Wireless Communications and Mobile Computing

Intelligent Network Management for 5G and Beyond

Lead Guest Editor: Taras Maksymyuk
Guest Editors: Juraj Gazda, Longzhe Han, Axel Sikora, and Tarcisio Maciel

Intelligent Network Management for 5G and Beyond

Wireless Communications and Mobile Computing

Intelligent Network Management for 5G and Beyond

Lead Guest Editor: Taras Maksymyuk
Guest Editors: Juraj Gazda, Longzhe Han, Axel Sikora, and Tarcisio Maciel

Contents

A Deep Learning-Based Power Control and Consensus Performance of Spectrum Sharing in the CR Network
Muhammad Muzamil Aslam, Liping Du, Zahoor Ahmed, Muhammad Nauman Irshad, and Hassan Azeem
Research Article (16 pages), Article ID 7125482, Volume 2021 (2021)

Behavior Anomaly Detection in SDN Control Plane: A Case Study of Topology Discovery Attacks
Li-Der Chou, Chien-Chang Liu, Meng-Sheng Lai, Kai-Cheng Chiu, Hsuan-Hao Tu, Sen Su, Chun-Lin Lai, Chia-Kuan Yen, and Wei-Hsiang Tsai
Research Article (16 pages), Article ID 8898949, Volume 2020 (2020)

A Blockchain Token-Based Trading Model for Secondary Spectrum Markets in Future Generation Mobile Networks
Mubbashar A. Khan, Mohsin M. Jamali, Taras Maksymyuk, and Juraj Gazda
Research Article (12 pages), Article ID 7975393, Volume 2020 (2020)

Energy Harvesting and Information Transmission Mode Design for Cooperative EH-Abled IoT Applications in beyond 5G Networks
Hongyuan Gao, Shibo Zhang, Yumeng Su, and Ming Diao
Research Article (17 pages), Article ID 6136298, Volume 2020 (2020)

Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7125482, 16 pages
https://doi.org/10.1155/2021/7125482

Research Article

A Deep Learning-Based Power Control and Consensus Performance of Spectrum Sharing in the CR Network

Muhammad Muzamil Aslam[1], **Liping Du**[1,2], **Zahoor Ahmed**[3,4], **Muhammad Nauman Irshad**[1] and **Hassan Azeem**[1]

[1]*School of Computer & Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China*
[2]*Shunde Graduate School of University of Science and Technology Beijing, Foshan 528000, China*
[3]*Department of Automation, Shanghai Jiaotong University, Shanghai 200240, China*
[4]*Department of Electronics, GC University Lahore, 54000, Pakistan*

Correspondence should be addressed to Liping Du; dlp2001@ies.ustb.edu.cn

Received 21 December 2019; Revised 5 November 2020; Accepted 23 January 2021; Published 19 February 2021

Academic Editor: Longzhe Han

The cognitive radio network (CRN) is aimed at strengthening the system through learning and adjusting by observing and measuring the available resources. Due to spectrum sensing capability in CRN, it should be feasible and fast. The capability to observe and reconfigure is the key feature of CRN, while current machine learning techniques work great when incorporated with system adaption algorithms. This paper describes the consensus performance and power control of spectrum sharing in CRN. (1) CRN users are considered noncooperative users such that the power control policy of a primary user (PU) is predefined keeping the secondary user (SU) unaware of PU's power control policy. For a more efficient spectrum sharing performance, a deep learning power control strategy has been developed. This algorithm is based on the received signal strength at CRN nodes. (2) An agent-based approach is introduced for the CR user's consensus performance. (3) All agents reached their steady-state value after nearly 100 seconds. However, the settling time is large. Sensing delay of 0.4 second inside whole operation is identical. The assumed method is enough for the representation of large-scale sensing delay in the CR network.

1. Introduction

Intelligent processing is one of the major advantages of CRN [1]. Due to this feature, these systems possess the capability to learn their environment, increase awareness, and reconfigure themselves accordingly. High environmental awareness of these systems, due to their ability to perceive, gives them an enormous advantage in a wireless communication environment. Therefore, a device's wireless operations can be effectively adapted to its surroundings, and thus, it uses the maximum resource to deliver the best results. Some techniques, e.g., intelligent reflecting surface (IRS), for secure communication of multiantenna have been studied in [2] which contain reflector elements that are reconfigurable and controlled by software that is communication-oriented [3, 4]. In most cases, spectrum sensing is used to achieve perception capability in these systems [5]. Therefore, weak spectrum sensing ability would limit their operations and, as a result, decrease the efficiency. In this context, many researchers have proposed techniques to improve spectrum sensing [6, 7], which contain, but are not limited to, wideband spectrum sensing, cooperative spectrum sensing [8], and sequential spectrum sensing. It is a concept widely used for perception improvement in a wireless networking environment [9]. In this context, spectrum measurements, multidimensional spectrum sensing, and interference sensing [10] have been studied extensively.

Primary users' (PU) role is considered an active and passive user model in academia for spectrum sharing. There was a cooperative or noncooperative relationship between PU and SU in the active models [11]. To enhance the transmission performance of the system, information interaction

has been performed, although, in research, SU performs spectrum sensing in a passive user model for finding power allocation or idle spectrum. When there is a passive PU role, then PU assigns its transmit power relying on its power controlling scheme [12].

CRN transmit power selection must be properly controlled for the transmissions of the CR users [13, 14]. This is the factor that can help to achieve high efficiency of the spectrum through reuse of the spectrum bands of the power units which are affected by the PU limitations. For spectrum sensing, resources are required at all nodes [15]. Hence, effective development of spectrum sensing is of high significance to achieve optimal bandwidth, time, and power spent in between the transmission and sensing [16–18]. Intelligent [19] wireless communication devices must be adaptive to their environment to be able to perceive their surroundings. This can be achieved through the optimization of working limitations and dynamic spectrum access.

In recent years, fast developments in machine learning have been noted. Applied in different problems of wireless communication [20], both supervised and unsupervised learning [21] have been considered. Different from supervised learning, *Q*-learning or reinforcement learning [22] is found to be valuable to maximize the long-term performance of a system and to achieve a balance between exploitation and exploration. Furthermore, deep learning has proven to be a predominant method of achieving sound and higher performance, which is unaffected by massive data sets, and also seems promising in wireless communication [22]. In this context, reinforcement learning is a machine learning technique inspired by biological phenomena, where knowledge is acquired to an agent by repeated trial-and-error communication with its surroundings. Environment feedback to an agent's action is used to optimize the machine for future behaviour. Dynamic interaction between the agent and its surroundings and the resulting response are two of the key topographies which make reinforcement learning techniques appealing for cognitive radio ad hoc network applications, mostly for the tasks of routing and spectrum decision [23]. In several cases, operations based on reinforcement learning are better than traditional solutions [24].

For better performance of CRN, awareness about surroundings (radiofrequency (RF)) is a must for CR. CR should sense and check all around the location for keeping an eye on the RF activities. There was a basic identification of spectrum sensing in CRs. Some sensing techniques, e.g., energy detection, based on a matched filter and cyclostationary process, have been studied in [25]. For better results of sensing accuracy, cooperative sensing was studied in [26] to focus on wireless network hidden terminal problems. Recently researchers studied cooperative CRS [27].

Cooperative communication is an advanced communication-in-future technology that optimizes signal transmissions w.r.t both medium contact control and physical coating as presented in [28]. There is an important role of *Q*-learning in CR application, for instance, dynamic spectrum access. Users iteratively improve their planning to attain their tasks through knowledge and recompense from their surroundings. In these days, there are several research techniques to solve the problems of power control and power allocation [29], for example, game theory [30], optimization theory [31], and machine learning; among these, deep reinforcement learning which is a kind of machine learning has got a lot of attention because of its fast speed over the complex problem. The contribution of its challenging ability in many applications can be found such as Atari Games [32]. Agents [15] that are trained through reinforcement learning are intelligent enough to absorb their act value policies from high-dimensional rare data. An example of such applications would be videos or images [33] also shown in our tentative result.

Deep ReLU (rectified linear unit) learning can assist in learning an operative action charge policy even when the state notes are contaminated by measurement errors or arbitrary noise. Using deep reinforcement learning power control-based spectrum sharing in wireless networks showing the relation between PU and SU, wireless sensors have been studied in [34] and channel selection policy for PU and SU has been studied in [35]. On the other hand, the conventional ReLU method is unfeasible for such issues because of an inadequate number of states in the presence of arbitrary noise. Therefore, deep reinforcement learning is appropriate for wireless communication applications, where municipal dimensions are usually arbitrary. In this paper, we described the consensus performance [4, 36] and the power control of spectrum sharing in CRN. We also introduced a sensing delay. In CR, power control is investigated for spectrum sharing to assure the quality of service for PU and SU. First of all, CRN users are considered noncooperative users such that the power control policy of primary users (PUs) is predefined, while the secondary user (SU) is unaware of PU's power control policy. Secondly, a model was constructed for cooperation between wireless sensors, PU, and SU. In this model, wireless sensors' received signal strength is spatially distributed to assist PU for power transmission information with SU. The performance criterion of sharing in CRN has been improved in the presence of sensing delay in a communication network. The studied model can get minibatch updates for several iterations. The network becomes able to converge quickly and meet a user's quality of services. The overall contributions of this work or difference with other works can be summarized as follows:

(i) The proposed method is different from the traditional method in the sense that there is a performance criterion of sharing, and consensus of CRN has been improved in the presence of sensing delay (communication delay) through the communication topology

(ii) An agent-based approach is introduced for the CR user's consensus performance for the first time

(iii) Besides, a deep learning power control strategy (reinforcement learning) has been developed along with an agent-based approach altogether for a more efficient performance of sharing control

In the rest of the paper, we discussed the following: overview of CRN, intelligent power control of spectrum sharing,

delay performance using a primary sensor network, and CRN consensus criteria under communication delay.

2. An Overview of CRN

We suppose a CRN which contains licensed users and unlicensed users; in this, the unlicensed user is willing to segment spectrum resources which are common with the licensed user, without any danger to harm the licensed user activity. The licensed user contains the primary transmitter (P_{T1}) and receiver P_{R1}; similarly, the unlicensed user consists of a secondary transmitter S_{T2} and receiver S_{R2}. The receiver and transmitter of both licensed and unlicensed users are shown in Figure 1 in a multicluster network number of sensors that are present for direct communication with the central station; the central station can be the Cluster Head (CH), the name cluster is because of network similarity with neural network. In our supposed work, there is no working cooperation between the licensed and unlicensed users, where a licensed user does not know the presence of the unlicensed user in the same surroundings even licensed users adjust its conduct power built on its particular power policy. The next step of the licensed user will be affected in a couched way by the sudden action of the unlicensed user; it should be highlighted that for the licensed user, the power control policy depends on the network of licensed users and the number of sensors used. Yet licensed and unlicensed users are without communication, and both have no information about the power transmitter and power control policy of each other. For an easy understanding, we consider that licensed and unlicensed users synchronously update their transmit power, where it is based on the time frame base.

Here, in terms of SNR_c, it is used for showing the quality of the system of licensed and unlicensed users. ∂_a and ∂_b show the transmit power of licensed and unlicensed users. The receiver value of SNR_c is given as

$$\mathrm{SNR}_c = \frac{|H_{xx}|^2 \partial_a}{\sum_{y \neq x} |H_{yx}|^2 \partial_b + N_x}, \quad \forall\, x = 1, 2. \tag{1}$$

In Equation (1), H_{xy} is the frequency expansion from P_{T1} (primary transmitter) to S_{T2} (secondary transmitter), while N_x is the noise power from P_{R1} (primary receiver). Here, the purpose is to assist the unlicensed users in acquiring sufficient control power policy, so that after some time of power adjustment, both licensed and unlicensed users are intelligent for the successful transmission of data with the essential quality of services. If unlicensed or licensed users are only aware of their transmit power, then such a task is complicated. To assist unlicensed users, in a wireless communication network using different locations, sensor nodes are hired for the dimension of received signal strength (RSS). Received signal strength measurements are related to both (licensed and unlicensed) users' transmit power, close-fitting the information system from the state. Here, we suppose that RSS information is helpful for the unlicensed user. Zigbee technology [7] and CRNs generally function with different frequencies, and there is no interference in transmission because of sensor node transmission in CRNs. Tocollect the received signal strength information from spatially scattered sensors, node is a basic condition, for example, source localization [17]. In the proposednumerical method, only once per time frame is needed to report RSS informationwhichconsists of low data rate.

We consider P_{R1} and S_{R2}; here, S_{R2} mollify the lowest SNR_c aimed at an effective response. For getting the quality of system condition, the licensed operator or user is considered to deceptively regulate its power transmission that is power control policy-based. We used "2" controlling power policies that supposed for licensed users, even if the licensed user (PU) adopts any other power control policy, our supposed method will also work. For the first policy, the transmit power of the licensed user is updated following the classical power control algorithm. Transmit power is adjusted on a time framed bases K(x) that maps the continuous values level, known as discretization operation.

$$\partial_a(r+1) = K\left(\frac{\omega'_c \partial_a(r)}{\mathrm{SNR}_c(r)}\right). \tag{2}$$

Here, $\mathrm{SNR}_c(r)$ is the SNR_c measured at P_{R1} at the r^{th} time frame. $\partial_a(r)$ is the transmit power at the r^{th} time frame.

Now, we explain the PU policy of our supposed work. Let us see

$$\partial_a(r+1) = \begin{cases} \partial^{\partial}_{y+1}, & \text{if } \partial_y{}^{\partial} \leq \tilde{\eta} \leq \partial^{\partial}_{y+1} \text{ and } y+1 \leq E_1, \\ \partial^{\partial}_{y-1}, & \text{else } \tilde{\eta} \leq \partial^{\partial}_{y-1} \text{ and } y-1 > 1, \\ \partial^{\partial}_{y}, & \text{otherwise}, \end{cases} \tag{3}$$

Here, $\tilde{\eta} := \acute{\omega}_c\, \partial_a(k)/\mathrm{SNR}_c(r)$; compared with Equation (2), Equation (3) has more unadventurous behaviour for power control policy. Transmit power is updated step by step. The power is increased by a single step when $\mathrm{SNR}_c(r) \leq \acute{\omega}_c$ and $\acute{\omega} \geq \acute{\omega}_c$; also, the power is reduced by a single phase, when $\mathrm{SNR}_c(r) \geq \acute{\omega}_c$ and $\acute{\omega} \geq \acute{\omega}_c$. If not increasing or decreasing, then it will stay at the present power level. And here, $\acute{\omega} := \mathrm{SNR}_c(r)\, \partial_a(r+1)/\partial_a(r)$ is the forecast SNR at the $(r+1)$ time frame. Now,

$$P_1 \approx \left\{ \partial^{\partial}_a \cdots \partial^{\partial}_{E_1} \right\}. \tag{4}$$

Here, $\partial^{\sigma}_a \leq \cdots \partial$. Furthermore, precisely we suppose that $K(x)$ is very equal to a discrete step, which is not lesser than x, and suppose that $k(x) = \partial$ if $x > \partial$. Now, assume the second power control policy and consider that the transmit power at the r^{th} time frame is $\partial_a(r) = \partial^{\sigma}_y$, where $\partial^{\sigma}_y \in P$. Equation (3) shows the transmit power of PUs updating.

A suggested control channel is used to notify the sensors about the available users. Scheming a CRN without a suggested control user can be studied in [9]. Sensor node switch between various users depends upon the available user; also,

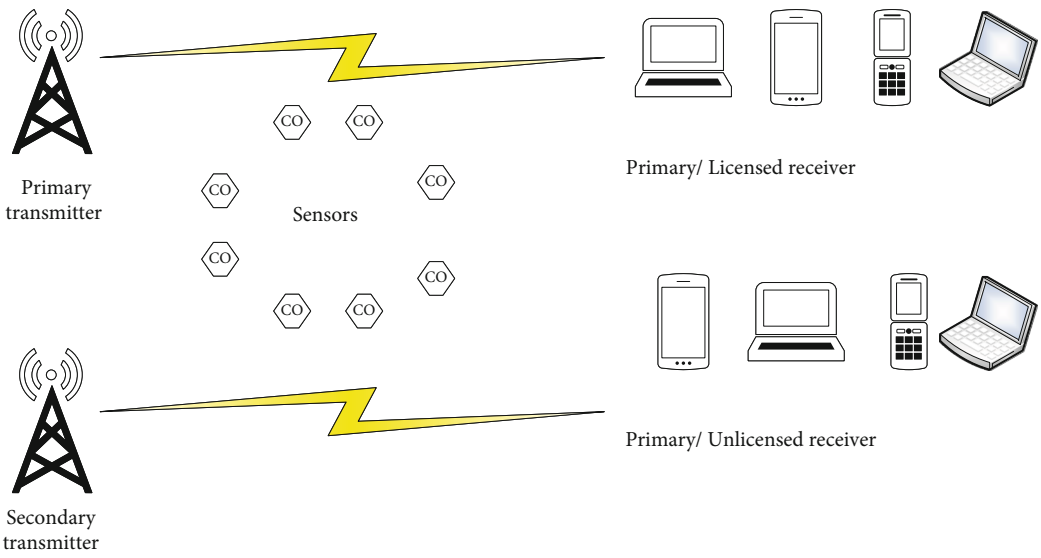


Figure 1: Spectrum sharing in CRNs (sharing the same channel).

there is a possibility of the sensor to operate on different users. Further info about this can be found in [30]. A multi-radio wireless sensor network has been explained in some literature which can be found in [18]. In this research article, consider that N sensors are installed to sample the received signal strength info. Consider that S_n represents node n and $\partial_n^m(r)$ represents the receiving power at the r^{th} frame at sensor n. We simulate the state of received signal strength observations.

And now, the SU policy is explained in the form of the following equation:

$$\partial_n^m(r) = \partial_a(r)J_{cn} + \partial_b(r)J_{dn} + L_n(r). \tag{5}$$

In Equation (5), $\partial_a(r)$ represents the transmit power of PUs, $\partial_b(r)$ represents the transmit power of SUs, J_{cn} is the path loss between sensor n and the primary transmitter, J_{dn} is the path loss between the unlicensed transmitter and sensor n, and $L_n(r)$ is the zero mean Gaussian random variables with variance φ_n^2.

Here, according to free-space broadcast, with regard to the Friis law, J_{cn} and J_{dn} are given below:

$$\begin{aligned} J_{cn} &= \left(\frac{\gamma}{4\pi D1_n}\right)^2, \\ J_{dn} &= \left(\frac{\gamma}{4\pi D2_n}\right)^2. \end{aligned} \tag{6}$$

Usually, the Friis equation is used for the calculation of ideal power which was received at an antenna from elementary info about the transmission. Here, "γ" is the wavelength of the signal $D1_n$ and $D2_n$ showing detachment between nodes and primary and secondary transmitters. The Friis equation is used only for the calculation of single frequency, even though transmission characteristically contained many. Here, also to suppose, from a finite set, we choose the transmit power of the unlicensed user:

$$P_2 \coloneqq \left\{\partial_a^m \cdots \partial_{E_2}^m\right\}, \tag{7}$$

where $\partial_a^m \leq \cdots \leq \partial_{E2}^m$; the SU goal is to see how power transmission can be settled depending on the received signal strength info $\{\partial_n^m(r)\}_n$; after rare rounds of power transmission at every time frame, both PU and SU can see their respective quality of system necessities. At least transmit power is present in our supposed work, which transmits power $\{\partial_{s_1}^{\sigma}, \partial_{s_2}^m\}$; here, P_{R1} and S_{R2} mollify their quality of system SNR necessities.

2.1. Intelligent Power Control of Spectrum Sharing. The distribution of system time is uniform, denoted as channel switching intermission. The presently employed channel develops unattainable culmination of the cooperative spectrum interval. When the sensor senses the second user, the channel might delay till the start of the next cooperative spectrum pause (if at least one user is available). In that case, both the user and sensor jump to the new channel. This process is termed as periodic switching (PS). Another cause may be that the central head may highlight the sensor if a new channel is available, as the prior channel is misplaced; this channel loss is stated as triggered switching (TS). The number of channel switching may be zero or one for periodic switching (PS); it depends whether the channel is present or not at the start of the cooperative spectrum interval. For a satisfactory transmission, the central station sends back an Automatic Control Key (ACK) to the sensors. If the sensor does not receive an ACK after a data packet transmission, it is assumed that the present channel is unavailable, and the transmission is stopped immediately. However, this kind of failure may happen due to channel fading which would cause a transmission failure on the CR sensor net. It is not a good practice to stop the transmission in this case, which would instigate needless

interference with the PN (Primary Network). Therefore, BE (Best-Effort) data traffic and actual-time data traffic can be assisted in this case. A good example of this channel switching and traffic resource allocation can be found in [2].

In the proposed work, SUs are assumed to take any action at each time step, such as for choosing transmit power from the earliest quantified power set P_2 based on its present state. To analyse this, the Markov decision process (MDP) [37] is considered. Particularly, for instance, consider an action $a(r) = \partial(r+1)$ by SU in the state $m(r)$. Then, the next new state $m(r+1)$ is dependent on the present state $m(r)$ and action of a decision-maker $a(r)$. Here, $m(r)$ and $a(r)$ are known. The following states will be independent of all earlier states and actions with the condition that

$$m(k) \approx \left[\overset{\vee}{\partial_a^n}(r) \cdots \overset{\vee}{\partial_N^n}(r) \right]^T. \tag{8}$$

For a new state, a corresponding reward $n(r) := n(m(r))$ is assumed for the decision-maker $a(r)$ which can be mathematically expressed as

$$\overset{\vee}{n}(r) \approx \begin{cases} z, & \text{if } \mathrm{SNR}_c(r+1) \geq \mathrm{SNR}_d(r+1) \geq \acute{\omega}_d, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

In this work, "z" is assumed to be 10. It can be seen through simulation that the value of z should be sufficiently large to avoid the learning disturbance. Communication between a SU, a PU, and the sensors is shown in Figure 2(a). It is to be highlighted that the unlicensed user is supposed to distinguish whether communication between P_{T1} and P_{R1} is successful or not. This kind of information can be gained to monitor an acknowledged signal transferred P_{R1} for an indication of successful transmission from P_{T1}. Figure 2(b) [15] shows the detached control architecture and is an object control system design in which there is a linear time-invariant (LTI) and multiple input-multiple output (MIMO) system.

The basic problem for the Markov decision process is to understand the policy for the decision-maker.

$$U^{\pi}(m(r)) \coloneqq \sum_{a=r}^{T-1} \gamma^{a-r} \overset{\vee}{n}(a), \tag{10}$$

where π stipulates the action $\pi(m)$ from the decision-maker. It is important to say that the purpose of SU is to understand a strategy π for the selection of its action $a(r)$ which is constructed on the present state $m(r)$ in a method that discounted cumulative reward reduced maximum as studied in [38]. In Equation (10), "γ" is the discount factor and T^{-1} is representing the time frame, at which the targeted state is touched. In this case, the targeted state is well defined, where $\mathrm{SNR}_c(r) \geq \acute{\omega}_i$ for $i = 1, 2$. The target is for studying optimal policy π^* that maximizes U^{π}, i.e.,

$$\pi^* = \arg\max U^{\pi}(m), \quad \forall m, \tag{11}$$

It is complicated to directly learn π^*. In the case of reinforcement learning, there is an alternative approach to solve Equation (11) through Q-learning. This alternative approach has been studied in [13]. Q-function (action value) is defined to determine the predictable reduced growing reward after the action $a(k+n)$, for state m. In such cases, the optimal policy would be built by choosing the action by the largest value in each state.

By the updated Bellman equation,

$$Q(m, a) = \overset{\vee}{n}(m, a) + \gamma \max\ Q\left(m', a'\right) m \tag{12}$$

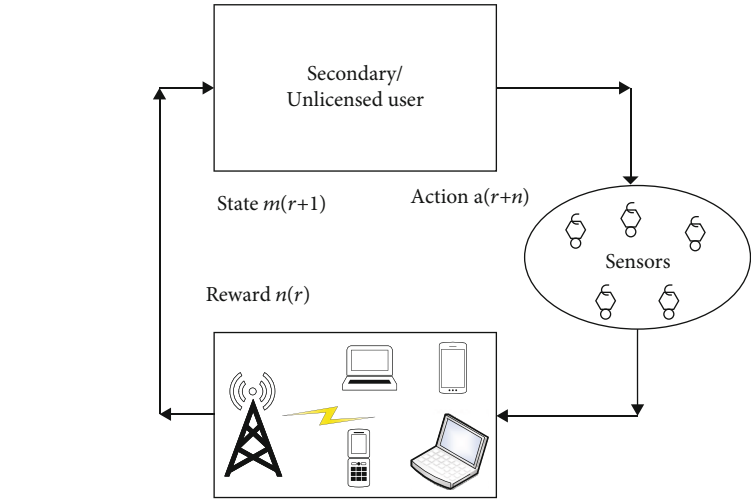
The elementary purpose of Q-learning and several other reinforcement learning algorithms is to iteratively update the action value according to the value updating instruction. In Equation (12), "m'" is the state followed by smearing of action "a" to the present state "m." The value iteration algorithm defined in Equation (12) approaches the best action value purpose. The number of states is limited for Q-learning, and at each state, the action value function is guessed separately. The Q-table is designed such that rows represent the states and columns represent the probable action state. One can choose an action from the Q-table as the maximum value of $Q(m, a)$ through an optimal action state. However, the value of state function m is continuous due to arbitrary variation in the received signal. Therefore, the Q-learning method is unfeasible in this work since it is not convenient to consider an unlimited figure of states. To overcome this issue, DQN (Deep Q-Network) is opted. A brief study and proposed work of DQN are presented in [39].

Different from the traditional Q-learning technique, which is used to generate unlimited action value functions, "ϑ" is used here to express weights of Q-learning, while the Q-table is substituted by the deep neural network $Q(m, a : \vartheta)$ that would act as the action value function. Particularly, for an input "m," DQN produces an E_2-dimensional vector for selecting action $a = \mathcal{A}_i^m$ from $\mathcal{P}_2$. When $m(r)$ is given, the data used for Q-network training is generated as follows: we explore amorphous selected action with ϵ_k having the largest output $Q(m(r), a(r): \vartheta_0)$, where ϑ_0 is used for denotation of the present iteration. Subsequently, the action an (r), SU gets a reward Υ(r), a new state$m(r+1)$is obtained. This transit, $d(r) := \{m(r), a(r), m(r), m(r+1)\}$, is held in the reiteration memory K. When enough transitions are collected in K, the training of the Q-network begins. Say $O =$ 300 evolutions, we casually choose a minibatch of transitions $\{d(i) i \in \Omega_r\}$ from K, and with adjustment of ϑ, DQN can be trained as given loss function is minimized:

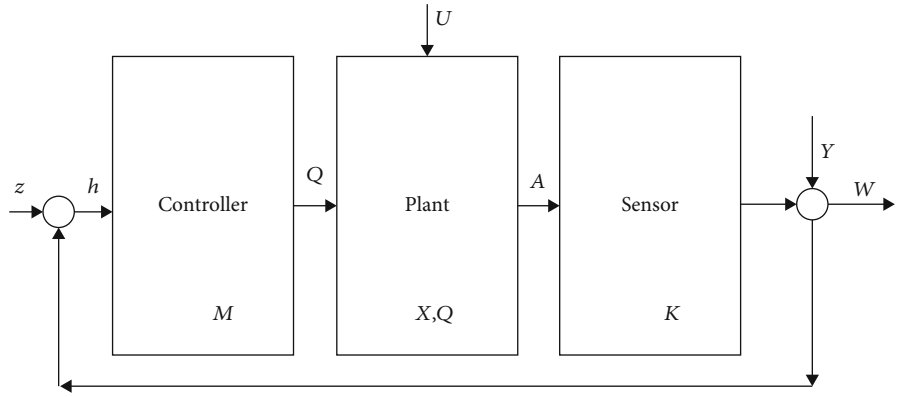
$$E(\vartheta) \approx \frac{1}{\Omega_r} \sum_{i \varepsilon \mho_r} Q'(i) - Q\left(m(i), \dddot{a}(i); \vartheta\right)^2 m. \tag{13}$$

Ω_r is the index set of the irregular minibatch used at the r^{th} iteration. The estimated value of the Bellman equation $Q'(i)$ for the i^{th} iteration is given by

$$Q'(i) = \overset{\vee}{n}(i) + \gamma \max\ Q\left(s(i+1), \ddot{a}; \vartheta_0\right), \quad \forall i \varepsilon \Omega_r m. \tag{14}$$



(a) Primary/licensed network



(b) Standard multiuser close loop

Figure 2: Unlicensed user and licensed network interactions and standard multiuser close loop.

Different from the conventional supervised learning, the goal for DQN knowledge is different from that of the traditional approach.

SU can select action after training, which gives a maximum projected value $Q(m, a, \vartheta^*)$. For simplicity, the DQN scheme is summarized in the power control scheme for SU. It would be clearer to highlight that during the training process of DQN, an SU needs the knowledge that satisfies the QoS (quality-of-service) necessities for the PU, and therefore, the SU is mollified. However, the only purpose of SU, after DQN training completion, is feedback from the sensors for the decision of the next transmit power. The junction problem of the supposed policy for power control is hence studied. Let "m" be a target state; then, if SU transmit power stays unchanged, it will be informal to say that the next state m' is also a goalmouth state. Transmit power can result from either Equation (2) or (3), and as a result, PU will be updated. In another way, SU will ultimately acquire to select a diffuse power as m' persists the goalmouth state. It can be concluded that once s achieves the target state, it will be in the goalmouth state till broadcast information is done. Now assume that there is a possibility of data loss because of the irregularity of data transmission and SU aims to take up for an innovative transmission; no long learning is required in this case. Rendering to control power policy, SU can select transmit power.

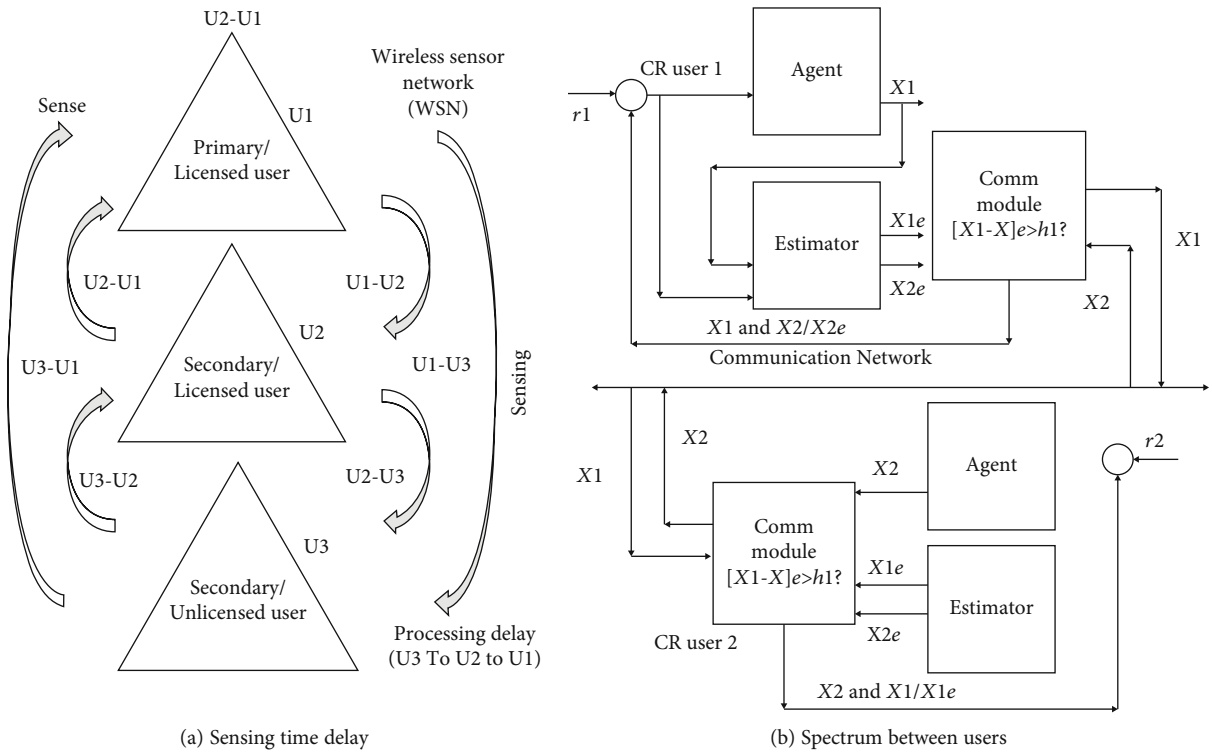
Figure 2(b) shows a standard multiuser closed loop representing power control in this case. Standard discrete time is considered for control system design. For a multi-input multioutput system called the MIMO system with linear time-invariant known as LTI, there are z input, w inputs, and h states.

$$\begin{aligned} Z(q+1) &= X_A(l) + Qv(l) + W(l)m, \\ C(q) &= K_A(q) + H(q)z, \end{aligned} \tag{15}$$

where "q" represents the time index linked with T time sampling in the discrete time domain. $X \in , K \in \ddot{R}^{z*n}$ and $\ddot{R}^{n*n}$, $Q \in \ddot{R}^{n*m}$ are output matrices and system input, and $W(.)$ and $H(.)$ are system measurements and disturbances, respectively. $U(.) = [u1(.), \cdots, uf(.)]^T$ and $Y(.) = [y1(.), \cdots, yz(.)]^T$ are expected to be barred, for example, $|ui(.)| \leq g_{iu}$, $i = 1, \cdots, f$ and $|yj(.)| \leq g_{jv}$, $j = 1, \cdots, z$, in which $g_{iu}{}^{'}s$ and $g_{jv}{}^{'}s$ are positive constant. For simplicity, all states are considered measurable, i.e., $K = I$. There is a possibility to design a

Initialize D with capacity O
Initialize network $Q(m, a, \vartheta)$ with variable weights $\vartheta = \vartheta_0$
Initialize $\mathcal{O}_a$ and $\mathcal{O}_b$ and then obtain m (1)
For $r = 1$, r do
Update $\mathcal{O}_a(r+1)$ via power control policy of PU (2) or (3)
With iterations ε_k, choose an arbitrary action $a(k)$; otherwise, select $a(k) = \max_a Q(m(r), a; \vartheta_0$
Obtain $m(r+1)$ via arbitrary model (5) and detect reward $n(r)$
Store transition $d(r) := \{m(r), a(r), m(r), m(r+1)\}$, in D
Sensing delay
Repeat sensing delay
If $r \geq O$ then
Sample random minibatch of iterations from $D\{d(i)i \in \Omega_r\}$,
Here, the Ω_r index is uniformly selected at independent
Minimize loss function of 12, in which goal is given by (15) $Q'(i)$
Adjust $\vartheta_0 = \arg\ \min_\vartheta E_\vartheta$
End if
$m(r)$ is the target state and then initialize $\partial_a(r+1)$ and $\partial_b(r+1)$ and then gain $m(r+1)$
end if
end for

Algorithm 1: Power control policy-deep learning training based.



(a) Sensing time delay

(b) Spectrum between users

Figure 3: Sensing delay (sensing time) and spectrum sharing between users.

feedback controller at state (1) through maximum input maximum output, which is shown as follows.

$$V(M) = M_k(q)m,$$
$$M[(M(q) - A(C) + H(q)]m. \tag{16}$$

That is, Figure 3(b) shows the spectrum sharing of the control system, which is well designed. So, the stability and performance of the system (z) could be definite through an approximate selection of V sampling time in Equation (16) and K as gain feedback design in Equation (15). A dynamic system is to be supposed (Equation (15)), there is a designed controller that works as a designed framework baseline.

