristic parameters of the multi-channel state.

$$\eta(h)^{s_r d'} = \check{\eta}(h) \left(\eta_1 \eta_2 + \sqrt[l_1]{\eta_3 (\hat{x} - |l_2|^{s_r d'})} - \eta_4 + \eta_5 \right) \quad (10)$$

where, $\eta(h)^{s_r d'}$ denotes the multi-channel state prediction model; $\check{\eta}(h)$ denotes the set of multi-channel randomly distributed states, and

$$\begin{aligned} \check{\eta}_1 &\rightarrow \text{Idle Channel Status} \\ \check{\eta}_2 &\rightarrow \text{Change channel state} \\ \check{\eta}(h) = [\check{\eta}_1, \check{\eta}_2, \check{\eta}_3, \check{\eta}_4, \check{\eta}_5] &= \check{\eta}_3 \rightarrow \text{Hold channel state} \\ &\check{\eta}_4 \rightarrow \text{Hold channel state} \\ &\check{\eta}_5 \rightarrow \text{Interference channel state} \end{aligned}$$

2.3 Multi-Channel Scheduling for Short-Range Wireless Communication Networks Based on the Q-Learning Feedback Mechanism

On the basis of obtaining $\check{\eta}(h)$ (the availability, load and potential interference factors of different channels have been clarified), in order to reduce the inter-channel interference in the communication scheduling process, make full use of the available spectrum resources to the maximum extent, and realize stable multi-channel scheduling, this study realizes multi-channel scheduling for short-range wireless communication networks based on the Q-learning feedback mechanism and the reinforcement learning capability of the Q-learning algorithm.

Since the channel state and selection at each moment are usually only affected by the current moment in Q-learning, the learning process is considered as a Markov Decision Process (MDP) model to synthesize the learning process, i.e., the next state of the system is considered to depend only on the current state [26] and is independent of the nature of the past state. The specific scheduling learning steps are as follows:

Step 1: Definition of state space and behavior space

The MDP was utilized to reflect the channel state and selection behavior. In order to make the decision process more convincing, the proposed channel is in a disturbed state (which is not presented in the specific decision-learning process). The multi-channel scheduling MDP is shown in Figure 3.

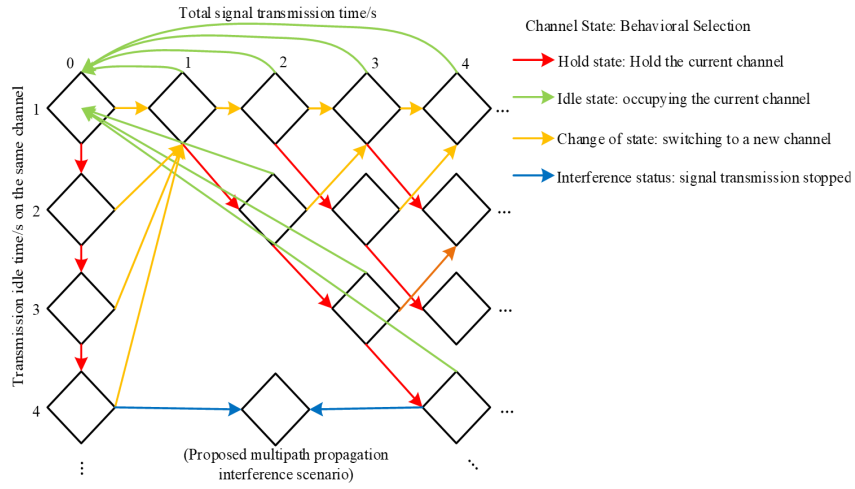


Figure 3. Multi-channel scheduling MDP

As can be seen from Figure 3, the diamond-shaped nodes in the decision-making process represent the states of each channel in the multi-channel scheduling process, and the directional lines between the nodes represent the corresponding data transmission and signal adjustment behaviors taken by the channel in each transmission state. Based on the decision-making principle in Figure 3, the state space $\check{\eta}$ and behavior space \check{A} were mapped to contain the channel states and the corresponding behavioral choices, respectively. Figure 3 shows the specific state and choice results.

