Securing Radio Resources Allocation with Deep Reinforcement Learning for IoE Services in Next-generation Wireless Networks

Yuhuai Peng, Xiaojing Xue, Ali Kashif Bashir, Xiaogang Zhu, Yasser D. Al-Otaibi, Usman Tariq, Keping Yu

Abstract—**The next generation wireless network (NGWN) is undergoing an unprecedented revolution, in which trillions of machines, people, and objects are interconnected to realize the Internet of Everything (IoE). with the emergence of IoE services such as virtual reality, augmented reality, and industrial 5G, the scarcity of radio resources becomes more serious. Moreover, there are hidden dangers of untrusted terminals accessing the system and illegally manipulating interconnected devices. To tackle these challenges, this paper proposes a securing radio resources allocation scheme with Deep Reinforcement Learning for IoE services in NGWN. First, the solution uses a BP neural network based on multi-feature optimized Firefly Algorithm (FA) for spectrum prediction, thereby improving the prediction accuracy and avoiding interference between unauthorized and authorized users with efficient radio utilization. Then, a spectrum sensing method based on deep reinforcement learning is proposed to identify the untrusted users in system while fusing the sensing results, to enhance the security of the cooperative process and the detection accuracy of spectrum holes. Extensive simulation results show that the proposal is superior to the traditional solutions in terms of prediction accuracy, spectrum utilization and energy consumption, and is suitable for deployment in future wireless systems.**

Index Terms—**Deep reinforcement learning, firefly algorithm, Internet of Everything, next-generation wireless networks, radio resources allocation.**

I. Introduction

WITH the rapid development of the 5G [1], [2] and Internet of Things (IoT) [3]–[5], the era of the Internet of Everything(IoE) is coming. The next generation wireless network (NGWN) [6]–[8] is undergoing an unprecedented revolution, in which trillions of machines, people, and objects are interconnected. Besides, with the emergence of IoE services such as virtual reality, augmented reality, and industrial 5G, etc., the scarcity of radio resources becomes more serious. Cognitive radio technology [9], [10] allows unauthorized users to access the idle licensed spectrum, which can effectively predict, sense and utilize spectrum holes, realize the reasonable radio resources allocation, and greatly improve the spectrum utilization. The advanced radio resources allocation method can provide a promising platform for future wireless systems in high reliability, high data rates and low energy consumption.

Massive IoE connection has become one of the main characteristics in next generation wireless systems [11], [12]. Among them, enhanced Machine Type Communications (eMTC) will support richer IoE applications and take up a lot of radio resources. In order to achieve greater openness, interconnection and integration, the access layer of future wireless systems is facing serious security threats [13]. A large number of devices in NGWN come from different manufacturers with different standards and working principles. Among them, potential malicious users may access the system and try to disrupt the normal operation. Malicious occupation of radio resources, refusal to participate in cooperation, and sending wrong results will affect the accuracy of spectrum sensing, and even cause node failure in severe cases. Therefore, improving the utilization efficiency of radio resources in untrustworthy environments is an urgent problem to address in NGWN.

Energy detection, cyclostationary feature detection and matching filtering have been widely used in wireless systems. However, in the actual environment, complex scenes seriously affect the performance of classical sensing algorithms [14]–[16]. Machine learning based methods [17]–[19] realizes intelligent spectrum sensing by modeling and reasoning the channel state information and user historical behavior characteristics. These methods have good performance, but cause higher energy consumption. Spectrum prediction can filter and screen the spectrum in advance, reducing the processing delay and energy consumption in sensing process. Among them, neural network based methods [20] have attracted wide attention because of its accuracy. To solve the efficient utilization of wireless radio in untrustworthy environment, a securing radio resources allocation scheme with Deep Reinforcement Learning for IoE services is proposed. The main contributions are summarized as follows:

1) A cognitive network system architecture for untrustworthy environments is developed, and the interference among primary users, secondary users, and potentially

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1702000, and the National Natural Science Foundation of China under Grant 61871107. (Corresponding author: Xiaogang Zhu)

Yuhuai Peng and Xiaojing Xue are with the School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China. (Email: pengyuhuai@mail.neu.edu.cn; xuexiao0028@163.com)

Ali Kashif Bashir is with Department of Computing and Mathematics, E-154, John Dolton, Chester Street, M15 6H, Manchester Metropolitan University, Manchester, United Kingdom. (Email: dr.alikashif.b@ieee.org)

Xiaogang Zhu is with School of Management, Nanchang University, Nanchang 330031, Jiangxi, China. (Email: ncuzxg@ncu.edu.cn)

Yasser D. Al-Otaibi is with Department of Information Systems, Faculty of Computing and Information Technology in Rabigh, King Abdulaziz University, Jeddah 21589, Saudi Arabia. (Email: yalotaibi@kau.edu.sa)

Usman Tariq is with College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia. (Email: u.tariq@psau.edu.sa)

Keping Yu is with Global Information and Telecommunication Institute, Waseda University, Shinjuku, Tokyo 169-8050, Japan. (Email: yukeping@fuji.waseda.jp)

untrustworthy users are analyzed. Meanwhile, a sensing frame structure, which consists of a prediction-sensing-interaction sub-channel and an uninterrupted transmission sub-channel, is designed to improve spectrum utilization and increase network throughput.

2) A spectrum prediction method based on multi-feature-optimized firefly algorithm (FA) and BP neural network is proposed. We optimize the FA using chaotic mutation, dynamic step size and bulletin board mechanism to avoid the algorithm converging to a local optimum. The optimized FA is used to train a BP neural network-based prediction model, which significantly improves the spectrum prediction accuracy.
3) A spectrum sensing scheme with deep reinforcement learning is proposed. Based on the established double deep Q network (DDQN) model, the convergence speed of reinforcement learning accelerates by using experience recovery mechanism. Moreover, this scheme uses the reputation model to identify the untrustworthy users in system, and selects the trustworthy ones for distributed cooperative sensing and information consensus fusion, which alleviates the interference in harsh environment and improve sensing accuracy.
4) Extensive simulation results demonstrate that our scheme can identify malicious users and generate consensus within 20 interactions. It can reduce 72% of spectrum-sensing energy consumption and increase spectrum utilization by 20% compared with traditional schemes.

The rest of this paper is organized as follows. Section II introduces the related work, and section III gives the system model. In section IV, the spectrum prediction and spectrum sensing methods are described in detail. Section V gives the experimental results and analysis. Finally, section VI summarizes this paper.

II. Related Works

In recent years, researchers are committed to enhancing the performance of spectrum sensing in many aspects, including accelerating speed, improving accuracy, and reducing energy consumption. Düzenli *et al.* [21] used a dynamic programming method to accelerate the calculation of the least square method, and developed a dynamic spectrum sensing strategy. Mu *et al.* [22] gave two spectrum sensing models with constraints, and respectively calculated the optimal sensing operation under the constraints. Muhammad *et al.* [23] proposed an intelligent adaptive spectrum sensing method that minimized the impact of minimal denial of service interference. Abhijit *et al.* [24] used a dual-threshold decision-making method to sense the licensed spectrum, and combined local difficult decisions in the fusion center to obtain a global decision, which improved network flexibility. Anastassia *et al.* [25] proposed a multi-band multi-user cooperative spectrum sensing scheme based on distributed learning. The secondary users use the consensus fusion method to make a judgment on the spectrum occupancy. Amirhosein *et al.* [26] proposed a method based on distributed diffusion, which improves the reliability of spectrum sensing through cooperation between users. Tong *et al.* [27] proposed a blind cooperative spectrum sensing method based on soft fusion. Each secondary user uses the prior knowledge of the channel and the primary signal to transmit the soft information to the fusion center for decision. Lee *et al.* [28] studied a deep cooperative sensing framework based on Convolutional Neural Network (CNN), using multiple secondary users to cooperate with each other to jointly detect a primary user, which greatly improved the accuracy of spectrum sensing. Sun *et al.* [29] proposed a multi-channel spectrum access scheme based on reinforcement learning. Users can use multiple channels for transmission, which improves the spectrum access capability of the cognitive Internet of Things. RRajaguru *et al.* [30] combined clustering with expectation maximization algorithm and reinforcement learning technology, and proposed a feature-based clustering classifier-based cooperative spectrum sensing technology to minimize energy consumption. Xu *et al.* [31] proposed a framework based on Bayesian machine learning for large-scale heterogeneous networks, which uses multiple secondary users to collect spectrum sensing data and coordinately derive the global spectrum state.