In this paper, there is a consideration that the PU and SU update their power of transmission synchronously. However, we would like to highlight that our proposed scheme's synchronous assumption is not a must. Consider that there is

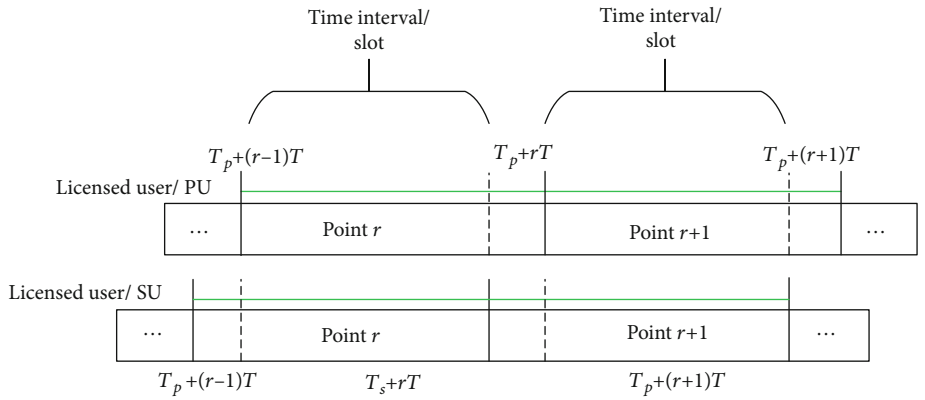


Figure 4: Licensed and unlicensed users updated transmit power.

no strict synchronous occurrence in the time frame between PU and SU; both synchronize transmit power at the start of a timeframe. Figure 4 represents the transmit power of a licensed and unlicensed user. PU adjusts its transmit power at time $t_p, t_p + T, t_p + 2T \cdots$, whereas SU's transmit power is updated at $t_s, t_s + T, t_s + 2T \cdots$, where T is representing the time of each setting. Deprived of loss simplification, we consider that $T > t_p - t_s > 0$. Our scheme, termed intelligent power control scheme, would work in matching through the in-synchronous form if both PU and SU achieve their respective goal. The essential info for PU can be $\mathrm{SNR}_c(r)$, and for SU, it can be $\mathrm{SNR}_c(r)$ and $\mathrm{SNR}_d(r)$, while $m(r)$ is the decision made during the time window $[t_p + (r-1)T.t_s + rt]$.

2.2. Delay Performance Using a Primary Sensor Network. A promising approach for CRN design can be the use of transmission data with the quality of services in wireless sensor networks. Designing a proper physical CRN, with effective and efficient spectrum allocation, sensing, and minimum potential interference, poses several challenges. Since license-free spectra are crowded in CRN, they are affected by uncontrolled interference, and hence, real-time traffic is recommended with the CRNs for the future. Figure 3(a) presents the sensing delay that occurred in the sensing network. It can be seen that there are several SUs on a single PU node. The sensing delay for a U2-U1 connection is less than that for a U3-U2-U1 connection. This is due to the extra sensing delay to sense and get information for an extra link.

The direct U3-U1 link is not obvious for the system. It checks the optimal link for U3 to connect with U1, and during that process, it depends on the response from U2. Similarly, the sensing delay will increase when there are several users on the primary node, and hence, information may be late.

Figure 3(b) represents a network of two cognitive radio users with a communication module. Here, we supposed that CR users use an estimated state for action and communicate its present state to the other user in case of an unsuccessful state. Both estimator 1 and estimator 2 estimate their corresponding $X1e$ and $X2e$. In a normal situation, there is no need for communication. CR user 1 is working based on its

Initialize $\mathcal{O}_b(1)$ and then obtain m (1)
for $r = 1, r$ do
Sensing delay
choose $a(k) = \max_a Q(m(r), a; \vartheta_0$
obtain $m(r+1)$
end for

Algorithm 2: Control power policy (DQN based).

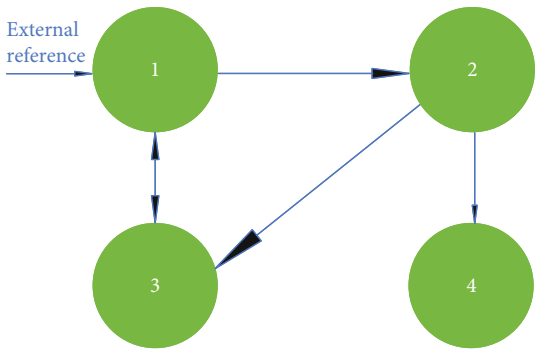


Figure 5: CR users in CRN.

own state $X1$, and CR user 2 estimated state $X2e$. When $X2$ is received from CR user 2, comm module 1 broadcasts X1 to CR user 2 if $|X1 - X1e| > h1$, where $h1$ is the defined threshold. Similarly, a communication mechanism and estimation are designed at CR user 2 and other CR users.

2.3. CRN Consensus Criteria under Communication Delay. Let transfer of message by the nearest agent m be received by agent n with a sensing delay T, which is similar to a network with a sensing delay of fixed one-hop communication. The algorithm of such consensus is given in the equation below.

$$\chi_m(t) = \sum_{n \varepsilon Z_m} h_{mn} \left(\chi_m \left(\tau - \dot{T} \right) - \chi_m \left(\tau - \dot{T} \right) \right) m. \quad (17)$$

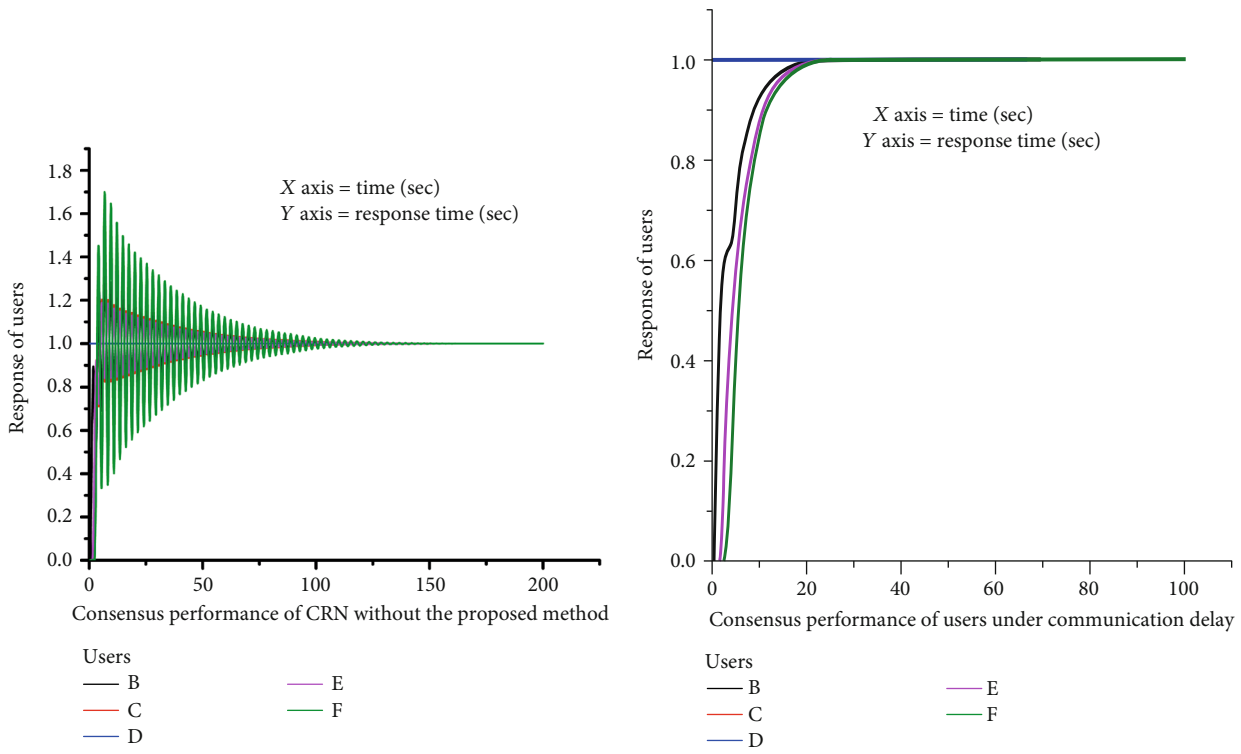


Figure 6: Consensus performance of CRN without the proposed method and with communication delay.

For getting an average consensus, a weight undirected graph G has been supposed in Equation (17). Therefore, the algorithm is

$$\chi_m(t) = \sum_{n\varepsilon Z_m} h_{mn}\left(\chi_n\left(\tau - \dot{T}\right) - \chi_m(\tau)\right)m. \tag{18}$$

The following form shows the collective dynamic of the network:

$$\chi(t) = -L_x\left(t - \dot{T}\right)m. \tag{19}$$

Taking Laplace to transform,

$$A(s) = \frac{W(s)}{s}\chi(0)m. \tag{20}$$

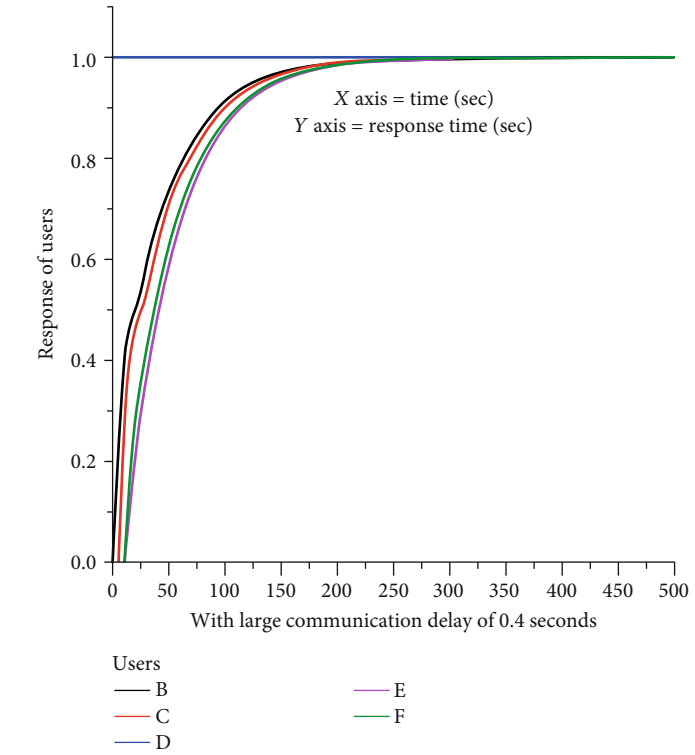
Here, function $W(s) = ((I_n + (I/s))\exp(-sT)L)^{-1}$. For stability verification of $H(s)$, the Nyquist criterion can be used. We can refer to the study as an approach that is similar in [40]. There is an upper bound sensing delay method for network stability that can maintain during time delay. The Equation (10) algorithm clarifies the problem of average consensus asymptotically for all early states or levels, if and only if $0 \le \dot{T} < \pi/2\gamma n$, where $\gamma n < 2\underline{\Delta}$. Refer to [36] for proof. A sufficient condition for the average consensus algorithm (17) is $\tau < \pi/4\underline{\Delta}$. There may be a trade-off between sturdiness to sensing time delay and having a large max degree. A network having a great degree is normally considered a free-scale network. To study [41], small networks and random graphs are impartially active for sensing delay, in the absence of several degrees. consensus achievement of a buildup engineering network is a good example.

2.4. Simulation. Suppose four CR users in CRN after their traffic management; the communication topology in their information sharing is shown in Figure 5.

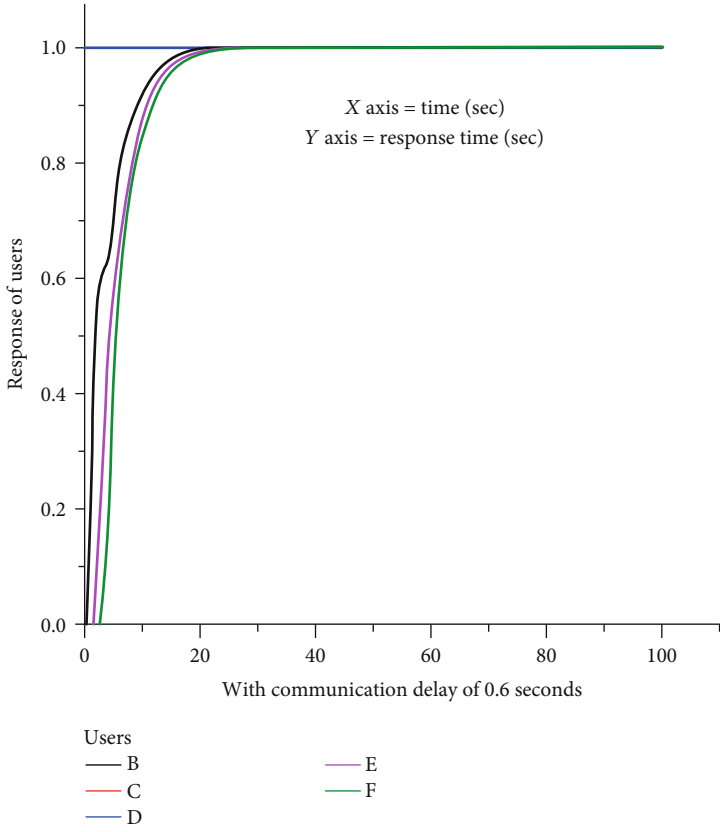
If we consider that all users have integrator dynamics, then by using $m_x = -L_x$, in which values of m changes regarding network dynamics $m_x(t)$, where $L = L_{xy}$ is the graph Laplacian of network, the performance without delay is shown in Figure 6(a). When delay has seemed to be in the network, it affects the performance of the CRN. Hence, consensus performance of the CRN could be seen in Figure 6(b) after communication delay without the proposed method. Although these responses are in consensus, their transient responses are very poor. Hence, the settling time is very large. So, for better performance, the designer should follow this proposed method to minimize the effects of time delay and sensing delay as well. Let a communication delay of 0.4 second between all users be the same; then, their consensus performance can be calculated by using the algorithm shown in Figure 7(a). This proposed method is also sufficient for large arbitrary communication delay of 0.4 sec. Thus, the consensus performance of the CR users with a large communication delay of 0.6 sec is shown in Figure 7(b).

3. Results and Discussion

The transmit power (in watts) for PU and SU is selected from a predefined set $P_c = P_d = \{1/20, 0.1/1, \cdots, 0.4/1\}$, and the

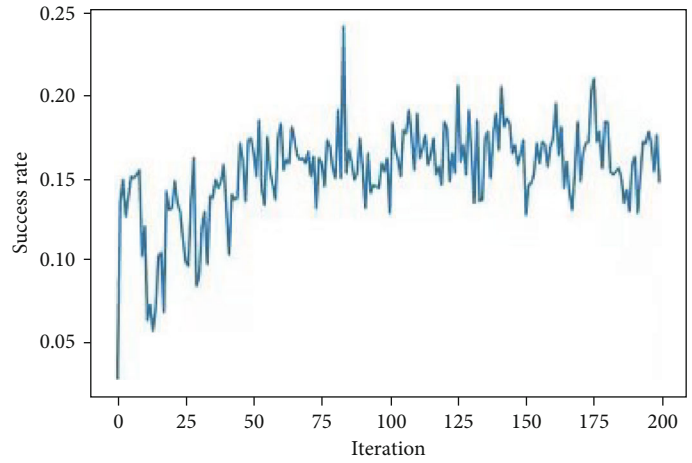


(a) With large communication delay of 0.4 second

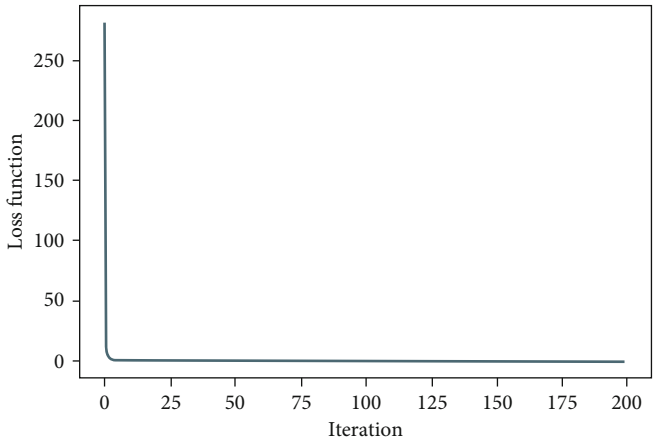


(b) With communication delay of 0.6 sec

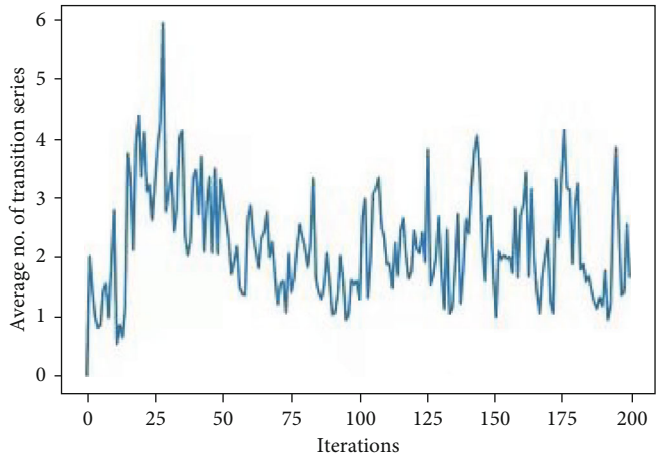
Figure 7: Communication delay in CRN.



(a) Success rate vs. no. of iterations



(b) Loss functions vs. no. of iterations



(c) Average no. of transition series vs. no. of iterations

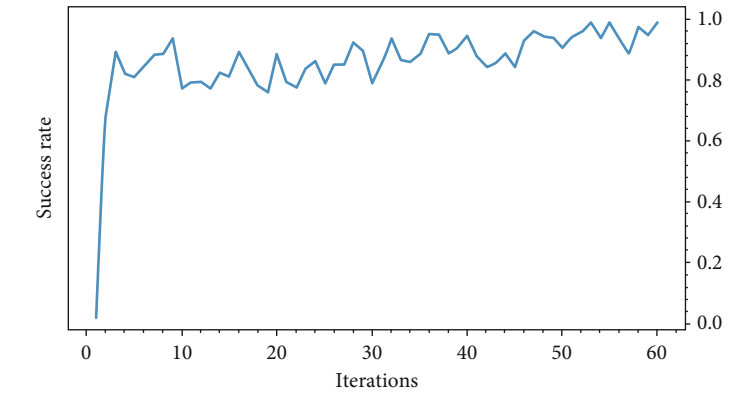
Figure 8: Success rate, loss function, and average number of transitions vs. iterations using no. of sensors (3).

noise power N_1 and N_2 at P_{R1} and S_{R2}, respectively, are set to 0.01 W. For simplicity, channel advancement from the primary transmitter and secondary transmitter to P_{R1} and S_{R2} is considered to be $h_{xy} = 1, \forall x, y$. To achieve the lowest SNR (signal-to-noise ratio), productive reception for PU and SU is adjusted to $\acute{\omega}_c = 11/9$ and $\acute{\omega}_d = 7/10$, respectively. The presence of transmit power $\{\partial_a, \partial_b\}$ can be easily noted, which means that the quality of the PU-SU system is satisfactory. Additionally, several active sensors remain N, which would accept the acknowledged signal strength information to help SU to adapt to the power control policy. There is uniformly distributed distance d_{xy} between the sensor node and T_{xi} transmitter in a pause range of [100,300] in meters.

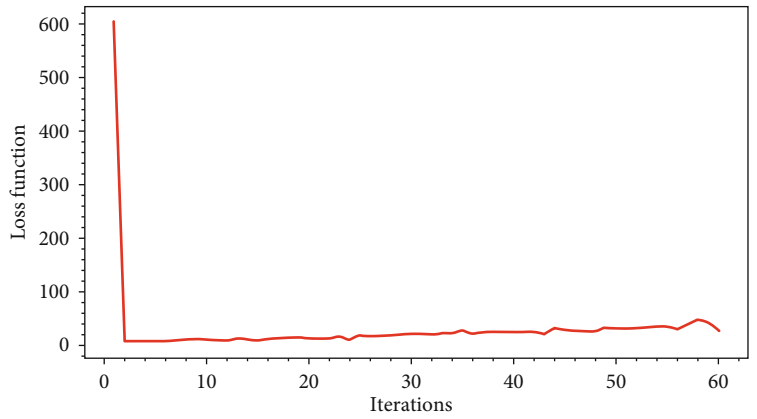
For the action value approximation function, we used the deep neural network (DNN), which contains feedforward layers. The number of these layers is three and is entirely linked. These hidden layers contain neurons 256, 256, and 512, respectively. For the 1st and 2nd layers, rectified linear units are working as an action function, where the ReLU layer output will be 0 or raw for negative input. The activation function is used in the case of the 3rd layer, and the Adam algorithm is used to update weight φ, while the mini-batch size is set to 256. Here, it is considered that rerun memory K holds$N_K = 400$. In each iteration, training of φ starts, and 300 evolutions/transitions are considered for each K. The number of transitions is adjusted to $r = 10^5$. Exploring the probability of a new-fangled action, action reduces with an increase in the number of iterations, and as the iterations increase to a sufficiently large value, it reduces from 0.8 to 0. Additionally, at reiteration r, it is supposed that $\underline{\epsilon}_r = 0.8(1 - r)/K$.

There is access to deep reinforcement learning in our article, specifically, success rate, loss function, and average number. The transition step is being used as transition number function r has been analysed. Here, the presentation is judged through two metrics, which are transition steps and average number. Success rate computation is in standings of the figure of successful trails to total figure of easily turn's a ratio. Here, assume a productive trail if "m" transfers to a target in 20 time frames. In the case of a productive trail, the average number of time frames needs to reach a target which termed as transition step average number.

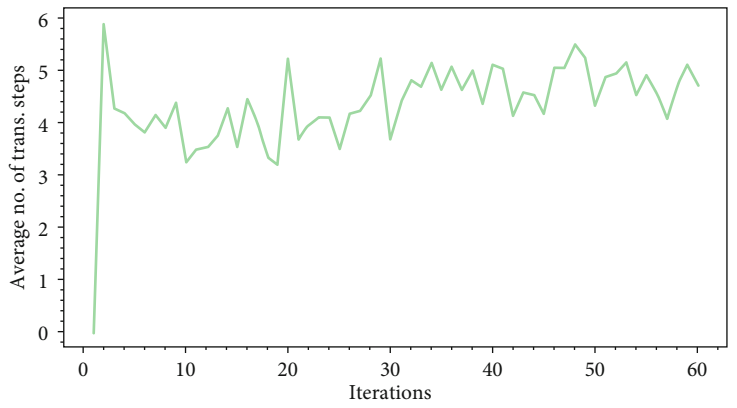
During the exercise process, loss function can be calculated as presented in [14]. After learning iteration r, SU



(a) Success rate vs. iterations



(b) Loss functions vs. iterations

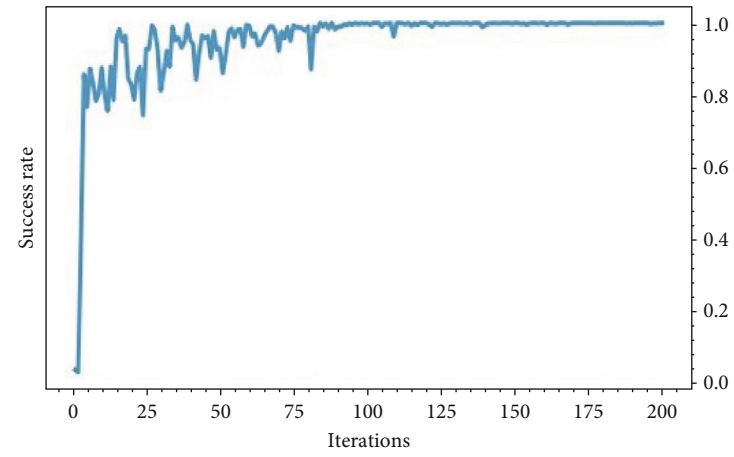


(c) Average no. of transition series vs. iterations

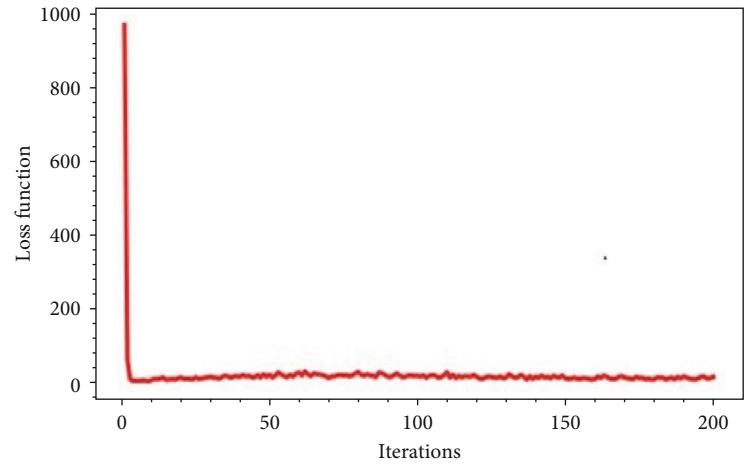
Figure 9: Success rate, loss function, and average number of transitions vs. iterations using no. of sensors (5) and iterations (10000).

may use the trained network for interaction with PU. Network performance is determined standings of an average number of transitions and success rates. Outcomes are average over 10^3 freely turns, where the chance start state is chosen for each track. The numerical results have been divided into two cases.

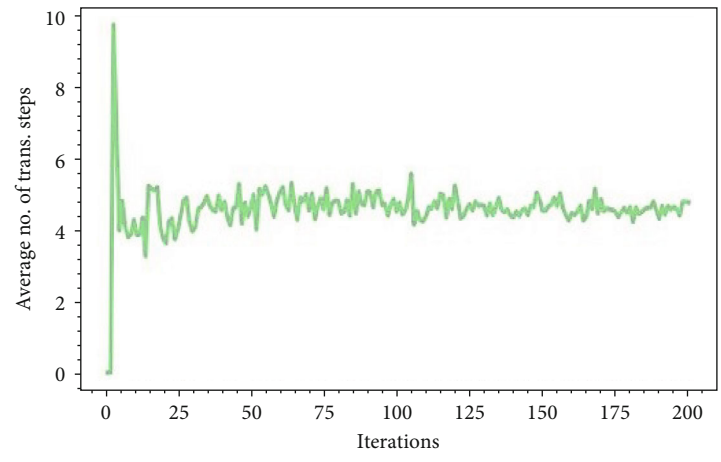
3.1. Case 1 (without Delay). Case one shows that by the guidance of learned power control policy, transmit power of SU/unlicensed user can intelligently be adjusted like final target may be touched since somewhat preliminary target indoors rare figure of transmission steps. In Figure 8(a) (success rate vs. no. of iterations for case 1), the number of sensors used is 3, in which we set according to the number of iterations k, $\sigma_n = (\partial_a^{\sigma} J_{cn} + \partial_b^m) J_{dn}/3$. With the increase in σ_n to reduce the number of sensors, the loss function value becomes larger. It is good policy to achieve the average number of transition steps and loss function vs. no. of iterations similar to those shown in Figures 8(b) and 8(c). Figure 8 reveals the strength of the deep reinforcement learning approach.



(a)



(b)



(c)

Figure 10: Success rate, loss function, and an average number of transitions vs. iterations using several sensors (10) and iterations (30000).

3.2. Case 2 (with Delay). Case 2 shows the sensing time delay (r+k) as shown in Figure 2. This sensing time delay has been found in the SU policy as in Equation (5). This sensing time delay occurs when PU shares transmit power or spectrum with SU; then, there is the penetration of SUs on the primary node that causes sensing delay during the sharing of transmit power or spectrum. In this case, power control performance reduces. Figure 9(a) shows the success rate vs. no. of iterations. Figure 9(b) shows the loss function vs. no. of iterations, and Figure 9(c) shows the average no. of transitions vs. no. of iterations. We showed the recital of the DQN-based power control method in which a secondary power control policy

is employed by the PU as in Equation (4) which seems additionally traditional. The target of learning a suitable power control policy is in Figure 9 which shows the loss function, success rate, and an average number of transition steps vs. k. Here, we used the number of sensors which is 10. $\sigma_n = (\partial_a^{\sigma} J_{cn} + \partial_b^m) J_{dn}/10$. It is to be observed that a greater number of iterations are required for getting a success rate. In our policy, the average number of transition steps is 2.5 for achieving the target state. Here, the PU used a second policy which only permits transmit power for an increase or decrease in the signal level at each step. For this purpose, we need more steps. The explained method sustains a supposed performance loss in the case of fewer sensors deployed; this happened due to random variations in observation states which built various conditions less unique from one another and save the agents from learning of an active strategy. Random variation effect can be neutralized by the use of more sensors.

Figure 10(a) shows the success rate vs. iterations, Figure 10(b) shows the loss function vs. iterations, and Figure 10(c) shows the average no. of transition steps vs. no. of iterations. For the purpose of training, we used a number of sensors (10). $\sigma_n = (\partial_a^{\sigma} J_{cn} + \partial_b^m) J_{dn}/10$. In this sample, the plot is run with 30000 iterations and a delay of 2 samples. Here, in the success rate plot, the effect of delay is clear. This sensing delay is because of the penetration of SUs on the primary node during the sharing of spectrum. In this case, power control performance will reduce.

Figure 6(a) shows the consensus performance of CRN showing results without the proposed model sensing delay. However, these are poor transit responses of consensus performance. All agents reached their steady-state value after nearly 100 seconds. However, the settling time is large. So, for progressing performance, one should track this projected technique to lessen the effects of communication and sensing time delay as well. If there is the consideration of communication and sensing delay amid entirely CR operators, suppose that sensing delay of 0.4 second inside whole operations is identical. The assumed method is enough for the representation of large-scale sensing delay in the CR network. Figures 7(a) and (7)(b) show the communication delay of 0.4 sec and 0.6 sec, respectively, in a CRN.

4. Conclusions

Machine learning and CR intelligence have a good capability to understand and adapt to the wireless environment. In wireless communication, machine learning techniques are linked with CR technology. The cognitive radio system which consists of PU and SU is well studied; in this paper, we explained the problem of spectrum sharing of PU and SU in CRNs, and also, we introduced sensing delay in communication. Because there was a common concept that there is no cooperation between PU and SU, PUs adjust their transmit powers on their own transmit control power policy. In this article, we discussed the consensus performance and introduced a Q-learning or deep reinforcement learning method for SU to study for the adjustment of its transmit power so that ultimately both the PU and SU have the ability to transmit respective data fruitfully with the essential qualities of services. Our numerical results showed that the considered learning method is healthy against the random variations in the state variation and within few numbers of steps. We can get our target from an initial state, and also, in numerical results, we showed sensing delay that is because of the number of SUs when PU transmits spectrum with the SU; then, there is a rush of SU on the primary node that caused sensing delay; it is because of a lot of several users. All agents reached their steady-state value after nearly 100 seconds. However, the settling time is large. The given method is enough for the representation of large-scale sensing delay in the CR network.

Data Availability

Data can be provided on request.

Conflicts of Interest

We hereby confirm that there is no conflict of interest between authors to declare.

Acknowledgments

This work was supported by the Scientific and Technological Innovation Foundation of Shunde Graduate School, University of Science and Technology Beijing (the major project no. is BK19CF002) of China.

References

[1] Z. Qin, X. Zhou, L. Zhang, Y. Gao, Y. Liang, and G. Y. Li, "20 years of evolution from cognitive to intelligent communications," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 6–20, 2020.

[2] Z. Chu, W. Hao, P. Xiao, and J. Shi, "Intelligent reflecting surface aided multi-antenna secure transmission," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 108–112, 2019.

[3] M. Di Renzo, K. Ntontin, J. Song et al., "Reconfigurable intelligent surfaces vs. relaying: differences, similarities, and performance comparison," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 798–807, 2020.

[4] Z. Ahmed, M. M. Khan, M. A. Saeed, and W. Zhang, "Consensus control of multi-agent systems with input and communication delay: a frequency domain perspective," *ISA Transactions*, vol. 101, pp. 69–77, 2020.

[5] N. Tadayon and S. Aïssa, "A multichannel spectrum sensing fusion mechanism for cognitive radio networks: design and application to IEEE 802.22 WRANs," *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 4, pp. 359–371, 2015.

[6] J. Adu Ansere, G. Han, H. Wang, C. Choi, and C. Wu, "A reliable energy efficient dynamic spectrum sensing for cognitive radio IoT networks," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6748–6759, 2019.

[7] A. Kumar and K. Kumar, "Multiple access schemes for cognitive radio networks: a survey," *Physical Communication*, vol. 38, article 100953, 2020.

[8] M. A. Al-Jarrah, A. Al-Dweik, S. S. Ikki, and E. Alsusa, "Spectrum-occupancy aware cooperative spectrum sensing using

adaptive detection," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2225–2236, 2020.

[9] Y. R. Kondareddy, P. Agrawal, and K. Sivalingam, "Cognitive radio network setup without a common control channel," in *MILCOM 2008 - 2008 IEEE Military Communications Conference*, pp. 1–6, San Diego, CA, 2008.

[10] A. A. Khan, M. H. Rehmani, and M. Reisslein, "Cognitive radio for smart grids: survey of architectures, spectrum sensing mechanisms, and networking protocols," *IEEE Communication Surveys and Tutorials*, vol. 18, no. 1, pp. 860–898, 2016.

[11] I. Mitliagkas, N. D. Sidiropoulos, and A. Swami, "Joint power and admission control for ad-hoc and cognitive underlay networks: convex approximation and distributed implementation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4110–4121, 2011.

[12] Z. Xiao, X. Shen, F. Zeng et al., "Spectrum resource sharing in heterogeneous vehicular networks: a noncooperative game-theoretic approach with correlated equilibrium," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9449–9458, 2018.

[13] M. H. Islam, Y. Liang, and A. T. Hoang, "Distributed power and admission control for cognitive radio networks using antenna arrays," in *2007 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 250–253, Dublin, 2007.

[14] H. V. Vu, N. H. Tran, and T. Le-Ngoc, "Full-duplex device-to-device cellular networks: power control and performance analysis," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3952–3966, 2019.

[15] M. M. Aslam, L. Du, Z. Ahmed, H. Azeem, and M. Ikram, "Consensus performance of traffic management system for cognitive radio network: an agent control approach," in *Cyberspace Data and Intelligence, and Cyber-Living, Syndrome, and Health*, vol. 1, Springer, Singapore, 2019.

[16] Y. Kuo, J. Yang, and J. Chen, "Efficient swarm intelligent algorithm for power control game in cognitive radio networks," *IET Communications*, vol. 7, no. 11, pp. 1089–1098, 2013.

[17] S. Tomic, M. Beko, and R. Dinis, "RSS-based localization in wireless sensor networks using convex relaxation: noncooperative and cooperative schemes," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 2037–2050, 2015.

[18] J. Gummeson, D. Ganesan, M. D. Corner, and P. Shenoy, "An adaptive link layer for range diversity in multi-radio mobile sensor networks," in *IEEE INFOCOM 2009*, pp. 154–162, Rio De Janeiro, Brazil, 2009.

[19] M. M. Aslam, M. N. Irshad, H. Azeem, and M. Ikram, "Cost effective & energy efficient intelligent smart home system based on IoT," *Afyon Kocatepe University International Journal of Engineering Technology and Applied Sciences*, vol. 3, pp. 10–20, 2020.

[20] F. Azmat, Y. Chen, and N. Stocks, "Analysis of spectrum occupancy using machine learning algorithms," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 6853–6860, 2016.

[21] N. Sizochenko, M. Syzochenko, N. Fjodorova, B. Rasulev, and J. Leszczynski, "Evaluating genotoxicity of metal oxide nanoparticles: application of advanced supervised and unsupervised machine learning techniques," *Ecotoxicology and Environmental Safety*, vol. 185, article 109733, 2019.

[22] G. Ciaparrone, F. L. Sánchez, S. Tabik, L. Troiano, R. Tagliaferri, and F. Herrera, "Deep learning in video multi-object tracking: a survey," *Neurocomputing*, vol. 381, pp. 61–88, 2020.

[23] X. Zhou, M. Sun, G. Y. Li, and B. H. Fred Juang, "Intelligent wireless communications enabled by cognitive radio and machine learning," *China Communications*, vol. 15, no. 12, pp. 16–48, 2018.

[24] R. Barto and A. G. Sutton, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, 1998.

[25] G. Ding, Y. Jiao, J. Wang et al., "Spectrum inference in cognitive radio networks: algorithms and applications," *IEEE Communication Surveys and Tutorials*, vol. 20, no. 1, pp. 150–182, 2018.

[26] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, part II: multiuser networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 6, pp. 2214–2222, 2007.

[27] Z. Liu, M. Zhao, K. Y. Chan, Y. Liu, and K. Ma, "Resource allocation strategy against selfishness in cognitive radio ad-hoc network based on Stackelberg game," *IET Communications*, vol. 13, no. 13, pp. 1962–1970, 2019.

[28] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961-1976," *IEEE Transactions on Information Theory*, vol. 23, no. 1, pp. 1–37, 1977.

[29] H. Zhang, N. Yang, K. Long, M. Pan, G. K. Karagiannidis, and V. C. M. Leung, "Secure communications in NOMA system: subcarrier assignment and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 7, pp. 1441–1452, 2018.

[30] Q. Yu, J. Chen, Y. Fan, X. Shen, and Y. Sun, "Multi-channel assignment in wireless sensor networks: a game theoretic approach," in *2010 Proceedings IEEE INFOCOM*, pp. 1–9, San Diego, CA, 2010.

[31] F. Fang, H. Zhang, J. Cheng, and V. C. M. Leung, "Energy-efficient resource allocation for downlink non-orthogonal multiple access network," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3722–3732, 2016.

