Distributed Spectrum and Power Allocation for D2D-U Networks: A Scheme based on NN and Federated Learning

Rui Yin · Zhiqun Zou · Celimuge Wu · Jiantao Yuan · Xianfu Chen

Abstract In this paper, a *Device-to-Device communication on unlicensed bands* (D2D-U) enabled network is studied. To improve the *spectrum efficiency* (SE) on the unlicensed bands and fit its distributed structure while ensuring the fairness among D2D-U links and the harmonious coexistence with WiFi networks, a distributed joint power and spectrum scheme is proposed. In particular, a parameter, named as price, is defined, which is updated at each D2D-U pair by a online trained *Neural network* (NN) according to the channel state and traffic load. In addition, the parameters used in the NN are updated by two ways, unsupervised self-iteration and federated learning, to guarantee the fairness and harmonious coexistence. Then, a non-convex optimization problem with respect to the spectrum and power is formulated and solved on each D2D-U link to maximize its own data rate. Numerical simulation results are

Rui Yin
School of Information and Electrical Engineering, Zhejiang University City College, Hangzhou 310015, China
E-mail: yinrui@zucc.edu.cn

Zhiqun Zou
College of Information Science and Electrical Engineering, Zhejiang University, Hangzhou, China

Celimuge Wu (Corresponding author)
Graduate School of Informatics and Engineering, The University of Electro-Communications,1-5-1, Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Jiantao Yuan
Institute of Ocean Sensing and Networking of the Ocean College, Zhejiang University, Hangzhou, China

Xianfu Chen
VTT Technical Research Centre of Finland, Finland

demonstrated to verify the effectiveness of the proposed scheme.

1 Introduction

The large-scale commercialization of the *fifth generation* (5G) mobile networks has brought us better communication experience with low latency and high data rate As a key technology in 5G networks, *device-to-device* (D2D) communication allows direct transmission between D2D terminals instead of relaying through the base stations, which improves both system *spectrum efficiency* (SE), *energy efficiency* (EE) and *quality-of-service* (QoS) of D2D pairs [2].

The conventional D2D communication mainly reuses the licensed channels with *long-term evolution* (LTE) cellular networks to increase system SE and EE in the licensed spectrum [3]. However, the licensed spectrum is basically managed by mobile communication operators and expensive. In addition, with the explosive growth of the number of smart terminals, the spectrum resources on the licensed bands are becoming more scarce and D2D communications may cause severe interference to the original cellular users. In order to guarantee the transmission performance of the cellular users as well as improve the QoS of D2D users, *Device-to-Device on unlicensed bands* (D2D-U) is proposed to enable D2D communication on unlicensed spectrum [4]. As the spectrum resources of unlicensed bands are abundant and free to use, D2D-U may significantly increase the SE and EE of D2D systems while guaranteeing the QoS of cellular users [5].

Most existing works have studied on the mode selection, power and spectrum allocation mechanisms for D2D enabled cellular networks. In [6], the impact of mode selection on effective capacity has been investigated via the Markov service process model. Authors in [7] have proposed a centralized optimal mode selection and resource allocation for D2D enabled cellular networks. A distributed joint spectrum sharing and power allocation problem has been modeled as a non-convex optimization problem in [8] and the suboptimal solution is obtained by convex approximation techniques. Similar problem is solved by a price-based model in [9], where a game-theoretic approach is proposed to mitigate interference among D2D pairs in a distributed way. Many machine learning-based methods have also been used to solve related problems in recent years. The authors of [10] have designed a transmit power control strategy to D2D pairs based on a *deep neural network* (DNN) structure, where the SE and interference are taken into account. A deep reinforcement learning-based method is utilized in [11] to maximize the sum rates of D2D links.

Recently, *long-term evolution on unlicensed bands* (LTE-U) system is introduced into the unlicensed spectrum. *listen-before-talk* (LBT) and *duty cycle method* (DCM) access mechanisms have been proposed for LTE-U based cellular users to access the unlicensed spectrum while ensuring the fair coexistence with WiFi networks [12–16]. In [17], the back-off window size based on LBT mechanism is adaptively adjusted according to the WiFi traffic load and available bandwidth on licensed spectrum, which improves the system spectrum efficiency. A Q-learning based scheme has been also proposed to adjust the back off window size of LBT in [18]. The performance of DCM mechanism has been analyzed in [19,20], where the reinforcement learning based methods have been employed to achieve fair coexistence between LTE-U and WiFi networks. A hybrid mechanism has been designed in [21], both LBT and DCM have been utilized and the flexible handoff between two mechanisms is achieved to meet fairness constraint.

Only a small amount of work has focused on D2D transmission on unlicensed bands. The conclusion of [22] has proven that D2D-U can significantly mitigate the congestion, conflicts and improve the system throughput. In [23], the sub-channel allocation of D2D-U enabled LTE cellular networks has been formulated as a many-to-many matching problem with externality and an iterative sub-channel swap algorithm has been proposed to improve the system performance. A reinforcement learning based scheme is proposed in [24], where a deep Q-network has been utilized to learn the traffic load on the unlicensed spectrum. It allows D2D-U system to model the joint allocation problem as a convex optimization problem. A decentralized joint spectrum and power allocation approach for D2D-U has been proposed in [25], which can meet the global minimization of power consumption among the D2D-U pairs.

After thorough investigation, most above mentioned power and spectrum allocation schemes are centralized, which may bring large signaling overhead to the base station and lead to high latency. Besides, most of the work concentrates only on maximizing system throughput without considering the different traffic loads of different D2D pairs. In this paper, we first define the unlicensed channel traffic load according to the number of competing WiFi users when DCM scheme is applied at D2D-U network. A price based model where D2D-U users need to pay for the channel resources is then built and a *Neural network* (NN) is applied to estimate the price to use unlicensed spectrum at each D2D-U link adaptively. In order to guarantee the harmonious coexistence with WiFi networks and the fairness among D2D-U pairs, the loss function is designed specifically to realize the online unsupervised learning of NN.

However, the training of NN in a distributed system is always unstable. If individual D2D-U links experience excessive noise or over fitting problem, it will cause severe interference to the performance of the whole system. In addition, when new D2D-U links join the network, the randomly initialized NN parameters will also cause serious fluctuations to the system which has converged. To solve above problem, the federated learning method [26] is utilized to update the parameters of the networks. In the scenario of federated learning, distributed terminals train NNs themselves, and then integrate gradients or parameters at the center base station periodically [27]. In our model, the integrated parameters can help correct those networks with large deviations and also initialize parameters for new D2D-U users. Afterwards, with the corresponding prices of channels, the spectrum and power allocation can be optimized jointly to maximize the transmission rates at each D2D-U pair.

The main contributions of the paper are summarized as follows.

1. A DCM based channel access model is built for D2D-U networks to share the unlicensed spectrum with WiFi networks. According to the *carrier sensing multiple access with collision avoidance* (CSMA/CA) mechanism adopted in WiFi, a novel traffic load on the unlicensed channel is defined.
2. To balance the SE and traffic load of D2D-U pairs while ensuring the fairness, a virtual variable, named as price, is defined, which is related to the traffic load, channel state of D2D-U links and the total

traffic load on the unlicensed channels. With the price, the transmission power and unlicensed spectrum allocation can be optimized jointly in a distributed way at each D2D-U pair.