Prediction algorithms can be used to assist and improve the accuracy of spectrum sensing. SUs can learn the behavior characteristics of PUs from historical data, and predict the occupancy before sensing. As a result, SUs no longer sense the spectrum with a higher probability of occupancy, reducing the energy consumption. Sung *et al.* [32] considered the actual on/off traffic model and proposed an optimal strategy for determining the transmission power of the auxiliary user, which maximizes the spectrum utilization of the auxiliary user. Ding *et al.* [33] developed an online spectrum prediction framework based on historical observation data. By effectively integrating time series prediction technology, the prediction problem was defined as a joint optimization problem, and the alternate direction optimization method was used to effectively solve the problem. Eltom *et al.* [34] proposed a cooperative spectrum prediction algorithm based on soft fusion. Compared with the spectrum prediction based on local and hard fusion, the prediction accuracy of this method is significantly improved. Tumuluru *et al.* [35] designed a spectrum prediction model based on neural network, which can accurately identify spectrum holes in cognitive networks. Ding *et al.* [36] first preprocessed historical spectrum occupancy data, and then designed a deep learning-based fusion network to predict spectrum occupancy. This network effectively combines multiple prediction results to improve prediction accuracy. Xu *et al.* [37] considered a scenario with multiple independent channels and multiple heterogeneous primary users, and proposed a deep reinforcement learning model based on dynamic spectrum access. However, the above method does not propose a corresponding security solution for the untrustworthy environment, cannot identify potential untrustworthy users, and cannot guarantee the security of massive access scenarios.

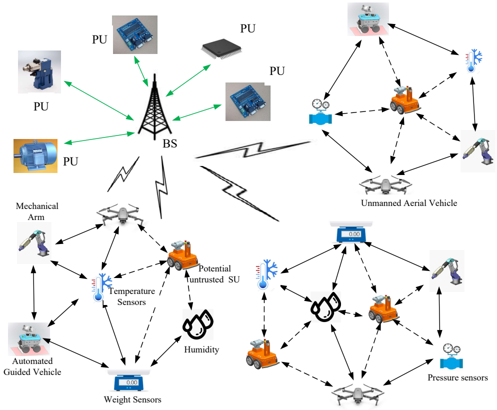


Fig. 1: System architecture.

III. System Model

A. The System Model

In NGWN, the number of devices will grow exponentially. Devices vary greatly in manufacturer and function, making it difficult to manage them consistently and reliably. Malicious devices may access the network and damage the data information in the network. We consider a NGWN system in an untrustworthy environment as shown in Figure 1. This is a distributed cooperative sensing network consisting of the primary user (PU), the secondary user (SU) and base station (BS). PU represents authorized user, and SU denotes cognitive users including potentially untrustworthy users. PUs can employ licensed spectrum to communicate with BS. In the NGWN, PU and SU cannot communicate with each other. So, SU must always keep a sense of the spectrum status while using the licensed spectrum to transmit data. Spectrum sensing refers to the use of energy detection, matched filter detection, and cyclostationary feature detection to determine whether the state of a channel is idle or occupied, and to provide available spectrum resources for dynamic spectrum access.

Firstly, SUs predict and sense spectrum to judge whether the target spectrum is free, and then take actions according to the result of the judgment. The SU can use the free authorized channel for D2D communication or access to the base station. When it is detected that the PU is using the spectrum, the SU will exit in time and stop the communication. Assuming the communication range of the BS can cover all SUs, and the network topology will not change. The SUs in a fixed area form a set. Each SU in the set can independently predict, sense, and share the results with neighbor SUs. The cognitive network can be mapped to an undirected graph. Among them, $V = \{1, 2, ..., k\}$ represents the number of SUs, Q denotes the connection relationship between different SUs. V and Q are the input of the sensing algorithm.

B. The structure of sensing frame

In the traditional sensing frame, the interaction process occupies the spectrum and interrupts the data transmission. And continuous data transmission cannot be realized, which reduces network throughput. Therefore, a dual sub-band sensing frame is designed to ensure that spectrum sensing and data transmission are implemented at the same time. The structure of sensing frame is shown in Figure 2. All cognitive users must follow the sensing frame.

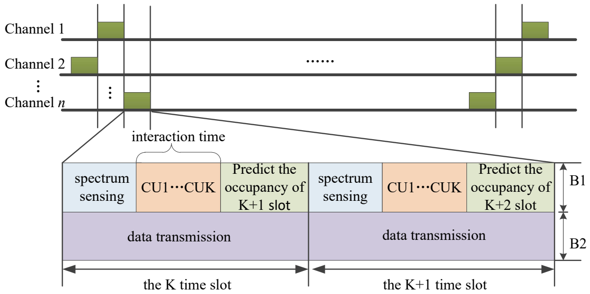


Fig. 2: The structure of sensing frame.

Each sensing frame is a time slot, which is divided into two sub-bands in the frequency domain. The B1 sub-band is used for spectrum prediction, sensing and interaction, and the B2 sub-band is used for data transmission. This can ensure the continuity of data transmission and improve network throughput. IEEE standard stipulates that SUs must exit occupied authorized spectrum within two seconds when PUs access. Therefore, the length of each time slot T is set to $2s$. In addition, before the end of each sensing frame, a sub slot is used to predict the spectrum occupation of the next slot to assist the spectrum sensing. The occupation states are saved to the database and will be used for spectrum prediction.

IV. Securing Radio Resources Allocation with Deep Reinforcement Learning

This section proposes a securing radio resources allocation with deep reinforcement learning for IoE services in next-generation wireless networks, which is divided into two parts: spectrum prediction and spectrum sensing. Firstly, SU predicts the occupancy state of the authorized spectrum, then senses the authorized spectrum that is predicted to be idle. Finally, the decision of spectrum occupancy state is obtained based on the sensing results. The spectrum is only considered accessible when both spectrum prediction and sensing results are idle. If it is inconsistent, users are not allowed to access. This process improves the accuracy of spectrum sensing and reduce energy consumption.

A. Spectrum Prediction Algorithm Based on BPNN

This section introduces the spectrum prediction algorithm based on back-propagation neural network (BPNN). Because the initial weights and thresholds in a BPNN are random, it may result in a slow convergence in the training process or fall into a local optimum solution. Therefore, this paper adopts the improved Firefly Algorithm (FA) to optimize the weights and thresholds.

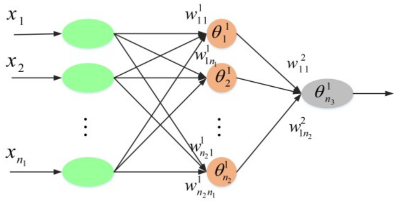


Fig. 3: Neural network topology.

1) BPNN model: The BPNN contains input layer, hidden layer and output layer, and its topology is shown in Figure 3. Where w^1 and θ^1 are the weights and thresholds in the hidden layer, w^2 and θ^2 represent corresponding parameters of the output layer. n_2 and n_3 denote the number of neurons in the hidden and output layers. After several experiments, the best performance is acquired when the number of neurons is $n_1 = 10, n_2 = 21, n_3 = 1$ respectively.

The channel state of i is defined as time sequence $X^1 = (x^{i,1}, x^{i,2}, ...x^{i,k})k = 1, 2, ...T$, where T is the time slot sequence number. If the occupation of channel i in time slot t is predicted, it is necessary to input the historical channel states $x^{i,t-10}$ to $x^{i,t-1}$ to BPNN, and the network will output the predicted result $x^{i,t}$. Since the range of spectrum prediction values should be between [0,1], the decision formula is set as follows:

$$Z_k = \begin{cases} 1 & Z'_k > 1, \\ Z'_k & otherwise, \\ 0 & Z'_k < 0, \end{cases} \tag{1}$$

where Z_k is the final judgment result, and Z'_k is the actual output of the prediction model.

2) BPNN parameter optimization based on improved FA: In this section, we adopt chaotic variation, dynamic step and bulletin board mechanism for multi-feature optimization of FA, which is used to obtain relevant parameters of the BPNN. The FA takes the output error of the BPNN as the objective function to seek the optimal position of fireflies in multi-dimensional space, which is the optimal initial weights and thresholds of the BPNN.

$$ObjectiveFunction : Minimize \frac{1}{2}(Y_k - Z_k)^2 \tag{2}$$

where Y_k is the real value and Z_k is the BPNN output value.