[32] V. Mnih et al., "Playing Atari with deep reinforcement learning," pp. 1–9, 2013, https://arxiv.org/abs/1312.5602.

[33] Y. Duan, X. Chen, R. Houthooft, J. Schulman, and P. Abbeel, "Benchmarking deep reinforcement learning for continuous control," *International Conference on Machine Learning, PMLR*, vol. 48, pp. 1329–1338, 2016.

[34] H. Zhang, N. Yang, W. Huangfu, K. Long, and V. C. M. Leung, "Power control based on deep reinforcement learning for spectrum sharing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 4209–4219, 2020.

[35] F. Hu, B. Chen, X. Zhai, and C. Zhu, "Channel selection policy in multi-SU and multi-PU cognitive radio networks with energy harvesting for Internet of everything," *Mobile Information Systems*, vol. 2016, 12 pages, 2016.

[36] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[37] B. Givan and R. Parr, "Introduction POMDP.pdf," pp. 1–23, 2001.

[38] Z. Ahmed, M. A. Saeed, A. Jenabzadeh, and Z. Weidong, "Frequency domain analysis of resilient consensus in multi-agent systems subject to an integrity attack," *ISA Transactions*, 2020.

[39] D. Choi, "Model-based-RL-deepmind," pp. 1–13, 2016.

[40] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465–1476, 2004.

[41] P. Erdos and A. Rényi, "On the evolution of random graphs," *Structure in Dynamic Networks*, vol. 9781400841, pp. 38–82, 2011.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 8898949, 16 pages
https://doi.org/10.1155/2020/8898949

Research Article

Behavior Anomaly Detection in SDN Control Plane: A Case Study of Topology Discovery Attacks

Li-Der Chou,[1] **Chien-Chang Liu**,[1] **Meng-Sheng Lai**,[1] **Kai-Cheng Chiu**,[1] **Hsuan-Hao Tu**,[1] **Sen Su**,[2] **Chun-Lin Lai**,[2] **Chia-Kuan Yen**,[2] **and Wei-Hsiang Tsai**[2]

[1]*Department of Computer Science and Information Engineering, National Central University, Taoyuan 32001, Taiwan*
[2]*Information & Communication Research Division, National Chung-Shan Institute of Science and Technology, Taoyuan 32546, Taiwan*

Correspondence should be addressed to Li-Der Chou; cld@csie.ncu.edu.tw

Received 22 July 2020; Revised 13 September 2020; Accepted 2 November 2020; Published 21 November 2020

Academic Editor: Juraj Gazda

Software-defined networking controllers use the OpenFlow discovery protocol (OFDP) to collect network topology status. The OFDP detects the link between switches by generating link layer discovery protocol (LLDP) packets. However, OFDP is not a security protocol. Attackers can use it to perform topology discovery via injection, man-in-the-middle, and flooding attacks to confuse the network topology. This study proposes a correlation-based topology anomaly detection mechanism. Spearman's rank correlation is used to analyze the network traffic between links and measure the round-trip time of each LLDP frame to determine whether a topology discovery via man-in-the-middle attack exists. This study also adds a dynamic authentication key and counting mechanism in the LLDP frame to prevent attackers from using topology discovery via injection attack to generate fake links and topology discovery via flooding attack to cause network routing or switching abnormalities.

1. Introduction

In recent years, with the rapid development of smart devices, mobile devices, and network technologies, the number of network devices has exploded, and numerous network services have emerged. Network services connect many devices to the Internet; thus, users' real-time processing capabilities and diversified service requirements for network services have also increased. However, traditional networks use a decentralized architecture, and network packets are difficult to monitor and manage. Therefore, a software-defined networking (SDN) architecture based on centralized management was proposed [1].

The SDN architecture separates the control layer of the traditional network device from the data layer and manages the routing and forwarding of network packets through one or several network controllers. SDN controllers can adjust network rules for centralized management with the topology connection status of the entire network. SDN integrates the link layer discovery protocol (LLDP) [2], which is called the OpenFlow discovery protocol (OFDP) [3], for topology discovery. The controller generates LLDP packets in a specific format to detect links between switches, and it maintains network topology information to facilitate optimal routing decisions for each service.

However, most SDN controllers perform topology discovery through OFDP without security mechanisms, such as RYU [4], Floodlight [5], OpenDaylight [6], and POX [7]. Therefore, an attacker can easily confuse the current network topology through a topology discovery attack. SDN controllers do not implement an authentication mechanism on OFDP. Any attacker can forge LLDP packets and send them to the switch to perform a topology discovery attack. The OFDP cannot determine whether the SDN controller has generated LLDP data packets; thus, making judgments becomes more difficult for the SDN controller. Attackers can use topology discovery attacks to generate multiple false links. As a result, the SDN controller cannot determine the routing and switching paths of each service or even forward normal traffic through the attacker's computer.

Related research has no complete detection mechanism for topology discovery attacks. The contributions of the present study can be highlighted as follows:

(1) This study designs and implements a topology anomaly detection mechanism based on correlation analysis, which can detect an abnormal LLDP packet

(2) The mechanism generates an authentication key in the LLDP packets and calculates the amount of received LLDP packets of each switch to avoid topology discovery via injection and flooding attacks

(3) For the topology discovery via man-in-the-middle attack, the mechanism uses Spearman's rank correlation coefficients [8] to analyze the network traffic between links and measure the round-trip time (RTT) of each LLDP packet

The remainder of this paper is organized as follows. Section 2 presents the related works. Section 3 describes the proposed methods. Section 4 presents the experimental results. Section 5 concludes the study.

2. Related Works

2.1. SDN. The characteristic of SDN is that physical equipment is used as a basic resource and is only responsible for data processing. SDN abstracts the control components of network packet processing and provides a unified management and control interface for controlling the upper-level components of packets. Network management engineers can use software programs to replan and define network topology, resource allocation, and flow control mechanisms. In addition to key OpenFlow technology based on this architecture, in recent years, many integrated SDN technologies have been discussed and developed, such as Programming Protocol-independent Packet Processors [9], Stratum [10], and ONOS [11].

The SDN architecture is divided into three layers, namely, infrastructure, control, and application layers. The top application layer includes various unique services and applications. The middle control layer is responsible for the configuration of network resources and network management by the SDN controller. The lowest basic equipment layer processes and forwards data based on protocols, such as OpenFlow [12].

2.2. OpenFlow. OpenFlow is a communication protocol that can access the data plane of a switch or router. The SDN controller [13] can communicate with the switch or router through OpenFlow. The network equipment supporting OpenFlow includes the following two components: (1) a medium called the OpenFlow Channel, which can communicate with the SDN controller, and (2) an OpenFlow flow table for recording network packet processing principles.

Flow table is composed of many flow entries. Each entry defines the actions required for each network flow. Flow entry includes fields, such as match and action fields. The match field is used to identify and classify each corresponding flow, including multiple identification fields in the second to fourth layers of the seven layers of the network OSI. The action field defines the actions that must be performed after each packet is matched, including forward to port, drop, and send to controller.

The OpenFlow switch only needs to compare the rules of the flow table to forward the data and communicate with the SDN controller through OpenFlow Packet-In and Packet-Out messages. The OpenFlow switch compares each received packet to the flow table. If a corresponding rule is found, then the packet is modified or forwarded according to the rule. If no corresponding rule can be found, then the received packet is forwarded to the SDN controller, which is called Packet-In. The SDN controller analyzes each packet from the OpenFlow Packet-In and issues rules to the OpenFlow switch for each flow, which is called Packet-Out.

2.3. OFDP. The OpenFlow switch does not have an LLDP mechanism; thus, the SDN controller performs topology discovery through the OFDP. OFDP sends LLDP packets to the switch and disassembles those return packets to obtain the current network topology. The OFDP cycle is different for each controller platform. The topology discovery cycle of Ryu controller is one second [14]. Figure 1 shows the operational flow of the OFDP.

Step 1. The SDN controller inquires the number of ports being used on each switch and generates fresh LLDP packets for each port. The packet includes the switch number and the port number and transmits the LLDP to the designated switch through an OpenFlow Packet-Out message.

Step 2. After receiving the LLDP packet, the switch sends it according to the specified port number.

Step 3. The switch that received the LLDP packet will send it back to the SDN controller through the Packet-In. After the SDN controller disassembles the packet, it can know which of the two switches and of the two ports are connected to each other.

Through the above process, the SDN controller performs the same action for each switch in the network and obtains the topology information of the entire network [15]. To maintain the latest network topology, the SDN controller performs the OFDP regularly.

2.4. LLDP. The LLDP is the second-layer network protocol in the seven-layer architecture of the OSI network [16]. It is an optional protocol in the IEEE 802 LAN protocol. The switch exchanges device and link information with each other through the LLDP protocol; such information includes functions supported by the device, management address, and port status.

Type/length/value (TLV) is the unit that forms the LLDP data unit (LLDPDU). Each TLV represents a message. The switch composes device and link information into different TLVs and encapsulates them into LLDPDUs. The LLDPDU is transmitted through the 802.3 Ethernet frame. Figure 2 shows the LLDP frame [17].

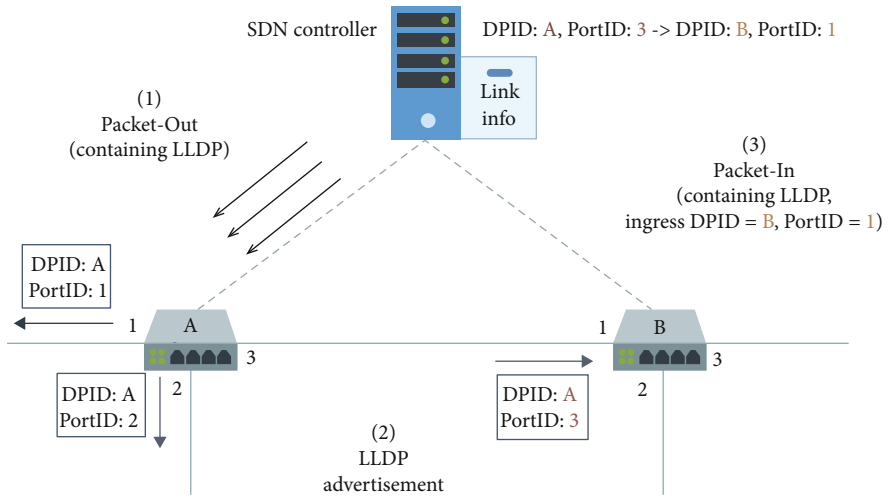


Figure 1: Operational flow of OFDP: encapsulate and send LLDP packets and disassemble them to obtain the current network topology.

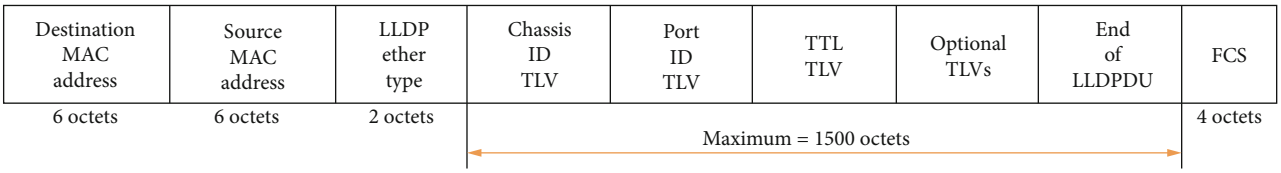


Figure 2: LLDP frame format. The destination MAC address is a fixed set of MAC addresses 01-80-C2-00-00-0E. The source MAC address is the MAC address of the sending LLDP device. The LLDP ether-type code is 88-CC. The chassis ID TLV is the identification code of the switch; port ID TLV is the port number; TTL TLV records the lifetime of this LLDPDU; optional TLV is the selected TLV in the LLDPDU; and end of LLDPDU is used to inform the receiver the body of this LLDP has ended.

The LLDPDU contains the mandatory and optional TLVs. The mandatory and optional fields can add up to 1500 octets. The mandatory TLVs are TLV types 0–3, which are end of LLDPDU, chassis ID, port ID, and time to live. The optional TLVs are TLV types 4–8 and 127, which are port description, system name, system description, system capabilities, management address, and organizationally specific TLVs. TLV types 9–129 are reserved fields and are not currently defined in the standard. Therefore, data types can be defined here to avoid errors in access to network devices using standard LLDPDUs. The timestamp field added to the LLDPDU in this study is also defined in the reserved field.

2.5. Topology Discovery Attacks. The SDN controller collects network topology information through the OFDP. However, the LLDP is not a secure mechanism. Given the lack of authentication messages, attackers can confuse the current network topology status by forging LLDP packets [18]. The three common topology attacks are injection, man-in-the-middle, and flooding attacks [19]. The purpose of the first two is to confuse the current network topology, and that of the third is to consume the computing resources of the SDN controller.

Figure 3(a) shows the topology discovery via injection attack. The attacker fabricates the LLDP packet and fills the chassis ID and port ID in the LLDPDU into the identification codes of other OpenFlow switches on the network. The attacker sends the fabricated LLDP packet to the interface card between the attacker and the switch. After the switch receives the packet, it transmits the packet to the controller through a Packet-In message. The attacker injects a fake link to the topology information because the controller cannot recognize whether the LLDP packet is sent by the controller.

In [20], the optional TLV in the LLDPDU is added to the authentication information. If the fabricated LLDP packet does not contain authentication information, then it will be considered an attack. However, authentication information is a fixed value; thus, the attacker can intercept the LLDP packet sent by the controller to acquire the authentication information and then fabricate the packet containing the authentication field. Thus, in [21], authentication information (KHMAC) is dynamically generated to increase the difficulty for attackers to fabricate the authentication information.

Figure 3(b) shows the topology discovery via man-in-the-middle attack. From the figure, the network has multiple attackers. Instead of spoofing the LLDP packet, Attacker 1 eavesdrops the LLDP packets sent from the SDN controller. Then, Attacker 1 transfers the LLDP packets into other formats, specifically in the pcap format, and transmits it to Attacker 2 who is connected to other switches. After receiving the pcap file, Attacker 2 will convert the format back and send the packet to the switch. After the switch receives the packet, it will transmit the packet to the controller through a Packet-In message. A fake link is added between Attackers 1 and 2. Given that the packet played back is the

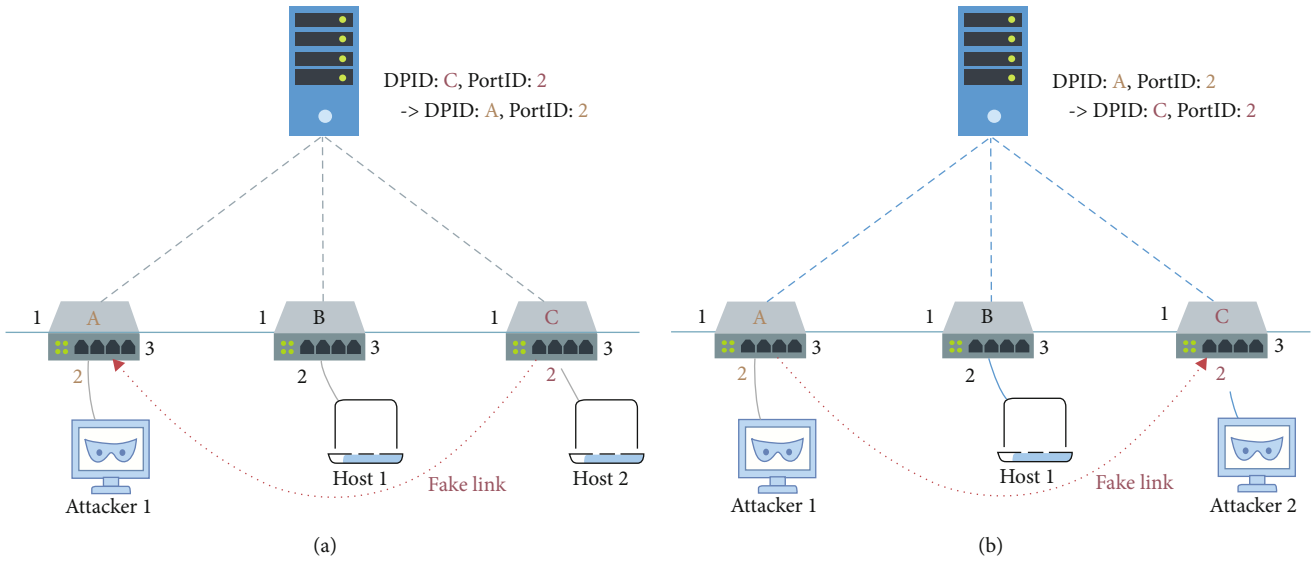


Figure 3: (a) Topology discovery via injection attack. The attacker fabricates the LLDP packet and injects a fake link. (b) Topology discovery via man-in-the-middle attack. Attacker 1 eavesdrops the LLDP packet and transmits it to Attacker 2 to add a fake link between them.

LLDP packet generated by the controller, the authentication field in the LLDPDU cannot be used to avoid the attack.

Literature [22] proposes a statistical analysis of link latencies (SALL) solution to detect topology discovery via man-in-the-middle attack through the latency time of each received LLDP packet. However, if an attacker quickly plays back the LLDP packets, then the latency time will be close to normal, and the attack cannot be detected.

Figure 4 shows the topology discovery via flooding attack. Multiple attackers on the network simultaneously generate multiple LLDP packets and transmit them to the switch. The switch transmits all the received LLDP packets to the controller for analysis, which consumes the computing resources of the SDN controller. Thus, the requests of the normal users cannot be served.

In [23], a secure OpenFlow topology discovery mechanism (sOFDP) is proposed to mitigate the topology discovery via flooding attack by observing the state of the switch. The controller performs OFDP topology discovery only when the connection state of the network changes. Therefore, if the state of the port is unchanged, then even if the attacker sends many LLDP packets, they will be discarded by the switch, and the controller will not be affected. However, the disadvantage of this method is that the attacker can change the state of the port before launching the attack to bypass the detection mechanism. In [24], the secure and lightweight link discovery protocol mechanism (SLDP) is proposed to avoid attacks by creating an eligible port list. Only the LLDP packets received from the port in the list will be accepted. All ports of all switches are on the list when the network starts. However, if an attacker attacks while it is still in the legal list, then the controller cannot avoid being attacked.

Based on the above-mentioned related research, the feature of this study is to propose a behavior anomaly detection mechanism based on correlation analysis and time difference analysis, which is applied to the defense and detection of OpenFlow topology discovery attacks. Compare the four literatures related to the motivations and problems to be solved in this study. The comparison table is shown in Table 1.

2.6. Correlation Coefficient. The correlation coefficient is a statistical technique used to reflect the correlation between two sets of variables. The correlation between the two sets of variables does not fully represent the relationship between them. Common algorithms for calculating the correlation of the variables include Spearman's rank correlation coefficient and Pearson correlation coefficient. Correlation coefficient with absolute values of 0–0.35, 0.36–0.67, and 0.68–1.0 is considered low or weak, modest or moderate, and strong or high correlation, respectively [25].

Spearman's rank correlation coefficient is an algorithm used to calculate the correlation between two sets of variables. It is not affected by outliers in the data, and the calculated data do not need to conform to the normal distribution. However, the data from the two sets of variables must be ordered. The data used to calculate the correlation of traffic between the switches in this study are stored sequentially, and outliers may sometimes exist due to external factors of network users. Thus, this study uses Spearman's rank correlation coefficient. The value of Spearman's rank correlation coefficient is between −1 and 1. A value greater than 0 indicates a positive correlation, and a negative correlation otherwise [26].

3. Proposed Methods

3.1. System Architecture and Design. Figure 5 shows the system architecture in this study, which is divided into three major modules and nine submodules. The three major modules are topology management module, LLDP handling module, and correlation-based topology anomaly detection (CTAD) module. The topology management module and the LLDP handling module detect topology discovery via injection and attacks with verification code information and packet statistical analysis. The CTAD module analyzes the

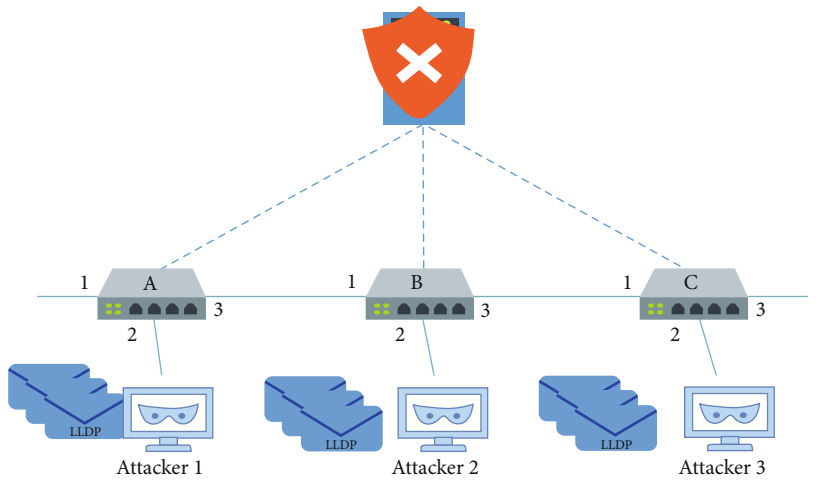


Figure 4: Topology discovery via flooding attack. Multiple attackers simultaneously generate multiple LLDP packets and transmit them to the switch.

Table 1: Related research comparison.

Features	Detection of topology discovery via injection attack	Detection of topology discovery via man-in-the-middle attack	Detection of topology discovery via flooding attack
KHMAC [21]	Supported	Supported	Not supported
SALL [22]	Not supported	Supported	Not supported
sOFDP [23]	Not supported	Not supported	Supported
SLDP [24]	Supported	Not supported	Supported
Proposed	Supported	Supported	Supported

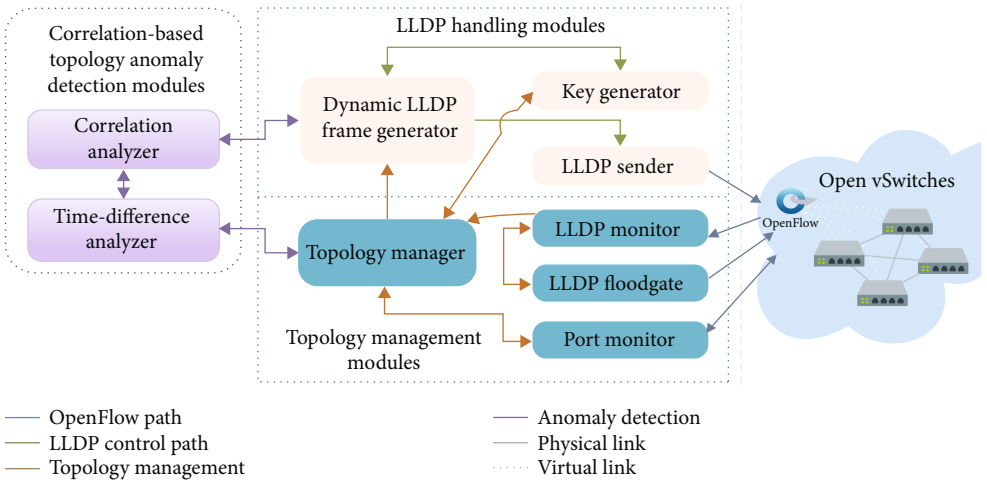


Figure 5: System architecture. It is divided into three major modules and nine submodules.

RTT of LLDP packets and link traffic to detect topology discovery via man-in-the-middle attacks. The nine submodules are topology manager, LLDP monitor, LLDP floodgate, port monitor, dynamic LLDP frame generator, key generator, LLDP sender, correlation analyzer, and time-difference analyzer.

3.1.1. Topology Management Module. This module monitors LLDP packets from switch, records real-time network topology information, and verifies whether a topology discovery was made by injection or flooding attack. This module includes four submodules, namely, topology manager, LLDP monitor, LLDP floodgate, and port monitor.

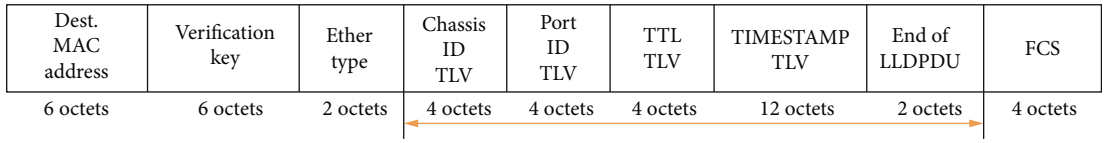


Figure 6: Proposed LLDP packet format. The packet content contains 6 bytes of authentication information (verification key) and 12 bytes of timestamp (TIMESTAMP).

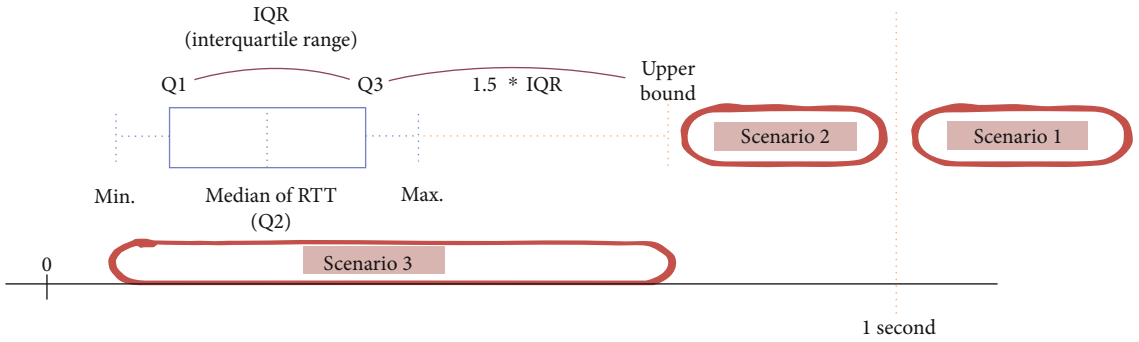


Figure 7: Scenarios of topology discovery via man-in-the-middle attack. Scenario 1: the delay time is greater than one second; scenario 2: the delay time is less than one second, but greater than the upper bound; and scenario 3: the delay time is less than the upper bound.

The topology manager monitors and records the topology status of the entire network by disassembling the LLDP collected by the LLDP monitor. In this manner, the link status between the switches can be obtained.

The LLDP monitor is responsible for parsing whether a Packet-In message received from a switch belongs to the LLDP packet. The LLDP packet is transmitted to the topology manager for further analysis. The LLDP monitor maintains an LLDP packet that receives the quantity table. After each topology discovery period ends, it calculates the number of LLDP packets to determine if any port suffers from topology discovery via flooding attack. Then, this module notifies the LLDP floodgate to block the attacker's network traffic.

The LLDP floodgate controls the flow of LLDP packets on the switch. When the LLDP monitor detects a topology discovery via flooding attack, the LLDP floodgate receives the chassis ID and port ID from the LLDP monitor. This module will perform the flow modification to the related switch through OpenFlow, and the switch will block all LLDP packets from this port.

The port monitor collects the number of switch ports and transmits the information to the topology manager for topology discovery. Moreover, it collects the current network traffic of each port. The correlation analyzer module will notify the port monitor to collect the network traffic of the specified port.

3.1.2. LLDP Handling Module. This module makes LLDP packets in a specific format, generates verification information for the LLDP packets, encapsulates the verification information into the LLDP packets, and sends the LLDP packets to a specific switch. This module includes three submodules, namely, dynamic LLDP frame generator, key generator, and LLDP sender.

The dynamic LLDP frame generator generates LLDP packets of different sizes and encapsulates LLDP verification information into the packets. The dynamic LLDP frame generator receives a list of ports from the topology manager at the beginning of each topology discovery cycle and generates LLDP packets in a specific format according to each port in the list. The packet contains 6 bytes of verification key and 12 bytes of timestamp, as shown in Figure 6. To verify whether a topology discovery via man-in-the-middle attack occurs, the dynamic LLDP frame generator generates 1500 MTU LLDP packets and sends numerous LLDP packets at the specified port to verify the correlation among different link traffic rates.

The key generator generates a set of random MAC addresses at the beginning of each topology discovery cycle to verify whether each LLDP packet returned from the switch is a topology discovery via injection attack. The key generator computes a set of MAC addresses through the current timestamp. The MAC addresses are encapsulated by the dynamic LLDP frame generator into the source MAC address field.

The LLDP sender propagates the LLDP packet encapsulated by the dynamic LLDP frame generator to the specified switch by an OpenFlow Packet-Out message.

3.1.3. CTAD Module. This module measures the time difference of the RTT of each LLDP packet and analyzes the correlation of network traffic on each link to determine a topology discovery via man-in-the-middle attack. This module includes two submodules, namely, correlation and time-difference analyzer.

Figure 7 illustrates that topology discovery via man-in-the-middle attacks can be divided into three scenarios. The first scenario is that the delay for the attacker to replay LLDP packets is greater than one second. In this attack scenario, the

SDN controller adopts a new set of authentication keys for the new topology discovery cycle because the attack delay time is greater than one second. Thus, the attack cannot take effect. The second scenario is that the delay time for the attacker to replay LLDP packets is greater than the upper quartile calculated by the box-and-whisker plot [27] plus 1.5 times interquartile range (IQR), but less than one second. The LLDP packet delay time in this attack scenario is greater than the usual value; thus, it will be judged as abnormal. The third scenario is that the delay time for the attacker to replay LLDP packets is less than the upper bound calculated by the box-and-whisker plot. This type of situation is difficult to detect because the delay is normal. In this study, the correlation analyzer is used to confirm the topology discovery via man-in-the-middle attack.

The correlation analyzer analyzes the correlation of network traffic on each link. When the fake link produced by the topology discovery via man-in-the-middle attack does not exist, the LLDP packets replayed by the attacker to establish this link must pass through other links in the network topology. If the SDN controller sends numerous LLDP packets to a fake link, then the traffic on this fake link and other links to the destination switch will increase simultaneously. The correlation analyzer sends numerous LLDP packets to the source of the new link through the dynamic LLDP frame generator and collects the network traffic of each port from the destination switch through the port monitor. In this manner, the correlation analyzer can calculate whether the link traffic to the destination switch is correlated. If the traffic to the destination switch of the new link is highly correlated with the traffic of other links on the destination switch after the Spearman's correlation analysis (correlation is greater than 0.7), then the topology manager will be notified to delete this added link.

The time-difference analyzer extracts the timestamp in the LLDP packet when the LLDP packet is received to obtain the RTT of the LLDP packet. If any new link is generated, then this module will filter out the historical RTT record of the link associated with this new link and calculate the upper bound of the RTT. If the RTT is greater than the upper bound, then it will be considered a topology discovery via man-in-the-middle attack.

3.2. System Operation Process and Mechanism. This section explains the operation process of the proposed topology management module, LLDP handling module, and CTAD module. Table 2 presents the variables and definitions of topology discovery and CTAD mechanism.

3.2.1. CTAD Operation Process. This section describes the OpenFlow topology discovery process before and after starting the CTAD mechanism and the operation process of the CTAD mechanism. The OpenFlow topology discovery process is divided into two phases. After Phase 1 is completed, the topology manager has the current network state T (Figure 8), and Phase 2 starts the CTAD mechanism (Figure 9). The process of the Phase 1 of OpenFlow topology discovery is described as follows.

Table 2: Variables and definitions of topology discovery and CTAD mechanism.

Variable	Definition
n	nth time slot of topology discovery
Key^n	Verification key for time slot n
Key^{lldp}	Verification key extracted from the received LLDP packet
Switch_i	ith switch
Switch_l	lth switch ($l \neq i$)
Port_j^i	jth port of the ith switch
Port_m^l	mth port of the lth switch ($l \neq i$)
timeout	Signal that informs the topology manager that the cycle of topology discovery has ended
P_i	List of ports of the ith switch
P_l	List of ports of the lth switch ($l \neq i$)
$\text{Link}_{(l,m)}^{(i,j)}$	Link between the jth port of ith switch and the mth port of lth switch
lldp_j^i	LLDP packet sent from the jth port of the ith switch
T	Topology information of the entire SDN
$N_{i,j}^{\text{lldp}}$	Number of LLDP packets received from Port_j^i
N^{lldp}	List of $N_{i,j}^{\text{lldp}}$
key_element	String containing number 0–9 and character A–F
sPath_l^i	Dijkstra's shortest path between the ith switch and lth switch
Skip_j^i	Tag representing if Port_j^i bypass the LLDP flooding check
$\text{Rtt}(i, j, n)$	RTT of lldp_j^i at time slot n
$Rtt(s)$	List of all RTT records in last s second(s)
$\text{IRQ}_{T(s)}$	IQR of $\text{RTT}(s)$ in the last s second(s)
$Q_{T(s)}^3$	The third quartile of $\text{RTT}(s)$ in the last s second(s)
send_rate	The rate of dynamic LLDP frame generator generating LLDP packets
Rate_j^i	The traffic rate of Port_j^i
Rate_m^l	The traffic rate of Port_m^l ($l \neq i$)
R_j^i	The list of traffic data of Port_j^i
R_m^l	The list of traffic data of Port_m^l ($l \neq i$)
c	The correlation coefficient between R_j^i and R_m^l
duration	The execution time of function lldp_gen_random (Port_j^i, *duration*)

Step 1. The topology manager collects the port list of each OpenFlow Switch_i on the network through port monitor.

Step 2. The port monitor requests the port list P_i for each Switch_i through OpenFlow. After receiving the request, the switch returns the port list to the port monitor and then sends it to the topology manager.

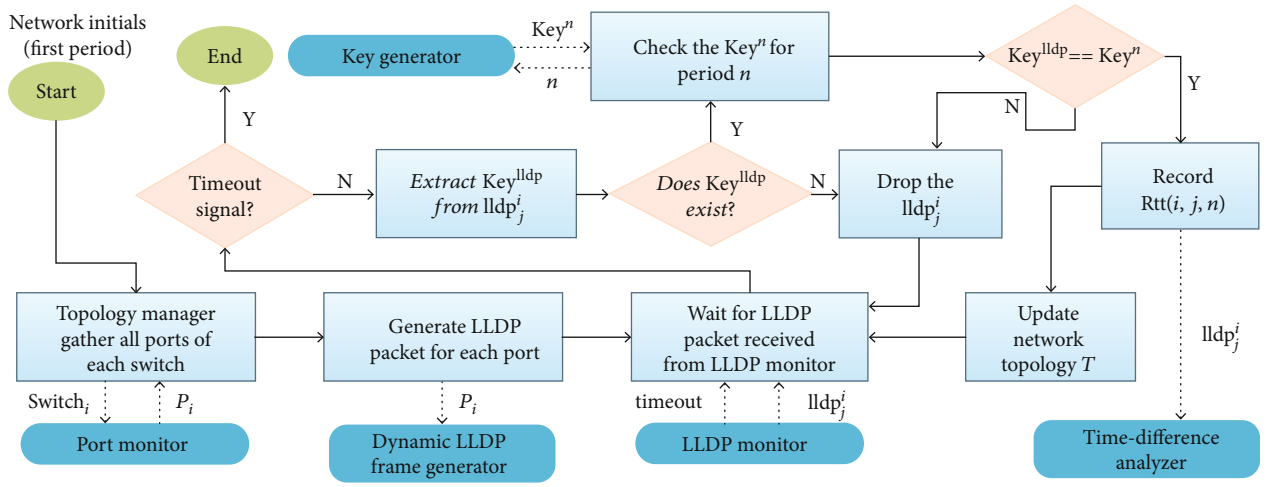


FIGURE 8: Phase 1 of the OpenFlow topology discovery.

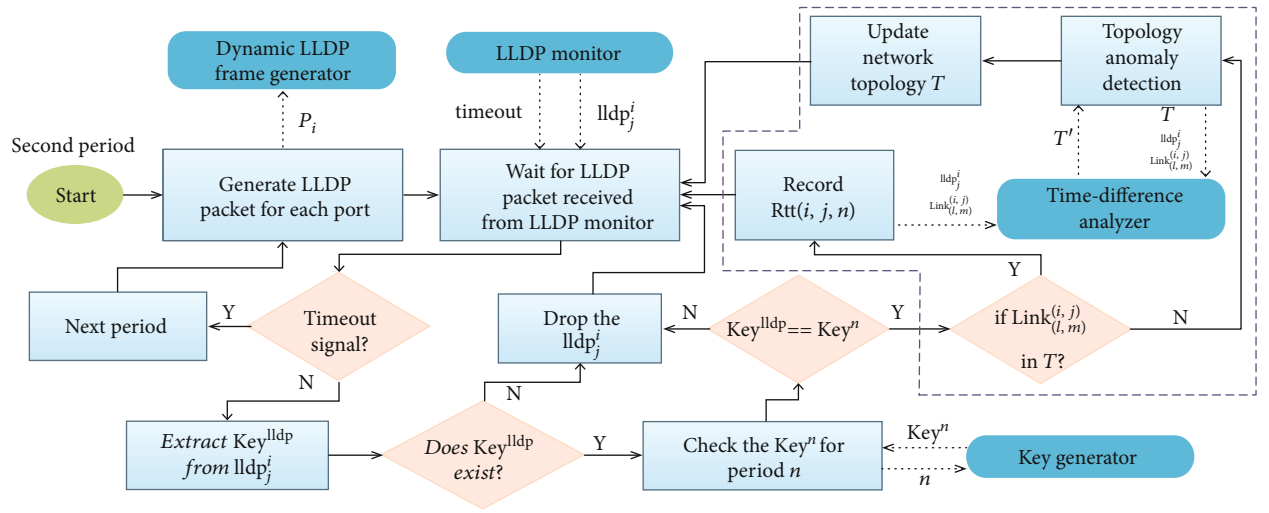


FIGURE 9: Phase 2 of the OpenFlow topology discovery.