3. Since it is hard to formulate an explicit function to model the relationship between the price, the traffic load and channel state information, an online trained NN with a specific loss function is applied to update the price at each D2D-U pair adaptively. The parameters of NN are updated via unsupervised self-iteration as well as federated learning to stabilize the system performance.

4. The centralized optimal solution is presented for comparison. Moreover, the simulation results are provided to verify the effectiveness of the scheme and the theoretical analysis.

The rest of this paper is organized as follows. Section 2 introduces the system model and a novel definition of WiFi traffic load on unlicensed channels. The price based learning model is proposed in Section 3 and Section 4, respectively. We analyze the simulation results in Section 5 and summarize the paper in Section 6

2 System Model

In this paper, we study the scenario where multiple D2D-U links share the unlicensed channels with WiFi *Access points* (APs), as shown in Fig. 1, where D2D-U links are able to simultaneously use multiple unlicensed channels and a single unlicensed channel can be shared by more than one D2D-U pair. *Macro base station* (MBS) can obtain the information on the achievable data rates of each D2D-U pair and the parameters of NNs via the control channel on licensed bands. It is worth mentioning that, in order to consider more realistic scenarios, the number of D2D-U links is not fixed, which means that D2D-U pairs may leave and join the network dynamically. To model the system mathematically, we use set $\mathcal{D} = \{d_0, d_1, ..., d_{N-1}\}$ to demonstrate N D2D-U links in the coverage of the MBS and N is not fixed. Moreover, there are M accessible unlicensed channels, denoted by set, $\mathcal{U}$, $\mathcal{U} = \{u_0, u_1, ..., u_{M-1}\}$, which are orthogonal with each other. To consider the fairness among D2D-U links and the harmonious coexistence with WiFi networks, $\mathcal{L}^{\mathcal{D}} = \{l_0^D, l_1^D, ..., l_{N-1}^D\}$ and $\mathcal{L}^{\mathcal{U}} = \{l_0^U, l_1^U, ..., l_{M-1}^U\}$ are used to denote the traffic loads of D2D-U links and WiFi systems on the unlicensed channels, respectively. In addition, WiFi APs adopt the CSMA/CA based *distributed coordination function* (DCF) to access the unlicensed channels while DCM mechanism is applied at D2D-U links to access the unlicensed channels.
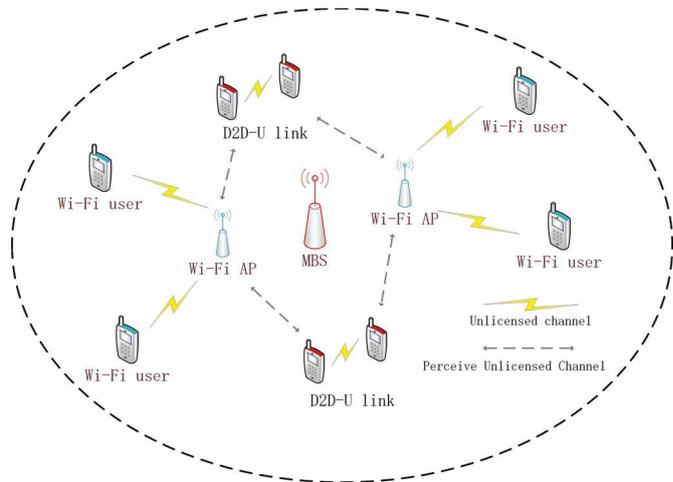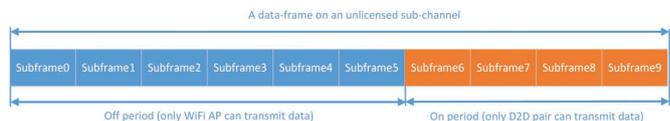

Fig. 1 System model.

Fig. 2 DCM mechanism.

2.1 Achievable data rates at D2D-U links

The transmission mode when applying DCM mechanism at D2D-U links is shown in Fig. 2, where the time frames on the unlicensed channels are divided into two parts. WiFi users can only transmit data during the first part, which is called 'off period', and the remaining section, named as 'on period', is occupied by D2D-U pairs. We further use $\theta_{i,j}$, $\theta_{i,j} \in [0, 1]$, to represent the proportion used by D2D-U link d_i on unlicensed channel u_j. Then, the achievable data-rate at d_i on unlicensed channels can be calculated by

$$R_i = \sum_{j=0}^{M-1} \theta_{i,j} B_j \log \left(1 + \frac{p_{i,j} h_{i,j}}{N_0 B_j} \right), \qquad (1)$$

where B_j is the bandwidth of unlicensed channel u_j, $h_{i,j}$ and $p_{i,j}$ are the channel power gain and corresponding transmission power of D2D-U pair d_i on u_j, respectively. N_0 is the noise power spectrum on unlicensed channel, which is fixed in the manuscript.

2.2 WiFi traffic load definition

To ensure the transmission requirements of WiFi users, D2D-U links must decide the duration of 'on period' based on the WiFi traffic load on the unlicensed channels, which means that D2D-U links need to obtain WiFi traffic load before accessing the channels.
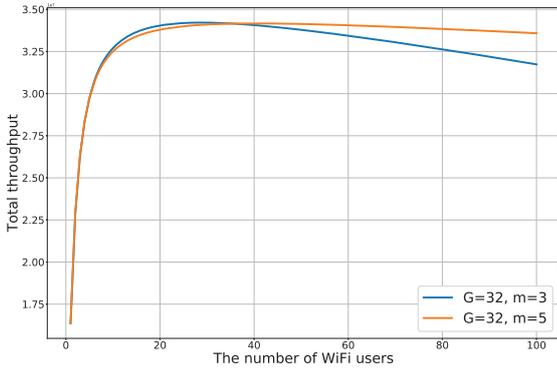
Fig. 3 Relationship between throughput and number of WiFi users.

Fig. 4 Relationship between $\hat{r}'_k$ and $\hat{r}^{\max}_k$.

The conventional traffic load of WiFi networks is mainly decided by the number of competing WiFi users. As demonstrated in [28,29], the extended Kalman filter has been used to achieve an accurate estimation on the number of active WiFi users, where D2D-U links filter the transmission collision probability in the channel and then obtain the number of WiFi users based on the estimated transmission collision probability. However, the impact of the number of WiFi users on the throughput of the WiFi system is non-linear and D2D-U links cannot directly determine the duration of 'on period' based on the number of WiFi users. Therefore, a novel WiFi traffic load definition is first defined when the DCM mechanism is employed at the D2D-U pairs.

As WiFi APs adopt binary slotted exponential back-off scheme in DCF, the relationship between the total WiFi throughput on an unlicensed channel and the number of WiFi users could be obtained according to [31], as illustrated in Fig 3. The size of back-off contention window, denoted as G, is 32 and the maximum back-off contention stage, denoted as m, is set to 3 and 5, respectively. We can observe that, as the number of WiFi users increases, the achievable throughput on the unlicensed channel increases first and then decreases. The reason is that when a large number of WiFi users compete for the same unlicensed channel, the transmission collision probability will increase, resulting in transmission failure and lower throughput.