Step 2: Initialization of the Q-Table

A Q-table was initialized, whose size is the product of the state space scale and the behavior space scale.

$$Q(\check{\eta}, a) = \frac{\check{\eta}}{\check{\eta}} \times \frac{a}{\check{A}} > q \quad (11)$$

where, $Q(\check{\eta}, a)$ denotes each entry in the Q-table; $\check{\eta}$ denotes the state space factor; a denotes the behavioral space factor; and q denotes the Q rate of the change convergence threshold. Each $Q(\check{\eta}, a)$ is the expected payoff of executing a in state $\check{\eta}$.

Step 3: Selection of behaviors

At each time step, actions were selected according to an ε -greedy strategy. A random action was selected with a certain probability ε to be explored in order to discover potentially better choices. The action with the largest Q value in the current state was selected with a probability of $1 - \varepsilon$ to be exploited in order to maximize the immediate payoff.

Step 4: Implementation of behaviors

The selected behavior a_1 was performed, i.e., selecting a certain channel for transmission and observing whether the transmission is successful or not. A successful transmission gives a positive reward while failure or high latency gives a negative reward.

Step 5: Update of the Q-Table

Considering the channel occupancy and the interference level of each channel, the Q-table update operation was performed according to the following equation to realize the dynamic allocation of channel resources:

$$Q''(\ddot{\eta}, a) = \frac{\ddot{\eta}}{\ddot{\eta}} \times \frac{a}{A} + \left[\varphi_2 \left(\frac{H}{\varphi_1} \right)^{L_H} - |\ddot{\eta}_1|^{\gamma_1} \right]^{\gamma_2} \quad (12)$$

where, $Q''(\ddot{\eta}, a)$ denotes the Q-table update result; φ_2 denotes the single-channel occupancy time; φ_1 denotes the total multi-channel occupancy time; H denotes the channel; L_H denotes the channel occupancy; $\ddot{\eta}_1$ denotes the new state observed after the execution of a_1 ; γ_1 denotes the discount factor used to weigh the importance of current and future rewards; and γ_2 denotes the learning rate that controls the extent to which new information affects old information.

Step 6: Algorithm convergence

Steps 2 to 5 were continuously repeated till the convergence of the Q-table, with the updated result being less than q . After convergence, the algorithm learns a stable policy, and quickly selects the best action for multi-channel assignment and interference reduction based on the current state, thereby realizing the multi-channel scheduling for short-range wireless communication networks.

3 Experimental Analyses

In order to verify the practical application performance of designing a multi-channel scheduling method for short-range wireless communication networks based on the Q-learning feedback mechanism, a localized short-range wireless communication network area was used as an experimental object and the Matlab platform was applied to conduct simulation experiments. About 140 communication nodes were deployed in the experiment. These nodes were randomly distributed in a 120m×120m area to simulate the actual short-range wireless communication network environment. Each node in the experimental environment has the same communication range of 30 m, i.e., each node can only communicate with other nodes whose distance is within 30 m. In the experimental area, there is a source node responsible for collecting information from other nodes throughout the area. The channel configuration information is based on the IEEE 802.11 b/g wireless network standard in the 5 GHz band, which defines multiple non-overlapping channels such as 36, 40, 44, 48, 149, 153, 157, 161, 165 and so on. The specific experimental equipment is shown in Table 1.