In the FA, the total number of the weights and thresholds in the BPNN is $w = n_1 * n_2 + n_2 + n_2 * 1 + 1$, and each firefly is expressed as

$$X(t) = (w^1_{11}, ..., w^1_{n1n2}, \theta^1_1, ..., \theta^1_{n2}, w^2_{11}, ..., w^2_{n1n2}, \theta^2) \tag{3}$$

Therefore, when generating the firefly population in the w-dimensional search space, the position of each firefly is demarcated by:

$$X_i(t) = (w^1_{i11}, ..., w^1_{in1n2}, \theta^1_{i1}, ..., \theta^1_{in2}, w^2_{i11}, ..., w^2_{in1n2}, \theta^2_i)^T \tag{4}$$

In the initial stage, the fluorescein value l_0,the decision radius r_0, and the maximum iteration number t_{max} are determined. During the iteration, the change of the fluorescein value for each firefly at certain position $X_i(t)$ is given by,

$$l_i(t) = (1 - \rho)l_i(t + 1) + \gamma f(X_i(t)) \tag{5}$$

where $l_i(t)$ and $l_i(t - 1)$ are the fluorescein values of i at moments t and $t - 1$ respectively, rho is the speed of play magnitude which taking values generally between $(0, 1)$, γ is the fluorescein update rate, which refers to the speed of fluorescein value change with each movement, $f(X_i(t))$ indicates the fitness function, which is the output error value of the BPNN obtained by solving for the current firefly position.

And within its own decision radius, each firefly will be attracted by others with higher fluorescein value and approach them. After determining the neighbor set $N_i(t)$, in order to decide the target moving object, firefly i calculates the attraction probability between itself and others by Eq. (6). Then firefly i selects j with the maximum probability value and move toward it. The new location of firefly i is described as Eq. (7).

$$P_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)} \tag{6}$$

$$X_i(t + 1) = X_i(t) + step(\frac{X_j(t) - X_i(t)}{\|X_j(t) - X_i(t)\|}) \tag{7}$$

where j is the moving target neighbor of firefly i, $\|X_j(t) - X_i(t)\|$ is the Euclidean distance between i and j, and $step$ is the moving step, which can be expressed as,

$$step = \frac{w}{e^t} + \eta \tag{8}$$

where t is the number of iterations and η is the minimum step to prevent being 0. We set step as the dynamic moving step, which ensures that the firefly have a larger step in the first iteration.

To guarantee that the movement will not exceed the search boundary, the update position that fall outside the boundary is corrected according to Eq. (9).

$$X_i(t + 1) = min\{x^{max}_k, max\{x^{min}_k, x_{ik}(t)\}, k = 1, 2, ..., w \tag{9}$$

where $x_ik(t)$ is the k^{th} element in the position of the firefly i at the t^{th} iteration, x^{max}_k and x^{min}_k denote the upper and lower limits of the position.

When the firefly completes the position update, its new decision radius is as follows:

$$r_i(t + 1) = min\{r_s, max\{0, r_i(t) + \beta(n_t - |N_i(t)|)\}\} \tag{10}$$

where β is the decision radius update rate, r_s is the decision radius threshold, n_t is the threshold for the number of fireflies, and $|N_i(t)|$ is the number of fireflies in the neighborhood.

Considering that firefly is easy to gather at the boundary, we adds chaotic mutation method to optimize the firefly which move to the boundary. According to Eq. (12), the firefly at the boundary is mutated, and the mutation firefly with high fitness value are selected to replace the original ones, which effectively increases the diversity of the population. At the same time, the bulletin board mechanism is used to record the information of the optimal firefly and its value during the iteration. When the maximum number of iterations is met, the information of firefly recorded on the bulletin board is the optimal solution to the objective function.

$$X_{in}(t) = X_i(t) * [1 + k * Z(n)], n = 1, 2, ..., M \tag{11}$$

Where $X_{in}(t)$ is the new firefly that mutates at the t^{th} iteration, and each mutation will yield M new firefly, and k is the regulating factor, denoted as follows:

$$k = 1 - (w - 1)/step \tag{12}$$

In addition, $Z(n)$ is generated by using the chaotic variation principle. Remarkably, $Z(1)$ is the original value at $n = 1$, which is a random dimensional vector between $(-1, 1)$, as shown in Eq. (13). After continuous iterations, M chaotic sequences are obtained, as shown in Eq. (14).

$$Z(1) = 1 - 2 * rand(1, D) \tag{13}$$

$$Z(n+1) = 4 * Z(n)^3 - 3 * Z(n), n = 1, 2, ..., M - 1 \tag{14}$$

The pseudo code of the BPNN parameter optimization based improved FA is shown in Algorithm 1. The time complexity is related to the population number N and the maximum number of iterations t_{max}. The time complexity of CMFA-BP is $O(N * t_{max})$.

B. Spectrum sensing algorithm based on deep reinforcement learning

Considering the impact of complex environment, noise and shadow effects on SU spectrum sensing in NGWN systems, this paper proposes a multi-user cooperative distributed sensing algorithm for spectrum sensing. In order to avoid the misinformation generated by malicious users affecting the sensing results, we designed a Spectrum Sensing Algorithm based on Deep Reinforcement Learning (SSDRL). This algorithm removes untrustworthy SU in the network so that the influence of malicious users is reduced. In practice, only the spectrum predicted to be idle needs to be sensed, reducing the energy consumption of spectrum sensing.

1) Double deep Q network based on CNN: In this section, we combine deep learning and reinforcement learning to form a double deep Q network (DDQN). The input of the network is a two-dimensional state matrix, and output is the corresponding Q value. We construct a convolutional neural network which is made up of convolutional and fusion parts. The convolutional part includes three basic sub-blocks that are composed of convolutional layer, ReLU layer and max-pooling layer. It is responsible for extracting spatial features of the input data, and the fusion part is used to classify the input data by collecting the results of its feature. In addition, a ReLU layer is added between the FC layers to introduce non-linearity. The network structure is shown in Figure 4.

Algorithm 1 **CMFA-BP**

```
Initialize population $N$ and location $X_i(0)$
Calculation fitness $f_i(X_i(0))$
$f_{best} \leftarrow Max f_i(X_i(0))$;
for $t \leftarrow 1$ to $t_{max}$ do
    for $i \leftarrow 1$ to $N$ do
        $f_i(X_i(t-1)) \leftarrow f_i(X_i(t))$
        $l_i(t-1) \leftarrow l_i(t)$
        Search a neighbor set $N_i(t)$
        if $N_i(t)$ exists then
            $X_i(t-1) \leftarrow X_i(t)$
        end if
        $r_i(t-1) \leftarrow r_i(t)$
        if $X_i(t) \notin (x_i^{min}$ or $x_i^{max})$ then
            $X_i(t) \leftarrow X_i^{'}(t)$
            $f_i(X_i(t)) \leftarrow f_i(X_i^{'}(t))$
        end if
        if $f_i(X_i(t)) < f_{best}$ then
            $f_{best} \leftarrow f_i(X_i(t))$
        end if
        if $t = t_{max}$ then
            Get $X_{best}(t)$
            Train the BP neural network break
        end if
    end for
end for
```

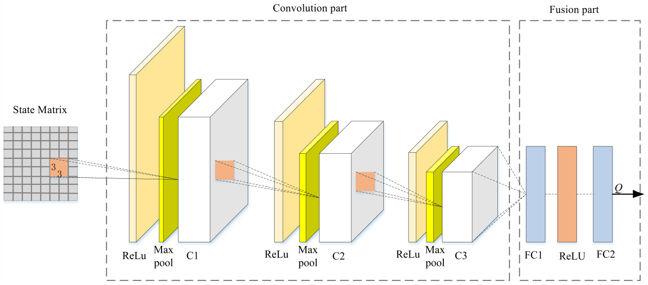


Fig. 4: Model of convolution neural network.

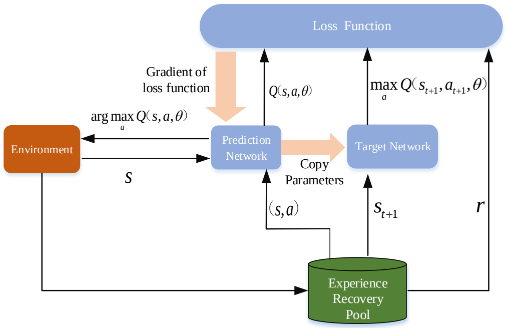


Fig. 5: The DDQN schematic.

We adopt an experience recovery pool and a DDQN to ensure that the model selects the most appropriate action. The DDQN is shown in Figure 5.