Step 3. The topology manager starts topology discovery by sending the port list P_i to the dynamic LLDP frame generator and calling the function lldp_gen_fix() to send LLDP packets to each port.

Step 4. The dynamic LLDP frame generator generates LLDP packets lldp_j^i for the ports in port list P_i and request Key^n of this topology discovery period n by calling the function gen_key(). The dynamic LLDP frame generator encapsulates Key^n and the current timestamp into this lldp_j^i and then uses OpenFlow Packet-Out message to send lldp_j^i to Switch_i through the LLDP sender.

Step 5. The LLDP monitor monitors whether the SDN controller has received an LLDP packet lldp_j^i and sends it to the topology manager.

Step 6. The topology manager unpacks and analyzes the received lldp_j^i. First, the LLDP authentication key is extracted. If an authentication key is present, then the LLDP conforms to the packet format generated by the dynamic LLDP frame generator. Otherwise, it represents a fake LLDP packet, which will be discarded.

Step 7. If Key^{lldp} exists in lldp_j^i, then the topology manager obtains the verification Key^n of the current topology discovery cycle n from key manager and compares it with Key^{lldp}. If the comparison result is different, then lldp_j^i is discarded. Otherwise, the topology manager stores the link information $\text{Link}_{(l,m)}^{(i,j)}$ recorded in lldp_j^i to the topology information T and sends lldp_j^i to the time-difference analyzer to calculate $\text{Rtt}(i, j, n)$.

Step 8. Steps 5–7 will continue until the end of this topology discovery cycle. When the LLDP monitor finds that the topology discovery cycle has ended, it sends a timeout to the topology manager to enter the Phase 2 of OpenFlow topology discovery.

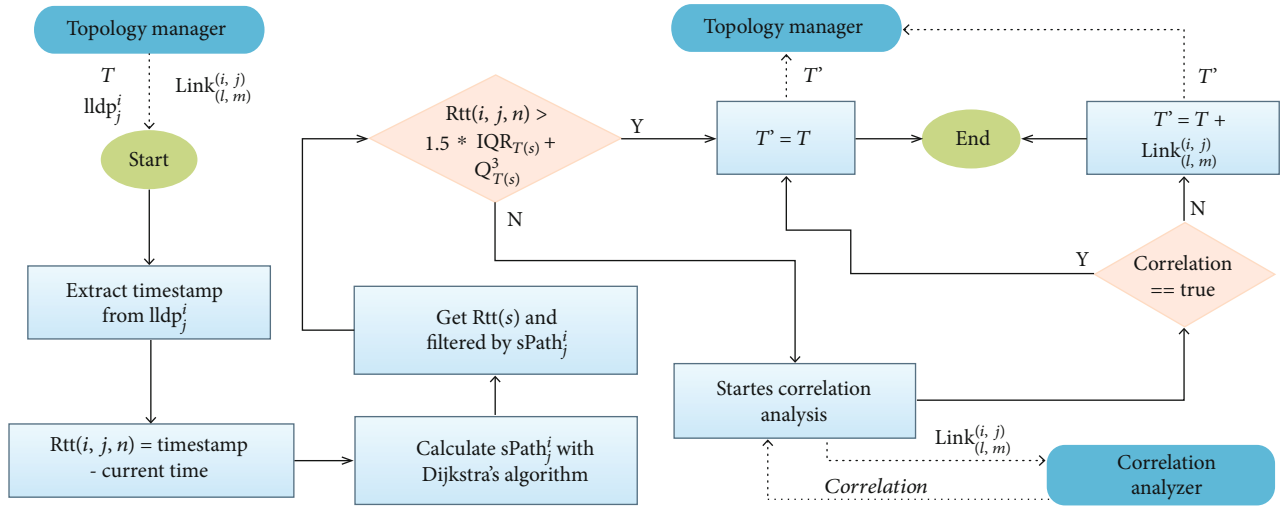


Figure 10: Time difference analysis.

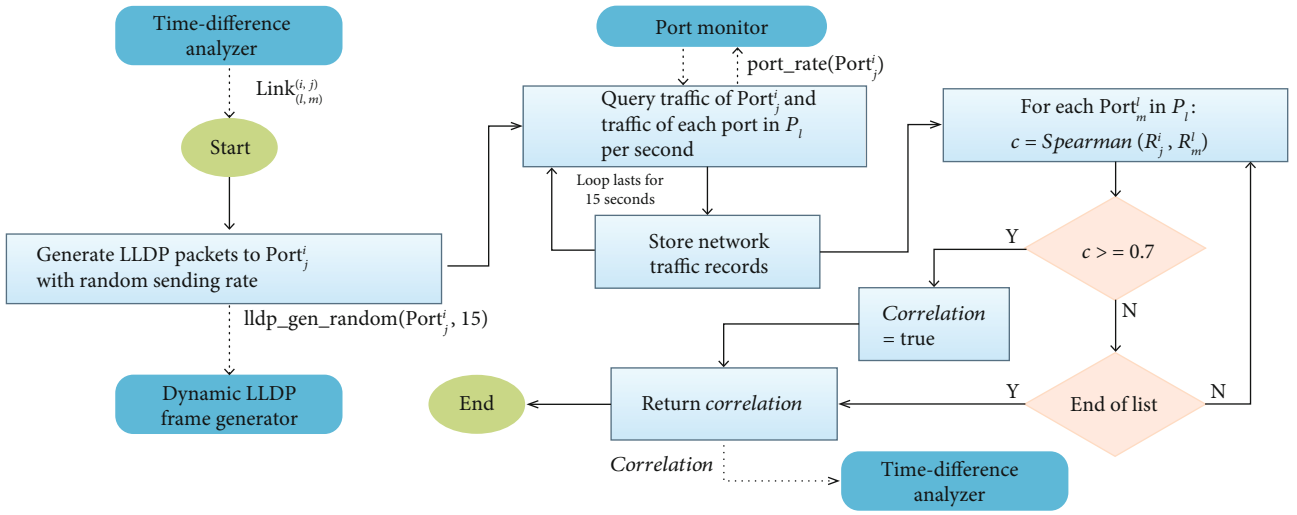


Figure 11: Correlation analysis.

Step 9. The LLDP monitor monitors the number of LLDP packets $N_{i,j}^{\text{lldp}}$ received on each port of each switch and records them in the statistical list N^{lldp}. The LLDP monitor will confirm the number of each $N_{i,j}^{\text{lldp}}$ in N^{lldp}. If the device connected to Port$_j^i$ sends numerous LLDP packets, function block_port() is called to the LLDP floodgate module to stop sending the LLDP packets received by Port$_j^i$ to the SDN controller.

After the Phase 1 of topology discovery is completed, the topology manager has the current topology T, and Phase 2 will start the CTAD mechanism. The process of the Phase 2 of OpenFlow topology discovery is described as follows.

Step 1. The topology manager starts the topology discovery and generates LLDP packets to the port list P_i of each Switch$_i$ through the dynamic LLDP frame generator.

Step 2. The LLDP monitor monitors whether the SDN controller has received an LLDP packet lldp$_j^i$ and sends it to the topology manager. After the topology discovery cycle ends, whether the device connected to Port$_j^i$ sends numerous LLDP packets will be confirmed.

Step 3. The topology manager unpacks and analyzes the received lldp$_j^i$. First, the LLDP authentication key is extracted. If a key exists, then the LLDP conforms to the packet format generated by the dynamic LLDP frame generator. Otherwise, it represents a fake LLDP packet, which will be discarded.

Step 4. If a Key$^{\text{lldp}}$ exists in lldp$_j^i$, then the topology manager acquires the verification Key^n of the current topology discovery cycle n from the key manager and compares it with Key$^{\text{lldp}}$. If the comparison result is different, then lldp$_j^i$ is

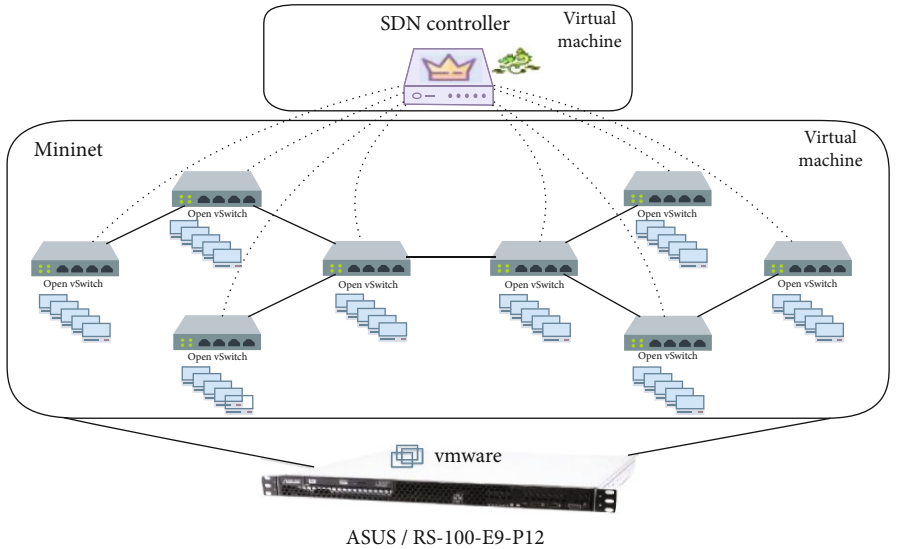


Figure 12: System implementation topology, including 8 Open vSwitches and 40 virtual computers (host).

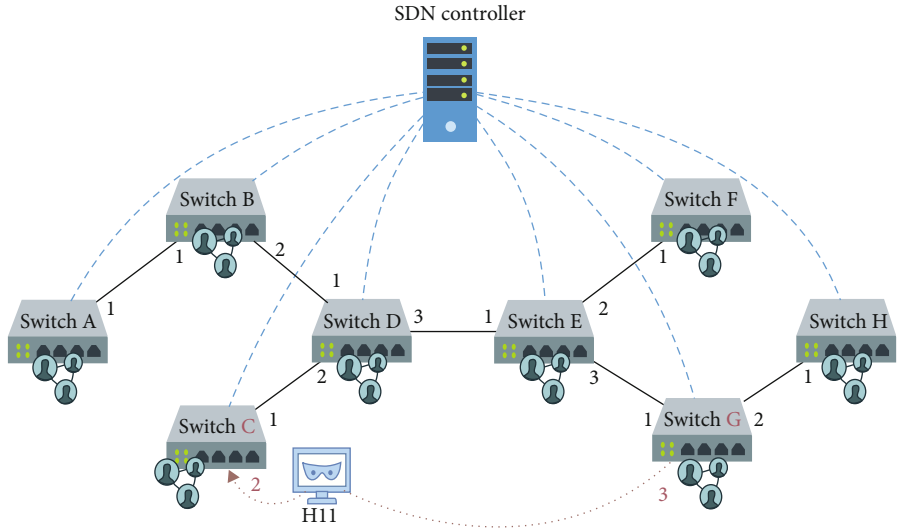


Figure 13: Topology discovery via injection attack. The attacker attempts to forge a (DPID: G, portID: 3→DPID: C, portID: 2) link.

discarded. Otherwise, lldp_j^i is converted to link information $\text{Link}_{(l,m)}^{(i,j)}$ and compared with the current topology information T. If $\text{Link}_{(l,m)}^{(i,j)}$ exists in T, then lldp_j^i is sent to the time-difference analyzer to calculate $\text{Rtt}(i, j, n)$; otherwise, the CTAD mechanism for topology discovery attack analysis will begin.

Step 5. Steps 1–4 are performed for each topology discovery cycle n.

The CTAD mechanism is divided into two phases. The first phase is the time difference analysis (Figure 10), and the second phase is the correlation analysis (Figure 11). This mechanism analyzes whether a topology discovery via man-in-the-middle attack exists. The process of the time difference analysis is described as follows.

Step 1. The topology manager triggers the time-difference analyzer and passes the current topology information T, the added link information $\text{Link}_{(l,m)}^{(i,j)}$, and the LLDP packet lldp_j^i that generates this link.

Step 2. After the time-difference analyzer receives the LLDP packet lldp_j^i from the topology manager, it reads the timestamp to obtain $\text{Rtt}(i, j, n)$ of this lldp_j^i.

Step 3. For Switch_i and Switch_l on the link information $\text{Link}_{(l,m)}^{(i,j)}$, the shortest distance sPath_l^i is calculated by Dijkstra's algorithm.

Step 4. All $\text{Rtt}(s)$ are filtered through sPath_l^i, and only the RTTs related to sPath_l^i are retained.

Table 3: Network topology.

Link	Source DPID	Source port ID	Destination DPID	Destination port ID
1	A	1	B	1
2	B	1	A	1
3	B	2	D	1
4	C	1	D	2
5	D	1	B	2
6	D	2	C	1
7	D	3	E	1
8	E	1	D	3
9	E	2	F	1
10	E	3	G	1
11	F	1	E	2
12	G	1	E	3
13	G	2	H	1
14	H	1	G	2

Step 5. The filtered $\mathrm{Rtt}(s)$ is calculated through the box-and-whisker plot to obtain the third quartile $Q^3_{T(s)}$ and the interquartile distance $\mathrm{IRQ}_{T(s)}$. If $\mathrm{Rtt}(i, j, n)$ of lldp^i_j is greater than $Q^3_{T(s)}$ plus $\mathrm{IRQ}_{T(s)}$, then this $\mathrm{Rtt}(i, j, n)$ is the outlier, and the process will skip Step 8; otherwise, this $\mathrm{Rtt}(i, j, n)$ is a normal value. To avoid man-in-the-middle attacks by replaying the topology of LLDP packets at a high rate, it must be verified through correlation analysis.

Step 6. $\mathrm{Link}^{(i,j)}_{(l,m)}$ is passed to the correlation analyzer for analysis.

Step 7. If the correlation reported by the correlation analyzer is true, then the next step is performed; otherwise, the process will skip to Step 9.

Step 8. $\mathrm{Link}^{(i,j)}_{(l,m)}$ is rejected to be added in topology information T.

Step 9. $\mathrm{Link}^{(i,j)}_{(l,m)}$ is added in the topology information T' to be returned to the topology manager.

The correlation analysis is described as follows.

Step 1. The time-difference analyzer triggers the correlation analyzer and passes the link information $\mathrm{Link}^{(i,j)}_{(l,m)}$ for analysis.

Step 2. The correlation analyzer calls the function lldp_gen_random() to trigger dynamic LLDP frame generator to send numerous random LLDP packets to Port^i_j.

Step 3. The dynamic LLDP frame generator sends random rate LLDP packets to Port^i_j.

Step 4. Skip^i_j is set as true to prevent Port^i_j from receiving a large number of LLDP packets from the dynamic LLDP frame generator and returns to the SDN controller.

Step 5. A countdown timer is set for the duration of the correlation analyzer.

Step 6. Send_rate is set to 10–100 Mbps.

Step 7. A large number of 1500 MTU-sized LLDP packets lldp^i_j are created and sent to Port^i_j at the transmission rate of send_rate.

Step 8. Steps 6 and 7 are repeated until the countdown timer ends.

Step 9. The dynamic LLDP frame generator transmits a large number of LLDP packets. At the same time, the correlation analyzer collects traffic rate for all ports P_l and Port^i_j on the destination switch Switch_l of $\mathrm{Link}^{(i,j)}_{(l,m)}$ and triggers the port monitor to collect data by calling port_rate() [28].

Step 10. The port monitor collects the traffic rate Rate^l_m from each port Port^l_m of all P_l and Port^i_j and obtains the aforementioned rate sets R^i_j and R^l_m.

Step 11. The correlation analyzer performs Spearman's rank correlation coefficient analysis on R^i_j collected on source port and R^l_m on all target ports. The analysis result is the correlation coefficient c. If multiple correlation coefficients are greater than 0.7, then $\mathrm{Link}^{(i,j)}_{(l,m)}$ is generated by the topology discovery via man-in-the-middle attack.

Step 12. The correlation analyzer reports the results of correlation analysis to the time-difference analyzer. If the correlation coefficients c is greater than 0.7, then the return correlation value is true and false otherwise.

4. Results and Discussion

The experimental platform uses the Open vSwitch to simulate a realistic OpenFlow switch. The Open vSwitch is built on a virtual machine to implement the overall OpenFlow switch environment through a Mininet simulator. The underlying architecture uses the server ASUS/RS100-E9-PI2 as the hypervisor, as shown in Figure 12. The SDN controller uses the Ryu controller provided by the Ryu SDN Framework Community, which communicates with the Open vSwitch on another virtual machine through OpenFlow. In this experiment, two attackers exist. Among the 40 virtual computers simulated by Mininet, 2 are used as attackers. LLDP generator [29], TCPdump [30], and TCPreplay [31] are utilized for topology discovery attacks. In this study, corresponding scenarios are designed for topology discovery via injection, flooding, and man-in-the-middle attacks.

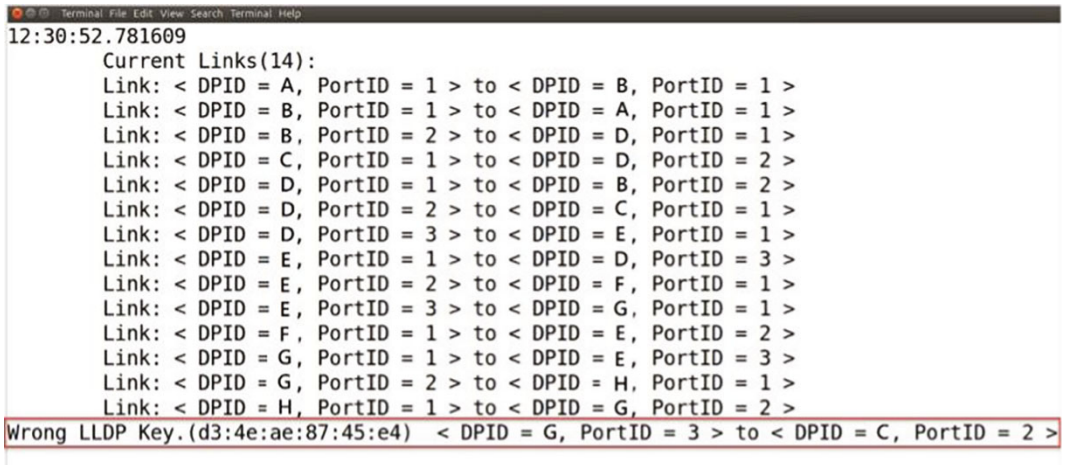


Figure 14: Topology discovery via injection attack. The fabricated LLDP packet has a wrong key.

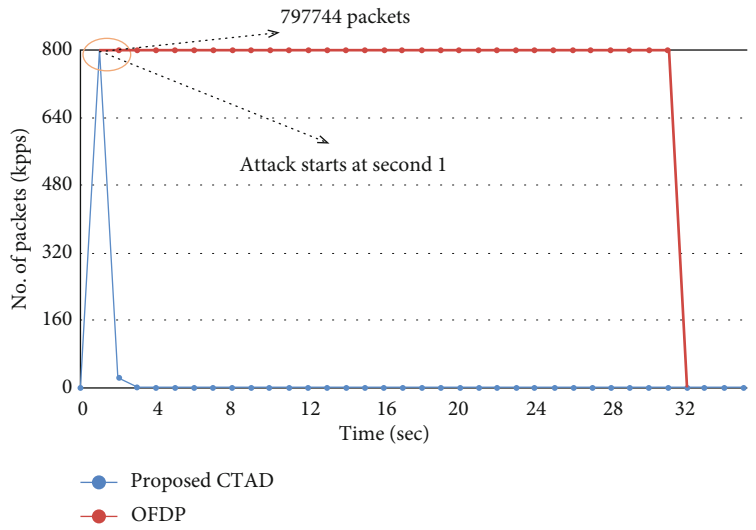


Figure 15: Topology discovery via flooding attack. The proposed mechanism can reduce the amount of attack packets.

4.1. Topology Discovery via Injection Attack Scenario. To simulate the topology discovery via injection attack, the LLDP packet is forged, the chassis ID and port ID are filled in the LLDPDU into Switch G and Port 3, and the fake LLDP packet is sent to Port 2 of Switch C. Switch C, which received the packet, is then transmitted to the SDN controller through a Packet-In message, and an attempt is made to forge a (DPID: G, portID: 3→DPID: C, portID: 2) link, as shown in Figure 13. The topology is shown in Table 3.

However, when the topology manager extracts the verification key from the LLDP packet that does not match the verification key of the current topology discovery period, this scenario indicates that the LLDP packet is fabricated by the attacker. Therefore, the fabricated LLDP packet will be discarded, as shown in Figure 14.

4.2. Topology Discovery via Flooding Attack Scenario. TCPreplay is used to generate numerous LLDP packets and conduct a 30-second attack for the simulation of topology discovery via flooding attack. The results show that compared with OFDP, after each topology discovery cycle ends, the LLDP monitor calculates the number of LLDP packets received on each port. Once it finds that a device connected to any port sends numerous fake LLDP packets, the port number will be sent to the LLDP floodgate for blocking. Thus, the topology discovery via flooding attack can be blocked, as shown in Figure 15.

4.3. Topology Discovery via Man-in-the-Middle Attack Scenario. In this scenario, a simulated attacker launches a topology discovery via man-in-the-middle attack, and the time difference and correlation analysis of the proposed CTAD mechanism will be discussed.

4.3.1. Latency of LLDP. This experiment compares the proposed CTAD mechanism with KHMAC [21] and OFDP [3]. Figure 16(a) shows the average RTT of LLDP packets under the network scale of 20 switches using CTAD,

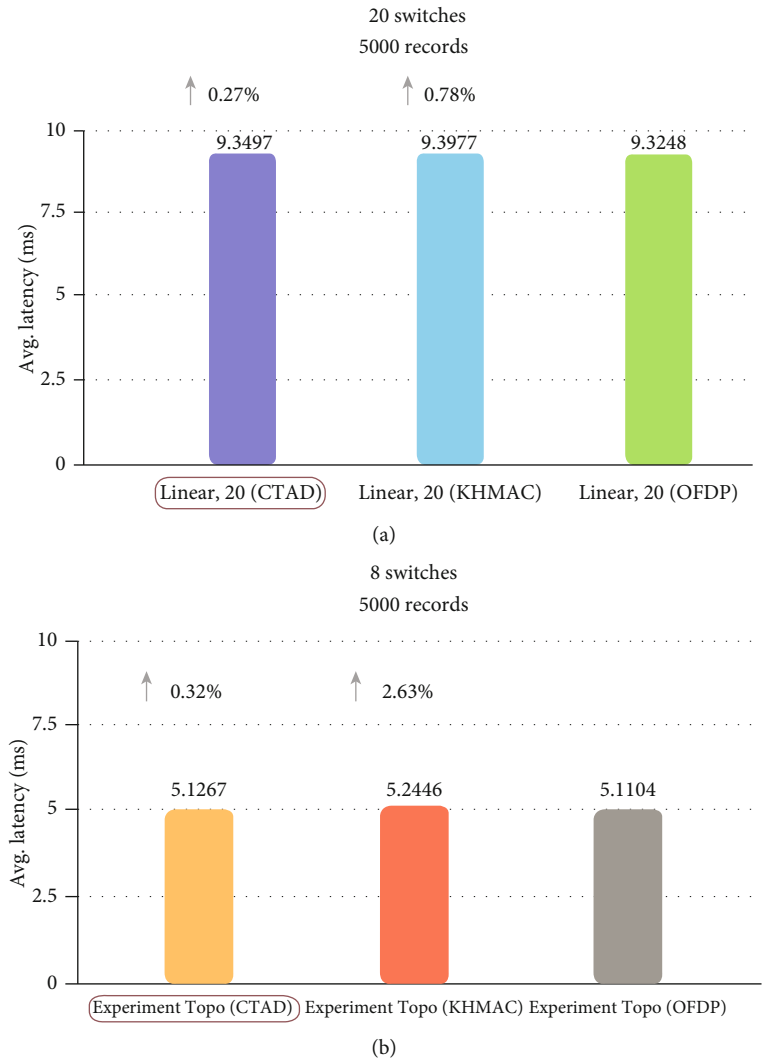


Figure 16: RTT of LLDP packets: (a) 20 switches; (b) 8 switches.

KHMAC, and OFDP. The CTAD adds authentication information and timestamps to the LLDP packets. The KHMAC only adds authentication information. Both will increase the overhead of the LLDP RTT due to the increase in LLDP length. Figure 16(b) shows the average RTT of LLDP packets under the network scale of 8 switches using CTAD, KHMAC, and OFDP. The proposed CTAD mechanism only increases the time by 0.32% compared with OFDP, whereas KHMAC increases the time by 2.63%.

4.3.2. Distribution Analysis of Latency. This experiment analyzes the RTT of LLDP packets. In this experiment, the SDN controller collects the data returned by the 5, 8, 31, and 127 virtual switches. Figure 17(a) illustrates that for the network scale of only 5 switches, the overall RTT distribution of LLDP packets is normally distributed. The distribution of the RTT of LLDP packets determines which outlier calculation method to be used for the time difference analysis. Common methods include the Z-score [33] and the box-and-whisker plot. The Z-score assumes that the data conforms to the normal distribution; thus, 99.7% of the data will be within the range of the maternal mean plus or minus 3 maternal standard deviations. Therefore, the data greater than the maternal mean of 3 maternal standard deviations will be regarded as outliers. The box-and-whisker plot does not assume that the data must conform to the normal distribution. It calculates the maximum, median, first to third quartiles, and IQR of the data to obtain the upper and lower bounds. The data outside the boundary will be regarded as outliers.

Figure 17(a) illustrates that the Z-score and box-and-whisker plot can be used. Figures 17(b)–7(d) indicate that the distribution of the data gradually becomes skewed and does not conform to the normal distribution; thus, it is not suitable to use the Z-score. Therefore, the time difference analysis proposed in this study uses the box-and-whisker plot to determine outliers.

4.4. CTAD Detection Rate. To simulate the topology discovery via man-in-the-middle attack, Attacker 1 (H11) taps and records the LLDP packet from the SDN controller at Port 2 of Switch C and then encapsulates the LLDP packet and sends it to Attacker 2 of Port 3 of Switch G (H31). Attacker

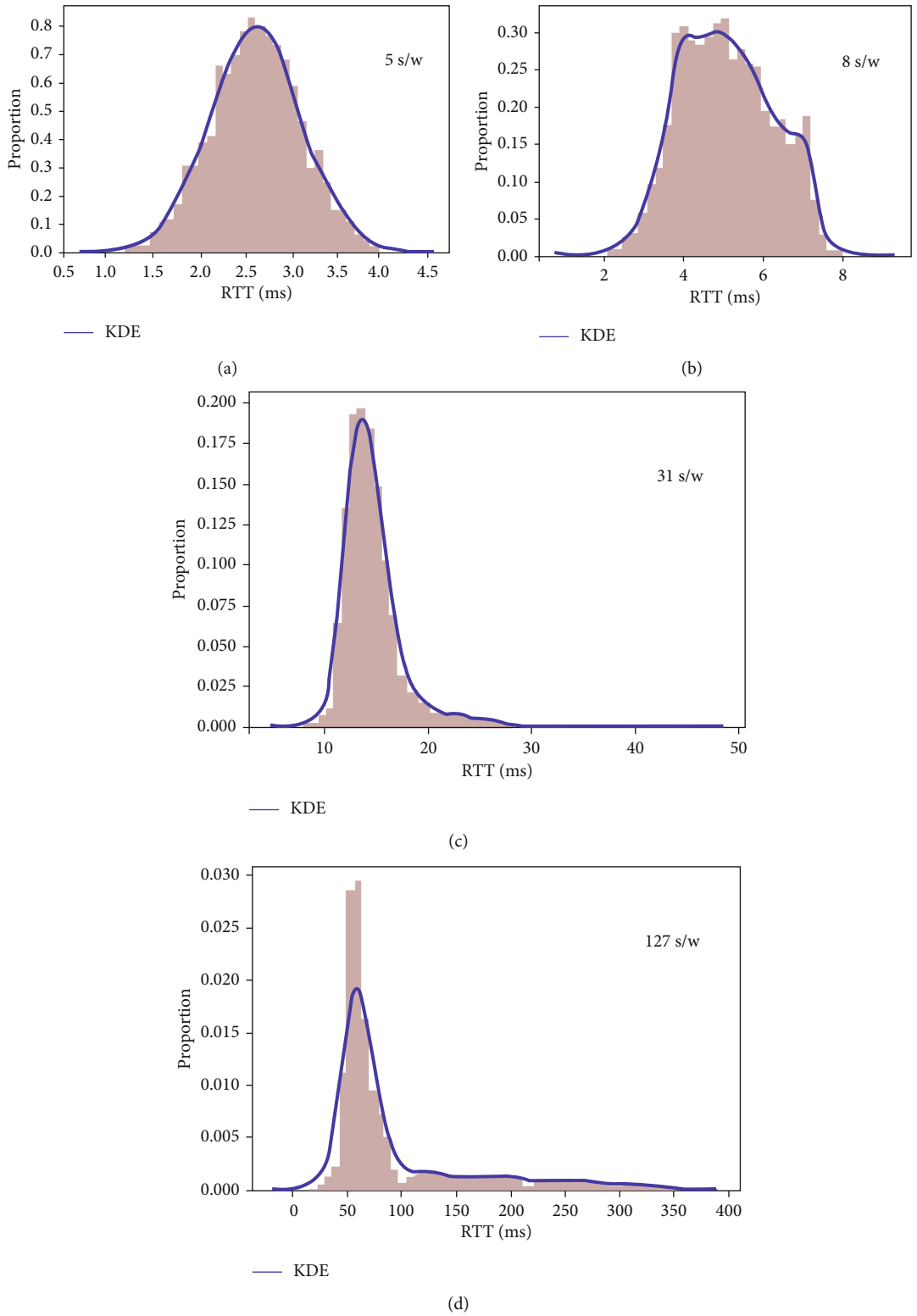


Figure 17: Overall RTT distribution: (a) 5 switches; (b) 8 switches; (c) 31 switches; (d) 127 switches.

2 (H31) changes back to the normal packet type and attempts to forge a (DPID: C, portID: 2→DPID: G, portID: 3) link, as shown in Figure 18.

To verify the topology discovery via man-in-the-middle attacks with high-speed replay, replay LLDP is generated with different delay times and compare them with SALL [22] and KHMAC [21]. The results show that the correlation analyzer can detect a topology discovery via man-in-the-middle attack that replays LLDP packets at high speeds as low as 40 ms (Figure 19).

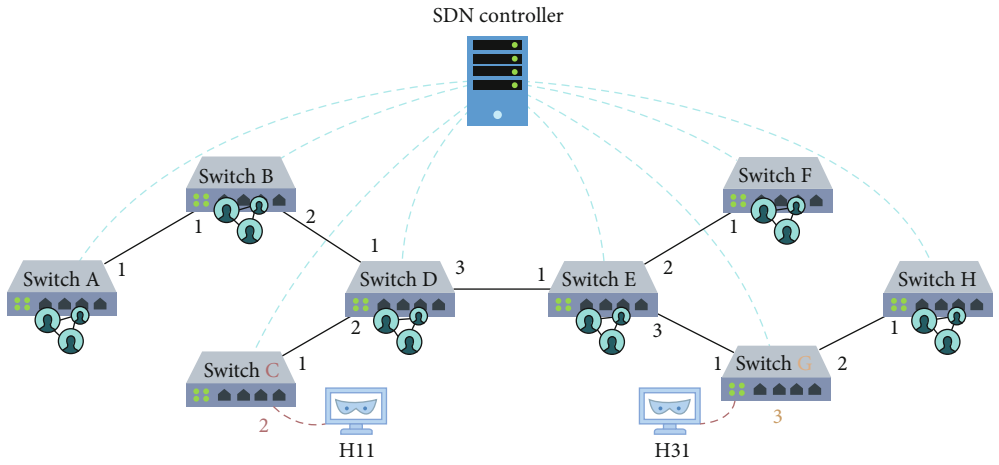


Figure 18: Topology discovery via man-in-the-middle attack. The attacker attempts to forge a (DPID: C, portID: 2→DPID: G, portID: 3) link.

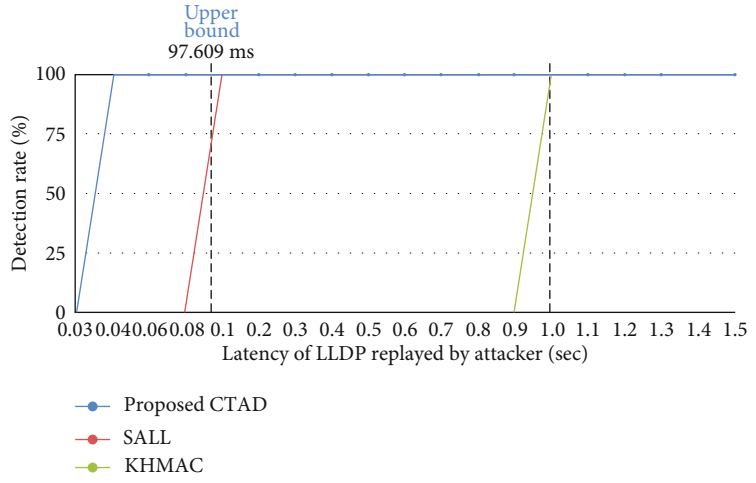


Figure 19: Comparison of topology discovery via man-in-the-middle attack detection rates.

5. Conclusions

This study adds verification information and timestamps to LLDP packets, detects each LLDP packet that is returned to the SDN controller, and successfully detects topology discovery via injection and man-in-the-middle attacks. It also calculates the amount of LLDP packets received on each port during each topology discovery cycle, which reduces the effect of topology discovery via flood attacks on the SDN controller. Spearman's rank correlation is used to analyze the high-speed and highly automated topology discovery via man-in-the-middle attacks and maintain correct topology information of the SDN controller.

The results show that compared with the KHMAC detection mechanism that uses only authentication information and the SALL detection mechanism that uses only timestamps, the proposed mechanism has a wider scope in detecting topology discovery via man-in-the-middle attacks.

Data Availability

The data used to support the findings of this study are included within the article.

Disclosure

This paper is a revised version of a paper originally presented at the 10th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, October 2019.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

Acknowledgments

The work described in this paper was supported by the National Chung-Shan Institute of Science & Technology under Grant NCSIST-ADV-V101 (108) and the Ministry of Science and Technology of the Republic of China under Grant 108-2221-E-008-033-MY3.

References

[1] "Software-defined networking – Wikipedia," https://en.wikipedia.org/wiki/Software-defined_networking.

[2] "Link layer discovery protocol - Wikipedia," https://en.wikipedia.org/wiki/Link_Layer_Discovery_Protocol.

[3] "OpenFlowDiscoveryProtocol - GENI: geni," https://groups.geni.net/geni/wiki/OpenFlowDiscoveryProtocol.

[4] Ryu SDN frameworkhttps://ryu-sdn.org/.

[5] "Project Floodlight - GitHub," https://github.com/floodlight.

[6] "Home - OpenDaylight," https://www.opendaylight.org/.

[7] "GitHub - brandonheller/riplpox: RipL-POX (Ripcord-Lite for POX): a simple network controller for OpenFlow-based data centers," https://github.com/brandonheller/riplpox.

[8] M. J. Allen and W. M. Yen, *Introduction to Measurement Theory*, Waveland Press, Long Grove, IL, 2001.

[9] "P4, [Online]," https://p4.org/.

[10] "Stratum – open networking foundation," https://www.opennetworking.org/stratum/.

[11] "Open network operating system (ONOS) SDN controller for SDN_NFV solutions," https://onosproject.org/.

[12] N. McKeown, T. Anderson, H. Balakrishnan et al., "OpenFlow," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 69–74, 2008.