Table 1. Experimental equipment and performance parameters

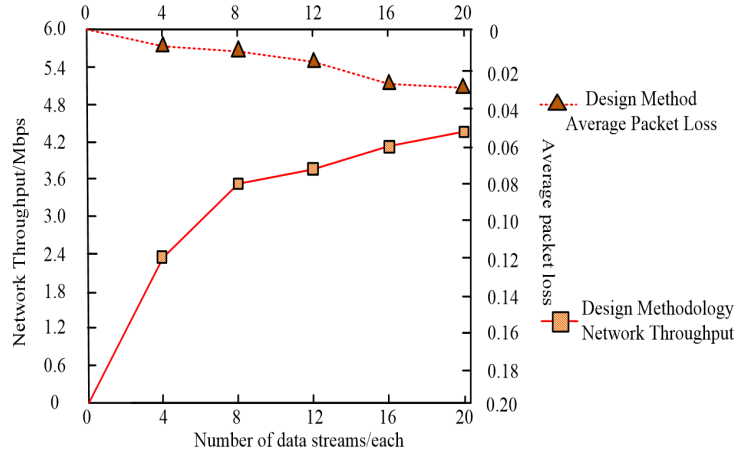
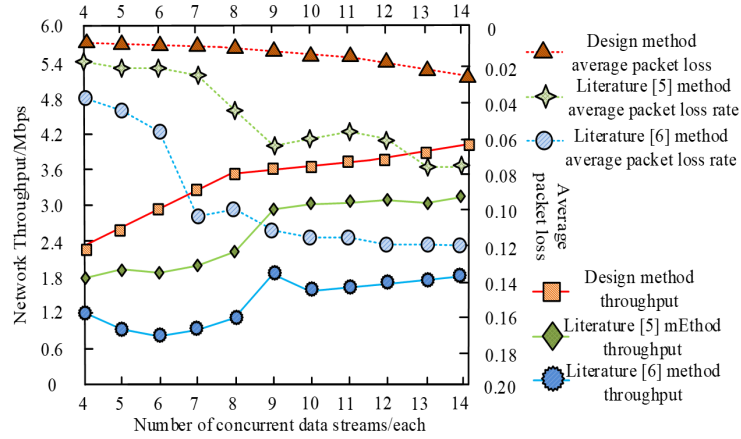
| Equipment (Software and Hardware) Model | Performance Parameters |
|--|---|
| Universal Software Radio Peripheral (USRP) B210 | Frequency range: 70MHz – 6GHz Bandwidth: 56 MHz Sample rate: 100 MSPS Support communication standards: LTE, Wi – Fi, GPS, etc. Software support: GNU Radio, MATLAB, etc. Power consumption: Approx. 8 W |
| Keysight N9344C Spectrum Analyzer | Frequency range: 9kHz – 20GHz (covering different models) RBW range: 10 Hz minimum DANL: - 161 dBm/Hz @ 1 GHz Real-time bandwidth: 25 MHz or 40 MHz selectable Connectivity: USB, LAN, VGA Software support: Keysight N934xC PC software |
| TelosB Sensor Node | Processor: TI MSP430F1611 Storage capacity: 10 KB RAM, 48 KB flash memory Communication interface: IEEE 802.15 .4 standard wireless communication interface Transmission rate: 250 kbps Operating frequency: 5.0 GHz |
| Omnet ++5.6 .2 | - |
| NS-3.33 Network Emulator | - |
| MATLAB R2022a | - |

The parameters of the simulation experiment are shown in Table 2.

In order to verify the channel resource utilization efficiency of the design method, the average packet loss rate and network throughput were taken as the evaluation indexes and were tested under different numbers of data streams. The results are shown in Figure 4.

Table 2. Parameters of the simulation experiment

| Simulation Parameters | Value/Mode Information |
|-----------------------------|-------------------------------------|
| Total number of iterations | 12000 |
| Intensive learning/phase | 22 |
| Transmit power/dBm | 18 Learning rate & 0.0016 |
| Switching delay/ms. | 2 |
| Switching interval/ms | 5 |
| Simulation time/ms | 200000 |
| Discount rate | 0.028 |
| Channel modulation mode | Non-coherent frequency shift keying |
| Number of data streams/each | 20 |

**Figure 4.** Average packet loss rate and throughput test results of the design approach**Figure 5.** Test results of the channel resource utilization efficiency of different methods in parallel data flow

As can be seen in Figure 4, the average packet loss rate and network throughput, generated during the multi-channel scheduling test, using the design method show an increasing trend as the number of data streams increases. When the number of data streams reaches the maximum value, the average packet loss rate and network throughput of the design method have the maximum value of 0.03 and 4.5 Mbps, respectively, which is able to maintain a lower packet loss rate and a higher throughput and obtain a high channel resource utilization efficiency. This is mainly due to the fact that the design method provides an in-depth understanding of the channel load and parallel processing by calculating the multi-channel state characteristic parameters before channel scheduling. It is shown that the method can make full use of each channel to avoid the loss of communication data due to channel conflict or congestion, thus optimizing the overall performance of the communication network.

The number of parallel data streams in the data stream ranges from 4 to 14. In order to further verify the practical

application performance of the design method, the methods proposed by Mohamed et al. [5] and Kumar et al. [6] were introduced as comparison methods. The message sending rate was uniformly set at 185 kbps. The average packet loss rate and network throughput of different methods under different numbers of concurrent data streams were verified, and the specific test results are shown in Figure 5.