2) Description of the SSDRL algorithm: In this section, we propose a cooperative spectrum sensing algorithm based on reinforcement learning to improve accuracy. This algorithm integrates reinforcement learning, reputation value model and consensus fusion model to evaluate the trustworthiness of SU. Firstly, we regard each SU as an agent and constantly search for neighbor SU with high trustworthiness to perform cooperative spectrum sensing. In order to achieve this process, each SU only cooperate with SU that are highly trustworthy, and the untrustworthy SU in the network is removed. At the same time, the set of neighbor SU is used as the model state and the neighbor user with a high reputation value is selected as the action. Significantly, the reputation value will increase when the sense result of SU is true. Then, the consensus fusion algorithm is used to continuously fuse information with neighbor SU until convergence is reached. Finally, we compare the result with the judgment threshold to get the corresponding result. The process of this algorithm is as follows.

Step 1: Initialize parameters.

We first initialize the energy and reputation values for the i^{th} SU to acquire different initial values.

$$x_i(0) = x_R \tag{15}$$

$$r_i(0) = R \tag{16}$$

where $x_i(0)$ is the initial energy value and $r_i(0)$ is the initial reputation value. Next, the set of neighbor users V_{ci} is initialized, ensuring that the SU only select cooperative users from V_{ci}.

Step 2: Judgement of SU credibility

This process adopts a deep reinforcement learning model to identify trustworthy SU, which is shown as follows.

(1) The initial values of the Q matrix $Q(s_t, a_t)$ are set to 0 and a state and action lookup table are generated. Simultaneously, the learning rate and discount factor are set to α and β respectively.

(2) SU gets the current state by querying the Q matrix, and uses simulated annealing algorithm to improve the greedy action selection strategy in DDQN. This algorithm regards the selection as a process of annealing and cooling to alleviate the problem that DDQN tends to fall into local optimum solutions. It contains three parameters: objective function, initial solution and solution space. During the process, the temperature T is first initialized and the increment of the objective function $\Delta = f(s') - f(s)$ is calculated by s'. When $\Delta T < 0$, the solution is considered as the latest solution. Otherwise, whether to accept s' as the latest solution is determined by p. We define the probability of selecting action a by the following.

$$P(a_t|s_t) = \frac{exp(Q_t(s_{t+1}, a_t)/\Gamma(t))}{\sum_{a_t \in A} exp(Q_t(s_{t+1}, a_t)/\Gamma(t))} \tag{17}$$

In equation (17), $\Gamma(t)$ is the initial temperature parameter. While the value of Q is in the initial state, the probability that SU selects neighbor users is zero, so it will randomly select cooperative users. As the agent goes through a period of learning, the Q value will be updated. By the time, the higher the Q value, the higher probability that the user will be selected.

(3) A deep reinforcement learning model is established to select action a and derived its instantaneous reputation value $r_{t+1}(s_{t+1})$. Based on the reputation value, we receive the corresponding action a and update the Q matrix. To construct the experience recovery pool, $e_t = (s_t, a_t, r_t, s_{t+1})$ will be stored until the number of samples is greater than the smallest number. If samples exceed the maximum capacity of the pool, it will replace the old samples with the new ones to ensure efficiency.

The above process will repeat until convergence. Afterwards, we just input the state matrix into the deep reinforcement learning model to get the action with the maximum Q value.

Step 3: Reputation Value Update

The reputation value of each SU will be updated during the deep reinforcement learning, at which the set of trustworthy neighbors is available by the size of the reputation value. For the i^{th} SU, the sense value $D_{i,j}(t)$ of the authorization signal for neighbor user J is:

$$D_{i,j}(t) = x_j(t) \tag{18}$$

where $j \in V_{ci}$ is the neighbor user of the j^{th} SU and $x_j(t)$ is the energy value at time t. Therefore, the verdict of neighbor user spectrum sensing $A_{i,j}(t)$ is given by:

$$A_{i,j}(t) = \begin{cases} 1 & D_{i,j}(t) \geq \sigma \\ -1 & D_{i,j}(t) < \sigma \end{cases} \tag{19}$$

The cooperative sensing result $B(t)$ for the i^{th} SU with trustworthy neighbors is as follows:

$$B(t) = \begin{cases} 1 & \sum_{j \in V_{ci}} A_{i,j}(t) > 0 \\ -1 & \sum_{j \in V_{ci}} A_{i,j}(t) \leq 0 \end{cases} \tag{20}$$

To sum up, the reputation value of the i^{th} SU is:

$$r_i(t) = r_i(t-1) \pm \sum_{j \in V_{di}} |B(t) + Z_i(k)| \tag{21}$$

$$V_{di} = \{j | r_j(t) > 100, j \in V_{ci}\} \tag{22}$$

in which $r_i(t)$ is the reputation value of the SU at t iterations while $r_i(t-1)$ is the value at $t-1$ iterations. $Z_i(t)$ is the local spectrum sensing judgement result, and V_{di} is the set of SU with reputation values greater than 100. When the cooperative spectrum sensing and the sensing of the SU are identical, the reputation value will increase by two each time, otherwise the reputation value will decrease.

Step 4: Information Fusion

With the deep reinforcement learning, we get the set V_{di} of trustworthy neighbor users. Afterwards, we interact with the information based on equation (23) to acquire the consistent verdict result.

$$x_i(t+1) = x_i(t) + \delta \sum_{j \in V_{di}} (x_j(t) - x_i(t)) \tag{23}$$

where $x_i(t+1)$ is the energy value of the SU at $t+1$ iterations.

Step 5: Convergence

When the SU information fusion result is identical or the number of iterations exceeds the maximum number of iterations, the iteration will stop. Otherwise, it is necessary to constantly repeat steps two to four again so as to get the final convergence value x^*.

$$x^* = \frac{\sum_{i=1}^{N} x_i(t)}{N} \tag{24}$$

Step 6: Channel state judgement

By comparing the SU information interaction results with judgement threshold, the final channel state judgement result D is found.

$$D = \begin{cases} H_0 & x^* < \sigma \\ H_1 & x^* \geq \sigma \end{cases} \tag{25}$$

The pseudo code of the spectrum sensing algorithm based on SSDRL is shown in Algorithm 2. Since the number of channels is much smaller than the maximum number of iterations, and the time complexity of SSDRL is $O(t_{max})$.

Algorithm 2 SSDRL

```
for $i \leftarrow 1$ to $N$ do
    Select the action $a_{(t)} \leftarrow S_{(t)}$
    Get $r_{(t)}$ and $S_{(t+1)} \leftarrow a_{(t)}$
    Store $e_t = (S_{(t)}, a_{(t)}, r_{(t)}, S_{(t+1)})$ and $S_{(t+1)} \leftarrow a_{(t)}$
    Calculate $Q$ value
end for
for $t \leftarrow 1$ to $t_{max}$ do
    if $t = t_{max}$ then
        Select $V_{ci} \leftarrow r_t$
        if $X_i = X^*$ then
            Make decision $D$
        else
            Reselect the action $a_t \leftarrow S_t$
        end if
    else
        Reselect the action $a_t \leftarrow S_t$
    end if
end for
```

V. Simulation experiment and performance analysis

A. Simulation environment and parameter settings

To verify the performance of this algorithm, firstly, we constructed the network architecture and set relevant parameters of three prediction algorithms. Secondly, the CMFA-BP prediction model was designed, and we compared it with FA-BP and GA-BP algorithms. Finally, the performance indexes of the algorithm were analyzed and the performance of the scheme in this paper was compared. The relevant parameter settings are shown in Table 1.