[13] A. Shalimov, D. Zuikov, D. Zimarina, V. Pashkov, and R. Smeliansky, "Advanced study of SDN/OpenFlow controllers," in *9th Central & Eastern European Software Engineering Conference in Russia*, pp. 24-25, Moscow, Russia, 2013.

[14] A. Azzouni, O. Braham, T. M. T. Nguyen, G. Pujolle, and R. Boutaba, "Fingerprinting OpenFlow Controllerscontrollers: The the First first Step step to Attack attack an SDN Control control Planeplane," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 4–8, Washington, DC, USA, 2016.

[15] H. Zhang, Z. Cai, Q. Liu, Q. Xiao, Y. Li, and C. F. Cheang, "A survey on security-aware measurement in SDN," *Security and Communication Networks*, vol. 2018, Article ID 2459154, 14 pages, 2018.

[16] H. Zimmermann, "OSI reference Model–The ISO model of architecture for open systems interconnection," *IEEE Transactions on Communications*, vol. 28, no. 4, pp. 425–432, 1980.

[17] V. Z. Attar and P. Chandwadkar, "Network discovery protocol lldp and lldp-med," *International Journal of Computer Applications*, vol. 1, no. 9, pp. 99–103, 2010.

[18] K. Benton, L. J. Camp, and C. Small, "OpenFlow vulnerability assessment," in *Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking*, pp. 151-152, Hong Kong, China, August 2013.

[19] T. H. Nguyen and M. Yoo, "Analysis of link discovery service attacks in SDN controller," in *2017 International Conference on Information Networking (ICOIN)*, pp. 11–13, Da Nang, Vietnam, January 2017.

[20] S. Hong, L. Xu, H. Wang, and G. Gu, *Poisoning network visibility in software-defined networks: new attacks and countermeasures*, National Down Syndrome Society, 2015.

[21] T. Alharbi, M. Portmann, and F. Pakzad, "The (in) security of topology discovery in software defined networks," in *Proceedings of 2015 IEEE 40th Conference on Local Computer Networks (LCN)*, pp. 26–29, Clearwater Beach, FL, USA, Oct. 2015.

[22] D. Smyth, S. McSweeney, D. O'Shea, and V. Cionca, "Detecting link fabrication attacks in software-defined networks," in *2017 26th International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–8, Vancouver, BC, Canada, August 2017.

[23] A. Azzouni, N. T. M. Trang, R. Boutaba, and G. Pujolle, "sOFTDP: secure and efficient topology discovery protocol for SDN," 2017, http://arxiv.org/abs/1705.04527.

[24] A. Nehra, M. Tripathi, M. S. Gaur, R. B. Battula, and C. Lal, "SLDP: a secure and lightweight link discovery protocol for software defined networking," *Computer Networks*, vol. 150, pp. 102–116, 2019.

[25] R. Taylor, "Interpretation of the correlation coefficient: a basic review," *Journal of Diagnostic Medical Sonography*, vol. 6, no. 1, pp. 35–39, 2016.

[26] L. Myers and M. J. Sirois, "Spearman correlation coefficients, differences between," *Encyclopedia of Statistical Sciences Sonography*, vol. 12, 2006.

[27] J. Laurikkala, M. Juhola, E. Kentala, N. Lavrac, S. Miksch, and B. Kavsek, "Informal identification of outliers in medical data," in *Fifth international workshop on intelligent data analysis in medicine and pharmacology*, Berlin, Germany, August 2000.

[28] D. J. Hamad, K. G. Yalda, and I. T. Okumus, "Getting traffic statistics from network devices in an SDN environment using OpenFlow," in *Proceedings of 2015 Information Technology and Systems (ITaS)*, pp. 951–956, Olympic Village, Sochi, Russia, 2015.

[29] A. Nicolae, L. Gheorghe, M. Carabas, N. Tapus, and C. L. Duta, "LLDP packet generator," in *2015 14th RoEduNet International Conference - Networking in Education and Research (RoEduNet NER)*, pp. 24–26, Craiova, Romania, September 2015.

[30] "GitHub - the-tcpdump-group/tcpdump: the TCPdump network dissector," https://github.com/the-tcpdump-group/tcpdump.

[31] "GitHub - appneta/tcpreplay:Pcap editing and replay tools for ∗NIX and Windows - Users please download source from," https://github.com/brandonheller/riplpox.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 7975393, 12 pages
https://doi.org/10.1155/2020/7975393

Research Article

A Blockchain Token-Based Trading Model for Secondary Spectrum Markets in Future Generation Mobile Networks

Mubbashar A. Khan,[1] **Mohsin M. Jamali**,[2] **Taras Maksymyuk**,[3] **and Juraj Gazda**[4]

[1]*Electrical and Computer Engineering Department at North Carolina Agricultural and Technical State University, Greensboro NC 27401, USA*
[2]*College of Engineering at the University of Texas of the Permian Basin, Texas, USA*
[3]*Department of Telecommunications, Lviv Polytechnic National University, Ukraine*
[4]*Department of Computers and Informatics, Technical University of Košice, 042 00, Slovakia*

Correspondence should be addressed to Juraj Gazda; juraj.gazda@tuke.sk

Received 2 January 2020; Revised 8 June 2020; Accepted 11 July 2020; Published 8 August 2020

Academic Editor: Miguel López-Benítez

Cognitive radio (CR) technology offers the possibility of an increase in spectrum utilization efficiency to resolve the prevalent spectrum scarcity problem. The economic survival of secondary spectrum markets (SSMs) is heavily dependent on the sharing of both the licensed spectrum and spectrum infrastructure by primary licensed operators (PLOs). In this research, an automated pricing model using a blockchain token called the spectrum dollar has been implemented for secondary radio spectrum trade. The use of spectrum dollars enables noncash-based secondary spectrum trade among PLOs based on a floor-and-trade rule. The pricing of spectrum dollars and the associated revenue shares are based on the underlying secondary spectrum trading behaviours of PLOs. PLOs that do not contribute enough secondary spectra to the SSM (to satisfy demand) suffer a loss proportional to the difference between their earned revenues and the specified floor value in the SSM. The secondary spectrum trade is assumed to be centrally managed by a spectrum broker, which announces the floor value for each bidding period while ensuring nonnegative revenue for the market itself. The use of the spectrum dollar along with the floor-and-trade methodology eliminates the possibilities for economic malpractice by PLOs that could increase spectrum reuse costs. In addition, the floor value provides automatic regulatory control to ensure the economic viability and prevent the technological hijacking of future SSMs.

1. Introduction

Currently, the utilization efficiency of numerous radio spectrum bands licensed by the Federal Communications Commission (FCC) is very low in general, while the use by existing wireless networks of specific licensed and unlicensed bands is highly congested [1–7]. One of the potential candidates to deal with this inequality is cognitive radio (CR) technology, which allows for dynamic spectrum access (DSA) [8]. DSA involves a reallocation of the underutilized portions in already-licensed radio spectrum bands to accommodate new secondary users (SUs) in an opportunistic manner [9]. The SUs coexist with the primary users (PUs) in the licensed radio spectrum on a noninterfering basis. Future secondary spectrum markets (SSMs) are heavily dependent on the licensed spectrum and communication infrastructure sharing by mutually competitive primary licensed operators (PLOs) [10]. SSMs are expected to provide a much cheaper alternative service to wireless users, because they do not require the purchase of licensed bands for their operation. If SSMs can offer a quality of service (QoS) that is comparable to that offered by PLOs, they may prove to be disruptive to the businesses of PLOs [2, 11]. PLOs might respond to this potential existential threat and indulge in economic malpractices to manipulate certain economic parameters and impede resource sharing through SSMs. Such malpractices may raise the cost of secondary use of the radio spectrum to the extent that it is no more affordable for end-users than use of the

licensed spectrum [12–14]. This could technically bring the emergence of CR technology and spectrum reuse to a halt. There exists a trade-off between the QoS provision and the associated pricing of the secondary radio spectrum that is enabled for reuse in SSMs [15–18]. The QoS [19–22] and economic robustness [23–28] of future SSMs have been two important topics of ongoing research on CR technology. Generally, revenue generation is core to the survival of SSMs [10].

This paper implements a secondary spectrum trade model based on a noncash approach. A new virtual token coined the spectrum dollar is proposed for the secondary spectrum trade. It is a unit incentive to be paid for the secondary radio spectrum trade-in SSMs. The spectrum dollar provides a methodology for revenue exchange among SSM entities without any involvement of a third party, such as banks, and helps to avoid the related incremental transaction costs. By its nature, the spectrum dollar is reminiscent of the application of the emerging tokenization concept in blockchain technology. A floor-and-trade rule is used to determine profitability shares based on the underlying secondary spectrum trading behaviours. This rule determines the exchange rate of spectrum dollars, the corresponding shares of trading PLOs, and participation fees to be charged by the SSM to maintain a no-loss situation for the spectrum broker managing the secondary radio spectrum trade. To maintain simplicity, in this research, the spectrum broker is assumed to be a nonprofit entity, and the focus of the study is on analysing the role of spectrum dollars as virtual tokens and their exchange based on the floor-and-trade rule for aiding secondary radio spectrum trade and pricing.

The floor is the threshold value determined by the spectrum broker to set the participation fees charged to trading PLOs for participation in the SSM, based on their adherence to the stipulated secondary spectrum trading behaviour for a given trading period. The participation fee not only forces the participating PLOs to enable spectrum reuse and avoid insurance of loss to the spectrum broker managing the trade but can also be used to generate revenues. It enlarges the profitability of the largest contributors and at the same time maximizes the loss of the lowest contributors in the SSM for a given period. This motivates participating PLOs to make their spectrum available for reuse in the SSM. The SSM generates revenues from the participation fee even if PLOs act maliciously to deny spectrum reuse. The model promises profit maximization to trading PLOs for sales maximization, subject to the satisfaction of spectrum demand in the SSM.

The rest of this paper is organized as follows. Section 2 discusses the related work. Section 3 introduces the system model for floor and trade rule and its advantages along with those offered by the use of spectrum dollars as a token for the secondary spectrum trade. The section is followed by the mathematical representation of the model in Section 4. The simulation results and analysis are provided in Section 5, and finally, conclusions are provided in Section 6.

2. Related Work

In general, we recognize two types of spectrum sharing schemes: centralized and distributed architectures. Within this paper, we focus our attention on the centralized scheme; although, the proposed scheme based on the application of the floor-and-trade rule can be easily extended to the distributed architecture. In centralized architecture, the spectrum broker is responsible for matching service requests from end-users with the spectrum bids of PLOs. The centralized architecture offers the prominent advantage of reduced complexity in terms of communication overhead, as there is no requirement of communication between end-users and PLOs. On the other hand, the centralized architecture could potentially suffer from a single point of failure, resulting in decreased efficiency of the SSM. Several solutions exist to overcome this issue, notably the proposal of a mirrored server architecture [29]. Instead of a single central server, for each SSM region, there could be multiple distributed servers, which are ready for the fail-over should the master server malfunction. Recently, blockchain technology took a step toward the wider application of such an architecture in various areas, including bioinformatics [30], the Internet of Things (IoT) [31, 32], and global payments [33]. In the spectrum sharing area, we have identified a few pioneering works proposing to use blockchain technology to enhance the efficiency of SSMs. Notably, [34] proposed the use of blockchain architecture for primary and secondary cooperative sharing, showing its advantages and drawbacks compared to the traditional database architecture. Security concerns related to spectrum sharing between aerial and terrestrial communications are addressed in [35]. The authors developed a secure spectrum trading and sharing scheme leveraging permissioned blockchain technology. In [36], the authors proposed a blockchain-based unlicensed spectrum sharing game, which allows a Nash equilibrium to be achieved among the operators by exchanging virtual cryptocurrency tokens. The blockchain technology used for spectrum sharing in the multiperator environment was considered in [37]. The authors proposed a smart contract to enable spectrum sharing while preserving the privacy of PLOs and the fairness of the system. Perhaps the conceptually closest paper to our contribution is this work presented in [38], where the authors proposed a two-stage privacy-preserving, incentive-compatible, and spectrum-efficient framework based on blockchain. The first stage involves the signing of a contract between the base station and end-users, for which the end-users receive a monetary bonus paid in cryptocurrency. The second stage involves mapping the available shared spectrum among the machine-to-machine (M2M) communication entities. Our paper also presents a two-stage mechanism. In the first stage, the spectrum broker collects the available spectrum resources from PLOs, while in the second stage, it managed the sealed bid auction process to match secondary spectrum requests and bids. It should be noted, however, that the revenues from the SSM are paid in noncash monetary units coined spectrum dollars. This approach naturally avoids the need for high transaction costs and the presence of intermediaries and eventually produces market efficiency. Due to its noncash nature, the latter part of our approach could be implemented by using a well-known public or application-specific permissioned blockchain, while the spectrum broker as the centralized entity would still be responsible for sustaining market

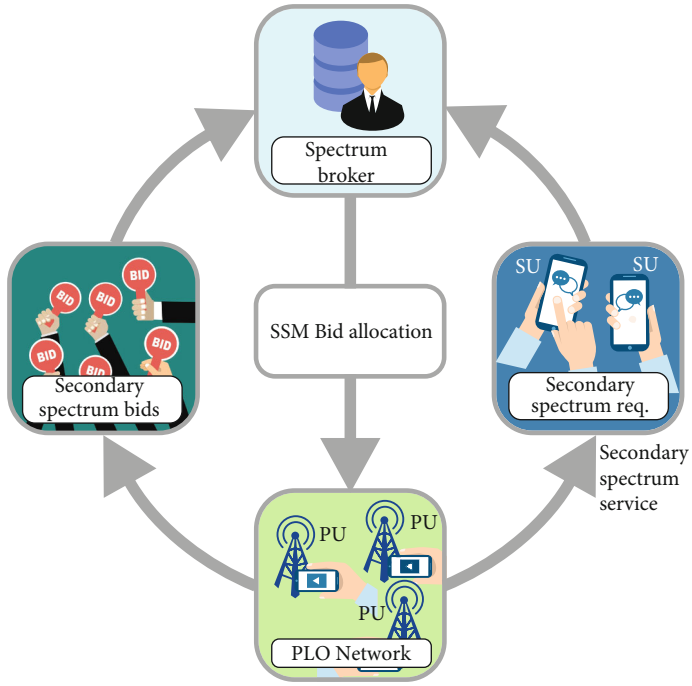


Figure 1: Secondary spectrum trade managed by a centralized spectrum broker.

equilibrium and providing a FIAT gateway exchange (spectrum dollar/USD) for PLOs.

3. The Proposed Trade Model

Figure 1 presents the scenario of a secondary radio spectrum trade model managed by a centralized spectrum broker. PLOs place their sealed bids specifying the amount of spectra and corresponding prices with the spectrum broker, which governs the auctioning process and takes responsibility for making the bid allocations. Based on supply and demand, PLOs trade their radio spectra to generate revenues for providing services to secondary users in the SSM. The spectrum auction process (i.e., secondary spectrum bids, spectrum broker, and SU requests) has been modelled as a sealed-bid spectrum auction as proposed in [7, 39]. The focus of this work is on revenue redistribution and facilitation among the SSM entities. The model developed in this research has been implemented over a trading horizon that comprises several periods of secondary radio spectrum trade. PLOs compete with each other for spectrum trade opportunities in the SSM. The floor value is a nonnegative number that determines the participation fee in spectrum dollars based on the total amount of money exchanged among participating PLOs. The floor value is also used to price the spectrum dollar and determine the corresponding profit-loss shares for participating PLOs to enable radio spectrum reuse. Spectrum dollars are proposed to be exchangeable among participating PLOs over the trade horizon. For this purpose, a token exchange block has been introduced to enable the exchange of spectrum dollars among PLOs to either buy (accumulate) enough spectrum dollars to meet the floor value or to sell spectrum dollars to convert their earnings as profit into FIAT money. The spectrum dollar exchange based on the floor value keeps the secondary spectrum trade fair [40].

The floor value set by the spectrum broker based on the underlying secondary spectrum trade is used to determine whether a participating PLO made a profit or a loss in a given secondary radio spectrum trading period. All trading PLOs must pay the participating fee equivalent to the floor value to the spectrum broker in spectrum dollars. Loss-making PLOs must purchase the spectrum dollars from profitable PLOs to make up the difference between their respective earnings and the set floor value (i.e., the participation fee) for a given trading cycle. This purchase enables loss-making PLOs to make their payments for trading in the SSM to the spectrum broker in spectrum dollars. The number of spectrum dollars to be exchanged by participating PLOs is used to set the price for the spectrum dollar using the demand-supply equilibrium principle. The number of spectrum dollars purchased by a loss-making PLO for a given trading cycle at the current spectrum dollar price in the SSM determines its total loss. Similarly, the number of excess spectrum dollars of profitable PLOs for a given trading period at the current spectrum dollar price in the SSM determines their respective profitability for the given trading period.

The key aspect of the proposed platform is the tokenization (with the spectrum dollar acting as the token) of the spectrum, which allows for the opening of the spectrum sharing market without intermediaries and with additional control by the regulator (the spectrum broker). In the proposed model, the token exchange block can be realized as the traditional FIAT gateway exchange block in blockchain, while the spectrum broker is a separate entity placed outside the blockchain model. This hybrid (blockchain application and

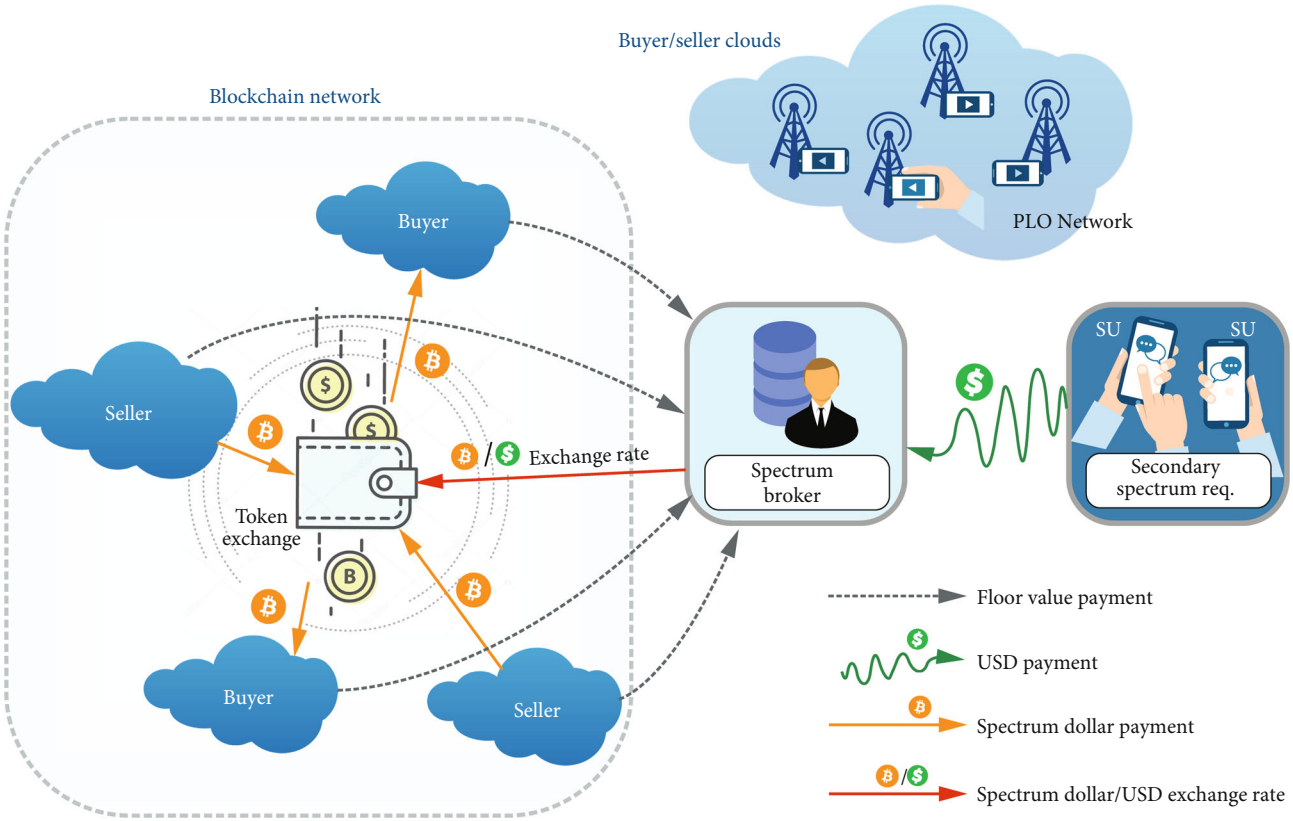


Figure 2: Block diagram of the proposed spectrum secondary market with USD and spectrum dollar transaction flows.

centralized control mechanism) scheme has already been considered in several studies, including [37]. It should also be noted that we reserve the examination of numerous details regarding the blockchain implementation (consensus mechanisms, selection of public (permissioned) blockchain, hyperledger, etc.) for our future research and place our main attention in this paper on the study of the entire demand/-supply chain at large.

3.1. Advantages of the Floor-and-Trade Rule. The proposed system model is based on the floor-and-trade rule, which provides several advantages to the trading process in the SSM, which are explained below.

(i) The floor-and-trade rule enables an automated methodology to impose a secondary spectrum withholding charge on PLOs and ensure their participation and spectrum utilization efficiency

(ii) The floor-and-trade methodology increases the chances of secondary spectrum availability by enforcing the participation of PLOs in the SSM to enable radio spectrum reuse

(iii) The automation of profit-loss calculations as revenue shares of the participating entities reduces the monitoring overhead for secondary spectrum trade reporting by PLOs. Further possibilities for false secondary spectrum trade values are minimized through the use of a spectrum withholding charge such as the participation fee in the SSM

(iv) There are no added operational costs (other than infrastructure sharing cost, if any) for the spectrum broker/SSMto enable and manage the secondary radio spectrum trade in a highly competitive manner using the floor-and-trade methodology

(v) The floor-and-trade methodology maximizes the profitability shares of the highest-contributing PLOs, and at the same time, it maximizes the loss of the lowest-contributing PLOs

(vi) The floor-and-trade rule minimizes the chances of possible economic malpractice involving participating PLOs manipulating secondary spectrum pricing. This avoids the establishment of a monopoly over the enabled secondary spectrum

(vii) The use of the spectrum dollar can enable benefits transfer among the entities belonging to SSMs of different sizes in a direct manner without requiring any involvement by third parties, such as banks, and hence avoids transaction costs incurred to the third parties

The block diagram of interaction among all SSM entities using the spectrum dollar and USD transactions is depicted in Figure 2.

Table 1: Notation for the different parameters in the floor-and-trade rule-based exchange methodology.

Notation	Parameters
h	Instant trading period in the SSM.
H	Trade horizon in the SSM.
n	Total number of participating PLOs in the SSM.
$Revenue_{i,h}^{S\$}$	Revenue in spectrum dollars for participating PLO i for a given h.
$Revenue_{j,h}^{S\$}$	Revenue in spectrum dollars for participating PLO j for a given h.
$Total\ Revenue_H^{S\$}$	Total revenue of n PLOs from secondary spectrum trade for a given H.
$Floor_H$	Floor value for exchange of spectrum dollars among n PLOs for a given H.
Seller	Spectrum dollar seller for a given trade horizon H.
Buyer	Spectrum dollar buyer for a given trade horizon H.
$\sum_{h=1}^{H} Revenue_{Seller,h}^{S\$}$	Total revenue of a spectrum dollar seller in the SSM for a given trade horizon H.
$\sum_{h=1}^{H} Revenue_{Buyer,h}^{S\$}$	Total revenue of a spectrum dollar buyer in the SSM for a given trade horizon H.
$P_H^{S\$}$	Spectrum dollar price for a given trade horizon H.
$Profitability_i^{USD}$	Profitability of participating PLO i in USD for a given trade horizon H.
$Total\ Sellers_H$	Number of spectrum dollar sellers in the SSM for a given trade horizon H.
$Total\ Buyers_H$	Number of spectrum dollar buyers in the SSM for a given trade horizon H.

4. Analytical Model for the Floor-and-Trade Rule

In this section, the fundamental relations between the spectrum dollar supply and demand and the corresponding equations are outlined. Let $Revenue_{i,h}^{S\$}$ and $Revenue_{j,h}^{S\$}$ represent the revenue in spectrum dollars for trading PLOi and PLOj, respectively, for a given instant trading period. Let $Total\ Revenue_H^{S\$}$ represent the total revenue of all trading PLOs in terms of spectrum dollars over a given trading horizon. Let $Floor_H$ represent the floor value for exchange of spectrum dollars among the nPLOs trading in the SSM. A PLO is a spectrum dollar seller if its revenue generation over a complete trading horizon, i.e., $\sum_{h=1}^{H} Revenue_{Seller,h}^{S\$}$, is equal to or greater than the floor value determined by the spectrum broker for a given bidding instant. Similarly, a PLO is a spectrum dollar buyer if its revenue generation over a complete trading horizon in spectrum dollars, i.e., $\sum_{h=1}^{H} Revenue_{Buyer,h}^{S\$}$, is less than the floor value set in the SSM for the given trading period. Let $P_H^{S\$}$ be the price calculated for the exchange of a unit spectrum dollar and $Profitability_i^{USD}$ be the profitability of a participating PLO in US dollars, for a given trading horizon. Let the total number of spectrum dollar seller be represented by $Total\ Sellers_H$. Also, let the number of spectrum dollar buyers be represented by $Total\ Buyers_H$ (i.e., $Total\ Buyers_H = n - Total\ Sellers_H$).

Table 1 provides the notation for the different parameters involved in the implementation of the floor-and-trade rule-based spectrum dollar exchange methodology. The revenue of an individual trading PLO in spectrum dollars is calculated as the sum of its revenue generation over all trading periods of a given trading horizon H, as given in Eq. (1)

$$Revenue_{i,H}^{S\$} = \sum_{h=1}^{H} Revenue_{i,h}^{S\$}. \tag{1}$$

$Total\ Revenue_H^{S\$}$ is the sum of revenues generated by individual PLOs over all trading periods of a given trading horizon H, as given in Eq. (2).

$$Total\ Revenue_H^{S\$} = \sum_{h=1}^{H} \sum_{i=1}^{n} Revenue_{i,h}^{S\$}. \tag{2}$$

The floor value is determined by the spectrum broker by dividing the total revenues generated by the total number of PLOs trading in the SSM, as given in Eq. (3)

$$Floor_H = \frac{Total\ Revenue_H^{S\$}}{n}. \tag{3}$$

The selling/buying status of each PLO trading in the SSM is determined by calculating the difference between its earnings and the specified floor value for a given trading horizon. This helps to determine the total number of spectrum dollar sellers and buyers for a given trade horizon H as follows:

$$\begin{aligned} Total\ Sellers_H &= Total\ Sellers_H + 1, \text{ if } Revenue_{i,H}^{S\$} - Floor_H \geq 0 \\ Total\ Buyers_H &= Total\ Buyers_H + 1, \text{ if } Revenue_{i,H}^{S\$} - Floor_H < 0 \end{aligned} \tag{4}$$

Upon the completion of a trade horizon, each trading PLO, whether a buyer or a seller, pays an amount equal to the current floor value to the SSM as a participation fee. At market equilibrium, the total supply of spectrum dollars is equal to their total demand in the SSM, as given in Eq. (5)

$$\sum_{Buyer=1}^{Total\,Buyers}\left(Floor_H - Revenue_{Buyer,H}^{S\$}\right) = \sum_{Seller=1}^{Total\,Sellers}\left(Revenue_{Seller,H}^{S\$} - Floor_H\right), \tag{5}$$

where $\sum_{Buyer=1}^{Total\,Buyers}(Floor_H - Revenue_{Buyer,H}^{S\$})$ is the spectrum dollar demand function, while the spectrum dollar supply function is given as $\sum_{Seller=1}^{Total\,Sellers}(Revenue_{Seller,H}^{S\$} - Floor_H)$. These functions provide the total number of spectrum dollars to be exchanged among trading PLOs for a given trading horizon. We would like to stress the fact that the expressions given in Eq. (5) comprise two important contributions. The cumulative sum of spectrum dollars, available for sale with $Total\,Sellers_H$ is given by $\sum_{Seller=1}^{Total\,Sellers} Revenue_{Seller,H}^{S\$}$. Also, there are $Total\,Sellers_H \times Floor_H$ spectrum dollars with the selling PLOs that are not available for sale but must be paid to the SSM as participation fees. Similarly, the cumulative sum of spectrum dollars to be purchased by the $Total\,Buyers_H$ in the SSM is given by $\sum_{Buyer=1}^{Total\,Buyers} Revenue_{Buyer,H}^{S\$}$. Additionally, $Total\,Buyers_H \times Floor_H$ spectrum dollars are to be paid by the buying PLOs to the SSM. Therefore, Eq. (5) can also be rewritten as Eq. (6)

$$\sum_{Seller=1}^{Total\,Sellers}\left(Revenue_{Seller,H}^{S\$}\right) - Total\,Sellers \times Floor_H = Total\,Buyers_H \times Floor_H - \sum_{Buyer=1}^{Total\,Buyers}\left(Revenue_{Buyer,H}^{S\$}\right). \tag{6}$$

The price of a unit spectrum dollar in USD can be determined by using either side (the demand or the supply functions) of Eq. (6)

$$P_H^{S\$} = \sum_{Seller=1}^{Total\,Sellers}\left(Revenue_{Seller,H}^{S\$}\right) - \left(Total\,Sellers_H \times Floor_H\right) = \left(Total\,Buyers_H \times Floor_H\right) - \sum_{Buyer=1}^{Total\,Buyers}\left(Revenue_{Buyer,H}^{S\$}\right), \tag{7}$$

where $Floor_H$ is given by Eq. (3) and Eq. (7) provides the optimum spectrum dollar price in the SSM based upon the total number of buyers/sellers and the floor price, while the floor price is based on total revenue and the total number of participants. The spectrum dollar price decreases with an increase in the quantity supplied and increases with an increase in the quantity demanded, and vice versa. Let $Profitability_i^{USD}$ represent the profitability in USD for a participating PLO i in the SSM. We can express its quantity as given by Eq. (8)

$$Profitability_i^{USD} = P_H^{S\$} \times \left(Revenue_{i,H}^{S\$} - Floor_H\right). \tag{8}$$

$Profitability_i^{USD}$ results in a negative value for PLOs that make a loss in the SSM for a given trade horizon. Figure 3 presents the flowchart for the spectrum dollar unit price and revenue calculation (in spectrum dollars). These metrics determine PLO profitability in USD at the end of the process.

5. Simulation Results and Analysis

5.1. Preliminaries. The proposed system is based on a periodic one-shot sealed-bid auction [41]. We consider two established spectrum-sharing techniques: (a) the carrot-and-stick-based allocation technique [17] and (b) the QoS-based allocation technique [18]. The revenue information of these two models is assumed to be released in spectrum dollars to enable the implementation of the floor-and-trade methodology in this work. In the simulations for both models, a uniform distribution of break-even prices has been used, and SSM participation will be valued for each individual participating PLO for a total of ten participating PLOs in the simulations. The optimum bidding price is calculated based on the sum of the probability of a participating PLO winning a bid either on or below the margin using the uniformly distributed break-even prices. The description of the spectrum auction bidding process is beyond the scope of this paper, and interested readers can find it in [17]. The model considers the PLO's will to participate in the SSM to determine its level of participation and its break-even cost as the reserve price below which it decides against a sale in the SSM. Hence, all participating PLOs do not offer all of the spectrum spaces available to the SSM due to the economic parameters. The allocations are performed after the auctioning process, that is, after the determination of the winners, the secondary spectrum pricing, and the related payments. Participating PLOs that win the bids either on or below the margin generate revenues in the SSM, while those that lose the bids do not get the opportunity to generate any revenue at that bidding instant but may participate in future bids. PLOs' profitability is calculated as the revenues that are generated beyond their break-even costs. The simulations were run for a complete bidding horizon containing 24 bidding instants, each corresponding to an hourly bid in the SSM.

5.2. Simulation and Results. Figure 4 presents the $Total\,Revenue_H^{S\$}$ calculated in spectrum dollars using Eq. (3) for ten PLOs participating in the SSM trade. The graph shows that the total revenue of the participating PLOs is higher for the QoS optimization-based secondary spectrum access model than that for carrot-and-stick model. This is because the carrot-and-stick model considers only quantity [17], while the QoS-based methodology considers both the quantity and quality of the traded spectrum during the bid allocation process. Pricing of the bids considering both quantity and

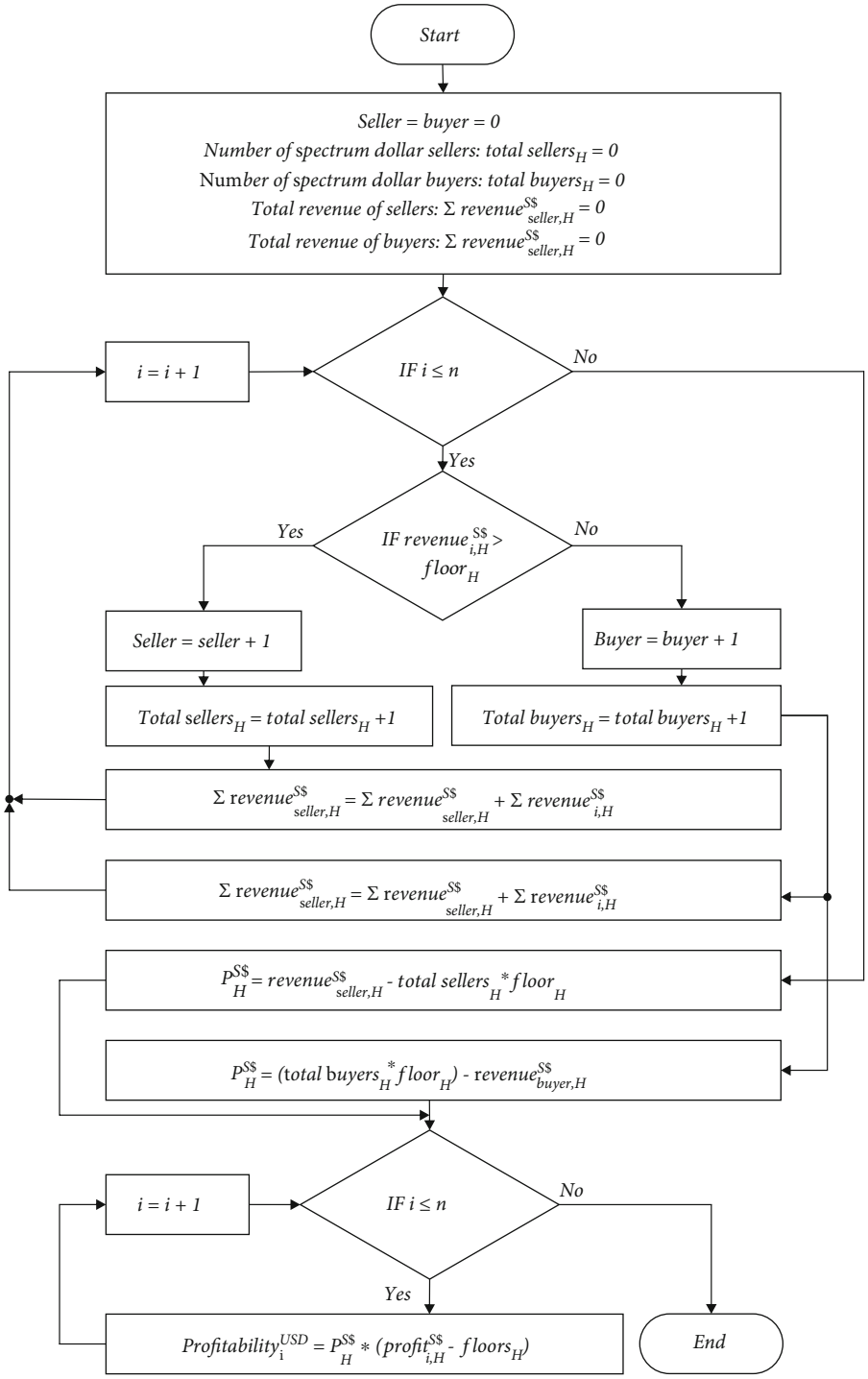


Figure 3: Flowchart for the spectrum dollar price and the revenue calculation in USD using the floor-and-trade rule.

quality increases the associated unit secondary spectrum costs in the SSM as proven in [18]. The corresponding revenues generated by each of the trading PLOs for the given trading horizon are shown in Table 2. PLOs 2, 5, 7, and 10 were unable to win any bids during the bidding horizon due to their high break-even costs, and hence, the bidding prices did not generate any revenue over the horizon. Moreover, it can be observed that for the QoS-based allocation model, the spectrum dollar price $P^{S\$}_H$ is higher, and at the same time, the number of spectrum dollars to be exchanged is higher as well. This confirms the theoretical foundation laid in Section 3.