As can be seen from Figure 5, with the continuous increase in the number of parallel data streams, the average packet loss rate of the design method is lower than that of the other methods, with the highest value of only 0.026. In addition, the network throughput is higher than that of the other methods, with the highest value of 4.0 Mbps. In the case of a gradual increase in the number of concurrent data streams, the design method is able to maintain a stable integrated scheduling performance (i.e., the packet loss rate does not rise significantly, and the throughput maintains a high level), indicating that the method has good scalability. This means that in practical applications, the design method can effectively allocate data streams to different channels, avoiding the situation that some channels are overloaded while others are idle, thereby effectively realizing the dynamic allocation and the high utilization efficiency of channel resources.

In order to verify the comprehensive response performance of the design method, the methods proposed by Mohamed et al. [5] and Kumar et al. [6] were introduced as comparison methods under different data traffic loads. In packet transmission under the number of channels 36, 40, and 44, the communication delay generated by the method was compared and analyzed. The results of the comprehensive response performance test are shown in Figure 6.

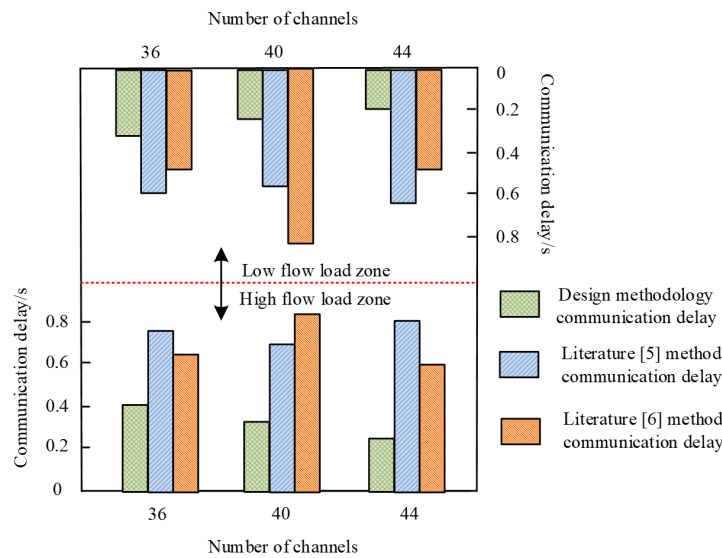


Figure 6. Results of the comprehensive response performance test

As can be seen from Figure 6, in the low traffic load area, using the design method for scheduling and processing of different channels under the 5GHz band, the communication delays generated by the data transmission during processing are all lower than 0.4 s, and the communication delay for processing data packets in channel 44 is only 0.2 s, which is lower than that of other methods. In the high traffic load area, due to the increase in the amount of data, the difficulty of parallel processing increases, and the communication delay generated by utilizing the design method is slightly higher than that in the low traffic load area, with the highest value of 0.4 s and the lowest value of 0.26 s, which are also lower than the other two methods. The channel optimization control using the design method can make the data in the transmission process produce a smaller time delay, which in turn ensures that the data can quickly arrive at the receiving end from the transmitting end, thereby meeting the needs of real-time communication, short-distance wireless communication and other application scenarios with rapid response characteristics and realizing the efficient and stable transmission of data.

4 Conclusion

In summary, this study innovatively proposes a multi-channel scheduling method for short-range wireless communication networks based on the Q-learning feedback mechanism to achieve efficient and stable multi-channel scheduling in short-range wireless communication networks. The method studied in this research optimizes the deployment of network channel nodes by analyzing the architecture of short-range wireless communication networks. The channel state characteristic parameters were calculated to predict the multi-channel state of short-range wireless communication networks. In addition, the high-efficiency and stable multi-channel scheduling of short-range wireless communication networks was carried out based on the Q-learning feedback mechanism. The experimental results

show that the multi-channel scheduling using the design method can maintain a high throughput and a low packet loss rate, obtain a high utilization efficiency of channel resources, and effectively realize the dynamic resource allocation. It enables the data to have a smaller time delay in the transmission process and thus ensures that the data can quickly reach the receiver from the sender. This method learns and adjusts the optimal channel selection strategy in real time according to different network environments and communication requirements, maximizes the use of available channel resources, improves the stability and efficiency of short-range wireless network communication, and has important research significance for promoting the innovative and modern development of the short-range wireless communication technology.