Herein, we use the number of WiFi users corresponding to the highest throughput, $n^{\max}_k$, to represent the maximum load that the WiFi network can handle on unlicensed channel u_k. If the number of WiFi users in the unlicensed channel is greater or equal to $n^{\max}_k$, the channel u_k is considered inaccessible to D2D-U pairs. Furthermore, in order to define the WiFi traffic load to fit the DCM mechanism, when the throughput of the WiFi network reaches maximum, the av-
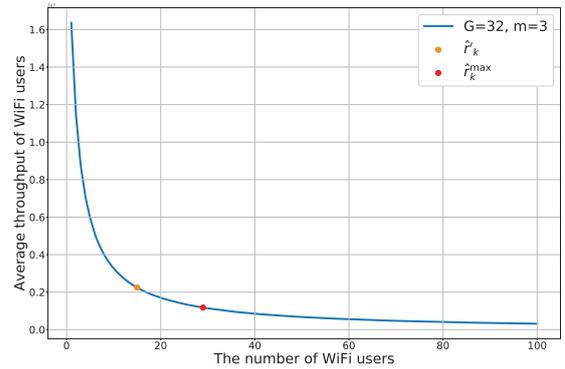
erage throughput of each WiFi user, $\hat{r}^{\max}_k$, is treated as the basic throughput guarantee of WiFi users. Let $R^{\max}_k$ denote the maximum system throughput on unlicensed channel u_k. Then, $\hat{r}^{\max}_k$ can be calculated as $\hat{r}^{\max}_k = \frac{R^{\max}_k}{n^{\max}_k}$. The basic throughput guarantee means that when D2D-U pairs reuse the unlicensed channels based on DCM mechanism, the average throughput of WiFi users should not be less than $\hat{r}^{\max}_k$.

Then we can calculate the minimum value of the 'off period' based on the above description. For unlicensed channel $u_k \in \mathcal{U}$, let $\hat{r}_k$ represent the average throughput of each WiFi user when no D2D-U links use u_k. On the other hand, when the number of WiFi users is less than $n^{\max}_k$, D2D-U pairs are allowed to use u_k with DCM mechanism and the average throughput of WiFi users is given by

$$\hat{r}'_k = \hat{r}_k(1 - \sum_{i=0}^{N-1} \theta_{i,k}). \tag{2}$$

According to the basic throughput guarantee, we can further achieve

$$\hat{r}^{\max}_k = \frac{R^{\max}_k}{n^{\max}_k} \leq \hat{r}'_k. \tag{3}$$

The relation ship between $\hat{r}'_k$ and $\hat{r}^{\max}_k$ is shown in Fig 4, when $\hat{r}'_k$ locates on the left side of $\hat{r}^{\max}_k$, u_k is available to D2D-U users. In order to adapt to DCM access mechanism, the WiFi traffic load l^U_k on u_k is defined as the minimum 'off period' duration that meets the basic throughput guarantee of WiFi users. Combining (2) and (3), l^U_k can be given by

$$l^U_k = \frac{\hat{r}^{\max}_k}{\hat{r}_k} \leq 1 - \sum_{i=0}^{N-1} \theta_{i,k}. \tag{4}$$

Since both $\hat{r}^{\max}_k$ and $\hat{r}_k$ can be calculated according to the physical layer parameters of the WiFi networks

[31] and the number of WiFi users can be accurately estimated by extended Kalman filter based methods [28, 29] and learning based method [30], D2D-U links are able to sense the traffic load in the unlicensed channel and then decide their own resource allocation scheme. In next section, the resource allocation model is built for D2D-U links and a priced-based solution is proposed.

3 Distributed price based Model

In this section, we first formulate a distributed optimization problem for each D2D-U link to maximize its own data rate. Then, in order to ensure fairness among D2D links, a priced-based solution is proposed to provide D2D-U links with different prices for using unlicensed spectrum under different traffic load and channel state conditions.

3.1 Problem formulation

For a single D2D link, $d_i \in \mathcal{D}$, to maximize its transmission rates while guaranteeing the performance of WiFi networks, an optimization problem can be formulated as

$$\max_{\{\theta_{i,j}, p_{i,j}\}} \{R_i\}, \tag{5}$$

subject to

$$C1 : \theta_{i,j} \leq 1 - l_j^U, \ \forall j \in [0, M-1], \tag{5a}$$

$$C2 : \sum_{j=0}^{M-1} \theta_{i,j} p_{i,j} \leq p_c, \tag{5b}$$

$$C3 : \theta_{i,j} p_{i,j} \leq p_u, \ \forall j \in [0, M-1], \tag{5c}$$

where $C1$ is to guarantee the fair coexistence with WiFi networks on the unlicensed channels, $C2$ is the total power constraint of d_i and $C3$ is the transmission power limit on the unlicensed channel according to the regulation.

Problem (5) is a non-convex problem but can be converted into a convex optimization problem and solved on each D2D link. In detail, a extra parameter $\eta_{i,j} = \theta_{i,j} p_{i,j}$ is introduced and (1) can be re-expressed as

$$R_i = \sum_{j=0}^{M-1} \theta_{i,j} B_j \log \left(1 + \frac{\eta_{i,j} h_{i,j}}{N_0 B_j \theta_{i,j}} \right), \tag{6}$$

then problem (5) is converted into

$$\max_{\{\theta_{i,j}, \eta_{i,j}\}} \{R_i\}, \tag{7}$$

subject to

$$C1 : \theta_{i,j} \leq 1 - l_j^U, \ \forall j \in [0, M-1], \tag{7a}$$

$$C2 : \sum_{j=0}^{M-1} \eta_{i,j} \leq p_c, \tag{7b}$$

$$C3 : \eta_{i,j} \leq p_u, \ \forall j \in [0, M-1]. \tag{7c}$$

Problem (7) is a convex optimization problem and can be solved by Lagrangian multiplier method. However, optimization problem (7) can only allow a D2D-U link to maximize its own throughput under the constraint of guaranteeing the fair coexistence with WiFi networks without considering the impact on other D2D-U pairs. When multiple D2D-U links share the same unlicensed channel, the possibility of transmission collision is extremely high, which leads to the lose on the performance of the D2D-U transmission. Therefore, the model needs to be improved based on the respective traffic load conditions of D2D-U links, where D2D-U links with heavy transmission tasks could use more spectrum resources while D2D-U links with light transmission tasks require only a small fraction of unlicensed spectrum resources. In next subsection, a priced-based solution is applied to achieve this goal.