The respective floor values $Floor_H$ for the total revenues generated by each model are shown in Figure 5. It can be seen that an increase in the floor value is proportionate to the respective total revenue values in spectrum dollars for each

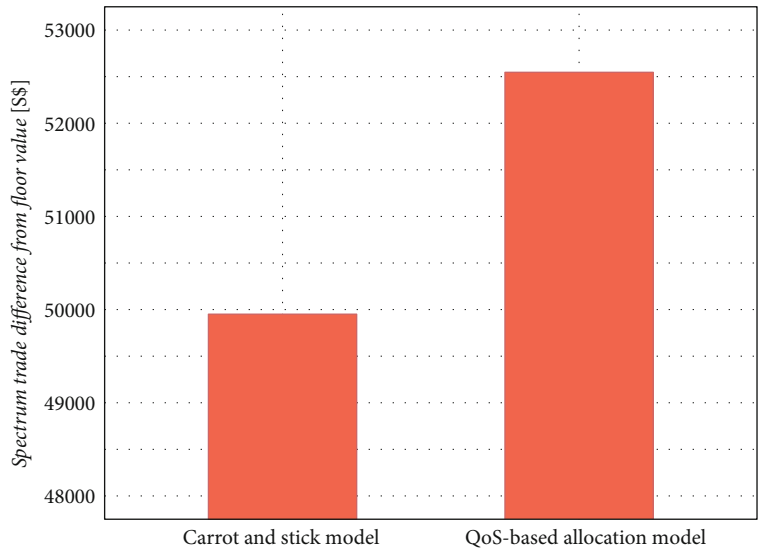


Figure 4: Total revenue of all participating PLOs in spectrum dollars for the secondary radio spectrum trade.

Table 2: Revenue for individual participating PLOs and total revenue for one trading horizon in the SSM.

PLO #	Revenue for carrot-and-stick model (S$)	Revenue for QoS-based allocation model (S$)
1	6925.09	5796.11
2	0	0
3	8605.64	7146.72
4	958.55	4626.51
5	0	0
6	12378.8	9830.56
7	0	0
8	9948.94	9909.52
0	11133.58	15236.64
10	0	0
Total revenue	**49950.6**	**52546.06**

trade model. This means that the higher the total revenue is, the higher the calculated floor value given the same number of PLOs trading in the SSM; thus, the QoS-based model has a higher floor value.

Comparing the revenues generated by each PLO to the corresponding floor value of each model, it can be seen that in both cases, 5 PLOs generated more revenues from SSM trade than the current floor value and are hence profitable, while the other 5 PLOs, which generated less revenue than the floor value, make a loss. Using the values for total revenue in spectrum dollars, floor values, and the number of spectrum dollar sellers and buyers, the spectrum dollar price is determined using either the demand or the supply function given in Eq. (8). The results are shown for both spectrum allocation methodologies in Table 3.

The corresponding revenue generated by each of the participating PLOs in the SSM trade for the trading horizon using the carrot and stick-based secondary spectrum trade model, with a floor value $Floor_H = 4995.06$ and respective spectrum dollar price $P_H^{S\$} = 24975.29$, is shown in Table 4. The sum of the profit/loss values for the secondary spectrum trade in the SSM for a given trade horizon is zero, which confirms the theoretical expectation introduced in Section 3.

Similarly, the same simulation scenario using the QoS optimization-based trade model is given in Table 5. In this case, the floor value $Floor_H = 5254$ and respective spectrum dollar price $P_H^{S\$} = 26273$, as shown in Table 5.

To obtain more generalized results, the number of trading PLOs in the SSM was increased to one hundred ($n = 100$), which trade different numbers of spectrum spaces. The floor value for $n = 100$ trading PLOs was calculated to be $Floor_H = 55$, with 46 PLOs as sellers and the other 54 PLOs as the buyers of spectrum dollars. As shown in Table 6, the unit spectrum trade price in the SSM is set to unity, generating $Total\ Revenue_H^{S\$} = 5072$ spectrum dollars, with the spectrum dollar price equal to $P_H^{S\$} = 2860$ USD.

Figure 6 shows PLOs' spectrum trade values relative to the floor value set in the SSM, while Figure 7 shows the corresponding relative revenue in USD for PLOs trading under the floor-and-trade methodology. The profitability shares of participating PLOs in USD depend upon their respective revenue differences in spectrum dollars from the floor value. The spectrum dollar price is calculated based on the demand and supply equilibrium value of spectrum dollars (for example, PLO 16 in Figure 7) at which participating PLOs make neither a profit nor a loss; in other words, the price trades the spectrum slots at a value exactly equal to the floor value in the SSM.

5.3. Sensitivity Analysis. A sensitivity analysis was performed on the proposed floor-and-trade model to analyse the effects of variation in the individual parameters. Some of the findings of the sensitivity analysis are as follows:

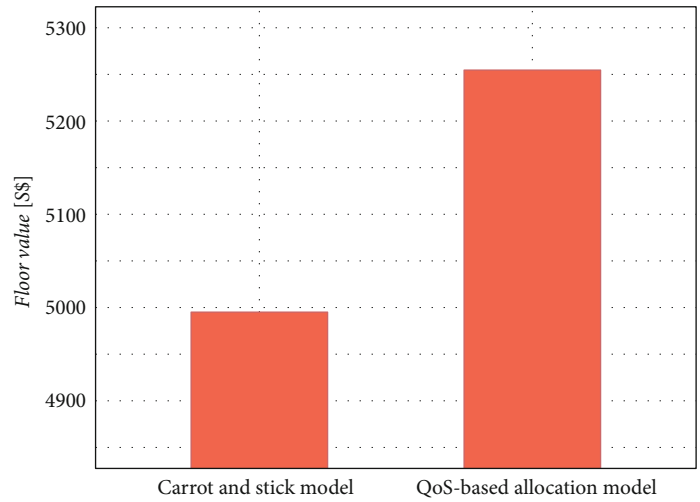


Figure 5: Floor values for each secondary spectrum trade methodology in the SSM.

Table 3: Revenue in spectrum dollars, respective floor values, and spectrum dollar prices for each secondary spectrum trade methodology.

Parameter	Carrot-and-stick model	QoS-based allocation model
Total spectrum dollars (S$)	49950.6	52546.06
PLOs participating in the SSM	10	10
Floor value	4995.06	5254.60
No.of spectrum dollar sellers	5	5
No.of spectrum dollar buyers	5	5
No.of spectrum dollar to exchange	24975.3	26273
Spectrum dollar price (USD)	24975.29	26273

Table 4: Revenue, respective difference with floor value, and profit/loss made by each participating PLO using the carrot and stick-based trade model.

PLO #	Revenue (S$)	Difference from floor (S$)	Profit/loss made (S$)
1	6925.09	1930.03	48203078.26
2	0	-4995.06	-124753122
3	8605.64	3610.58	90175318.67
4	958.55	-4036.51	-100813048.2
5	0	-4995.06	-124753122
6	12378.8	7383.74	184411121.6
7	0	-4995.06	-124753122
8	9948.94	4953.88	123724639.2
9	11133.58	6138.52	153311378.6
10	0	-4995.06	-124753122

Table 5: Revenue, respective difference with floor value, and profit/loss made by each participating PLO using the QoS optimization-based trade model.

PLO #	Revenue (S$)	Difference from floor (S$)	Profit/loss made (S$)
1	5796.11	541.504	14226950.84
2	0	-5254.606	-138054421.1
3	7146.72	1892.114	49711567.89
4	4626.51	-628.096	-16501985.05
5	0	-5254.606	-138054421.1
6	9830.56	4575.954	120224176.7
7	0	-5254.606	-138054421.1
8	9909.52	4654.914	122298695.2
9	15236.64	9982.034	262258278.7
10	0	-5254.606	-138054421.1

Table 6: Floor-and-trade simulation for 100 trading PLOs in the SSM.

No. of PLOs	No. of sellers	Total revenue (S$)	Floor value	Spectrum dollar price (USD)
100	48	5500	55	2860

(i) The floor value increases with an increase in total revenue earnings, while it decreases with an increase in the number of PLOs trading in the SSM

(ii) The spectrum dollar price increases with an increase in the floor value, and vice versa

(iii) The spectrum dollar price and the profitability shares of all participating PLOs in USD are zero, with all participating PLOs being sellers and none being buyers in the SSM. This is intuitive, as in this case, the earnings for each of the participating PLOs are equivalent to the floor value, and there are no spectrum dollars to exchange

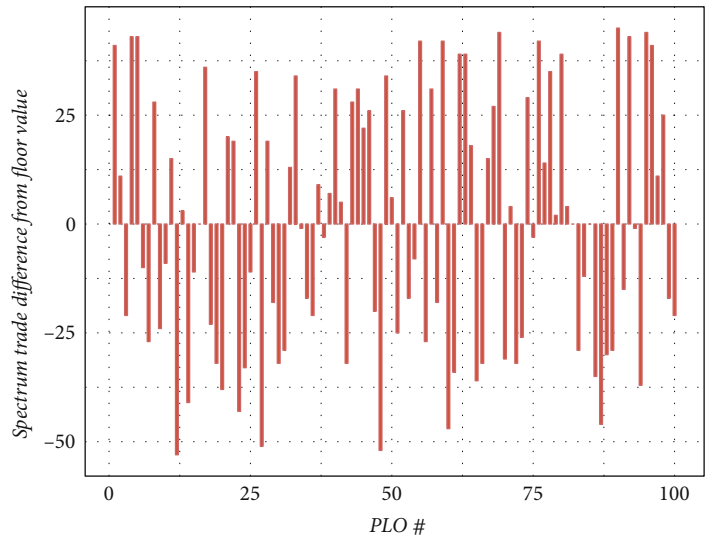


Figure 6: Spectrum trade by PLOs relative to the floor value.

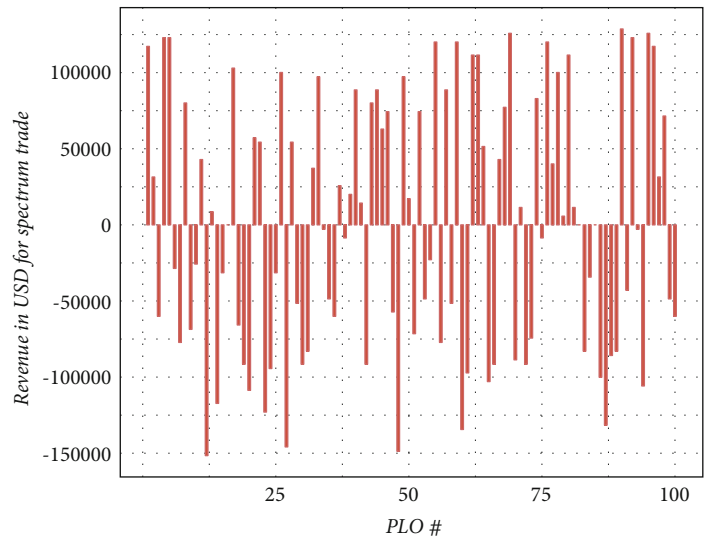


Figure 7: Revenue in USD for spectrum trade by PLOs relative to the floor value.

(iv) The spectrum dollar price is highest when there are no spectrum dollar sellers and all participating PLOs are spectrum dollar buyers

The floor-and-trade methodology automates a value-based reward-and-punishment mechanism for the underlying secondary spectrum trade. The highly competitive environment for the exchange of spectrum dollars using the floor-and-trade methodology maximizes the reward of the participating PLOs that bring the most value to the SSM. It simultaneously maximizes the punishment of those that withhold access to their available spectrum for secondary use. The participation fee and spectrum dollar price are calculated in real-time based on the underlying secondary spectrum trade behaviours in the SSM.

6. Conclusions

We have proposed a blockchain-based model of a secondary spectrum market based on spectrum dollar tokens. In the proposed model, the floor-and-trade rule is applied to regulate spectrum dollar pricing depending on the performance of the overall trade in the SSM rather than that of individual PLOs. The spectrum dollar price increases with a decrease in the number of spectrum dollar sellers or an increase in the number of spectrum dollar buyers, and vice versa. The floor-and-trade rule-based methodology maximizes the reward for the largest contributors and simultaneously maximizes the punishment for those making the smallest contributions to enabling radio spectrum reuse. The automated process of the floor-and-trade rule-based methodology

minimizes the monitoring overhead of secondary spectrum trade reporting. Spectrum dollar pricing and exchange minimize the control of participating PLOs over the economic parameters of the underlying secondary spectrum trade. PLOs cannot manipulate the economic parameters of the secondary spectrum trade, and hence, the chances for the establishment of a monopoly over the secondary spectrum resource are annulled. SSMs can base their secondary spectrum allocations purely on the QoS values offered to them. In the future, SSMs may decide to sell their earnings to meet their operating costs and/or make capital investments to enable radio spectrum reuse.

Data Availability

The data used to support the findings of this study have been deposited in the Hindawi repository on github repo (https://github.com/JurajGazda/Hindawi).

Conflicts of Interest

The authors declare no conflict of interest regarding this publication.

Acknowledgments

This research was funded by the Slovak Research and Development Agency, projects APVV-18-0214 and APVV-15-0055 and by the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of Slovakia under the contract VEGA 1/0268/19. This research was also supported by Ukrainian government project No.0120U100674 "Designing the novel decentralized mobile network based on blockchain architecture and artificial intelligence for 5G/6G development in Ukraine".

References

[1] E. Hossain, D. Niyato, and Z. Han, "Dynamic Spectrum Access and Management in Cognitive Radio Networks," *Books 24x7*, Cambridge university press, 1st edition, 2005.

[2] V. Mishra, L. C. Tong, and C. Syin, "QoS based spectrum decision framework for cognitive radio networks," in *18th IEEE International Conference on Networks (ICON)*, 2012.

[3] D. S. Aldar, "Centralized integrated Spectrum sensing for cognitive radios," *International Journal of Computer Science and Communication*, vol. 1, no. 2, pp. 227–229, 2010.

[4] M. Khan, B. Altaf, and M. Moghal, *Introduction, Decision Making Techniques for Cognitive Radios*, Lambert Acad. Publishing, 1st edition, 2012.

[5] M. R. Moghal, M. A. Khan, and H. A. Bhatti, "Spectrum optimization in cognitive radios using elitism in genetic algorithms," in *2010 6th International Conference on Emerging Technologies (ICET)*, pp. 49–54, IEEE, 2010.

[6] T. Maksymyuk, M. Brych, M. Klymash, and M. Jo, "Cooperative channels allocation in unlicensed spectrum for D2D assisted 5G cellular network," in *2017 2nd International Conference on Advanced Information and Communication Technologies (AICT)*, pp. 197–200, 2017.

[7] M. A. Khan and S. Ahmed, *Decision Making Techniques in Cognitive Radios*, Lambert Publishing, 2008.

[8] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, 2006.

[9] P. Pawelczak, *Opportunistic Spectrum Access: Designing link and transport layer*, Dept. of Telecommunications, TU Delft Univ. of Tech, 2009.

[10] P. R. Gronsund, O. Grondalen, and M. Lahteenoja, "Business case evaluations for LTE network offloading with cognitive femtocells," *Telecommunications Policy*, vol. 37, no. 2-3, pp. 140–153, 2013.

[11] P. R. Gronsund, *Cognitive Radio from a Mobile Operators Perspective: System Performance and Business Case Evalua- Tions*, Doctoral Dissertation, Mathematics and Natural Sciences Dept., Univ of Oslo, 2013.

[12] D. Horváth, V. Gazda, and J. Gazda, "Agent-based modeling of the cooperative spectrum management with insurance in cognitive radio networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, 2013.

[13] O. Raoof, Z. Al-Banna, and H. S. Al-Raweshidy, "Competitive Spectrum Sharing in Wireless Networks: A Dynamic Non-Cooperative Game Approach," in *Wireless and Mobile Networking*, pp. 197–207, IFIP Advances in Information and Communication Technology, 2009.

[14] G. Bugár, M. Vološin, T. Maksymyuk et al., "Techno-economic framework for dynamic operator selection in a multi-tier heterogeneous network," *Ad Hoc Networks*, vol. 97, p. 102007, 2020.

[15] X. Zhou and H. Zheng, *TRUST: A General Framework for Truthful Double Spectrum Auction*, INFOCOM IEEE, 2012.

[16] Y. N. Shnaiwer, S. A. Zummo, W. Mesbah, and S. A. Ahmadi, "Cost-efficient secondary users grouping for two-tier cognitive radio networks," *Physical Communication*, vol. 25, no. 1, pp. 1–13, 2017.

[17] M. A. Khan and M. M. Jamali, "Carrot and stick model for dynamic secondary radio spectrum trade with QoS optimization," *Physical Communication*, vol. 29, pp. 203–216, 2018.

[18] M. A. Khan and M. M. Jamali, "Quality of Service optimization-based dynamic secondary spectrum access model using Genetic Algorithms," *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 8, 2017.

[19] I. Anjum, M. A. Razzaque, M. M. Hassan, A. Alelaiwi, and S. M. Mizanur Rahman, "Quality-of-service-aware weighted-fair medium access control protocol for coexisting cognitive radio networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, 2016.

[20] Z. Tabakovic and M. Grgic, "Cognitive radio frequency assignment with interference weighting and categorization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, 2016.

[21] L. Zhai, H. Wang, and C. Gao, "A Spectrum Access Based on Quality of Service (QoS) in Cognitive Radio Networks," *PLoS One*, vol. 11, no. 5, 2016.

[22] A. Elarfaoui and N. Elalami, "Optimization of QoS parameters in cognitive radio using combination of two crossover methods in genetic algorithm," *International Journal of Communications, Network and System Sciences*, vol. 16, no. 11, pp. 478–483, 2013.

[23] M. Pan, J. Sun, and Y. Fang, "Purging the back-room dealing: secure spectrum auction leveraging paillier cryptosystem," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 866–876, 2011.

[24] Z. Ji and K. Liu, "Multi-stage pricing game for collusion-resistant dynamic spectrum allocation," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 182–191, 2008.

[25] S. Sengupta and M. Chatterjee, "Designing auction mechanisms for dynamic spectrum access," *Mobile Networks and Applications*, vol. 13, no. 5, pp. 498–515, 2008.

[26] L. Gao, Y. Xu, and X. Wang, "Map: multiauctioneer progressive auction for dynamic spectrum access," *IEEE Transactions on Mobile Computing*, vol. 10, no. 8, pp. 1144–1161, 2011.

[27] L. Guijarro, V. Pla, J. R. Vidal, and J. Martinez-Bauset, "Entry, competition, and regulation in cognitive radio scenarios: a simple game theory model," *Mathematical Problems in Engineering*, vol. 2012, 13 pages, 2012.

[28] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, "Towards real-time dynamic spectrum auctions," *The International Journal of Computer and Telecommunications Networking*, vol. 52, no. 4, pp. 879–897, 2008.

[29] E. Cronin, A. R. Kurc, B. Filstrup, and S. Jamin, "An efficient synchronization mechanism for mirrored game architectures," *Multi- media Tools and Applications*, vol. 23, no. 1, pp. 7–30, 2004.

[30] D. Randall, P. Goel, and R. Abujamra, "Blockchain applications and use cases in health information technology," *Journal of Health& Medical Informatics*, vol. 8, no. 3, pp. 8–11, 2017.

[31] S. Huckle, R. Bhattacharya, M. White, and N. Beloff, "Internet of things, blockchain and shared economy applications," *Procedia Computer Science*, vol. 98, pp. 461–466, 2016.

[32] N. Kshetri, "Can blockchain strengthen the internet of things?," *IT Professional*, vol. 19, no. 4, pp. 68–72, 2017.

[33] M. Pilkington, *Blockchain Technology: Principles and Applications*, Research handbook on digital transformations Edward Elgar Publishing, 2016.

[34] M. B. Weiss, K. Werbach, D. C. Sicker, and C. E. C. Bastidas, "On the application of blockchains to spectrum management," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 2, pp. 193–205, 2019.

[35] J. Qiu, D. Grace, G. Ding, J. Yao, and Q. Wu, "Blockchain-based secure spectrum trading for unmanned-aerial-vehicle-assisted cellular networks: an operator's perspective," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 451–466, 2019.

[36] T. Maksymyuk, J. Gazda, L. Han, and M. Jo, "Blockchain-Based Intelligent Network Management for 5G and Beyond," in *2019 3rd International Conference on Advanced Information and Communications Technologies (AICT)*, pp. 36–39, 2019.

[37] S. Zheng, T. Han, Y. Jiang, and X. Ge, "Smart contract-based spectrum sharing transactions for multi-operators wireless communication networks," *IEEE Access*, vol. 8, pp. 88547–88557, 2020.

[38] Z. Zhou, X. Chen, Y. Zhang, and S. Mumtaz, "Blockchain-empowered secure spectrum sharing for 5g heterogeneous networks," *IEEE Network*, vol. 34, no. 1, pp. 24–31, 2020.

[39] U. U. Khan, N. Dilshad, M. H. Rehmani, and T. Umer, "Fairness in cognitive radio networks: models, measurement methods, applications, and future research directions," *Journal of Network and Computer Applications*, vol. 73, pp. 12–26, 2016.

[40] D. Dranove, C. Garthwaite, and C. Ody, *A Floor-And-Trade Proposal to Improve the Delivery of Charity-Care services by U.S. Nonprofit Hospitals*, Hamilton Project, Northwestern Univ, 2015.

[41] C. Wang and P. Guo, "Behavioral models for first-price sealed-bid auctions with the one-shot decision theory," *European Journal of Operational Research*, vol. 261, no. 3, pp. 994–1000, 2017.

Hindawi
Wireless Communications and Mobile Computing
Volume 2020, Article ID 6136298, 17 pages
https://doi.org/10.1155/2020/6136298

Research Article

Energy Harvesting and Information Transmission Mode Design for Cooperative EH-Abled IoT Applications in beyond 5G Networks

Hongyuan Gao, Shibo Zhang, Yumeng Su, and Ming Diao

College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Hongyuan Gao; gaohongyuan@hrbeu.edu.cn

Received 16 November 2019; Accepted 9 January 2020; Published 12 February 2020

Guest Editor: Juraj Gazda

Energy harvesting (EH) technology is considered to be a promising approach to provide enough energy for energy-constrained Internet of Things (IoT). In this paper, we propose an energy harvesting and information transmission mode for the spectrum sharing system with cooperative EH-abled IoT applications in beyond 5G networks. Different from most existing IoT spectrum-sharing research studies, in our system, both primary user (PU) and IoT devices (IDs) collect energy for their information transmission. In addition, for all IDs, they should realize two communication functions: working as relays to help the information transfer process of PU and completing their own information transmission. We analytically derive exact expressions for the throughput of the primary system and IoT system and then formulate two objective functions. It is easy to see that power splitting ratio, dynamic EH ratio, power sharing ratio, and relay selection should be optimized to get the best performance for different communication circumstances. Actually, it is a hybrid NP-hard problem to optimize these parameters and traditional algorithms cannot solve it well. Therefore, a novel algorithm-quantum whale optimization algorithm (QWOA) is proposed to obtain the best performance. Simulation results show the good performance of QWOA in different simulation scenarios.

1. Introduction

Internet of Things (IoT) is an emerging technique that can provide intelligent information exchange opportunity for smart devices in the future communication scenarios and the beyond 5G networks [1, 2]. With the development of IoT, energy efficiency and spectrum efficiency have been two essential concerns due to the rapidly growing number of IoT devices [3, 4]. Since the technique of IoT spectrum sharing is able to make better use of the spectrum resource, many experts and scholars have a deep study on it then apply to IoT systems [5–8]. Some IoT spectrum-sharing protocols are proposed in [9–11]. In the investigated IoT system, the licensed primary user (PU) can use its own spectrum if it needs at any time, and IoT devices (IDs) can share the PU's spectrum in the case that IDs have no worse effect on the PU. In [10], Khan et al. studied the spectrum-sharing IoT structure. Some future application scenarios, such as smart grid, smart city, and wireless sensor are investigated. An arbitrarily-shaped underlay cognitive network was investigated by Guo et al. [11]. To satisfy interference constraint and meet outage probability of the PU, a cooperative protocol was proposed. In [12], Aslam et al. designed opportunistic spectrum sharing protocol, frame structure, and energy consumption function.

In order to prolong the life of the network, EH is seen as an effective technique which is widely applied in energy-limited environment, where the batteries of communication devices are hard to recharge or replace [13, 14]. EH technique can supply enough energy for cooperative communication, information transfer, and many other telecommunication works [15–17]. To explore potential of EH, many study works have been conducted in recent years [18, 19]. To satisfy energy demand for the communication system, a fixed harvesting energy ratio was considered, and the communication system could harvest energy in one time slot for information transfer

in [20–22]. In [21], Gao et al. investigated an energy harvesting cooperative communication protocol. Network users can collect energy from ambient environment and complete their own information transmission. In [22], relay nodes can harvest energy from wireless signals. Maji et al. presented the mathematical expressions for the system outage probability.

The combination of IoT and EH, which is aimed to exert advantage of them, has been a hot topic in recent years [23, 24]. It is necessary to investigate the EH-IoT applications because some IoT scenarios are limited by energy constraint and cannot have a good performance [25–27]. An EH-abled IoT system which consists of IoT devices and an access point were considered by Lee and Lee [26]. Hybrid energy resources, i.e., wireless signals and on-grid energy resource were combined to realize information transmission. Gurjar et al. [27] proposed an energy transfer approach in an overlay cognitive IoT system. The system throughput, outage probability, and energy efficiency were investigated. Human kinetic energy is also seen as an energy resource by Ju et al. [28]. The power allocation structure, sink selection, and power control blocks were discussed. Research studies [26–28] illustrated the effectiveness of EH technology in IoT applications.

For improving quality of service (QoS) and extend network communication range, relay cooperative communication is an essential technique in modern communication which can be combined with spectrum-sharing networks to further improve the performance [29–32]. Some works technically studied the relay selection problem for cooperative networks [33–35]. In [33], to measure performance of the system more comprehensively, Atapattu et al. considered the minimum achievable rate of the system. To further exploit benefits of relay cooperative communication, there are more and more studies on relay selection problem in recent years. For maximizing sum rate for the whole network in a multiuser network, Sharma et al. [35] presented a relay selection protocol for the cooperative ad hoc network. The optimization problem was assigning available relay nodes for different communication nodes. However, common channel interference (CCI) was not considered in [33, 35]. To find a relay selection scheme with CCI, Singh et al. [34] presented a new relay selection scheme considering CCI among multiple users. However, the performance of the proposed protocol is not optimal solution and only a suboptimal solution is obtained.

Most existing works as mentioned above mainly discussed about designing a better system of spectrum-sharing IoT [8, 10, 12] or EH-abled IoT [19, 27, 28]. The research studies about cooperative information exchange in EH-abled IoT with spectrum sharing are few. However, for some IoT distance transmission situations, cooperative relay transmission is necessary and can greatly improve network performance [29, 36, 37]. Developing the cooperative information transmission mode for the EH-abled IoT with spectrum sharing is essential. Besides, considering the problem of existing wireless EH instability (susceptible to the channel gain), it is necessary to design a dynamic EH time slot structure to collect energy (ensuring energy supply). A dynamic EH time slot structure can maintain the stability of system information transmission and EH.

In this paper, we consider a cooperative EH-abled IoT with spectrum sharing (CEH-IoT-SS), where each ID can not only share the licensed spectrum of the PU to transmit its own information but also can help PU's information transmission. For different practical requirements, two objective functions are formulated and the main contribution of this paper can be summarized as follows:

(i) We propose a novel CEH-IoT-SS model for existing IoT applications. Different from models in existence, there are multiple IDs and single PU in CEH-IoT-SS and some IDs can work as relays to help PU's information transfer. At the same time, each ID transmits its own information during one time slot. The PU is composed of a primary transmitter (PT) and a primary receiver (PR). The PT collects energy from ambient energy source while each IoT device transmitter (IDT) harvests energy from ambient energy source and a portion of received PT's signal. The energy harvested from PT's signal for each ID can be seen as a transmission energy supplement.

(ii) We derive analytical expressions for the throughput of primary network and IDs. Two objective functions are given for different communication situations. Analytical expression indicates that the scheme of power splitting ratio, dynamic EH ratio, power sharing ratio, and relay selection should be optimized.

(iii) A new algorithm-quantum whale optimization algorithm (QWOA) is designed to optimize problems which are proposed. Actually, finding the relay selection scheme is a discrete optimization problem. However, the process for obtaining the best power splitting ratio, power sharing ratio, and dynamic EH ratio is solving a continuous optimization problem. Since it is a hybrid optimization problem, traditional continuous or discrete algorithms cannot get the best solution. Therefore, QWOA is proposed to solve it. Through simulation results, we can see QWOA has advantages which traditional algorithms do not have.

Other parts for this paper are structured in the following way. In Section 2, we will give our proposed CEH-IoT-SS model and time slot structure. For Section 3, expressions of throughput for primary network and IDs are derived. For different communication scenarios, two objective functions are formulated. In Section 4, QWOA is proposed to find the best power splitting ratio, dynamic EH ratio, power sharing ratio, and relay selection scheme for different communication scenarios. Simulation results and conclusion will be provided in Sections 5 and 6.

2. System Model and Time Slot Structure

In this part, a CEH-IoT-SS is considered. As shown in Figure 1, the primary network consists of one PT and one PR. PT tries to deliver information to PR. We consider the situation where PT cannot transmit information to PR directly so PT needs some relays to complete information transfer [27]. M pairs of IDs, which consist of M IDTs and M IoT device receivers (IDRs), implement two communication functions: (1) some IDTs serve as relays to help PT transmit information; (2) each IDT has its own information for transmitting to its corresponding IDR. Both PT and IDTs can harvest energy from ambient energy sources to complete their own information transmission. Besides, IDTs can scavenge energy from a portion of received signal from PT as transmission energy supplement. In the proposed CEH-IoT-SS, energy storage equipment can be a super capacitor and there is no energy rested at the end of each time slot [21, 38].

In this model, each information channel follows Rayleigh distribution, and parameters of different Rayleigh distributions are related with distance between two devices. The channel state information (CSI) between two devices is the same during the same time slot. Meanwhile, two channels are mutual independent from each other in one time slot and CSI of the same channel in two different time slots is independent. CSI from PT to the mth IDT $(m = 1, 2, \ldots, M)$ is represented as $G_{\mathrm{PT,IDT}_m}$, CSI from PT to the mth IDR is represented as $G_{\mathrm{PT,IDR}_m}$, CSI from the mth IDT to the mth IDR is represented as $G_{\mathrm{IDT}_m,\mathrm{IDR}_m}$, and CSI from the mth IDT to PR is represented as $G_{\mathrm{IDT}_m,\mathrm{PR}}$.

To simplify representation of the distance between two communication devices, we make definition as follows: the distance between PT and the mth IDT is denoted as $d_{\mathrm{PT,IDT}_m}$. Similarly, the distance between the PT and the mth IDR is denoted as $d_{\mathrm{PT,IDR}_m}$, the distance between the mth IDT to the mth IDR is denoted as $d_{\mathrm{IDT}_m,\mathrm{IDR}_m}$, and the distance between the mth IDT and PR is denoted as $d_{\mathrm{IDT}_m,\mathrm{PR}}$. The CSI $G_{\mathrm{PT,IDT}_m}$, $G_{\mathrm{PT,IDR}_m}$, $G_{\mathrm{IDT}_m,\mathrm{IDR}_m}$, and $G_{\mathrm{IDT}_m,\mathrm{PR}}$ follow exponentially distribution with parameters $d^{-\ell}_{\mathrm{PT,IDT}_m}$, $d^{-\ell}_{\mathrm{PT,IDR}_m}$, $d^{-\ell}_{\mathrm{IDT}_m,\mathrm{IDR}_m}$, and $d^{-\ell}_{\mathrm{IDT}_m,\mathrm{PR}}$, respectively. ℓ is the factor of path loss. For any two different exponentially distribution, it is independent from each other. In proposed CEH-IoT-SS, each PT transmits its information to PR by cooperative communication with IDTs, and IDTs also transmit their own information to corresponding IDRs. In addition, both PT and IDTs collect energy for information transmission. Figure 2 shows one time slot of EH and information transmission mode for the proposed CEH-IoT-SS. Each time slot for the CEH-IoT-SS is composed of three phases: (1) in the first phase $(0, \alpha T]$, both PT and IDTs harvest energy from ambient energy source, α is dynamic EH ratio, $0 < \alpha < 1$, and T is the duration of one time slot. (2) In the second phase $(\alpha T, (1+\alpha)T/2]$, PT transmits its information to IDTs and IDRs. One portion of each IDT's received signal is used for cooperative communication. The other portion of received signal is the collected energy by each IDT as the transmission energy supplement. (3) For the third phase $((1+\alpha)T/2, T]$, each IDT which is selected as a relay will transmit its own signal and received signal to IDR and PR simultaneously. Other IDTs transmit their own signals to their corresponding IDRs.

3. Throughput Analysis and Objective Function Design

This part we will derive expressions for the throughput of primary network and IDs in CEH-IoT-SS. For different communication needs, we formulate two objective functions.

3.1. EH of PT and IDT in the First Phase. For the first phase, both PT and IDTs collect energy from ambient energy source. The length of the first phase is αT, so energy stored by PT and IDTs can be shown as

$$E_{\mathrm{PT}} = Y_{\mathrm{PT}} \alpha T, \tag{1}$$

where Y_{PT} is EH rate of PT [38, 39].

Similarly, the energy stored for the mth IDT can be calculated as

$$E^{T_1}_{\mathrm{IDT}_m} = \alpha T Y_{\mathrm{IDT}_m}, \tag{2}$$

where Y_{IDT_m} is EH rate of the mth IDT. The EH process of PT is ended and PT no longer collects energy in one time slot. However, IDTs will still gather energy in the second phase. Since all the harvested energy is used up in one time slot, so transmit power provided by PT can be shown as

$$P_{\mathrm{PT}} = \frac{E_{\mathrm{PT}}}{(1-\alpha)T/2} = \frac{2Y_{\mathrm{PT}}\alpha}{1-\alpha}. \tag{3}$$

3.2. Information Transmission of PT in the Second Phase. In the second phase for one time slot, PT transmits its signal x to IDTs and IDRs, where x is normalized as $E|x|^2 = 1$. The received signals at the mth $(m = 1, 2, \ldots, M)$ IDT and the mth IDR are shown as follows:

$$y_{\mathrm{IDT}_m} = \sqrt{P_{\mathrm{PT}}} G_{\mathrm{PT,IDT}_m} x + n_1, \tag{4}$$

$$y_{\mathrm{IDR}_m} = \sqrt{P_{\mathrm{PT}}} G_{\mathrm{PT,IDR}_m} x + n_2, \tag{5}$$

where n_1 is additive white Gaussian noise (AWGN), in which power is w_1, and n_2 is AWGN with power w_2.

On the basis of power splitting approach, the power splitting circuit of the mth IDT will divide the received signal into two parts, $\sqrt{\beta} y_{\mathrm{IDT}_m}$ and $\sqrt{1-\beta} y_{\mathrm{IDT}_m}$, which is used for EH and information transfer, respectively. $\beta\,(0 < \beta < 1)$ is the power splitting ratio. Hence, the portion of the received signal for EH can be given by

$$y^{\mathrm{EH}}_{\mathrm{IDT}_m} = \sqrt{\beta} y_{\mathrm{IDT}_m} = \sqrt{P_{\mathrm{PT}} \beta} G_{\mathrm{PT,IDT}_m} x + n_1 \sqrt{\beta}. \tag{6}$$

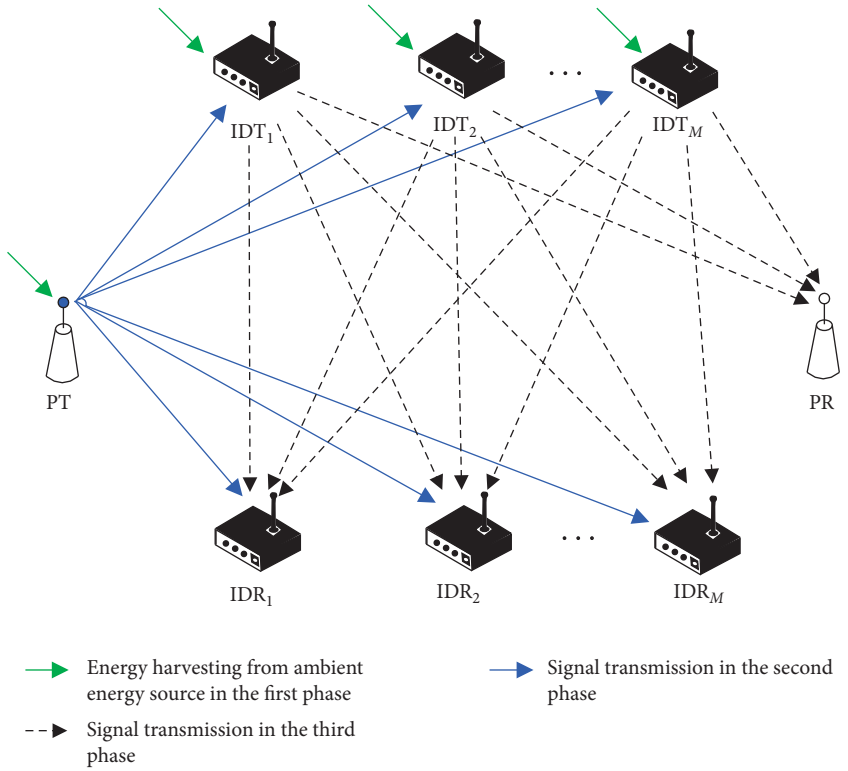


Figure 1: System model of CEH-IoT-SS.