Funding

The paper was funded by the National Natural Science Foundation of China (Grant No.: 61602216).

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] J. Du, B. Jiang, C. Jiang, Y. Shi, and Z. Han, "Gradient and channel aware dynamic scheduling for over-the-air computation in federated edge learning systems," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1035–1050, 2023. <https://doi.org/10.1109/JSAC.2023.3242727>
- [2] J. Xiong, H. Hu, P. Cheng, C. Yang, Z. Shi, and L. Gui, "Wireless resource scheduling for high mobility scenarios: A combined traffic and channel quality prediction approach," *IEEE Trans. Broadcast.*, vol. 68, no. 3, pp. 712–722, 2022. <https://doi.org/10.1109/tbc.2022.3141609>
- [3] L. Yang, Y. Xu, Z. Huang, H. Rao, and D. E. Quevedo, "Learning optimal stochastic sensor scheduling for remote estimation with channel capacity constraint," *IEEE Trans. Ind. Inform.*, vol. 19, no. 3, pp. 2565–2573, 2023. <https://doi.org/10.1109/TII.2022.3178806>
- [4] J. Wei and D. Ye, "On two sensors scheduling for remote state estimation with a shared memory channel in a cyber-physical system environment," *IEEE Trans. Cybern.*, vol. 53, no. 4, pp. 2225–2235, 2023. <https://doi.org/10.1109/tcyb.2021.3112677>
- [5] M. Mohamadi, B. Djamaa, M. R. Senouci, Y. Grine, and R. Laribi, "An effective channel selection solution for reliable scheduling in industrial IoT networks," *J. Netw. Syst. Manag.*, vol. 30, no. 4, p. 59, 2022. <https://doi.org/10.1007/s10922-022-09679-z>
- [6] B. S. Kumar, S. G. Santhi, and S. Narayana, "Optimal energy-delay scheduling using improved Beetle Antennae Search (BAS) for energy-harvesting WSNs," *Wirel. Pers. Commun.*, vol. 126, no. 3, pp. 2533–2556, 2022. <https://doi.org/10.1007/s11277-022-09828-2>
- [7] Y. Hu, H. Huang, and N. Yu, "Device scheduling and channel allocation for energy-efficient Federated Edge Learning," *Comput. Commun.*, vol. 189, pp. 53–66, 2022. <https://doi.org/10.1016/j.comcom.2022.03.008>
- [8] N. C. Luong, T. T. Anh, Z. Xiong, D. Niyato, and D. I. Kim, "Joint time scheduling and transaction fee selection in blockchain-based RF-powered backscatter cognitive radio network," *Comput. Netw.*, vol. 214, p. 109135, 2022. <https://doi.org/10.1016/j.comnet.2022.109135>
- [9] M. Mihelčić, "Redescription mining on data with background network information," *Knowl.-Based Syst.*, vol. 260, p. 110109, 2023. <https://doi.org/10.1016/j.knosys.2022.110109>
- [10] S. Xian, D. Ma, H. Guo, and X. Feng, "Route intelligent recommendation model and algorithm under the Pythagorean hesitant fuzzy linguistic environment," *Comput. Appl. Math.*, vol. 42, no. 3, p. 110, 2023. <https://doi.org/10.1007/s40314-023-02249-2>
- [11] E. A. Devi, S. Radhika, and A. Chandrasekar, "An energy-efficient MANET relay node selection and routing using a fuzzy-based analytic hierarchy process," *Telecommun. Syst.*, vol. 83, no. 2, pp. 209–226, 2023. <https://doi.org/10.1007/s11235-023-00995-6>
- [12] B. Ghosh, S. Adhikary, S. Chattopadhyay, and S. Choudhury, "Achieving energy efficiency and impact of SAR in a WBAN through optimal placement of the relay node," *Wirel. Pers. Commun.*, vol. 130, no. 3, pp. 1861–1884, 2023. <https://doi.org/10.1007/s11277-023-10361-z>
- [13] R. S. Kumaran and G. Nagarajan, "Mobile sink and fuzzy based relay node routing protocol for network lifetime enhancement in wireless sensor networks," *Wirel. Netw.*, vol. 28, no. 5, pp. 1963–1975, 2022. <https://doi.org/10.1007/s11276-022-02960-8>