3.2 Priced-based solution

Different from the Stackelberg game based method in [9,33,34] to minimize users' overhead, our method adjusts the price when assets are limited. In the proposed price-based model, each D2D-U link is considered as a consumer and spectrum resources in unlicensed bands are provided to consumers as commodities. The total money which each D2D-U link has are set to the same, which is represented by C. Define the price corresponding to the unlicensed channel u_j for the D2D-U link d_i as $c_{i,j}$ and when d_i transmits data on u_j, the price d_i needs to pay is written as $\theta_{i,j} \times c_{i,j}$. Accordingly, the optimization problem(5) can be expressed with an extra fairness constraint as

$$\max_{\{\theta_{i,j}, p_{i,j}\}} \{R_i\}, \tag{8}$$

subject to

$$C1 : \theta_{i,j} \leq 1 - l_j, \ \forall j \in [0, M-1], \tag{8a}$$

$$C2 : \sum_{j=0}^{M-1} \theta_{i,j} p_{i,j} \leq p_c, \tag{8b}$$

$$C3 : \theta_{i,j} p_{i,j} \leq p_u, \ \forall j \in [0, M-1], \tag{8c}$$

$$C4 : \sum_{j=0}^{M-1} \theta_{i,j} c_{i,j} \leq C. \tag{8d}$$

The above problem is also a non-convex optimization problem and can not be solved in its current formation. Same as in Problem (5), replacing $p_{i,j}$ with $p_{i,j} = \frac{\eta_{i,j}}{\theta_{i,j}}$ and the above problem is converted into

$$\max_{\{\theta_{i,j}, \eta_{i,j}\}} \{R_i\}, \qquad (9)$$

subject to

$$C1: \theta_{i,j} \leq 1 - l_j, \ \forall j \in [0, M-1], \qquad (9a)$$

$$C2: \sum_{j=0}^{M-1} \eta_{i,j} \leq p_c, \qquad (9b)$$

$$C3: \eta_{i,j} \leq p_u, \ \forall j \in [0, M-1], \qquad (9c)$$

$$C4: \sum_{j=0}^{M-1} \theta_{i,j} c_{i,j} \leq C. \qquad (9d)$$

Herein, one important key to solve 9 is to find $c_{i,j}$. If $c_{i,j}$ is known, the above problem is a convex optimization problem and the optimal solution can be obtained based on the Lagrangian multiplier method. The Lagrangian function of Problem 9 is constructed as

$$
\begin{aligned}
L(\theta_{i,j}, \eta_{i,j}, \mu_j^{(1)}, \mu^{(2)}, \mu_j^{(3)}, \mu^{(4)}) &= -R_i \\
+ \sum_{j=0}^{M-1} \mu_j^{(1)}(\theta_{i,j} + l_j - 1) &+ \mu^{(2)}(\sum_{j=0}^{M-1} \eta_{i,j} - C) \\
+ \sum_{j=0}^{M-1} \mu_j^{(3)}(\eta_{i,j} - p_u) &+ \mu^{(4)}(\sum_{j=0}^{M-1} \theta_{i,j} c_{i,j} - C),
\end{aligned}
\qquad (10)
$$

where $\mu_j^{(1)}$, $\mu^{(2)}$, $\mu_j^{(3)}$ and $\mu^{(4)}$ are the Lagrangian multipliers and the *Karush-Kuhn-Tucker* (KKT) conditions of $L(\cdot)$ are derived based on (10) as

$$\frac{\partial L}{\partial \theta_{i,j}} = 0, \ \forall j \in [0, M-1], \qquad (11)$$

$$\frac{\partial L}{\partial \eta_{i,j}} = 0, \ \forall j \in [0, M-1], \qquad (12)$$

$$\mu_j^{(1)}(\theta_{i,j} + l_j - 1) = 0, \ \forall j \in [0, M-1], \qquad (13)$$

$$\mu^{(2)}(\sum_{j=0}^{M-1} \eta_{i,j} - C) = 0, \qquad (14)$$

$$\mu_j^{(3)}(\eta_{i,j} - p_u) = 0, \ \forall j \in [0, M-1], \qquad (15)$$

$$\mu^{(4)}(\sum_{j=0}^{M-1} \theta_{i,j} c_{i,j} - C) = 0, \qquad (16)$$

$$\mu_j^{(1)} \geq 0, \mu^{(2)} \geq 0, \mu_j^{(3)} \geq 0, \mu^{(4)} \geq 0, \ \forall j \in [0, M-1]. \qquad (17)$$

On the basis of KKT conditions, the optimal solutions of $\theta_{i,j}$ and $\eta_{i,j}$ should satisfy

$$\eta_{i,j} = \theta_{i,j} B_j \left(\frac{\log e}{\mu^{(2)} + \mu_j^{(3)}} - \frac{N_0}{h_{i,j}}\right), \qquad (18)$$

and

$$\log(1 + \frac{\eta_{i,j} h_{i,j}}{N_0 B_j \theta_{i,j}}) - \frac{\eta_{i,j} h_{i,j} \log e}{N_0 B_j \theta_{i,j} + \eta_{i,j} h_{i,j}} = \frac{\mu_j^{(1)} + \mu^{(4)} c_{i,j}}{B_j}. \qquad (19)$$

Then, according to (18) and (19), $\eta_{i,j}$ and $\theta_{i,j}$ can be achieved for different d_i and u_j. Since (19) is a transcendental equation, numerical method can be applied to find the solution.

Based on above analysis, $\theta_{i,j}$ and $p_{i,j}$ can be obtained with the known price, $c_{i,j}$, which can be used to ensure the fairness among D2D-U pairs. To adjust prices adaptively to reach the fairness, each D2D-U pair needs to determine the $c_{i,j}$ based on its own traffic load, channel state information and the WiFi traffic load on the unlicensed channel. Accordingly, we denote $c_{i,j}$ as

$$c_{i,j} = F(l_i^D, l_j^U, h_{i,j} | s^c), \qquad (20)$$

where l_i^D represents the transmission task of d_i, $h_{i,j}$ represents the channel power gain on unlicensed channel u_j. The function $F(\cdot)$ is to model the relationship between the price and the traffic loads of the D2D links and the unlicensed channels. When the traffic load of D2D-U link, d_i, is heavy, d_i is encouraged to use unlicensed spectrum with a low price while the price to D2D-U link with less traffic load is high; Moreover, unlicensed channels with low traffic loads or strong channel power gain for the D2D link will need to be paid with low prices while under channels with high traffic loads or poor channel conditions will be expensive. In addition, to mitigate the channel access conflict among D2D-U links, a feedback signal s^c is set on d_i. If d_i collides with other D2D-U links on the channel, s^c is activated and the price should be enhanced accordingly. Therefore, function $F(\cdot)$ should have the following characteristics:

(1). $F(\cdot)$ should decrease monotonically with respect to l_i^D;
(2). $F(\cdot)$ should increase monotonically with respect to l_j^U;
(3). $F(\cdot)$ should decrease monotonically with respect to $h_{i,j}$;
(4). $F(\cdot)$ should increase with the activation of s^c.

As for the fairness among D2D-U links, we define *Expected transmission time* (ETT) as the ratio of a D2D-U link's traffic load to its achievable data rates.
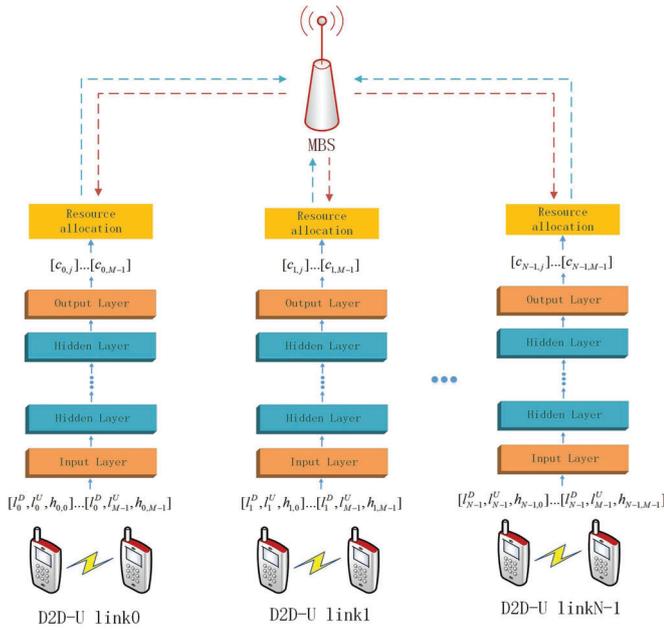
Fig. 5 The structure of NN.