TABLE I: Parameter settings

Items	**Parameters**
Transmit power of authorized users	75dB
Distance between authorized users and cognitive users	5km
Network coverage diameter	2km
Communication range of cognitive users	300m
Noise power	-85dB
Probability of false alarm	0.05
Sampling frequency	1MHz
Number of iterations	50
Learning rate	0.2
Discount factor	0.8
Cooperative sensing fusion parameters	0.1
Sensed spectrum	50MHz
Spectrum sensing time	0.05ms
Initial reputation value of cognitive users	100

B. Evaluation indexes

1) Spectrum efficiency: The spectrum efficiency is defined as the ratio of the number of time slots that sensed as idle to all idle time slots. The higher the index, the better the spectrum sensing performance of this algorithm.

$$SE = \frac{N(Correctly\ sensed\ idle\ time\ slots)}{N(All\ idle\ time\ slots)} \tag{26}$$

2) Probability of error prediction: The probability of error prediction, which generally ranges from 0 to 0.5, means the possibility that the algorithm prediction result is inconsistent with the channel usage state. The equation is given as (26).

$$\begin{aligned} P(all) = & \frac{\sum_{i=1}^{N}(Z_{k+1}^{i} = 1 | Y_{k+1}^{i} = 0)}{N} + \\ & \frac{\sum_{i=1}^{N}(Z_{k+1}^{i} = 0 | Y_{k+1}^{i} = 1)}{N} \end{aligned} \tag{27}$$

3) Probability of false alarm: The probability of false alarm prediction, which is defined as the possibility that an idle channel is predicted to be occupied, is generally between 0 and 0.5. The higher it is, the more the authorized spectrum is underutilized, and so the worse the spectrum prediction performance is. The probability of false alarm prediction $P(1|0)$ can be expressed as:

$$P(1|0) = \frac{\sum_{i=1}^{N}(Z_{k+1}^{i} = 1 | Y_{k+1}^{i} = 0)}{\sum_{i=1}^{N}(Y_{k+1}^{i} = 0)} \tag{28}$$

4) Energy consumption: With the help of spectrum prediction technology, cognitive users no longer sense the spectrum with a high probability of being occupied, saving the energy consumption of spectrum sensing. The energy consumption of the traditional algorithms in spectrum sensing over the entire spectrum is described as:

$$E_{CU_s} = N * E \tag{29}$$

And the energy consumption of spectrum sensing after prediction is:

$$E_{CU_p} = (N - N_{busy}) * E \tag{30}$$

where E is the energy consumption of each channel sensed by SU; N is the total number of channels; N_{busy} is the number of channels whose spectrum predicted to be occupied, and the energy consumption reduction rate of the prediction algorithm $E_{reduce(\%)}$ can be calculated by equations (30).

$$E_{reduce}(\%) = \frac{E_{CU_s} - E_{CU_p}}{E_{CU_s}} \tag{31}$$

C. Performance analysis of spectrum prediction algorithm

The occupation of spectrum is regular, which can be approximated by a model. Considering that it is very difficult to collect data on the spectrum occupation of each user, we adopt the $M/G/1$ queuing to model the spectrum usage and use the data as training samples. In this model, M denotes the time interval between each spectrum occupation and the next by an authorized user, which obeys the Poisson distribution with parameter two. And G represents the time interval of each authorized user to occupy the spectrum, it follows the geometric distribution with the mean of μ.

The mutation probability, crossover probability and generation gap of genetic algorithm are 0.05, 0.7 and 0.9, respectively. The parameters were used to construct three spectrum prediction models. The error between the real and predicted values of the CMFA-BPO and FA-BPO algorithms were compared. Since the spectrum prediction performance was affected by the intensity of the traffic, two situations where the traffic intensity $\Delta = 0.5$ and $\Delta = 0.8$ respectively were simulated. In addition, we compared the performance of each algorithm under the same traffic intensity. As shown in Figure 6 and Figure 7, the CMFA-BP algorithm had higher prediction accuracy than those of the FA-BP algorithm, while it had a higher accuracy in the testing set. Additionally, with the increase of the traffic, the test error of CMFA-BP algorithm had been reduced from 0.33 to 0.14, which is a reduction of 57.6%.

The prediction results were binarized, and the error probabilities and false alarm probabilities were evaluated by counting the errors at various traffic strengths. In Figure 8 and Figure 9, the prediction probability curves at different traffic are shown. λ is the user data transmission interval, which will affect the number of hops in the authorized channel. The spectrum prediction performance at $\lambda = 10$ and $\lambda = 20$ was validated.

As the traffic increased, it appears that the CMFA-BP algorithm had a lower prediction error probability and false alarm probability than others. This is because when the traffic increased, users tended to transmit data on the same channel. As λ continued to increase, the time interval between users would be longer, and the accuracy of spectrum prediction increased accordingly.

As shown in Table 2, compared to direct spectrum sensing, the energy consumption of the SU varies with λ when the traffic was 0.8. The total number of time slots was set to 50, and the information on different channels was collected. Likewise, the CMFA-BP algorithm correctly occupied more

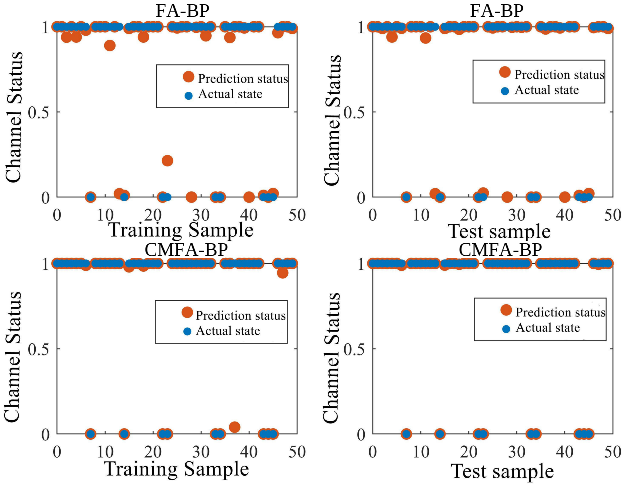


Fig. 6: Comparison of spectrum prediction($\Delta = 0.5$).

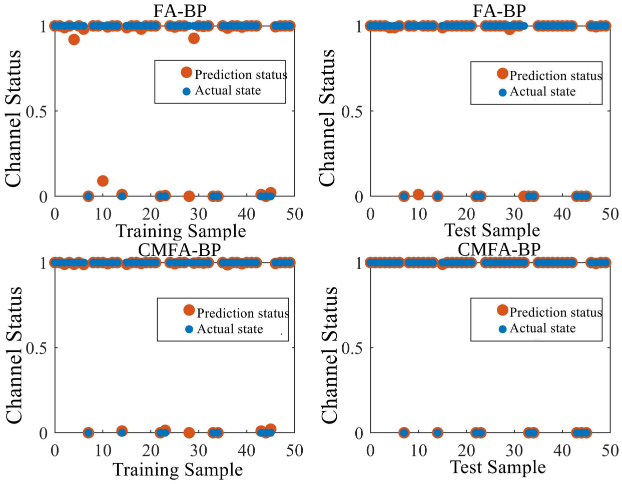


Fig. 7: Comparison of spectrum prediction($\Delta = 0.8$).

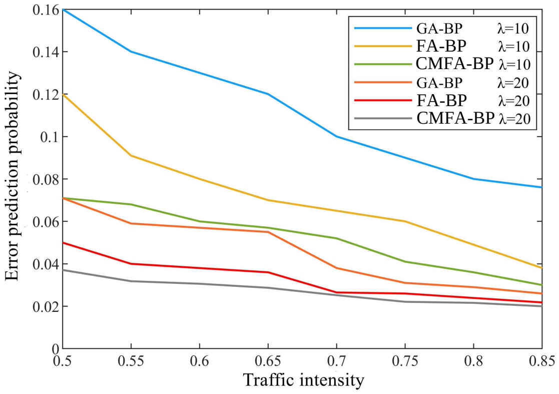


Fig. 8: Comparison of prediction error probability.

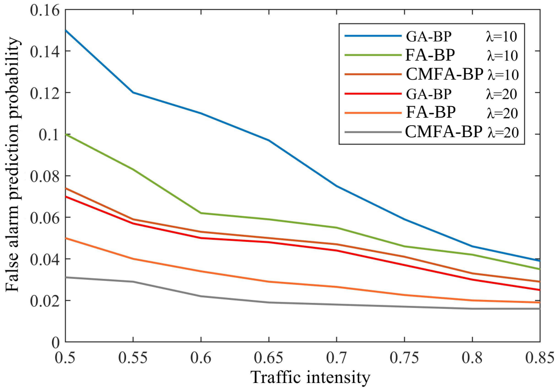


Fig. 9: Comparison of predicted false alarm probability.

time slots, and the energy improvement rate reached more than 50%. With the increase of λ, the energy consumption reduced as well. Therefore, it is proved that under the same conditions, the energy efficiency of CMFA-BP algorithm is significantly higher than others.