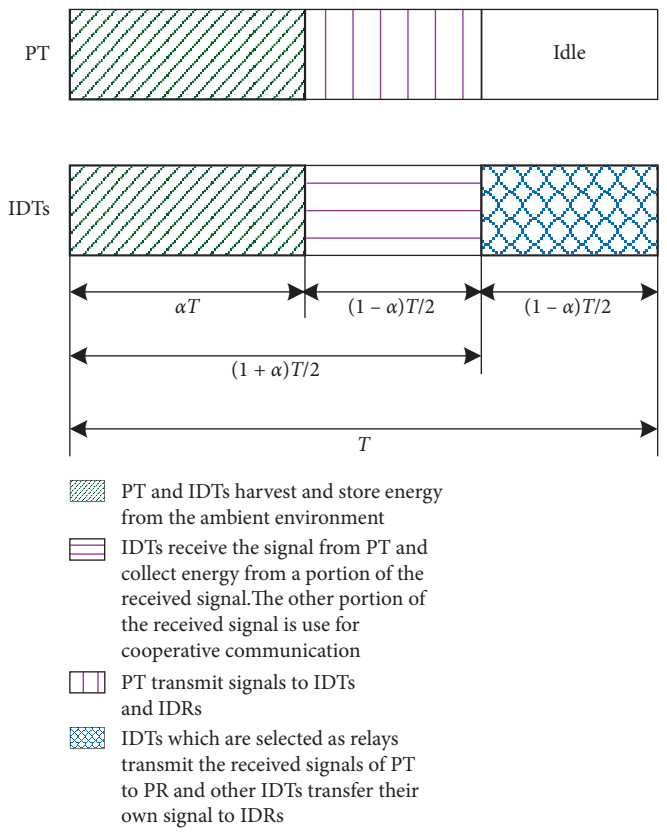


Figure 2: Structure of time slot.

Energy stored by the mth IDT at the second phase is calculated as

$$E_{\mathrm{IDT}_m}^{T_2} = \frac{(1-\alpha)T}{2}\rho P_{\mathrm{PT}}\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2 = T\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2, \tag{7}$$

where ρ is the efficiency of EH. Since collected energy from AWGN is small, it can be neglected. Therefore, total energy stored by the mth IDT in the first two phases can be shown as

$$E_{\mathrm{IDT}_m} = E_{\mathrm{IDT}_m}^{T_1} + E_{\mathrm{IDT}_m}^{T_2} = \alpha T Y_{\mathrm{IDT}_m} + T\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2. \tag{8}$$

The mth IDT utilizes all the harvested energy in the third phase to complete cooperative communication and its own information transfer, so the power at the mth IDT can be given by

$$P_{\mathrm{IDT}_m} = \frac{E_{\mathrm{IDT}_m}}{(1-\alpha)T/2} = \frac{2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2}{1-\alpha}. \tag{9}$$

On the other hand, the portion of the received signal at the mth IDT for information transfer is as follows:

$$\begin{aligned} y_{\mathrm{IDT}_m}^{\mathrm{IT}} &= \sqrt{1-\beta}\,y_{\mathrm{IDT}_m} + n_3 = \sqrt{P_{\mathrm{PT}}(1-\beta)}G_{\mathrm{PT,IDT}_m}x \\ &\quad + n_1\sqrt{(1-\beta)} + n_3, \end{aligned} \tag{10}$$

where n_3 is AWGN which is caused by the signal conversion from the passband to baseband [40] at the mth IDT and the power of n_3 is w_3.

The received signal at the mth IDT is needed to be normalized, and the power normalization factor ψ_m is in the following:

$$\psi_m = \frac{1}{\sqrt{P_{\mathrm{PT}}(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}}. \tag{11}$$

The received normalization signal of the mth IDT is

$$\begin{aligned} y_{\mathrm{IDT}_m}^{\mathrm{IT-Norm}} &= \psi_m y_{\mathrm{IDT}_m}^{\mathrm{IT}} = \frac{\sqrt{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)}G_{\mathrm{PT,IDT}_m}}{\sqrt{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}}x \\ &\quad + \frac{\sqrt{(1-\beta)}}{\sqrt{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}}n_1 \\ &\quad + \frac{1}{\sqrt{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}}n_3. \end{aligned} \tag{12}$$

3.3. Information Transmission of IDT in the Third Phase. In the third phase, some IDTs will be selected as relays. If the mth IDT is selected as a relay, the mth IDT transmits the received normalized-signal $y_{\mathrm{IDT}_m}^{\mathrm{IT-Norm}}$ and its own signal z_{IDT_m} by superposition coding in the following:

$$z_m = \sqrt{\sigma}\,y_{\mathrm{IDT}_m}^{\mathrm{IT-Norm}} + \sqrt{1-\sigma}\,z_{\mathrm{IDT}_m}, \tag{13}$$

where $\sigma\,(0<\sigma<1)$ is power sharing ratio, z_m is the combined signal which is divided into two parts. The first part $\sqrt{\sigma}\,y_{\mathrm{IDT}_m}^{\mathrm{IT-Norm}}$ is used to transmit PT's signal, and the second part $\sqrt{1-\sigma}\,z_{\mathrm{IDT}_m}$ is used to transmit IDT's own information. If the mth IDT is not selected as a relay, it just needs to transmit its own information to its corresponding IDR:

$$z_m = z_{\mathrm{IDT}_m}. \tag{14}$$

According to different communication functions of IDT, the signal transferred by the mth IDT can be shown as

$$z_m = \begin{cases} \sqrt{\sigma}\, y_{\mathrm{IDT}_m}^{\mathrm{IT-Norm}} + \sqrt{1-\sigma}\, z_{\mathrm{IDT}_m}, & \text{if IDT}_m \text{ is selected as relay,} \\ z_{\mathrm{IDT}_m}, & \text{if IDT}_m \text{ is not selected as relay.} \end{cases} \tag{15}$$

Therefore, the received signal at PR is given by

$$\begin{aligned} z_{\mathrm{PR}} &= \sum_{m=1}^{M} \sqrt{P_{\mathrm{IDT}_m}} G_{\mathrm{IDT}_m,\mathrm{PR}} z_m + n_4 \\ &= \sum_{m=1}^{M} b_m \frac{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))\sigma P_{\mathrm{IDT}_m}(1-\beta)}\, G_{\mathrm{IDT}_m,\mathrm{PR}} G_{\mathrm{PT,IDT}_m}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}} x \\ &\quad + \sum_{m=1}^{M} b_m \frac{\sqrt{\sigma P_{\mathrm{IDT}_m}(1-\beta)}\, G_{\mathrm{IDT}_m,\mathrm{PR}}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}} n_1 \\ &\quad + \sum_{m=1}^{M} b_m \frac{\sqrt{\sigma P_{\mathrm{IDT}_m}}\, G_{\mathrm{IDT}_m,\mathrm{PR}}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}} n_3 \\ &\quad + \sum_{m=1}^{M} b_m \sqrt{P_{\mathrm{IDT}_m}(1-\sigma)}\, G_{\mathrm{IDT}_m,\mathrm{PR}} z_{\mathrm{IDT}_m} \\ &\quad + \sum_{m=1}^{M} (1-b_m) G_{\mathrm{IDT}_m,\mathrm{PR}} \sqrt{P_{\mathrm{IDT}_m}}\, z_{\mathrm{IDT}_m} + n_4, \end{aligned} \tag{16}$$

where $b_m\,(m = 1, 2, \ldots, M)$ shows whether the mth IDT is selected as a relay for information transfer of PT. If the mth IDT is selected as a relay, $b_m = 1$. The relay selection scheme of IDTs can be shown as $\mathbf{b} = [b_1, b_2, \ldots, b_M]$. If the mth IDT is not chosen, $b_m = 0$. n_4 is AWGN with power w_4. $\sum_{m=1}^{M} b_m \left[\sqrt{\sigma P_{\mathrm{IDT}_m}(1-\beta)}\, G_{\mathrm{IDT}_m,\mathrm{PR}} / \sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)|G_{\mathrm{PT,IDT}_m}|^2 + (1-\beta)w_1 + w_3}\right] n_1$, $\sum_{m=1}^{M} b_m \left[\sqrt{\sigma P_{\mathrm{IDT}_m}}\, G_{\mathrm{IDT}_m,\mathrm{PR}} / \sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)|G_{\mathrm{PT,IDT}_m}|^2 + (1-\beta)w_1 + w_3}\right] n_3$, $\sum_{m=1}^{M} b_m \sqrt{P_{\mathrm{IDT}_m}(1-\sigma)}\, G_{\mathrm{IDT}_m,\mathrm{PR}} z_{\mathrm{IDT}_m}$, $\sum_{m=1}^{M} (1-b_m) G_{\mathrm{IDT}_m,\mathrm{PR}} \sqrt{P_{\mathrm{IDT}_m}}\, z_{\mathrm{IDT}_m}$, and n_4 are IDT signal interference. Therefore, the signal-to-interference plus noise ratio (SINR) of the PR is given by

$$\gamma_{\mathrm{PR}} = \frac{\left(\sum_{m=1}^{M} b_m \left(\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))\sigma P_{\mathrm{IDT}_m}(1-\beta)}\, G_{\mathrm{IDT}_m,\mathrm{PR}} G_{\mathrm{PT,IDT}_m} / \sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}\right)\right)^2}{\left(\sum_{m=1}^{M} b_m \left(w_1 \sigma P_{\mathrm{IDT}_m}(1-\beta)\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2 / (2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3\right) + \sum_{m=1}^{M} b_m \left(w_3 \sigma P_{\mathrm{IDT}_m}\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2 / (2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3\right) + \sum_{m=1}^{M} b_m P_{\mathrm{IDT}_m}(1-\sigma)\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2 + \sum_{m=1}^{M} (1-b_m)\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2 P_{\mathrm{IDT}_m} + w_4\right)}. \tag{17}$$

According to (9) and (17), the throughput of PR, i.e., the throughput of the primary network can be shown as follows:

$$R_{\mathrm{PR}} = \frac{1-\alpha}{2}\log_2\left(1+\gamma_{\mathrm{PR}}\right) = \frac{1-\alpha}{2}\log_2\left(1+\frac{\mathrm{Signal}_{\mathrm{PR}}}{\mathrm{Interference}_{\mathrm{PR}}}\right),$$

$$\mathrm{Signal}_{\mathrm{PR}} = \left(\sum_{m=1}^{M} b_m \frac{\sqrt{\left(\left(4\alpha Y_{\mathrm{IDT}_m} + 4\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2\right)/(1-\alpha)^2\right) Y_{\mathrm{PT}}\alpha\sigma(1-\beta)}\, G_{\mathrm{IDT}_m,\mathrm{PR}} G_{\mathrm{PT,IDT}_m}}{\sqrt{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3}}\right)^2,$$

$$\begin{aligned}
\mathrm{Interference}_{\mathrm{PR}} = {} & \sum_{m=1}^{M} b_m \frac{\left(\left(2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2\right)/(1-\alpha)\right) w_1\sigma(1-\beta)\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2}{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3} \\
& + \sum_{m=1}^{M} b_m \frac{\left(\left(2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2\right)/(1-\alpha)\right) w_3\sigma\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2}{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{PT,IDT}_m}\right|^2 + (1-\beta)w_1 + w_3} \\
& + \sum_{m=1}^{M} b_m \frac{2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2}{1-\alpha}(1-\sigma)\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2 \\
& + \sum_{m=1}^{M} (1-b_m)\left|G_{\mathrm{IDT}_m,\mathrm{PR}}\right|^2 \frac{2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2}{1-\alpha} + w_4.
\end{aligned} \tag{18}$$

Similarly, the average throughput of M IDT-IDR transmission pairs in CEH-IoT-SS is expressed as follows:

$$R_{\mathrm{IDT-IDR}_{\mathrm{average}}} = \frac{1-\alpha}{2M}\sum_{m=1}^{M}\log_2\left(1+\frac{\mathrm{Signal}_{\mathrm{IDR}_m}}{\mathrm{Interference}_{\mathrm{IDR}_m}}\right),$$

$$\mathrm{Signal}_{\mathrm{IDR}_m} = \begin{cases} \dfrac{2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2}{1-\alpha}\left|G_{\mathrm{IDT}_m,\mathrm{IDR}_m}\right|^2, & \text{if } \mathrm{IDT}_m \text{ is not selected as relay,} \\ \dfrac{2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2}{1-\alpha}(1-\sigma)\left|G_{\mathrm{IDT}_m,\mathrm{IDR}_m}\right|^2, & \text{if } \mathrm{IDT}_m \text{ is selected as relay,} \end{cases}$$

$$\begin{aligned}
\mathrm{Interference}_{\mathrm{IDR}_m} = {} & \sum_{i=1,i\neq m}^{M} b_i\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 \frac{2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2}{1-\alpha}(1-\sigma) \\
& + \sum_{i=1,i\neq m}^{M} b_i \frac{\sigma\left(\left(2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2\right)/1-\alpha\right)(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_1}{\left(2Y_{\mathrm{PT}}\alpha/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{PT,IDT}_i}\right|^2 + (1-\beta)w_1 + w_3} \\
& + \sum_{i=1,i\neq m}^{M} b_i \frac{\sigma\left(\left(2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2\right)/(1-\alpha)\right)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_3}{\left(2Y_{\mathrm{PT}}\alpha/1-\alpha\right)(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 + (1-\beta)w_1 + w_3} \\
& + \sum_{i=1,i\neq m}^{M} (1-b_i)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 \frac{2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2}{1-\alpha} + w_5,
\end{aligned} \tag{19}$$

where w_5 is the AWGN power.

Proof. See Appendix. □

3.4. Objective Function Design for Different Communication Requirements. According to the throughput results of the primary network and IDT-IDR transmission pairs, we design two objective functions for improving the data transmission rate in different communication scenarios.

3.4.1. Objective Function of IDT-IDR Transmission Pairs' Average Throughput with Primary Network Constraint. For this communication scenario, we aim to maximize the average throughput of IDT-IDR transmission pairs. At the same time, the performance of the primary network should satisfy requirement. That is to say, throughput of the primary network should be higher than a certain value. According to (18), we can find that the IDT-IDR transmission pairs' throughput is impacted by power splitting ratio, dynamic EH ratio, power sharing ratio, and relay selection scheme. Therefore, we focus on optimizing these parameters to maximize the primary network. Under this constraint of the primary network, we aim to maximize the average throughput of IDT-IDR transmission pairs. Combine (18) and (19), objective function can be given by

$$\begin{aligned} \max \quad & R_{\text{IDT-IDR}_{\text{average}}}(\mathbf{b}, \alpha, \beta, \sigma) \\ \text{subject to} \quad & (1)\, b_m \in \{0, 1\} \\ & (2)\, 0 < \alpha < 1 \\ & (3)\, 0 < \beta < 1 \\ & (4)\, 0 < \sigma < 1 \\ & (5)\, R_{\text{PR}}^{\text{require}} \le R_{\text{PR}}, \end{aligned} \tag{20}$$

where $R_{\text{PR}}^{\text{require}}$ is threshold that the throughput of the primary network should achieve.

3.4.2. Objective Function of Sum Throughput with Primary Network Constraint. In this situation, we aim to maximize sum throughput of the whole network with constraint of the primary network. Optimal power splitting ratio, dynamic EH ratio, power sharing ratio, and relay selection scheme should be found and objective function of sum throughput is formulated as follows:

$$\begin{aligned} \max \quad & R_{\text{sum}}(\mathbf{b}, \alpha, \beta, \sigma) = R_{\text{PR}} + R_{\text{IDT-IDR}_{\text{average}}} \\ \text{subject to} \quad & (1)\, b_m \in \{0, 1\} \\ & (2)\, 0 < \alpha < 1 \\ & (3)\, 0 < \beta < 1 \\ & (4)\, 0 < \sigma < 1 \\ & (5)\, R_{\text{PR}}^{\text{require}} \le R_{\text{PR}}. \end{aligned} \tag{21}$$

To obtain the best throughput of two communication circumstances proposed above is the process for solving the hybrid NP-hard optimization problem [41]. However, traditional intelligent algorithms cannot solve it well because of slow convergence speed and poor convergence accuracy. Hence, we proposed a novel algorithm-QWOA.

4. Optimization for Throughput Based on Quantum Whale Optimization Algorithm

For the problems which are proposed above, relay selection is a discrete optimization problem. However, getting the best power splitting ratio, dynamic EH ratio, and power sharing ratio is a progress of finding continuous solution. Since domains of them are different and dimensions are complex, traditional intelligent algorithms cannot be easily applied to get the optimal solution. Hence, a new intelligent algorithm-quantum whale optimization algorithm (QWOA) is proposed to solve it.

4.1. QWOA for Hybrid Optimization Problem. QWOA combines the advantages of whale optimization algorithm (WOA) [42] and quantum evolution theory of quantum computing [43–45]. Since relay selection is a discrete optimization problem and obtaining other parameters is a continuous optimization problem, for variables in different domains, we will design different evolution approaches. For the optimization problems (20) and (21), the number of dimension for the relay selection scheme is M (M represents the number of IDT-IDR transmission pairs). The number of continuous variables, i.e., power splitting ratio, dynamic EH ratio, and power sharing ratio, is three. Hence, the number of optimization problem dimension is $(M + 3)$. In an $(M + 3)$-dimensional space ($(M + 3)$ represents the number of dimension for the optimization problem), there are H quantum whales which are used to find the optimal solution. For each quantum whale, it consists of $(M + 3)$ quantum bits. The first M dimensions are used to optimize discrete variables, and the final three dimensions are used for optimizing continuous variables. Both discrete variables and continuous variables will evolve in different ways. The hth $(h = 1, 2, \ldots, H)$ quantum whale $\mathbf{x}_h^t$ in the tth iteration is given by

$$\mathbf{x}_h^t = \begin{bmatrix} \varphi_{h1}^t, \varphi_{h2}^t, \ldots, \varphi_{hi}^t, \ldots, \varphi_{hM}^t, \varphi_{h(M+1)}^t, \varphi_{h(M+2)}^t, \varphi_{h(M+3)}^t \\ \phi_{h1}^t, \phi_{h2}^t, \ldots, \phi_{hi}^t, \ldots, \phi_{hM}^t, \phi_{h(M+1)}^t, \phi_{h(M+2)}^t, \phi_{h(M+3)}^t \end{bmatrix}, \tag{22}$$

where $i = 1, 2, \ldots, M + 3$ and $|\varphi_{hi}^t|^2 + |\phi_{hi}^t|^2 = 1$. To lower computational complexity and improve evolutionary efficiency of QWOA, we define $0 \le \varphi_{hi}^t \le 1$ and $0 \le \phi_{hi}^t \le 1$. For each φ_{hi}^t, $\varphi_{hi}^t = \sqrt{1 - (\phi_{hi}^t)^2}$, each ϕ_{hi}^t can be computed by its corresponding φ_{hi}^t. Hence, $\mathbf{x}_h^t$ is simply expressed in the following:

$$\begin{aligned} \mathbf{x}_h^t &= \left[\varphi_{h1}^t, \varphi_{h2}^t, \ldots, \varphi_{hi}^t, \ldots, \varphi_{hM}^t, \varphi_{h(M+1)}^t, \varphi_{h(M+2)}^t, \varphi_{h(M+3)}^t\right] \\ &= \left[x_{h1}^t, x_{h2}^t, \ldots, x_{hi}^t, \ldots, x_{hM}^t, x_{h(M+1)}^t, x_{h(M+2)}^t, x_{h(M+3)}^t\right], \end{aligned} \tag{23}$$

where $i = 1, 2, \ldots, M + 3$ and $0 \le x_{hi}^t \le 1$. Each x_{hi}^t is the ith quantum bit of the hth quantum whale $\mathbf{x}_h^t$.

For the hth quantum whale $\mathbf{x}_h^t$, $x_{hi}^t (i = 1, 2, \ldots, M+3)$ should be measured to its own definition domain $\overline{x}_{hi}^t$ according to the following rules:

$$\overline{x}_{hi}^t = \begin{cases} 1, & \vartheta_{hi}^t > x_{hi}^t, \\ 0, & \vartheta_{hi}^t \le x_{hi}^t, \end{cases} \quad i = 1, 2, \ldots, M, \tag{24}$$

$$\overline{x}_{hi}^t = l_i + x_{hi}^t (u_i - l_i), \quad i = M+1, M+2, M+3, \tag{25}$$

where ϑ_{hi}^t is a uniform random number which is distributed from zero to one, l_i is the ith dimensional variant's lower bound for continuous variables, and u_i is the ith dimensional variant's upper bound for the continuous variables. After the process of mapping, we calculate fitness of each quantum whale according to the fitness function. Until the tth iteration, the global optimal quantum whale $\mathbf{p}_g^t$ of the whole population is shown as $\mathbf{p}_g^t = [p_{g1}^t, p_{g2}^t, \ldots, p_{gi}^t, \ldots, p_{gM}^t, p_{g(M+1)}^t, p_{g(M+2)}^t, p_{g(M+3)}^t]$, and its mapping state is $\overline{\mathbf{p}}_g^t = [\overline{p}_{g1}^t, \overline{p}_{g2}^t, \ldots, \overline{p}_{gi}^t, \ldots, \overline{p}_{gM}^t, \overline{p}_{g(M+1)}^t, \overline{p}_{g(M+2)}^t, \overline{p}_{g(M+3)}^t]$.

For the whole population, each quantum whale is evolved in different ways, which contains encircling prey, searching for prey, and bubble-net attacking. For each evolution mode, we will update their quantum rotation angle and quantum bit with different methods. During each iteration, we first calculate the coefficient vector $\mathbf{A}_h^t = [A_{h1}^t, A_{h2}^t, \ldots, A_{h(M+3)}^t]$ of the hth quantum whale, and the calculation rules can be shown by

$$A_{hi}^t = 2 \cdot \xi_1 \cdot \psi - \psi, \tag{26}$$

where ξ_1 is a uniform random number that is distributed from zero to one. $\psi = 1 - t/K$ and K is the maximum iteration number. For the hth quantum whale, we generate a uniform random number ξ_2 that is distributed from zero to one. When $\xi_2 < 0.5$ and $|\mathbf{A}_h^t| < 1$ ($|\bullet|$ is the modulus function), the hth quantum whale will encircle the prey and the ith ($i = 1, 2, \ldots, M+3$) quantum rotation angle and the ith ($i = 1, 2, \ldots, M+3$) quantum bit of the hth quantum whale are updated as

$$\theta_{hi}^{t+1} = \begin{cases} |\mathbf{A}_h^t| \cdot \xi_3 \cdot \left(\overline{p}_{gi}^t - \overline{x}_{hi}^t\right), & \text{if } i = 1, 2, \ldots, M, \\ |\mathbf{A}_h^t| \cdot \xi_3 \cdot \left(p_{gi}^t - x_{hi}^t\right), & \text{if } i = M+1, M+2, M+3, \end{cases} \tag{27}$$

$$\mu_{hi}^{t+1} = \begin{cases} \text{abs}\left(x_{hi}^t \cdot \cos\theta_{hi}^{t+1} - \sqrt{1 - \left(x_{hi}^t\right)^2} \cdot \sin\theta_{hi}^{t+1}\right), & \text{if } i = 1, 2, \ldots, M, \\ \text{abs}\left(x_{hi}^t \cdot \cos\theta_{hi}^{t+1} + \sqrt{1 - \left(x_{hi}^t\right)^2} \cdot \sin\theta_{hi}^{t+1}\right), & \text{if } i = M+1, M+2, M+3, \end{cases} \tag{28}$$

where ξ_3 is a uniform random number that is distributed from zero to one. abs $(\cdot)$ is the absolute value function. When $\xi_2 < 0.5$ and $|\mathbf{A}_h^t| \ge 1$, the hth quantum whale will enter searching for the prey mode. In this mode, the quantum rotation angle is given by

$$\theta_{hi}^{t+1} = \begin{cases} |\mathbf{A}_h^t| \cdot \xi_4 \cdot \left(\overline{x}_{ai}^t - \overline{x}_{hi}^t\right), & \text{if } i = 1, 2, \ldots, M, \\ |\mathbf{A}_h^t| \cdot \xi_4 \cdot \left(x_{ai}^t - x_{hi}^t\right), & \text{if } i = M+1, M+2, M+3, \end{cases} \tag{29}$$

where ξ_4 is a uniform random number, $a \in \{1, 2, \ldots, H\}$ is a random integer, and the quantum bit of each quantum whale is updated as (28).

If ξ_2 is no less than 0.5, the quantum whale will take bubble-net attacking behaviour. In this situation, the quantum rotation angle and the quantum bit can be shown as

$$\theta_{hi}^{t+1} = \begin{cases} c_2 \cdot \xi_5 \cdot \left(\overline{p}_{gi}^t - \overline{x}_{hi}^t\right) \cdot e^{b\xi_6}, & \text{if } i = 1, 2, \ldots, M, \\ c_1 \cdot \xi_5 \cdot \left(p_{gi}^t - x_{hi}^t\right) \cdot e^{b\xi_6}, & \text{if } i = M+1, M+2, M+3, \end{cases} \tag{30}$$

$$\mu_{hi}^{t+1} = \begin{cases} \text{abs}\left(x_{hi}^t \cdot \cos\theta_{hi}^{t+1} - \sqrt{1 - \left(x_{hi}^t\right)^2} \cdot \sin\theta_{hi}^{t+1}\right), & \text{if } i = 1, 2, \ldots, M, \\ \text{abs}\left(x_{hi}^t \cdot \cos\theta_{hi}^{t+1} + \sqrt{1 - \left(x_{hi}^t\right)^2} \cdot \sin\theta_{hi}^{t+1}\right), & \text{if } i = M+1, M+2, M+3, \end{cases} \tag{31}$$

where ξ_5 is a uniform random number that is distributed from zero to one, c_1, c_2, and b are constants which determine the radius of bubble-net, and ξ_6 is a uniform random number which is distributed from -1 to 1.

The quantum whale $\boldsymbol{\mu}_h^{t+1}$ $(h = 1, 2, \ldots, H)$ is measured to its definition domain $\overline{\boldsymbol{\mu}}_h^{t+1}$ according to (24) and (25). Then, the fitness of $\overline{\boldsymbol{\mu}}_h^{t+1}$ is computed by fitness function. If the fitness of $\overline{\boldsymbol{\mu}}_h^{t+1}$ is better than that of $\overline{\mathbf{x}}_h^t$, $\mathbf{x}_h^{t+1} = \boldsymbol{\mu}_h^{t+1}$; else,

$\mathbf{x}_h^{t+1} = \mathbf{x}_h^t$. At last, the best quantum whale of current iteration is used to update the global optimal quantum whale.

4.2. Throughput Optimization Process Based on QWOA. QWOA can optimize problem of maximizing IDT-IDR transmission pairs' average throughput or network sum throughput in CEH-IoT-SS. We convert the maximizing throughput problem into the process of finding global optimal quantum whale. The fitness function is set as $f(\overline{\mathbf{x}}_h^t) = \begin{cases} R(\overline{\mathbf{x}}_h^t), & \text{satisfy constraint condition} \\ 0, & \text{else} \end{cases}$. Hence, the throughput optimization process based on QWOA is presented in Algorithm 1:

5. Simulation Results and Analysis

In this part, the performance of proposed QWOA for different communication scenarios is investigated. The simulation results are divided into two parts. In the first part, we compare the optimization results of proposed QWOA and other intelligent algorithms for different communication needs. In fact, since existing algorithms cannot optimize both the continuous problems and discrete problems, we will give the comparison results of QWOA and the circumstances where the classical continuous algorithms are applied to the proposed problems directly. For the second simulation part, we aim to study the impact of different system parameters on the whole system according to different objective functions. In the simulation, PT and PR are located at (0, 1) and (2, 1) on X-Y plane. For the IDT-IDR transmission pairs, all IDTs are generated in a circle and the center of the circle is (1, 1) and the radius of the circle is 0.5, while IDRs are located in a circle in which center and radius are (1, 0) and 0.5, respectively. The factor of path loss ℓ is equal to 3 [40] and $\rho = 1$. For all nodes, the power of AWGN is the same, i.e., $w_1 = w_2 = w_3 = w_4 = w_5 = 10^{-1}$. In order to not lose generality, the EH rate of each IDT is the same. Unless otherwise stated, $Y_{\mathrm{PT}} = 10$, $Y_{\mathrm{IDT}} = Y_{\mathrm{IDT}_m} = 10$ $(m = 1, 2, \ldots, M)$, $R_{\mathrm{PR}}^{\mathrm{require}} = 0.3$, and the number of IDT-IDR pairs is 10. All simulation results are average of 500 trials.

5.1. Comparison Simulation Result of QWOA and Other Intelligent Algorithms. In this section, we show the performance of QWOA and other intelligent algorithms for the proposed two problems, which are sum throughput and IDT-IDR transmission pairs' average throughput. Figures 3–6 show the simulation results when the objective function is maximizing the average throughput of IDT-IDR transmission pairs, i.e., equation (20). Meanwhile, Figures 7–9 investigate comparison results for maximizing sum throughput, and the objective function is equation (21). In order to intuitively compare the performance of the traditional algorithm and the proposed algorithm, we apply some continuous intelligent algorithms to CEH-IoT-SS. The process for optimizing power splitting ratio, dynamic EH ratio, and power sharing ratio by continuous intelligent algorithms is the same as the course of optimizing continuous variables. However, for optimizing the relay selection scheme, we are going to round continuous variables up to integer for the compared continuous intelligent algorithms. Compared algorithms can be shown as follows: whale optimization algorithm (WOA), differential evolutionary algorithm (DEA), and the proposed QWOA. In all simulation results, the maximal iteration number is 500 and the number of population size H is set as 20. For QWOA, $c_1 = 2$, $c_2 = 2$, and $b = 1.5$. For WOA, the parameter settings can refer to [42]. For DEA, the parameter settings of DEA can refer to [46].

Figures 3 and 7 give the comparison results of QWOA and two classical continuous intelligent algorithms when considering equations (20) and (21) as optimization object, respectively. It is easy to see that QWOA has a faster convergence speed and higher convergence accuracy than WOA and DEA under the same simulation conditions. Since the proposed QWOA uses the thinking of quantum evolution theory and whale hunting characteristics, QWOA is more suitable to solve hybrid optimization problems. For continuous variables and discrete variables, we design different evolution strategies to obtain the solution. The resolving method of traditional WOA is updating individual in the population by some certain equations, and it is easy to get into a local optimum in the early searching stage. From the simulation results, we can see that WOA gets a better solution in the last dozens of iterations. This is because the traditional WOA will process an accurate search due to the searching step length which is generated in a fixed way. However, it is still a local optimum. Unlike WOA, since designed evolution strategies can make full use of the advantages of quantum theory and WOA, the best solution is obtained by QWOA. For DEA, it has disadvantages such as local convergence and randomness and it cannot have a good performance on the proposed problem. To summarize, QWOA can overcome the shortcomings of traditional algorithms and show its advantage in convergence speed and convergence accuracy.

In Figure 4, we investigate the impact of Y_{IDT} on the IDT-IDR transmission pairs' average throughput. $Y_{\mathrm{PT}} = 10$ and Y_{IDT} varies from 2 to 16 in simulation. It is easy to see that as Y_{IDT} increases, IDT-IDR transmission pairs' average throughput becomes larger and larger for QWOA. This is because as Y_{IDT} increases, each IDT can get more energy to store for information transfer. A larger Y_{IDT} can make each IDT use less time to store enough energy. However, due to the fact that WOA and DEA cannot find the best solution, IDT-IDR transmission pairs' average throughput for WOA and DEA is not increasing all the time as Y_{IDT} becomes larger. Simulation results show that QWOA can find the best solution for the whole network and have a better performance compared with WOA and DEA for different Y_{IDT}.

Figure 5 considers the circumstance where $Y_{\mathrm{IDT}} = 10$ and Y_{PT} increases from 2 to 16. For different simulation situations, QWOA has the best performance all the time. The average throughput of IDT-IDR transmission pairs becomes larger and larger when Y_{PT} obtains a higher value. It is easy to understand. As Y_{PT} increases, PT can collect more energy for the same time and PT has more energy for transmitting its own information. In addition, each IDT can store more energy from PT as Y_{PT} increases. More energy of IDTs can bring a bigger SINR for IDT-IDR transmission pairs and more easily satisfy the need of the primary network.

(1) **Input** system parameters of CEH-IoT-SS, the fitness function;
(2) **Initialize** the initial population of H quantum whales;
(3) $t = 1$ // the first iteration;
(4) Measure each quantum whale to the definition domain according to (24) and (25);
(5) Compute fitness of each quantum whale in accordance with the fitness function and find out the global optimal quantum whale $\mathbf{p}_g^t$;
(6) **while** $t \leq K$
(7) Calculate the coefficient vector of each quantum whale
(8) **if** $\xi_2 < 0.5$ and $|\mathbf{A}_h^t| < 1$
(9) Update the quantum whale through (27) and (28)
(10) **else if** $\xi_2 < 0.5$ and $|\mathbf{A}_h^t| \geq 1$
(11) Update the quantum whale through (29) and (28)
(12) **else**
(13) Update the quantum whale through (30) and (31)
(14) **end if**
(15) **end if**
(16) Measure each quantum whale and compute the fitness according to the fitness function;
(17) Update the quantum whale and the global optimal quantum whale according to the fitness;
(18) Set $t = t + 1$;
(19) **end while**
(20) According to the global optimal quantum whale, obtain the power splitting ratio, dynamic EH ratio, power sharing ratio, and relay selection scheme after K iterations;
(21) **Output**: Power splitting ratio, dynamic EH ratio, power sharing ratio, and relay selection scheme.

Algorithm 1: Throughput optimization process based on QWOA.

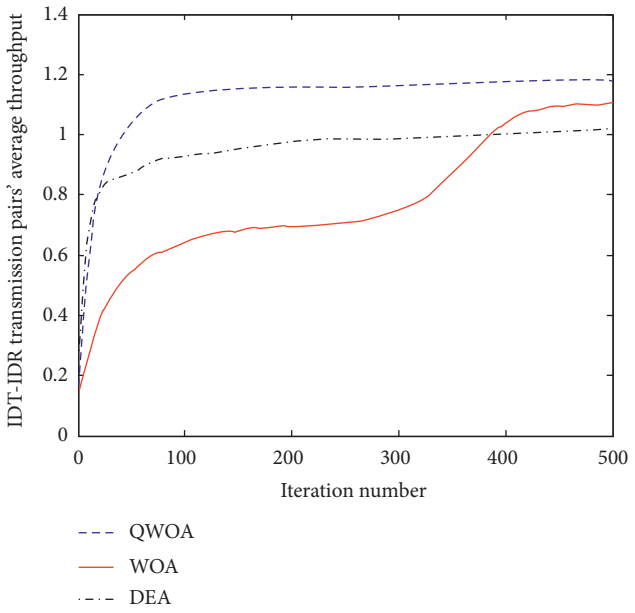


Figure 3: Comparison of convergence performance for QWOA, WOA, and DEA (IDT-IDR transmission pairs' average throughput).

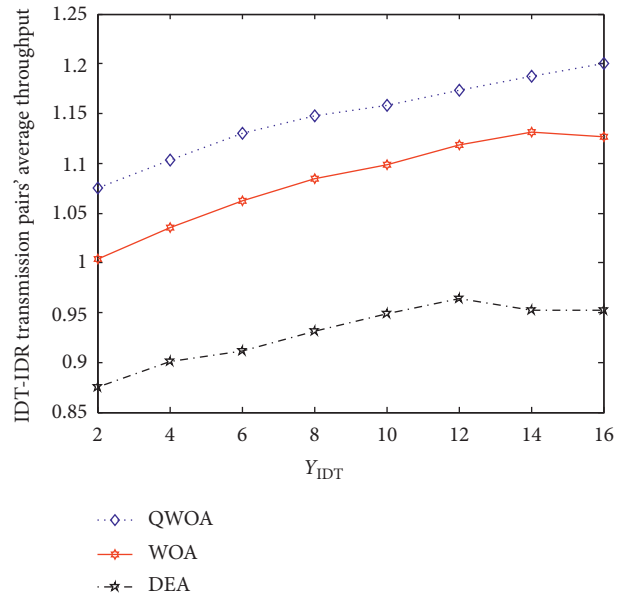


Figure 4: IDT-IDR transmission pairs' average throughput of three schemes with different Y_{IDT}.