Then, the fairness sharing on unlicensed channels among D2D-U links is denoted as that the ETT values of all D2D-U links are equal, which is written as

$$\frac{l_0^D}{R_0} = \frac{l_1^D}{R_1} = \cdots = \frac{l_{M-1}^D}{R_{M-1}}. \tag{21}$$

However, it is difficult to directly build an explicit mathematical model to formulate the function $F(\cdot)$ and achieve (21). To address this issue, an online training *neural network* (NN) architecture is exploited to implement function of $F(\cdot)$ and the loss function for all D2D-U links are provided based on s^c and the assistance of MBS. Specific details of the adopted NN will be given in the next section.

4 Learning based Method

Because of the strong fitting performance and robustness of NN, the online trained NN is utilized to achieve adaptive adjustment of the prices in a dynamic D2D-U environment, where the parameters of NN are updated according to the loss function as well as the federated learning. The structure of the whole distributed price system is illustrated in Fig. 5 where each D2D-U pair holds a NN itself and determines its own prices of different channels. In this section, the structure and the loss function of NN will be first introduced to train NN in an unsupervised online mode, then to improve the stability of the system and better accept new D2D-U users, the federated learning based mode will be proposed to adjust parameters in the networks.

4.1 Unsupervised online training method

As demonstrated in Fig. 5, the output of NN in D2D-U link d_i can be calculate as

$$\hat{c}_{i,j} = \hat{F}(l_i^D, l_j^U, h_{i,j}|\phi), \tag{22}$$

where $\hat{c}_{i,j}$ is the output as well as the price estimated by NN, the input of NN is $[l_i^D, l_j^U, h_{i,j}]$. In particular, before $[l_i^D, l_j^U, h_{i,j}]$ is feeded into NN, the input data needs to be normalized to avoid problems caused by different orders of magnitude. $\hat{F}(\cdot)$ indicates the forward propagation of NN and ϕ represents all the weights and bias parameters. At each iteration, all the parameters in NN are updated based on the gradient descent algorithm, which is denoted by

$$\phi = \phi - \alpha \frac{\partial Q}{\partial \phi}, \tag{23}$$

where α is the learning rate of NN and Q is the loss function. Since it is hard to achieve global optimal solution of problem (8), we cannot obtain labels and use the supervised learning method to train the network. Therefore, based on (21) and the collision detection, the loss function Q is formulated by two parts to train NN in an unsupervised way. Q_1 and Q_2 are used to represent these two parts, respectively, where Q_1 is defined as:

(1). if $\frac{l_i^D}{R_i}$ is larger than ETT values of $\frac{M+1}{2}$ D2D-U links (when M is odd) or $\frac{M}{2}$ D2D-U links (when M is even), $Q_1 = q$;

(2). if $\frac{l_i^D}{R_i}$ is smaller than ETT values of $\frac{M+1}{2}$ D2D-U links (when M is odd) or $\frac{M}{2}$ D2D-U links (when M is even), $Q_1 = -q$;

(3). else $Q_1 = 0$;

where q is the adjustment step size of the price and is set to a tiny positive value. Q_2 corresponds to conflict feedback s^c, which is defined as:

(1). if d_i collides with other D2D-U links, Q_2 is set to v_1;

(2). else $Q_2 = v_2$;

In order to mitigate the transmission collision among D2D-U links effectively in actual operation, v_1 is set to a larger positive value to significantly increase the price of the unlicensed channel when collision happens. v_2 is a negative value which aims at decreasing prices to allow D2D-U links to use more spectrum resources when no collision happens. Herein, the value Q_1 can be provided by MBS and Q_2 can be decided on d_i according to its transmission collision situation. Accordingly, the target

of NN output is denoted as $T = \hat{c}_{i,j} + Q_1 + Q_2$ and Q can be calculated by Mean-Squared Loss as

$$Q = (T - \hat{c}_{i,j})^2 = (Q_1 + Q_2)^2. \tag{24}$$

More specifically, in each time slot, d_i will obtain the prices of M unlicensed channels and get M loss values based on (24). Here M training data will be treated as a batch to jointly train the network. Furthermore, to keep the convergence of NN and the stability of output, we use Sigmoid function to limit the output value in a certain range according to the actual conditions. The activation function of the output layer is set to be $w \times$ Sigmoid$(\cdot)$, which limits the output in $[0, w]$.

It is noteworthy that each D2D-U link holds a NN independently to determine the price corresponding to the utilization on the unlicensed channels. When the neural networks of D2D-U links converge, the system has reached an equilibrium. When the traffic load of D2D-U links or WiFi system changes, neural networks will converge to a new equilibrium adaptively. However, in such an online dynamic environment, the network may be hard to converge due to environmental noise or over fitting. Due to the setting of the loss function, the un-convergence of a single NN may affect the updates of other D2D-U users, which leads to poor system performance. Additionally, when new D2D-U pairs join in the system, their NN parameters will be initialized randomly, which can also cause the instability of the system. To solved the mentioned problem, the federated learning based mode will be utilized to further adjust NN parameters.

4.2 Federated learning based method

Federated learning based schemes have achieved significant performance in distributed scenarios [32]. Without a large amount of data interaction, federated learning ensures safety and stability of the system by integrating the gradient or parameters information of distributed users at the central processor and then feedback to the distributed terminals. In our D2D-U model, each D2D-U pair is a distributed terminal and MBS plays the role as the center. Then MBS periodically collects NN parameters of D2D-U pairs in the region and averages them, which can be denoted as

$$\hat{\phi} = \frac{\sum_{i=0}^{N-1} \phi_i}{N}, \tag{25}$$

where ϕ_i is NN parameters of D2D-U link d_i. Since the traffic load and channel gain of different D2D-U links are not the same, the information contained in the neural network corresponding to the expected value

of $\hat{\phi}$, $E(\hat{\phi})$, is the average traffic load and transmission status of all D2D-U users in the region.

Then for new D2D-U pairs who join the system, instead of initializing NN parameters randomly, we use current $\hat{\phi}$ saved by MBS as the starting point of NN for training. It can effectively reduce the instability of the initial training of NN and accelerate the network convergence. For existing D2D-U pair d_i, in each iteration of parameter updates based on federated learning, d_i needs to adjust the value of ϕ_i according to $\hat{\phi}$. If NN in d_i has converged and output stable price value, ϕ_i should be kept or update little. On the other hand, if the output is unstable or the loss value is large, ϕ_i need correction. The update of ϕ_i is calculated by

$$\phi_i = \beta\hat{\phi} + (1 - \beta)\phi_i, \tag{26}$$

where $\beta \in (0, 1)$ is the update intensity parameter which is determined adaptively by the network convergence. In order to realize the judgment of network convergence, here the accumulated absolute loss value of NN is utilized to be the basis for judgment. Let $Q_{sum,i}$ be the accumulated absolute loss value of d_i, if the value of $Q_{sum,i}$ is tiny, which means that NN has converged and β should be small. When $Q_{sum,i}$ is a large number, β need to be set large accordingly. Here we use the sigmoid function again to express the relationship between $Q_{sum,i}$ and β as

$$\beta = \frac{1}{1 + e^{-(\epsilon x - \gamma)}}, \tag{27}$$

where γ and ϵ are constant value and decided by specific environment and the loss value. Figure 2 describes this relationship more intuitively, we can find that when $Q_{sum,i}$ is small, the value of β is close to 0, as $Q_{sum,i}$ increases, β also increases till close to 1.