TABLE II: Energy consumption reduction rate

Prediction algorithm	**Parameter**	**Number of correctly occupied time slots**	**Energy consumption reduction rate** (%)
GA-BP	$\lambda = 10$	20	40
	$\lambda = 20$	22	44
FA-BP	$\lambda = 10$	23	46
	$\lambda = 20$	25	50
CMFA-BP	$\lambda = 10$	26	72
	$\lambda = 20$	27	54

As shown in Figure 10 and Figure 11, the spectrum utilization of three spectrum prediction algorithms was compared, with λ is 10 or 20 respectively. Here, spectrum sensing algorithms were energy detection methods. The proposed algorithm is generally higher than the other. As the traffic and λ increased, the spectrum utilization rate gradually increased as well. And it leveled off after the traffic reaches 0.7. This is because as the traffic increased, the number of channel hops decreased. The predictability of the channel increased, and the prediction performance became better. When $\lambda = 10$, compared with the GA-BP and the FA-BP algorithm, the spectrum utilization rate of the CMFA-BP algorithm increased by 20.1% and 9.3%, respectively. As for $\lambda = 20$, the spectrum utilization rate of the CMFA-BP algorithm increased by 12.7% and 6.7%, respectively.

D. *Performance analysis of spectrum sensing algorithm*

Assuming that the location of all users in the network was constant, the network environment was stable, and there were no large fluctuations of sensing confidence. In this section, three experiments were designed to evaluate the performance of the SSDRL algorithm. Figure 12 illustrates the energy sensed by each SU with iterations. Starting from the 10th iterations, the sensing energy reached convergence. It is demonstrated that the SSDRL algorithm enabled users to complete data fusion rapidly.

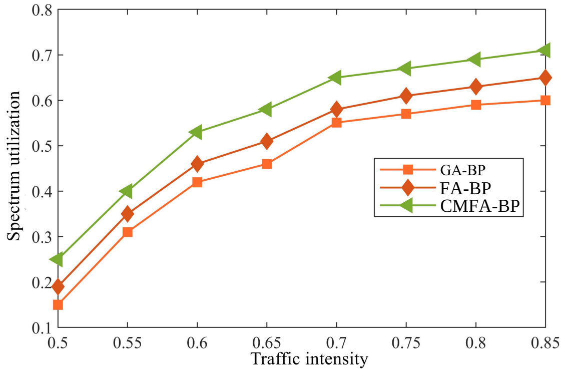


Fig. 10: Comparison of spectrum utilization ratio($\lambda = 10$).

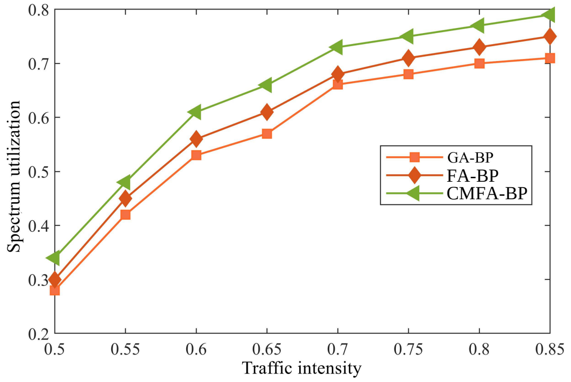


Fig. 11: Comparison of spectrum utilization ratio($\lambda = 20$).

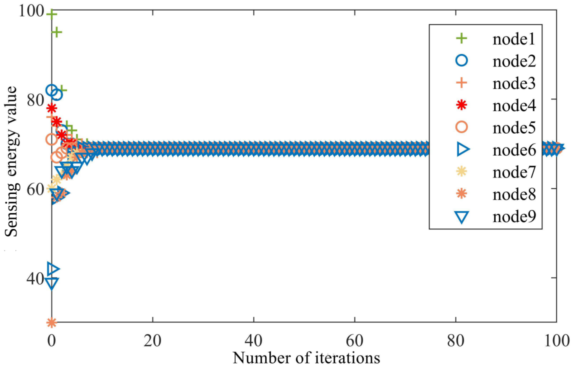
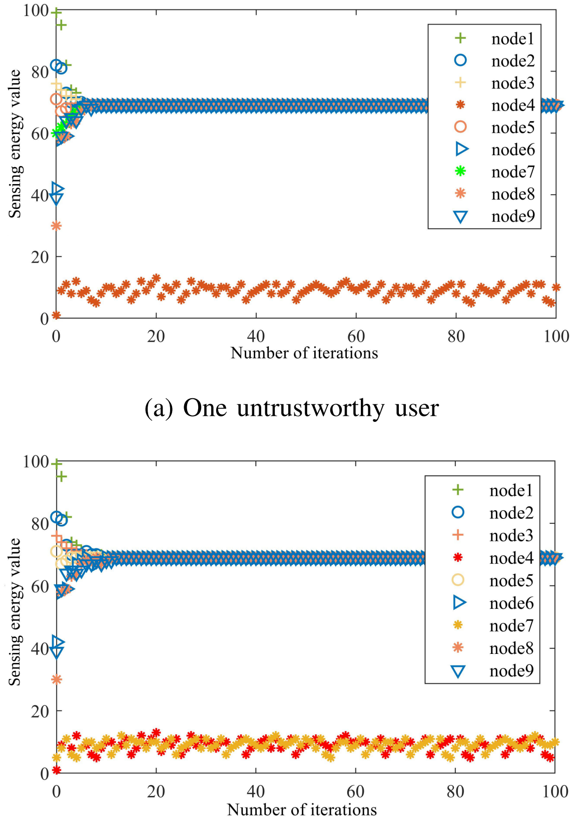


Fig. 12: Sensing energy values under different iteration times.

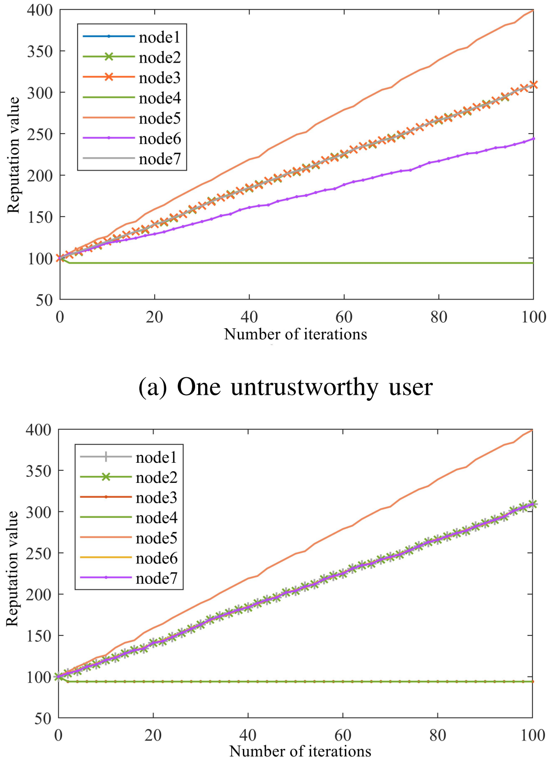
untrustworthy users can be identified through multiple iterations. Figure 13(b) shows the change of sensing energy when there are two untrustworthy users in the system. With the iteration, the difference between the sensing energy values of trustworthy users and untrustworthy users gradually becomes larger. After untrustworthy users are identified, they will no longer participate in cooperation, and the sensing results will



(a) One untrustworthy user

(b) Two untrustworthy users

Fig. 13: Energy values under different iteration times.



(a) One untrustworthy user

(b) Two untrustworthy users

Fig. 14: Reputation value under different iteration times.

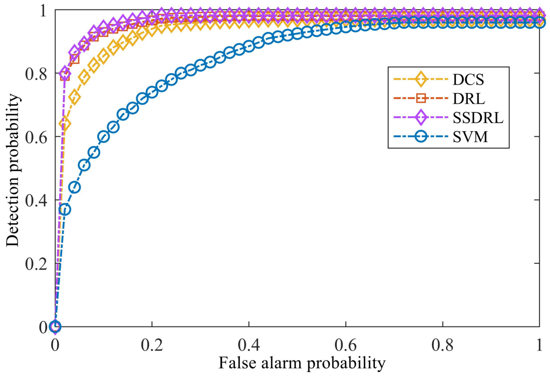


Fig. 15: Detection Probability of Different Spectrum Sensing Algorithms.

tend to be consistent. The network with two untrustworthy users needs 15 iterations to converge, while the network with one untrustworthy user only needs 10 iterations. Therefore, the fewer untrusted users, the faster the consensus fusion. Experiments show that the spectrum sensing algorithm proposed in this paper can effectively identify untrustworthy users and improve network performance.