Figure 6 shows the IDT-IDR transmission pairs' average throughput with different $R_{PR}^{require}$. $R_{PR}^{require}$ varies from 0.1 to 0.8. Simulation results give the performance of QWOA, WOA, and DEA when $R_{PR}^{require}$ increases. For any $R_{PR}^{require}$, the proposed QWOA can show the best performance all the time. From the simulation results, we can draw a conclusion that the average throughput of IDT-IDR transmission pairs decreases as $R_{PR}^{require}$ increases. This performance matches the theory. The power need of the primary network is increasing as $R_{PR}^{require}$ reaches a larger value. Since the power for transmission of the primary network becomes larger, the power that IDT can use becomes less. Therefore, the average throughput of IDT-IDR transmission pairs becomes less.

Figure 8 considers the sum throughput of the whole system as Y_{PT} increases. Y_{PT} increases from 2 to 16. It is clear that the sum throughput reaches a higher value when Y_{PT} increases. When Y_{PT} becomes larger, the energy collected by PT increases in one time slot. More energy of PT not only helps information transfer of the primary network but also provides more power for IDT-IDR transmission pairs and

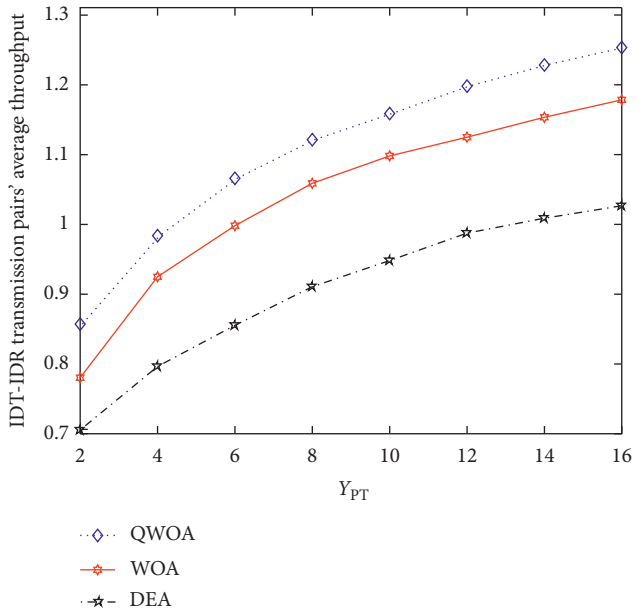


Figure 5: IDT-IDR transmission pairs' average throughput of three schemes with different Y_{PT}.

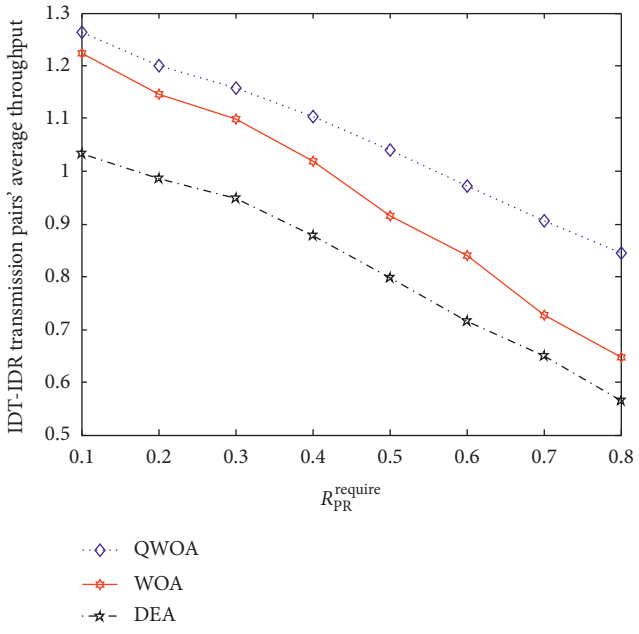


Figure 6: IDT-IDR transmission pairs' average throughput of three schemes with different $R_{PR}^{require}$.

each IDTs can share more energy to process cooperative communication and its own information transmission. Besides, by comparing with WOA and DEA, we can find that QWOA can get the best simulation results all the time.

In Figure 9, the impact of $R_{PR}^{require}$ on the whole system is presented. For a certain $R_{PR}^{require}$, QWOA can obtain the best performance compared with WOA and DEA. As $R_{PR}^{require}$ increase, the sum throughput of the whole system becomes smaller and smaller. This is because a higher $R_{PR}^{require}$ will bring more constraint for the primary network and IDT-IDR transmission

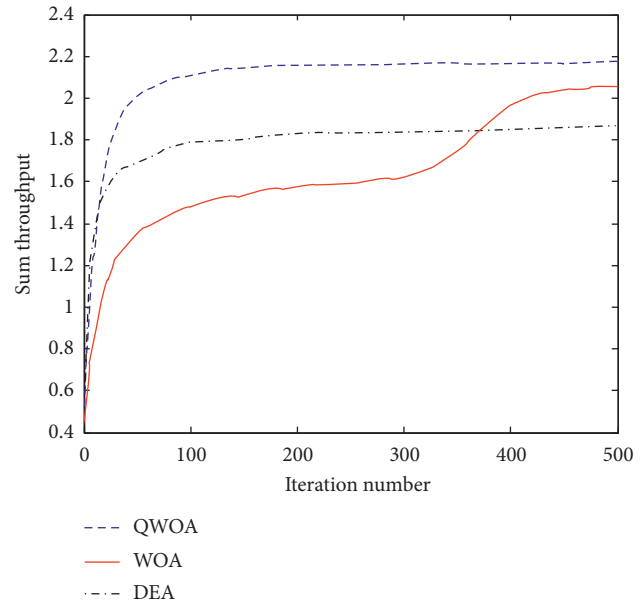


Figure 7: Comparison of convergence performance for QWOA, WOA, and DEA (sum throughput).

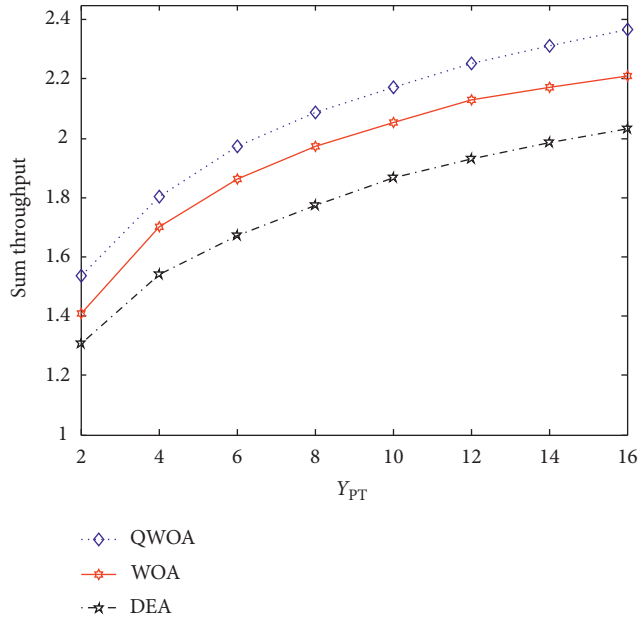


Figure 8: Sum throughput of three schemes with different Y_{PT}.

pairs. The number of solutions for the best power splitting ratio, dynamic EH ratio, power sharing ratio, and the relay selection decreases with a higher $R_{PR}^{require}$. Therefore, the sum throughput becomes smaller with a higher $R_{PR}^{require}$. From Figures 3 to 9, the proposed QWOA can get the best performance all the time. All the simulation results show the advantages of QWOA in convergence speed and convergence accuracy.

5.2. Simulations of Different Parameters Based on QWOA for CEH-IoT-SS. In this part, the impact of partial different parameters on two proposed objective functions for

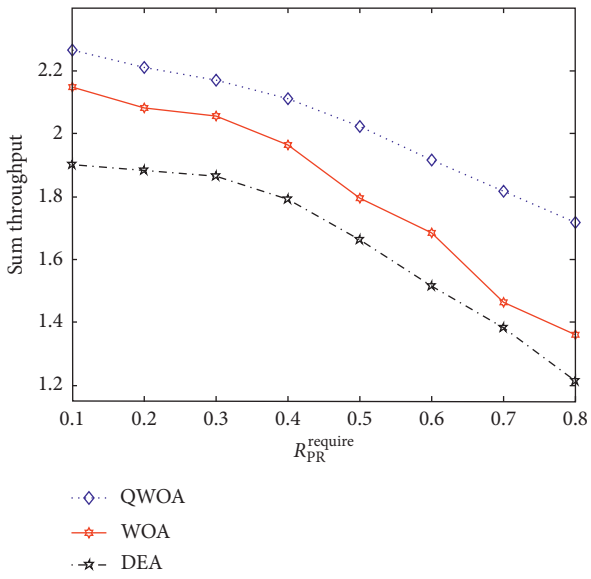


Figure 9: Sum throughput of three schemes with different $R_{PR}^{require}$.

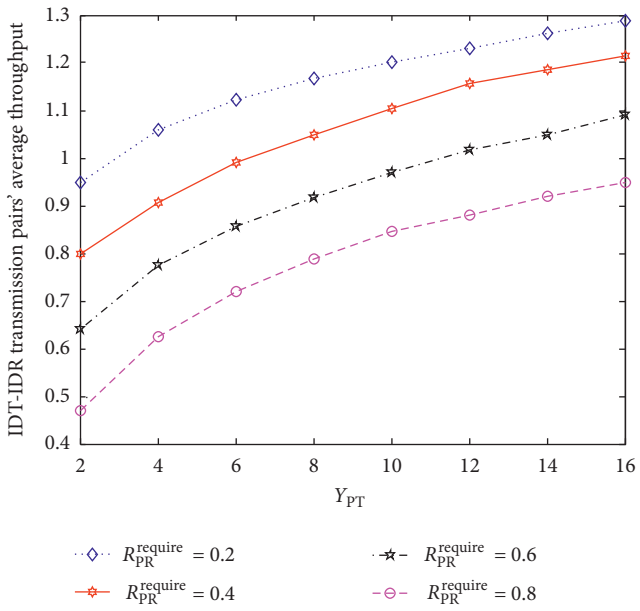


Figure 10: IDT-IDR transmission pairs' average throughput of different Y_{PT} and $R_{PR}^{require}$.

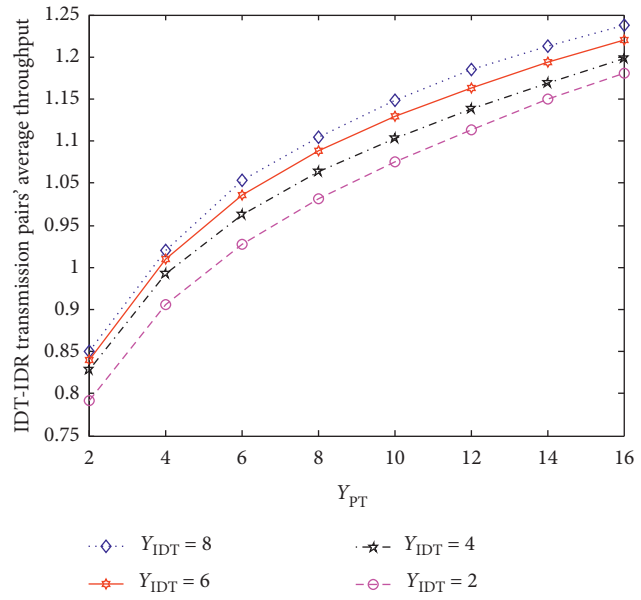


Figure 11: IDT-IDR transmission pairs' average throughput of different Y_{PT} and Y_{IDT}.

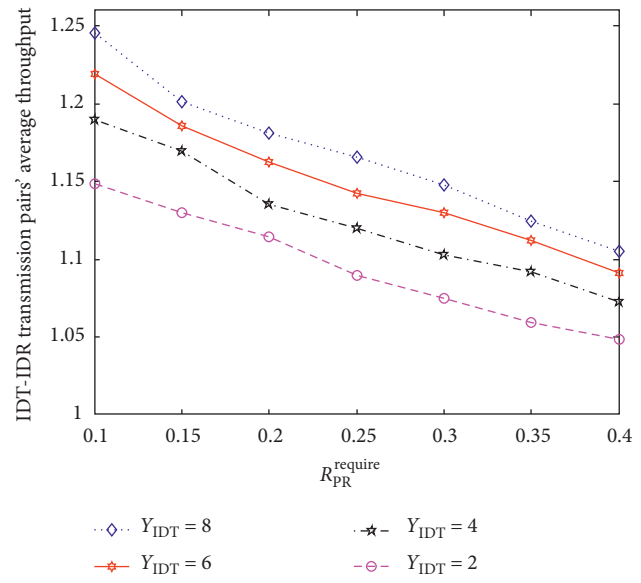


Figure 12: IDT-IDR transmission pairs' average throughput of different $R_{PR}^{require}$ and Y_{IDT}.

CEH-IoT-SS is presented. We show part of simulation results because of the similarity of parameter effects.

Figure 10 presents the IDT-IDR transmission pairs' average throughput when $R_{PR}^{require} = 0.2$, $R_{PR}^{require} = 0.4$, $R_{PR}^{require} = 0.6$, and $R_{PR}^{require} = 0.8$. The EH rate, i.e., Y_{PT} varies from 2 to 16. Due to the fact that a higher Y_{PT} can help PT store more energy at the same time and IDT can share more energy from PT, the average throughput of IDT-IDR transmission pairs becomes higher and higher when Y_{PT} increases. Meanwhile, the more need of the primary network will bring part of energy from IDT-IDR transmission pairs. Hence, IDT-IDR transmission pairs' average throughput will decrease with a higher $R_{PR}^{require}$.

Figure 11 considers the impact of Y_{IDT} and Y_{PT} on IDT-IDR transmission pairs' average throughput. Y_{PT} varies from 2 to 16. Y_{IDT} is 2, 4, 6, and 8, respectively. From the simulation results, we can easily find that IDT-IDR transmission pairs' average throughput increases larger with a higher Y_{IDT} and Y_{PT}. For a certain period of time, the higher Y_{IDT} and Y_{PT}, the more energy will be stored. More energy will bring a higher power to the primary network and IDT-IDR transmission pairs for information transmission and it can increase the throughput for the certain communication problem. Therefore, boosting Y_{IDT} and Y_{PT} can improve the performance of IDT-IDR transmission pairs.

In Figure 12, the impact of different $R_{PR}^{require}$ and Y_{IDT} on the IDT-IDR transmission pairs is investigated. $R_{PR}^{require}$ is increasing from 0.1 to 0.4, while Y_{IDT} is 2, 4, 6, and 8,

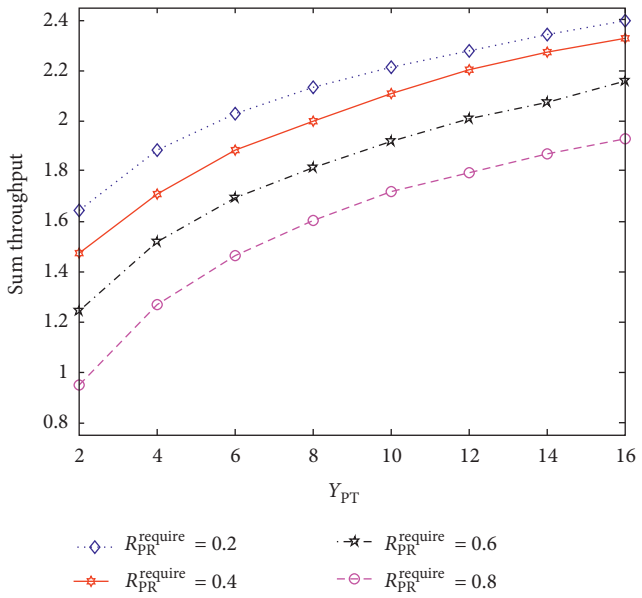


FIGURE 13: Sum throughput of different Y_{PT} and $R_{PR}^{require}$.

respectively. For a certain Y_{IDT}, the average throughput of IDT-IDR transmission pairs decreases with $R_{PR}^{require}$ becomes higher and higher. Meanwhile, the average throughput of IDT-IDR transmission pairs increases with a higher Y_{IDT}.

Figure 13 presents the simulation results which consider different Y_{PT} and $R_{PR}^{require}$ for the sum throughput of the primary network and IDT-IDR transmission pairs. In the simulation, $R_{PR}^{require}$ is set as 0.2, 0.4, 0.6, and 0.8, respectively. Y_{PT} varies from 2 to 16. It is easy to see that a higher Y_{PT} is conducive to the whole network and a higher $R_{PR}^{require}$ will reduce the throughput of the network. To improve Y_{PT} is a meaningful thing for the whole network. It is obvious that the proposed QWOA can satisfy different throughput requirements of the primary network.

6. Conclusion

In this paper, we design a novel EH and information transmission mode for CEH-IoT-SS. In the network, IDTs can work as relays for PU's information transmission and finish its own information transfer in one time slot. The expressions for the throughput of the primary network and IDT-IDR transmission pairs are mathematically derived. For different communication scenarios, we formulate two objective functions with different communication demands. Since the formulated objective functions are hybrid optimization problems, traditional intelligent algorithms cannot solve it well. A new algorithm-QWOA is proposed to solve it. Compared with the traditional algorithms, QWOA can get the best solution for hybrid optimization problems and various system parameters. Simulation results validate that QWOA has a wide application for different communication scenarios and hybrid problems. In our future work, we will apply our work to other specific IoT scenarios, such as A-IoT, fog computing, and blockchain.

Appendix

For the mth IDR, thereceived signal in the third phase is shown by

$$
\begin{aligned}
z_{\mathrm{IDR}_m} &= \sqrt{P_{\mathrm{IDT}_m}} G_{\mathrm{IDT}_m,\mathrm{IDR}_m} z_{\mathrm{IDT}_m} + \sum_{i=1,i\neq m}^{M} \sqrt{P_{\mathrm{IDT}_i}} G_{\mathrm{IDT}_i,\mathrm{IDR}_m} z_{\mathrm{IDT}_i} + n_5 \\
&= \sqrt{P_{\mathrm{IDT}_m}} G_{\mathrm{IDT}_m,\mathrm{IDR}_m} z_{\mathrm{IDT}_m} + \sum_{i=1,i\neq m}^{M} b_i G_{\mathrm{IDT}_i,\mathrm{IDR}_m} \sqrt{P_{\mathrm{IDT}_i}(1-\sigma)} z_{\mathrm{IDT}_i} \\
&\quad + \sum_{i=1,i\neq m}^{M} b_i \frac{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))\sigma P_{\mathrm{IDT}_i}(1-\beta)} G_{\mathrm{IDT}_i,\mathrm{IDR}_m} G_{\mathrm{PT},\mathrm{IDT}_i}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT},\mathrm{IDT}_i}\right|^2 + (1-\beta)w_1 + w_3}} x \\
&\quad + \sum_{i=1,i\neq m}^{M} b_i \frac{\sqrt{\sigma P_{\mathrm{IDT}_i}(1-\beta)} G_{\mathrm{IDT}_i,\mathrm{IDR}_m}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT},\mathrm{IDT}_i}\right|^2 + (1-\beta)w_1 + w_3}} n_1 \\
&\quad + \sum_{i=1,i\neq m}^{M} b_i \frac{\sqrt{\sigma P_{\mathrm{IDT}_i}} G_{\mathrm{IDT}_i,\mathrm{IDR}_m}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT},\mathrm{IDT}_i}\right|^2 + (1-\beta)w_1 + w_3}} n_3 \\
&\quad + \sum_{i=1,i\neq m}^{M} (1-b_i) G_{\mathrm{IDT}_i,\mathrm{IDR}_m} \sqrt{P_{\mathrm{IDT}_i}} z_{\mathrm{IDT}_i} + n_5,
\end{aligned}
\tag{A.1}
$$

where n_5 is the AWGN with power w_5. $\sqrt{P_{\mathrm{IDT}_m}} G_{\mathrm{IDT}_m,\mathrm{IDR}_m} z_{\mathrm{IDT}_m}$ is a useful signal part for the mth IDR and the other parts are interference signals. According to (5), the mth IDR has received information x of PT in the

second phase. Therefore, the signal interference from PT, i.e., $\sum_{i=1,i\neq m}^{M} b_i [\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))\sigma P_{\mathrm{IDT}_i}(1-\beta)}\, G_{\mathrm{IDT}_i,\ \mathrm{IDR}_m} G_{\mathrm{PT,IDT}_i}/\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)|G_{\mathrm{PT,IDT}_i}|^2 + (1-\beta)w_1 + w_3}\,]x$ can be removed from (A.1) and the expression for the residual part is shown by

$$\begin{aligned} z_{\mathrm{IDR}_m} = {} & \sqrt{P_{\mathrm{IDT}_m}} G_{\mathrm{IDT}_m,\mathrm{IDR}_m} z_{\mathrm{IDT}_m} + \sum_{i=1,i\neq m}^{M} b_i G_{\mathrm{IDT}_i,\mathrm{IDR}_m} \sqrt{P_{\mathrm{IDT}_i}(1-\sigma)} z_{\mathrm{IDT}_i} \\ & + \sum_{i=1,i\neq m}^{M} b_i \frac{\sqrt{\sigma P_{\mathrm{IDT}_i}(1-\beta)} G_{\mathrm{IDT}_i,\mathrm{IDR}_m}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_i}\right|^2 + (1-\beta)w_1 + w_3}} n_1 \\ & + \sum_{i=1,i\neq m}^{M} b_i \frac{\sqrt{\sigma P_{\mathrm{IDT}_i}} G_{\mathrm{IDT}_i,\mathrm{IDR}_m}}{\sqrt{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_i}\right|^2 + (1-\beta)w_1 + w_3}} n_3 \\ & + \sum_{i=1,i\neq m}^{M} (1-b_i) G_{\mathrm{IDT}_i,\mathrm{IDR}_m} \sqrt{P_{\mathrm{IDT}_i}} z_{\mathrm{IDT}_i} + n_5. \end{aligned} \tag{A.2}$$

The SINR of the mth IDR is shown as follows:

$$\gamma_{\mathrm{IDR}_m} = \begin{cases} \dfrac{P_{\mathrm{IDT}_m}\left|G_{\mathrm{IDT}_m,\mathrm{IDR}_m}\right|^2}{\left(\sum_{i=1,i\neq m}^{M} b_i \left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 P_{\mathrm{IDT}_i}(1-\sigma) + \sum_{i=1,i\neq m}^{M} b_i \left(\sigma P_{\mathrm{IDT}_i}(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_1/(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_i}\right|^2 + (1-\beta)w_1 + w_3\right) + \sum_{i=1,i\neq m}^{M} b_i \left(\sigma P_{\mathrm{IDT}_i}\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_3/(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 + (1-\beta)w_1 + w_3\right) + \sum_{i=1,i\neq m}^{M}(1-b_i)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 P_{\mathrm{IDT}_i} + w_5\right)}, & \text{if } \mathrm{IDT}_m \text{ is not selected as relay,} \\ \dfrac{(1-\sigma)P_{\mathrm{IDT}_m}\left|G_{\mathrm{IDT}_m,\mathrm{IDR}_m}\right|^2}{\left(\sum_{i=1,i\neq m}^{M} b_i \left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 P_{\mathrm{IDT}_i}(1-\sigma) + \sum_{i=1,i\neq m}^{M} b_i \left(\sigma P_{\mathrm{IDT}_i}(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_1/(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_i}\right|^2 + (1-\beta)w_1 + w_3\right) + \sum_{i=1,i\neq m}^{M} b_i \left(\sigma P_{\mathrm{IDT}_i}\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_3/(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 + (1-\beta)w_1 + w_3\right) + \sum_{i=1,i\neq m}^{M}(1-b_i)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 P_{\mathrm{IDT}_i} + w_5\right)}, & \text{if } \mathrm{IDT}_m \text{ is selected as relay.} \end{cases} \tag{A.3}$$

According to (9) and (A.3), the throughput of the mth IDR is given by

$$\begin{aligned} R_{\mathrm{IDR}_m} &= \frac{1-\alpha}{2}\log_2\left(1+\gamma_{\mathrm{IDR}_m}\right) = \frac{1-\alpha}{2}\log_2\left(1+\frac{\mathrm{Signal}_{\mathrm{IDR}_m}}{\mathrm{Interference}_{\mathrm{IDR}_m}}\right), \\ \mathrm{Signal}_{\mathrm{IDR}_m} &= \begin{cases} \dfrac{2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2}{1-\alpha}\left|G_{\mathrm{IDT}_m,\mathrm{IDR}_m}\right|^2, & \text{if } \mathrm{IDT}_m \text{ is not selected as relay,} \\ \dfrac{2\alpha Y_{\mathrm{IDT}_m} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_m}\right|^2}{1-\alpha}(1-\sigma)\left|G_{\mathrm{IDT}_m,\mathrm{IDR}_m}\right|^2, & \text{if } \mathrm{IDT}_m \text{ is selected as relay,} \end{cases} \\ \mathrm{Interference}_{\mathrm{IDR}_m} &= \sum_{i=1,i\neq m}^{M} b_i\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 \frac{2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2}{1-\alpha}(1-\sigma) \\ &\quad + \sum_{i=1,i\neq m}^{M} b_i \frac{\sigma\left(\left(2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2\right)/(1-\alpha)\right)(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_1}{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{PT,IDT}_i}\right|^2 + (1-\beta)w_1 + w_3} \\ &\quad + \sum_{i=1,i\neq m}^{M} b_i \frac{\sigma\left(\left(2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2\right)/(1-\alpha)\right)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 w_3}{(2Y_{\mathrm{PT}}\alpha/(1-\alpha))(1-\beta)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 + (1-\beta)w_1 + w_3} \\ &\quad + \sum_{i=1,i\neq m}^{M}(1-b_i)\left|G_{\mathrm{IDT}_i,\mathrm{IDR}_m}\right|^2 \frac{2\alpha Y_{\mathrm{IDT}_i} + 2\rho Y_{\mathrm{PT}}\alpha\beta\left|G_{\mathrm{PT,IDT}_i}\right|^2}{1-\alpha} + w_5. \end{aligned} \tag{A.4}$$

Similarly, the average throughput of M IDT-IDR transmission pairs in CEH-IoT-SS is expressed by

$$R_{\text{IDT-IDR}_{\text{average}}} = \frac{1}{M}\sum_{m=1}^{M} R_{\text{IDR}_m} = \frac{1}{M}\sum_{m=1}^{M} \frac{1-\alpha}{2}\log_2\left(1+\gamma_{\text{IDR}_m}\right)$$

$$= \frac{1-\alpha}{2M}\sum_{m=1}^{M} \log_2\left(1+\gamma_{\text{IDR}_m}\right). \tag{A.5}$$

Combining (A.4) with (A.5), we can obtain the expression of IDT-IDR transmission pairs' average throughput as given in (19). This completes the proof.

Data Availability

The data (figures) used to support the findings of this study are included within the article. Further details can be provided upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61571149), Special China Postdoctoral Science Foundation (no. 2015T80325), Fundamental Research Funds for the Central Universities (nos. HEUCFP201808 and HEUCF190801), and China Postdoctoral Science Foundation (no. 2013M530148).

References

[1] N. Javaid, A. Sher, H. Nasir, and N. Guizani, "Intelligence in IoT-based 5G networks: opportunities and challenges," *IEEE Communications Magazine*, vol. 56, no. 10, pp. 94–100, 2018.

[2] J. Iannacci, "RF-MEMS for high-performance and widely reconfigurable passive components - a review with focus on future telecommunications, Internet of Things (IoT) and 5G applications," *Journal of King Saud University—Science*, vol. 29, no. 4, pp. 436–443, 2017.

[3] F. Benkhelifa, K. Tourki, and M.-S. Alouini, "Proactive spectrum sharing for SWIPT in MIMO cognitive radio systems using antenna switching technique," *IEEE Transactions on Green Communications and Networking*, vol. 1, no. 2, pp. 204–222, 2017.

[4] A. Aijaz and A. H. Aghvami, "Cognitive machine-to-machine communications for Internet-of-Things: a protocol stack perspective," *IEEE Internet of Things Journal*, vol. 2, no. 2, pp. 103–112, 2015.

[5] P. Rawat, K. D. Singh, and J. M. Bonnin, "Cognitive radio for M2M and Internet of Things: a survey," *Computer Communications*, vol. 94, pp. 1–29, 2016.

[6] G. Ding, Q. Wu, L. Zhang, Y. Lin, T. A. Tsiftsis, and Y.-D. Yao, "An amateur drone surveillance system based on the cognitive Internet of Things," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 29–35, 2018.

[7] Z. Li, B. Chang, S. Wang, A. Liu, F. Zeng, and G. Luo, "Dynamic compressive wide-band spectrum sensing based on channel energy reconstruction in cognitive Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2598–2607, 2018.

[8] K. Zaheer, M. Othman, M. H. Rehmani, and T. Perumal, "A survey of decision-theoretic models for cognitive Internet of Things (CIoT)," *IEEE Access*, vol. 6, pp. 22489–22512, 2018.

[9] S. Solanki, P. K. Sharma, and P. K. Upadhyay, "Adaptive link utilization in two-way spectrum sharing relay systems under average interference-constraints," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3461–3472, 2018.

[10] A. A. Khan, M. H. Rehmani, and A. Rachedi, "Cognitive-radio-based Internet of Things: applications, architectures, spectrum related functionalities, and future research directions," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 17–25, 2017.

[11] J. Guo, S. Durrani, and X. Zhou, "Performance analysis of arbitrarily-shaped underlay cognitive networks: effects of secondary user activity protocols," *IEEE Transactions on Communications*, vol. 63, no. 2, pp. 376–389, 2015.

[12] S. Aslam, W. Ejaz, and M. Ibnkahla, "Energy and spectral efficient cognitive radio sensor networks for Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3220–3233, 2018.

[13] T. P. Do, I. Song, and Y. H. Kim, "Simultaneous wireless transfer of power and information in a decode-and-forward two-way relaying network," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1579–1592, 2017.

[14] J. Yan and Y. Liu, "A dynamic SWIPT approach for cooperative cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 11122–11136, 2017.

[15] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Wireless-powered relays in cooperative communications: time-switching relaying protocols and throughput analysis," *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1607–1622, 2015.

[16] T. Li, P. Fan, and K. B. Letaief, "Outage probability of energy harvesting relay-aided cooperative networks over Rayleigh fading channel," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 972–978, 2016.

[17] N. Zhao, S. Zhang, F. R. Yu, Y. Chen, A. Nallanathan, and V. C. M. Leung, "Exploiting interference for energy harvesting: a survey, research issues, and challenges," *IEEE Access*, vol. 5, pp. 10403–10421, 2017.

[18] B. V. Nguyen, H. Jung, D. Har, and K. Kim, "Performance analysis of a cognitive radio network with an energy harvesting secondary transmitter under Nakagami-m fading," *IEEE Access*, vol. 6, pp. 4135–4144, 2018.

[19] Z. Yang, W. Xu, Y. Pan, C. Pan, and M. Chen, "Energy efficient resource allocation in machine-to-machine communications with multiple access and energy harvesting for IoT," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 229–245, 2018.

[20] D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and Z. Han, "Ambient backscatter: a new approach to improve network performance for RF-powered cognitive radio networks," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3659–3674, 2017.

[21] H. Gao, W. Ejaz, and M. Jo, "Cooperative wireless energy harvesting and spectrum sharing in 5G networks," *IEEE Access*, vol. 4, pp. 3647–3658, 2016.

[22] P. Maji, S. D. Roy, and S. Kundu, "Secrecy outage analysis in a hybrid cognitive relay network with energy harvesting,"

International Journal of Communication Systems, vol. 30, no. 10, Article ID 3228, 2017.

[23] A. O. Ercan, M. O. Sunay, and I. F. Akyildiz, "RF energy harvesting and transfer for spectrum sharing cellular IoT communications in 5G systems," *IEEE Transactions on Mobile Computing*, vol. 17, no. 7, pp. 1680–1694, 2018.

[24] X. M. Huang, R. Yu, J. W. Kang, Z. Q. Xia, and Y. Zhang, "Software defined networking for energy harvesting Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1389–1399, 2018.

[25] A. Caruso, S. Chessa, S. Escolar, X. del Toro, and J. C. Lopez, "A dynamic programming algorithm for high-level task scheduling in energy harvesting IoT," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2234–2248, 2018.

[26] H.-S. Lee and J.-W. Lee, "Resource and task scheduling for SWIPT IoT systems with renewable energy sources," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2729–2748, 2019.

[27] D. S. Gurjar, H. H. Nguyen, and H. D. Tuan, "Wireless information and power transfer for IoT applications in overlay cognitive radio networks," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3257–3270, 2019.

[28] Q. Ju, H. Li, and Y. Zhang, "Power management for kinetic energy harvesting IoT," *IEEE Sensors Journal*, vol. 18, no. 10, pp. 4336–4345, 2018.

[29] S. S. Kalamkar and A. Banerjee, "Secure communication via a wireless energy harvesting untrusted relay," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 3, pp. 2199–2213, 2017.

[30] A. Mukherjee, T. Acharya, and M. R. A. Khandaker, "Outage analysis for SWIPT-enabled two-way cognitive cooperative communications," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 9032–9036, 2018.

[31] X. Zhang, Z. Zhang, J. Xing, R. Yu, P. Zhang, and W. Wang, "Exact outage analysis in cognitive two-way relay networks with opportunistic relay selection under primary user's interference," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2502–2511, 2015.

[32] B. Lyu, H. Guo, Z. Yang, and G. Gui, "Throughput maximization for hybrid backscatter assisted cognitive wireless powered radio networks," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2015–2024, 2018.

[33] S. Atapattu, Y. D. Yindi Jing, H. Hai Jiang, and C. Tellambura, "Relay selection and performance analysis in multiple-user networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 8, pp. 1517–1529, 2013.

[34] K. Singh, A. Gupta, and T. Ratnarajah, "Energy efficient resource allocation for multiuser relay networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1218–1235, 2017.

[35] S. Sharma, Y. Shi, Y. T. Hou, and S. Kompella, "An optimal algorithm for relay node assignment in cooperative ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 19, no. 3, pp. 879–892, 2011.

[36] W. Zhong, G. Chen, S. Jin, and K.-K. Wong, "Relay selection and discrete power control for cognitive relay networks via potential game," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5411–5424, 2014.

[37] A. Zappone, S. Atapattu, M. Di Renzo, J. Evans, and M. Debbah, "Energy-efficient relay assignment and power control in multi-user and multi-relay networks," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 1070–1073, 2018.

[38] S. Yin, E. Zhang, Z. Qu, L. Yin, and S. Li, "Optimal cooperation strategy in cognitive radio systems with energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4693–4707, 2014.

[39] O. Elnahas, M. Elsabrouty, O. Muta, and H. Furukawa, "Game theoretic approaches for cooperative spectrum sensing in energy-harvesting cognitive radio networks," *IEEE Access*, vol. 6, pp. 11086–11100, 2018.

[40] Z. Wang, Z. Chen, B. Xia, L. Luo, and J. Zhou, "Cognitive relay networks with energy harvesting and information transfer: design, analysis, and optimization," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2562–2576, 2016.

[41] H. Gobjuka and Y. J. Breitbart, "Ethernet topology discovery for networks with incomplete information," *IEEE/ACM Transactions on Networking*, vol. 18, no. 4, pp. 1220–1233, 2010.

[42] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016.

[43] H. Gao, S. Zhang, Y. Su, M. Diao, and M. Jo, "Joint multiple relay selection and time slot allocation algorithm for the EH-abled cognitive multi-user relay networks," *IEEE Access*, vol. 7, pp. 111993–112007, 2019.

[44] L. Yang, B. Y. Yang, and Y. F. Chen, "Full quantum treatment of Rabi oscillation driven by a pulse train and its application in ion-trap quantum computation," *IEEE Journal of Quantum Electronics*, vol. 49, no. 8, pp. 641–651, 2013.

[45] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, Santa Fe, NM, USA, November 1994.

[46] N. N. Qin and J. L. Chen, "An area coverage algorithm for wireless sensor networks based on differential evolution," *International Journal of Distributed Sensor Networks*, vol. 14, no. 8, Article ID 155014771879673, 2018.