Based on the above interpretation on the proposed NN and the federated learning based mode, the process of the joint power and spectrum allocation algorithm for a single D2D-U link, $d_i \in \mathcal{D}$, is summarized in Algorithm 1. T is the system clock recorded by MBS and T_{fl} is the length of the federated learning period.

5 Numerical Results

In this section, the simulation results are demonstrated to verify the performance of the proposed distributed D2D-U communication scheme. In the simulation setup, the relevant parameters of NN and federated learning are demonstrated in Table 1 and the parameters related to the D2D-U and WiFi networks are given in Table 2. Since the real-time performance of the algorithm is required in practice, we use shallow fully connected
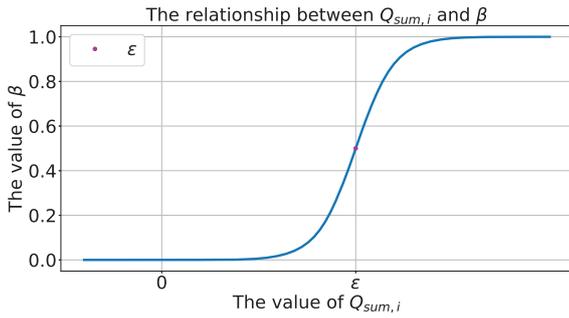
Fig. 6 The relationship between $Q_{sum,i}$ and β.

Algorithm 1 Distributed joint spectrum and power allocation at d_i

---

1: Initialize the structure and the parameters of d_i;
2: Initialize the structure and the parameters of NN from MBS;
3: **while** d_i is transmitting data **do**
4:    Estimate the number of WiFi users with EKF based method;
5:    Estimate the traffic load $\mathcal{L}^{\mathcal{U}}$ on unlicensed spectrum based on (4);
6:    Normalize the input of NN;
7:    The prices of all unlicensed channels are calculated based on (22);
8:    Problem(8) is solved to get the resources allocation scheme of d_i;
9:    Calculate the loss value on the basis of (24);
10:   ADD the absolute loss value to $Q_{sum,i}$;
11:   d_i trains the NN based on gradient descent algorithm in (23);
12:   **if** T can divide T_{fl} evenly **then**
13:      MBS collects data and calculates $\hat{\phi}$ based on (25);
14:      d_i updates ϕ_i based on (26);
15:      Reset $Q_{sum,i}$ to 0;
16:   **end if**
17: **end while**

---

neural networks to reduce algorithm complexity and for quickly convergence during online training processing. Then during an iteration time slot, D2D-U link will first estimate WiFi traffic load based on EKF method mentioned before, then the price of the unlicensed channels can be derived based on the proposed method. With the unlicensed channels' price as well as WiFi traffic load, D2D-U link can decide its spectrum and power allocation scheme. Finaly NN is trained based on the transmission transmission of D2D-U link and the little signaling interaction with MBS.

5.1 Effectiveness on the proposed NN

We first verify the effectiveness on the proposed scheme, assuming that there are two D2D-U links, d_1 and d_2, and two independent unlicensed channels, u_1 and u_2. The traffic load of d_1 is set to be larger than d_2 and

Table 1 Parameters of NN.

Parameters	Value
Learning rate α	0.0001
Number of hidden layers	2
Number of neurons	32/32
Active function	tanh/tanh/tanh/$w \times$ sigmoid
Max value of output w	10
Fairness-based loss q	0.01
Conflict multiplication loss v_1	0.03
Maximum throughput loss v_2	-0.03
Federated learning param γ	1.2
Federated learning param ϵ	0.4
Federated learning period T_{fl}	100

Table 2 Parameters of D2D-U links.

Parameters	Value
Total power control p_c	35dBm
Power control on one unlicensed channel p_u	23dBm
AWGN noise power N	-95dBm
Total assets C	1

the WiFi traffic load on unlicensed channel u_1 is set to be less than that on u_2. Then the output prices from the NNs of two D2D-U links on two unlicensed channels are illustrated in Fig. 7. It can be observed that after about 250 iterations, the D2D-U system has achieved convergence, which means that the loss value of NN is tiny and each D2D-U link can stably estimate the price of unlicensed channels.

Here we can find that the prices of u_1 and u_2 is much more cheaper for d_1 with heavy transmission tasks, which implies that d_1 is encouraged to use more unlicensed spectrum resources. In addition, since the WiFi traffic load on u_1 is lighter than that on u_2, the value of $c_{1,1}$ is smaller than $c_{1,2}$ and d_1 will select the unlicensed channel u_1 in the first place. As for d_2, since u_1 has priority to be selected by d_1 with much less price, to avoid transmission collision, $c_{2,1}$ is trained larger than $c_{2,2}$ and d_2 will mainly use u_2 for transmission. Fig. 8 shows the fairness between d_1 and d_2, where we can observe that as the prices converge, $\frac{l_1^D}{R_1}$ is equal to $\frac{l_2^D}{R_2}$ and the fairness between d_1 and d_2 is achieved.

5.2 Verification on the fairness

To further illustrate the fairness of the proposed algorithm, the performance of D2D-U system with more D2D-U links and unlicensed channels is analyzed and the result is compared with centralized algorithm which concentrates on maximizing system throughput. In the centralized algorithm, price based model is not applied and all the parameters are collected at MBS. A joint
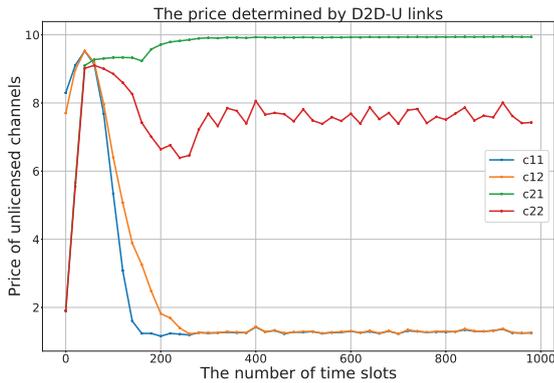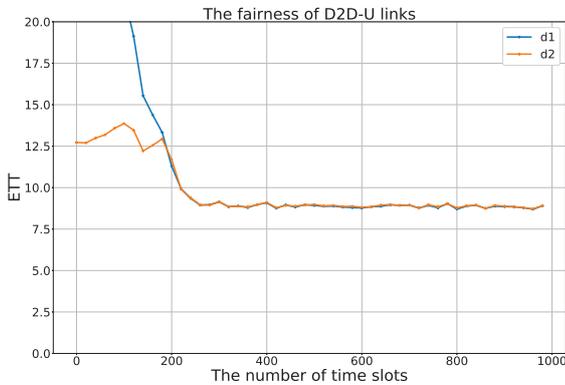
Fig. 7 The price determined by D2D-U links.

Fig. 8 The transmission fairness among D2D links.