Spectrum sensing algorithm based on deep reinforcement learning was used to identify untrusted SUs. The reward in reinforcement learning was used to adjust the reputation of SUs. As shown in Figure 14(a) and Figure 14(b), the growth rate of reputation value of different SUs was different. This is because when the neighbors were trusted users, SU achieved better performance and the convergence speed. In the subsequent cooperation process, SUs also tended to choose users with high reputation value. If SUs sent an error message, it would be gradually recognized and the reputation value would be reduced. Subsequently, it would no longer participate in cooperative spectrum sensing, and the reputation value would not change.

In order to evaluate the performance of the SSDRL algorithm, it was compared with other cooperative spectrum sensing algorithms, including DCS, SVM and DRL. In Figure 15, compared with the other three machine learning algorithms, the SSRDL algorithm had better performance under different false alarm probability. When the false alarm probability was large, the performance of three algorithms was close to each other.

It is noted from Figure 16 and Figure 17 that, as the transmit power and the number of samples increase, the sensing error of each algorithm tends to decrease and then stabilize. The sensing error of the SSDRL algorithm is much lower, because cooperative spectrum sensing eliminates the interference of noise and influence of untrustworthy users.

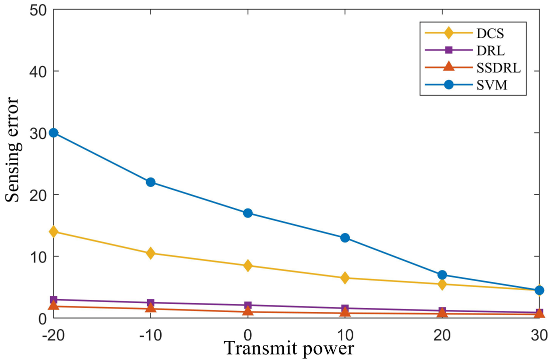


Fig. 16: Sensing error under different transmit power.

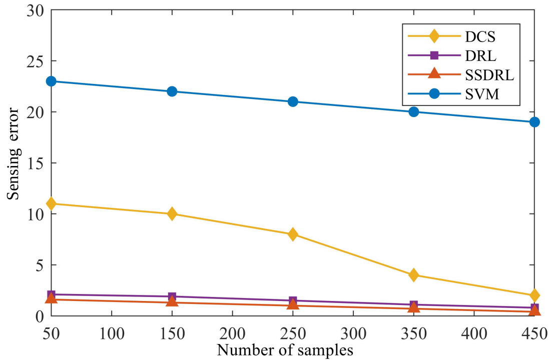


Fig. 17: Sensing error under different sampling numbers.

VI. Conclusion

To improve the efficiency and security of wireless radio resources allocation in NGWN, this paper proposes a secure radio resources allocation scheme based on deep reinforcement learning. The optimized firefly algorithm is used to initialize the weights and thresholds of the BP neural network, which effectively improves the convergence speed during training. In the process of spectrum sensing, SU uses the optimized BP neural network to predict the occupancy. The energy detection method is also used to sense the occupancy state of the target spectrum that is predicted to be idle. Then, each SU interacts with neighboring SUs, and uses deep reinforcement learning algorithms and multi-user cooperation mechanisms to fuse the results, and finally form a consensus across the network. The reputation mechanism is used to remove malicious users for securing cooperative sensing. Simulation results show compared with the traditional solution, the proposal can effectively improve the accuracy of spectrum sensing, significantly reduce the energy consumption in untrustworthy environment.

References

[1] M. Kumar, P. Mukherjee, K. Verma, S. Verma, and D. B. Rawat, “Improved deep convolutional neural network based malicious node detection and energy-efficient data transmission in wireless sensor networks,” *IEEE Transactions on Network Science and Engineering*, pp. 1–1, 2021.

[2] L. Liu, C. Chen, Q. Pei, S. Maharjan, and Y. Zhang, “Vehicular edge computing and networking: A survey,” *Mobile Networks and Applications*, vol. 26, p. 1145–1168, 2021.

[3] D. Wang, Y. He, K. Yu, G. Srivastava, L. Nie, and R. Zhang, “Delay sensitive secure noma transmission for hierarchical hap-lap medical-care iot networks,” *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2021.

[4] H. Yang, W.-D. Zhong, C. Chen, A. Alphones, and X. Xie, “Deep-reinforcement-learning-based energy-efficient resource management for social and cognitive internet of things,” *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5677–5689, 2020.

[5] T. Guo, K. Yu, M. Aloqaily, and S. Wan, “Constructing a prior-dependent graph for data clustering and dimension reduction in the edge of aiot,” *Future Generation Computer Systems*, vol. 128, pp. 381–394, 2022.

[6] S. Bakri, B. Brik, and A. Ksentini, “On using reinforcement learning for network slice admission control in 5g: Offline vs. online,” *International Journal of Communication Systems*, vol. 34, p. e4757, 2021.

[7] L. Liu, J. Feng, Q. Pei, C. Chen, Y. Ming, B. Shang, and M. Dong, “Blockchain-enabled secure data sharing scheme in mobile-edge computing: An asynchronous advantage actor–critic learning approach,” *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2342–2353, 2021.

[8] F. Ding, K. Yu, Z. Gu, X. Li, and Y. Shi, “Perceptual enhancement for autonomous vehicles: Restoring visually degraded images for context prediction via adversarial training,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2021.

[9] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.

[10] B. Brik and A. Ksentini, “Toward optimal mec resource dimensioning for a vehicle collision avoidance system: A deep learning approach,” *IEEE Network*, vol. 35, no. 3, pp. 74–80, 2021.

[11] T. Pamuklu and C. Ersoy, “Grove: A cost-efficient green radio over ethernet architecture for next generation radio access networks,” *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 1, pp. 84–93, 2021.

[12] H.-V. Tran, G. Kaddoum, and C. Abou-Rjeily, “Collaborative rf and lightwave power transfer for next-generation wireless networks,” *IEEE Communications Magazine*, vol. 58, no. 2, pp. 27–33, 2020.

[13] F. Ding, G. Zhu, M. Alazab, X. Li, and K. Yu, “Deep-learning-empowered digital forensics for edge consumer electronics in 5g het-nets,” *IEEE Consumer Electronics Magazine*, vol. PP, no. 99, pp. 1–1, 2020.

[14] L. Yang, K. Yu, S. X. Yang, C. Chakraborty, Y. Lu, and T. Guo, “An intelligent trust cloud management method for secure clustering in 5g enabled internet of medical things,” *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2021.

[15] A. Mariani, S. Kandeepan, and A. Giorgetti, “Periodic spectrum sensing with non-continuous primary user transmissions,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1636–1649, 2015.

[16] B. Brik, A. Ksentini, and M. Bouaziz, “Federated learning for uavs-enabled wireless networks: Use cases, challenges, and open problems,” *IEEE Access*, vol. 8, pp. 53841–53849, 2020.

[17] C. Feng, B. Liu, Z. Guo, K. Yu, Z. Qin, and K.-K. R. Choo, “Blockchain-based cross-domain authentication for intelligent 5g-enabled internet of drones,” *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[18] M. Kim, N.-I. Kim, W. Lee, and D.-H. Cho, “Deep learning-aided scma,” *IEEE Communications Letters*, vol. 22, no. 4, pp. 720–723, 2018.

[19] F. Obite, A. D. Usman, and E. Okafor, “An overview of deep reinforcement learning for spectrum sensing in cognitive radio networks,” *Digital Signal Processing*, vol. 113, p. 103014, 2021.

[20] W. Lee, M. Kim, and D.-H. Cho, “Deep power control: Transmit power control scheme based on convolutional neural network,” *IEEE Communications Letters*, vol. 22, no. 6, pp. 1276–1279, 2018.

[21] T. Düzenli and O. Akay, “A new spectrum sensing strategy for dynamic primary users in cognitive radio,” *IEEE Communications Letters*, vol. 20, no. 4, pp. 752–755, 2016.

[22] J. Mu, D. Xie, H. Huang, and X. Jing, “Computation-constrained spectrum sensing in iot-based scenarios,” *IET Communications*, vol. 14, pp. 3631–3638, 2020.

[23] M. Amjad, H. Afzal, H. Abbas, and A. Subhani, “Ads: An adaptive spectrum sensing technique for survivability under jamming attack in cognitive radio networks,” *Computer Communications*, vol. 172, 2021.

[24] Bhowmick, A., Chandra, Dhar, Roy, S., and Kundu, “Double threshold-based cooperative spectrum sensing for a cognitive radio network with improved energy detectors,” *Communications, IET*, vol. 9, no. 18, pp. 2216–2226, 2015.