- [14] P. Joshi, A. S. Raghuvanshi, and S. Kumar, "An intelligent delay efficient data aggregation scheduling for distributed sensor networks," *Microprocess. Microsyst.*, vol. 93, p. 104608, 2022. <https://doi.org/10.1016/j.micpro.2022.104608>
- [15] J. Hua, R. Liu, and F. Hao, "Two-channel false data injection attacks on multi-sensor remote state estimation," *Asian J. Control*, vol. 25, no. 5, pp. 3776–3791, 2023. <https://doi.org/10.1002/asjc.3067>
- [16] Y. Qi, J. Dang, Z. Zhang, L. Wu, B. Zhu, and L. Wang, "Filter optimization for non-orthogonal CP-FBMA system based on statistical channel state information," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 2, pp. 839–855, 2023. <https://doi.org/10.1109/TWC.2022.3198710>
- [17] D. Ojha and S. Dwarkadas, "Preventing coherence state side channel leaks using TimeCache," *IEEE Trans. Comput.*, vol. 72, no. 2, pp. 374–385, 2023. <https://doi.org/10.1109/TC.2022.3209922>
- [18] Y. Shang, F. Liu, P. Qin, Z. Guo, and Z. Li, "Research on path planning of autonomous vehicle based on RRT algorithm of Q-learning and obstacle distribution," *Eng. Comput.*, vol. 40, no. 5, pp. 1266–1286, 2023. <https://doi.org/10.1108/EC-11-2022-0672>
- [19] H. Damgacioglu and N. Celik, "A two-stage decomposition method for integrated optimization of islanded AC grid operation scheduling and network reconfiguration," *Int. J. Electr. Power Energy Syst.*, vol. 136, p. 107647, 2022. <https://doi.org/10.1016/j.ijepes.2021.107647>
- [20] J. He, C. Chadha, S. Kushwaha, S. Koric, D. Abueidda, and I. Jasiuk, "Deep energy method in topology optimization applications," *Acta Mech.*, vol. 234, no. 4, pp. 1365–1379, 2023. <https://doi.org/10.1007/s00707-022-03449-3>
- [21] Q. Ma, E. C. Demeter, and S. Basu, "Learning topology optimization process via convolutional long-short-term memory autoencoder-decoder," *Int. J. Numer. Methods Eng.*, vol. 124, no. 11, pp. 2571–2588, 2023. <https://doi.org/10.1002/nme.7221>
- [22] Y. Shimizu, S. Morimoto, M. Sanada, and Y. Inoue, "Automatic design system with generative adversarial network and convolutional neural network for optimization design of interior permanent magnet synchronous motor," *IEEE Trans. Energy Convers.*, vol. 38, no. 1, pp. 724–734, 2023. <https://doi.org/10.1109/TEC.2022.3208129>
- [23] G. Leclerc, "Julia sets of hyperbolic rational maps have positive Fourier dimension," *Commun. Math. Phys.*, vol. 397, no. 2, pp. 503–546, 2023. <https://doi.org/10.1007/s00220-022-04496-6>
- [24] A. S. Tsagkaris, N. Kalogiouri, V. Hrbek, and J. Hajslova, "Spelt authenticity assessment using a rapid and simple Fourier transform infrared spectroscopy (FTIR) method combined to advanced chemometrics," *Eur. Food Res. Technol.*, vol. 249, no. 2, pp. 441–450, 2023. <https://doi.org/10.1007/s00217-022-04128-2>
- [25] M. Kankashvar, H. Bolandi, and N. Mozayani, "Multi-agent Q-learning control of spacecraft formation flying reconfiguration trajectories," *Adv. Space Res.*, vol. 71, no. 3, pp. 1627–1643, 2023. <https://doi.org/10.1016/j.asr.2022.09.034>
- [26] P. Rahul and B. Kaarthick, "Proficient link state routing in mobile ad hoc network-based deep Q-learning network optimized with chaotic bat swarm optimization algorithm," *Int. J. Commun. Syst.*, vol. 36, no. 1, p. e5324, 2023. <https://doi.org/10.1002/dac.5324>