Table 3 The comparison of transmission fairness.

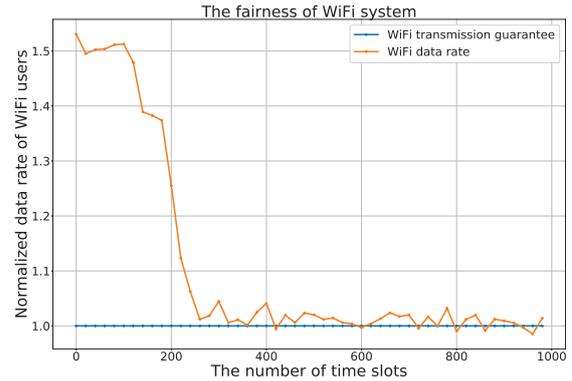| Users number	Normalized traffic load	ETT(price-based)	ETT(max throughput)
1	0.8	8.779	14.010
2	0.6	8.788	10.508
3	0.4	8.750	7.005
4	0.2	8.810	3.503

Fig. 9 Achieved fairness of WiFi system.

5.3 Effectiveness of federated learning

To further verify the effectiveness of federated learning, the convergence results of new D2D-U pairs with different NN parameters are illustrate in Fig. 10 and Fig. 11, respectively. Here the new D2D-U link in Fig. 10 and Fig. 11 are the same instead of NN parameters and system has converged before new users coming, then ETT value of new D2D-pairs and the average ETT value of all D2D-U pairs are provided to prove the performance of the proposed scheme. In Fig. 11, NN parameters are initialized randomly where new D2D-U pair spends more time to achieve the optimal output and the fluctuation of system performance is also greater. On the contrary, with the federated learning parameters saved in MBS, new D2D-pair in Fig. 10 completes convergence faster and the system performance is more stable.

5.4 Achievable data rates

Fig. 12 demonstrates the total achievable data rates of D2D-U links with centralized and proposed scheme, respectively. From the figure, we can observe that the data rates obtained by the proposed are close to that by the centralized scheme, which indicates that the proposed method can almost maximize the system performance while ensuring the fairness of D2D-U pairs.
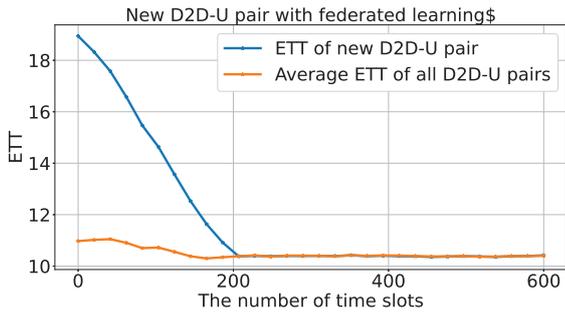
optimization problem is then built and solved to maximize the throughput of D2D-U system while guaranteeing the harmonious coexistence with the WiFi networks. The ETT value comparison of two schemes is shown in Table 3. When the normalized traffic load of D2D-U links is '1', the actual traffic load is $10e8$bits. It can be found that the proposed price based distributed method can adaptively allocate resources at the D2D-U pairs with respect to the traffic load and channel conditions to guarantee the fairness. On the other hand, the centralized algorithm is far from achieving the fairness.

The fair coexistence with WiFi networks by the proposed scheme is depicted in Fig. 9, where the actual WiFi traffic load is normalized based on the the basic WiFi throughput guarantee calculated in (3). Simulation result shows that after the convergence of the price, the WiFi throughput is basically equal to the basic WiFi throughput guarantee. Due to the set of v_2 which encourages D2D-U links to use more spectrum resources with less collision, there are a little transmission collision which leads to the tiny impairment to the WiFi system throughput.

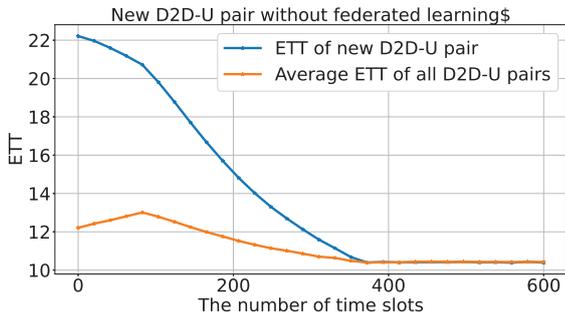**Fig. 10** The ETT of new D2D-U pair with federated learning.

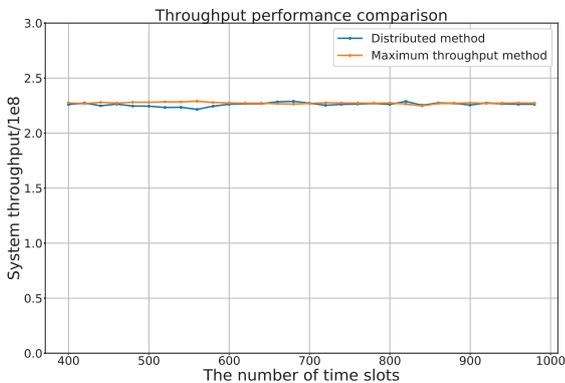**Fig. 11** The ETT of new D2D-U pair without federated learning.

Fig. 12 The comparison of system throughput between price based method and maximum throughput method.

Besides, the details of D2D-U link power allocation is illustrated in Fig. 13, where the unlicensed channel traffic load on u_0 and u_1 is low and on u_2 and u_3 is high. The abscissa of Fig. 13 is different unlicensed channels and the ordinate is $\eta_{i,j}$ of related D2D-U pairs and unlicensed channels. It can be observed that in the price based model, the D2D-U link with more traffic load reuses the spectrum resources of more ideal channels and the D2D-U link with less traffic load chooses to reuse more crowded channels. While in the centralized method, the change of D2D-U traffic load has no effect
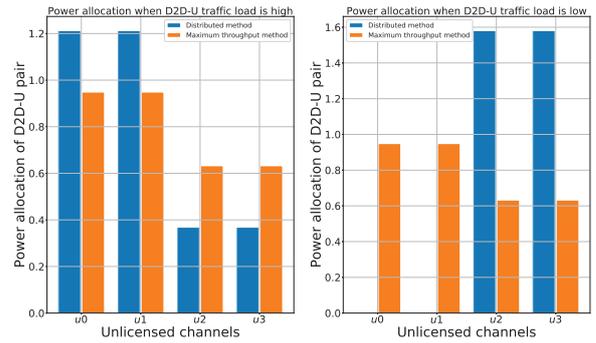

Fig. 13 The comparison of power allocation between price based method and maximum throughput method.

on its power allocation scheme. Therefore, simulation results justify that the proposed method can guarantee the fairness among D2D-U pairs with least lose on the data rates comparing with the centralized optimal solution.

6 Conclusion

In this paper, in order to reuse the spectrum resources on the unlicensed spectrum to improve the performance of D2D-U system, a distributed power and spectrum allocation mechanism with adaptive price adjustment scheme is proposed. An unsupervised online learning structure is employed on each D2D-U link to estimate the prices of all perceived unlicensed channels, a federated learning based approach is utilized to improve system stability and performance. Then the power and spectrum optimization models can be formulated and solved by D2D-U pairs to access the unlicensed spectrum. Numerical simulation proves that the proposed algorithm allows D2D-U link to maximize data-rate while ensuring the fairness of WiFi system and the fairness among different D2D-U users.