[25] A. Gharib, W. Ejaz, and M. Ibnkahla, “Distributed learning-based multi-band multi-user cooperative sensing in cognitive radio networks,” in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2018.

[26] A. H. Gazestani and S. A. Ghorashi, “Distributed diffusion-based spectrum sensing for cognitive radio sensor networks considering link failure,” *IEEE Sensors Journal*, vol. 18, no. 20, pp. 8617–8625, 2018.

[27] J. Tong, M. Jin, Q. Guo, and Y. Li, “Cooperative spectrum sensing: A blind and soft fusion detector,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2726–2737, 2018.

[28] W. Lee, M. Kim, and D.-H. Cho, “Deep cooperative sensing: Cooperative spectrum sensing based on convolutional neural networks,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 3005–3009, 2019.

[29] C. Sun, H. Ding, and X. Liu, “Multichannel spectrum access based on reinforcement learning in cognitive internet of things,” *Ad Hoc Networks*, vol. 106, p. 102200, 2020.

[30] R. Rajaguru, V. Rajendran, and P. Marichamy, “A hybrid spectrum sensing approach to select suitable spectrum band for cognitive users,” *Computer Networks*, vol. 20, p. 107387, 2020.

[31] Y. Xu, P. Cheng, Z. Chen, Y. Li, and B. Vucetic, “Mobile collaborative spectrum sensing for heterogeneous networks: A bayesian machine learning approach,” *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5634–5647, 2018.

[32] K. W. Sung, S.-L. Kim, and J. Zander, “Temporal spectrum sharing based on primary user activity prediction,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3848–3855, 2010.

[33] G. Ding, F. Wu, Q. Wu, S. Tang, F. Song, A. V. Vasilakos, and T. A. Tsiftsis, “Robust online spectrum prediction with incomplete and corrupted historical observations,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8022–8036, 2017.

[34] H. Eltom, S. Kandeepan, Y.-C. Liang, and R. J. Evans, “Cooperative soft fusion for hmm-based spectrum occupancy prediction,” *IEEE Communications Letters*, vol. 22, no. 10, pp. 2144–2147, 2018.

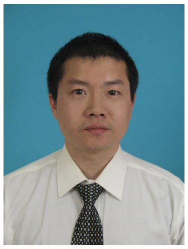
[35] V. K. Tumuluru, P. Wang, and D. Niyato, “A neural network based spectrum prediction scheme for cognitive radio,” in *2010 IEEE International Conference on Communications*, pp. 1–5, 2010.

[36] X. Ding, L. Feng, Y. Zou, and G. Zhang, “Deep learning aided spectrum prediction for satellite communication systems,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 16314–16319, 2020.

[37] Y. Xu, J. Yu, and R. M. Buehrer, “The application of deep reinforcement learning to distributed spectrum access in dynamic heterogeneous environments with partial observations,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4494–4506, 2020.

List of Figures

List of Tables



Yuhuai Peng received the Ph.D. degree in communication and information systems from Northeastern University in 2013.

He is currently an associate professor in the same university. His research interests include Internet of Things (IoT), industrial communication networks, health monitoring, etc.

Xiaojing Xue received the M.E. degree in the school of Computer Science and Engineering at Northeastern University, Shenyang, China. Her research interests include artificial intelligence, industrial Internet of things and cognitive networks.

Ali Kashif Bashir (Senior Member, IEEE) is a Reader/Associate Professor and Program Leader of BSc (H) Computer Forensics and Security at the Department of Computing and Mathematics, Manchester Metropolitan University, United Kingdom. He is also with School of Electrical Engineering and Computer Science, National University of Science and Technology, Islamabad (NUST) as an Adjunct Professor and School of Information and Communication Engineering, University of Electronics Science and Technology of China (UESTC) as an Affiliated Professor and Chief Advisor of Visual Intelligence Research Center, UESTC. He is a senior member of IEEE, member of IEEE Industrial Electronic Society, member of ACM, and Distinguished Speaker of ACM. His past assignments include Associate Professor of ICT, University of the Faroe Islands, Denmark; Osaka University, Japan; Nara National College of Technology, Japan; the National Fusion Research Institute, South Korea; Southern Power Company Ltd., South Korea, and the Seoul Metropolitan Government, South Korea. He has worked on several research and industrial projects of South Korean, Japanese and European agencies and Government Ministries. In his career, he has obtained over 2.5 Million USD funding. He received his Ph.D. in computer science and engineering from Korea University South Korea. He has authored over 180 research articles; received funding as PI and Co-PI from research bodies of South Korea, Japan, EU, UK and Middle East; supervising/co-supervising several graduate (MS and PhD) students. His research interests include internet of things, wireless networks, distributed systems, network/cyber security, network function virtualization, machine learning, etc. He is serving as the Editor-in-chief of the IEEE FUTURE DIRECTIONS NEWSLETTER. He is also serving as area editor of KSII Transactions on Internet and Information Systems; associate editor of IEEE Internet of Things Magazine, IEEE Access, Peer J Computer Science, IET Quantum Computing, Journal of Plant Disease and Protection. He is leading many conferences as a chair (program, publicity, and track) and had organized workshops in flagship conferences like IEEE Infocom, IEEE Globecom, IEEE Mobicom, etc.

Xiaogang Zhu received his B.S. degree from Nanchang University, China in 2000, and he received his M.S. degree in Software Engineering from Tongji University, Shanghai, China. He is currently a professor in the School of Management and the director of the Institute of Big Data and Cybersecurity at Nanchang University. His main research areas are data security, big data analysis and applications.

Yasser D. Al-Otaibi is currently an Assistant Professor in the Department of Information Systems at the Faculty of Computing and Information Technology in Rabigh, King Abdulaziz University, Jeddah, Saudi Arabia. He received a PhD degree in Information Systems from Griffith University, Australia in 2018. His current research interests include IT adoption and acceptance, wireless sensor networks, and IoT.

Usman Tariq received the Ph.D. degree in information and communication technology in computer science from Ajou University, South Korea. He is currently an Associate Professor with the College of Computer Engineering and Science, Prince Sattam Bin Abdulaziz University. He is also a skilled Research Engineer. He has a strong background in ad hoc networks and network communications. He is also experienced in managing and developing projects from conception to completion. He has worked in large international scale and long-term projects with multinational organizations.

Keping Yu received the M.E. and Ph.D. degrees from the Graduate School of Global Information and Telecommunication Studies, Waseda University, Tokyo, Japan, in 2012 and 2016, respectively. He was a Research Associate and a Junior Researcher with the Global Information and Telecommunication Institute, Waseda University, from 2015 to 2019 and 2019 to 2020, respectively, where he is currently a Researcher.

Dr. Yu has hosted and participated in more than ten projects, is involved in many standardization activities organized by ITU-T and ICNRG of IRTF, and has contributed to ITU-T Standards Y.3071 and Supplement 35. He received the Best Paper Award from ITU Kaleidoscope 2020, the Student Presentation Award from JSST 2014. He has authored 100+ publications including papers in prestigious journal/conferences such as the IEEE Wireless Communications, ComMag, NetMag, IoTJ, TFS, TII, T-ITS, TVT, TNSE, TGCN, CEMag, IoTMag, ICC, GLOBECOM etc. He is an Associate Editor of IEEE Open Journal of Vehicular Technology, Journal of Intelligent Manufacturing, Journal of Circuits, Systems and Computers. He has been a Lead Guest Editor for Sensors, Peer-to-Peer Networking and Applications, Energies, Journal of Internet Technology, Journal of Database Management, Cluster Computing, Journal of Electronic Imaging, Control Engineering Practice, Sustainable Energy Technologies and Assessments and Guest Editor for IEICE Transactions on Information and Systems, Computer Communications, IET Intelligent Transport Systems, Wireless Communications and Mobile Computing, Soft Computing, IET Systems Biology. He served as general co-chair and publicity co-chair of the IEEE VTC2020-Spring 1st EBTSRA workshop, general co-chair of IEEE ICCC2020 2nd EBTSRA workshop, general co-chair of IEEE TrustCom2021 3nd EBTSRA workshop, session chair of IEEE ICCC2020, TPC co-chair of SCML2020, local chair of MONAMI 2020, Session Co-chair of CcS2020, and session chair of ITU Kaleidoscope 2016. His research interests include smart grids, information-centric networking, the Internet of Things, artificial intelligence, blockchain, and information security.