References

1. Z. Zou, R. Yin, C. Wu, J. Yuan, and X. Chen, "Distributed spectrum and power allocation for D2D-U Networks,"

EAI. MONAMI., Nov. 2020.

2. K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Deviceto-device communication as an underlay to LTE-Advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42-49, Dec. 2009.

3. G. Yu, L. Xu, D. Feng, R. Yin, G. Li, and Y. Jiang, "Joint mode selection and resource allocation for device-to-device communications," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3814-3824. Nov. 2014.

4. Y. Wu, W. Guo, H. Yuan, L. Li, S. Wang, X. Chu, and J. Zhang "Device-to-device meets lte-unlicensed," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 154-159, May. 2016.

5. R. Liu, G. Yu, F. Qu, Z. Zhang, "Device-to-device communications in unlicensed spectrum: mode selection and resource allocation," *IEEE ACCESS*, vol. 4, pp. 4720-4729, Aug. 2016.

6. S. Shah, M. Rahman, A. Mian, A. Imran, S. Mumtaz, and O. Dobre, "On the impact of mode selection on effective capacity of device-to-device communication," *IEEE Wireless Commun. Letters*, vol. 8, no. 3, pp. 945-948, Jun. 2019.

7. G. Yu, L. Xu, D. Feng, R. Yin and G. Li, "Joint mode selection and resource allocation for device-to-device communications," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3814-3824, Nov. 2014

8. R. Yin, C. Zhong, G. Yu, Z. Zhang, K. Wong, and X. Chen, "Joint Spectrum and Power Allocation for D2D Communications Underlaying Cellular Networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2182-2195, Apr. 2016.

9. R. Yin, G. Yu, H. Zhang, Z. Zhang, and G. Li, "Pricing-Based Interference Coordination for D2D Communications in Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1519-1532, Mar. 2015.

10. W. Lee, M. Kim and D. Cho, "Transmit Power Control Using Deep Neural Network for Underlay Device-to-Device Communication," *IEEE Commun. Letters*, vol. 8, no. 1, pp. 141-144, Feb. 2019.

11. A. Moussaid, W. Jaafar, W. Ajib and H. Elbiaze, "Deep Reinforcement Learning-based Data Transmission for D2D Communications," *IEEE Int. Conf. WiMob.*, 2018, pp. 1-7.

12. QualComm, "Extending LTE advanced to unlicensed spectrum," *White Paper*, San Diego, CA, USA, Dec. 2013.

13. Huawei, "U-LTE: Unlicensed spectrum utilization of LTE," *White Paper*, Shenzhen, China, 2014.

14. "Study on licensed-assisted access to unlicensed spectrum (Release 13)," *document 3GPP*, Sophia Antipolis Cedex, France, TR 36.889, Jul. 2015.

15. X. Sun, and L. Dai, "Towards fiar and efficient spectrum sharing between LTE and WiFi in unlicensed bands: fairness-constrained throughput maximization," *IEEE Trans. Wireless Commun.*, pp. 1-1, Jan. 2020, early access.

16. Q. Cui, W. Ni, S. Li, B. Zhao, R. Liu, and P. Zhang, "Learning-assisted clustered access of 5G/B5G networks to unlicensed spectrum," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 31-37, Feb. 2020.

17. R. Yin, G. Yu, A. Maaref, and G. Li, "LBT-based adaptive channel access for LTE-U systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6585-6597, Oct. 2016.

18. V. Maglogiannis, D. Naudts, A. Shahid and I. Moerman, "A Q-Learning Scheme for Fair Coexistence Between LTE and Wi-Fi in Unlicensed Spectrum," *IEEE Access*, vol. 6, pp. 27278-27293, May. 2018.

19. P. Santana, V. Sousa, F. Abinader, and J. Neto, "DM-CSAT: a LTE-U/Wi-Fi coexistence solution based on reinforcement learning," *Telecommun Syst*, vol. 71, no. 4, pp. 615-626, Aug. 2019.

20. J. Neto, S. Neto, P. Santana, and V. Sousa, "Multi-cell LTE-U/Wi-Fi coexistence evaluation using a reinforcement learning framework," *Sensors*, vol. 20, no. 7, pp. 1855-1877, Mar. 2020.

21. S. Liu, R. Yin and G. Yu, "Hybrid Adaptive Channel Access for LTE-U Systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9820-9832, Oct. 2019.

22. S. Andreev, O. Galinina, A. Pyataev, K. Johnsson, and Y. Koucheryavy, "Analyzing assisted offloading of cellular user sessions onto D2D links in unlicensed bands," *IEEE J. Selected Areas in Commun.*, vol. 33, no. 1, pp. 67-80, Jan. 2015.

23. H. Zhang, Y. Liao and L. Song, "D2D-U: Device-to-Device Communications in Unlicensed Bands for 5G System," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3507-3519, June 2017.

24. Z. Zou, R. Yin, X. Chen and C. Wu, "Deep Reinforcement Learning for D2D transmission in unlicensed bands," *IEEE/CIC ICCC Workshops*, pp. 42-47. 2019.

25. R. Yin, Z. Wu, S. Liu, C. Wu, J. Yuan and X. Chen, "Decentralized Radio Resource Adaptation in D2D-U Networks," *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2020.3016019.

26. J. Konecny, H. Brendan McMahan, F. Yu, P. Richtarik, A. Suresh and D. Bacon, "Federated Learning: Strategies for Improving Communication Efficiency," *arXiv preprint arXiv:11610.05492*, 2017.

27. S. Niknam, H. S. Dhillon, and J. H. Reed, "Federated learning for wireless communications: Motivation, opportunities and challenges," *arXiv preprint arXiv:1908.06847*, 2019.

28. G. Bianchi and I. Tinnirello, "Kalman filter estimation of the number of competing terminals in an IEEE 802.11 network," *IEEE INFOCOM*, vol. 2, pp. 844-852, Feb. 2003.

29. F. Qin, X. Dai and J. Mitchell, "Effective-SNR estimation for wireless sensor network using Kalman filter," *Ad Hoc Networks*, vol. 11, no. 3, pp. 944-958, May. 2013.

30. R. Yin, Z. Zou, C. Wu, J. Yuan, X. Chen and G. Yu, "Learning-based WiFi Traffic Load Estimation in NR-U Systems," *IEICE Trans.*, doi: 10.1587/transfun.2020EAP1063, Aug. 2020.

31. G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535-547, Mar. 2000.

32. Q. Yang, Y. Liu, T. Chen and Y. Tong, "Federated Machine Learning: Concept and Applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, Feb. 2019.

33. Z. Zhou, B. Wang, B. Gu, B. Ai, S. Mumtaz and J. Rodriguez, "Time-Dependent Pricing for Bandwidth Slicing Under Information Asymmetry and Price Discrimination," *IEEE Trans. Commun.*, vol. 68, no. 11, Nov. 2020.

34. B. Gu, X. Yang, Z. Lin, W. Hu, M. Alazab and R. Kharel, "Multi-Agent Actor-Critic Network-based Incentive Mechanism for Mobile Crowdsensing in Industrial Systems," *IEEE Trans. Ind Inform.*, 2020.