Partially Overlapping Channel Selection in Jamming Environment: A Hierarchical Learning Approach

Lei Zhao[1,2](✉), Jincheng Ge[3], Kailing Yao[1,2], Yifan Xu[1,2], Xiaobo Zhang[1], and Menglan Fan[4]

[1] College of Communications Engineering, PLA Army Engineering University, Nanjing, China
leizhao365@163.com, kailing-yao@126.com, yifanxu@163.com, xb_zhang2008@126.com

[2] Key Embedded Technology and Intelligent System Laboratory, Guilin University of Technology, Guilin, China

[3] Unit 95965 of PLA, Hengshui, China
gejincheng@hotmail.com

[4] Unit 31102 of PLA, Nanjing, China
fanmenglan@126.com

Abstract. This paper solves the channel selection with anti-jamming problem using partially overlapping channel (POC) in limited spectrum environment. Since it is difficult for users to obtain global information of networks, this paper realizes the coordination of channel access by the local information interaction. The channel selection with anti-jamming problem is formulated as a Stackelberg game where the jammer acts as leader and users act as followers. We prove that the game model exists at least one Stackelberg equilibrium (SE) solution. To achieve the equilibrium, a hierarchical learning algorithm (HLA) is proposed. Based on the proposed method, the system can achieve the improvement of throughput performance by minimizing local interference. Simulation results show the proposed algorithm can achieve good performance under jamming environment, and the network throughput can maintain a stable state with the jamming intensity increasing.

Keywords: Intelligent anti-jamming · Stackelberg game · Channel selection · Partially overlapping channel

This work was supported by the National Natural Science Foundation of China under Grant No. 61771488, No. 61671473 and No. 61701533, in part by Natural Science Foundation for Distinguished Young Scholars of Jiangsu Province under Grant No. BK20160034, and the Guang Xi Universities Key Laboratory Fund of Embedded Technology and Intelligent System (Guilin University of Technology).

X. B. Zhai et al. (Eds.): MLICOM 2019, LNICST 294, pp. 130–144, 2019.
https://doi.org/10.1007/978-3-030-32388-2_12

1 Introduction

With the development of technology, the communication security has attracted great attention. Nowadays, there are mainly two challenges for this problem. For one thing, the internal interference of multi-node networks will become serious as the number of network nodes increases. For another, the malicious jammer has become more and more intelligent which is able to sense the strategy of users and select the optimal channel to maximize the damage to the network.

Studies on anti-jamming problems have been a hot topic these years which are discussed from the following aspects: power domain, spectrum domain, space domain and multi-domains. In power domain, users adjust the transmission power to combat with the smart jammer. In [1,2], authors studied the anti-jamming problem with one jammer and one user. They modeled the power control problem based on the Stackelberg game. Authors of [3] considered the observed errors in power control problems on basis of [1,2]. In [4], authors considered the power control problem in unmanned aerial vehicle networks. In spectrum domain, legitimate users selected proper channels to escape from attacks of the malicious jammer. Authors of [5] discussed non-overlapping channel (NOC) selection problem which considered outer jamming and internal interference at the same time. In [6], authors proposed an effective channel selection with anti-jamming problem in cognitive radio network. In space domain, the existing work mainly focused on beam forming, users attempted to deal with the smart jammer by changing the antenna direction. In [7], a muti-domains anti-jamming problem was studied. When strength of jamming was weak, users adjusted their transmission power to confront the smart jammer. While the strength was becoming stronger, users had to select proper channels to maximize their utility function. In [8,9], anthors investigated the anti-jamming problems by considering joint channel selection and transmission rate. Analysing these studies above, we can easily find most papers consider anti-jamming problems based on NOCs, assuming spectrum resource is sufficient. However, as the number of users becomes larger, network bandwidth becomes scarce in practical environment. To tackle this problem, the technology of Partially overlapping channel (POC) is developed in this paper. POC, which is proposed by Mishra Arunesh of [10,11], has attracted great attention in wireless systems with limited spectrum resource. Based on POC, authors of [13] proposed a greedy algorithm to improve throughput compared with NOC. The existing work mentioned above shows that most studies about POC focused on interference elimination, rarely considering jamming communication environment.

It is of great significance to investigate the channel selection problems based on POCs in limited resource environment. In this paper, malicious jamming and mutual interference are both considered. We optimize the channel selection problem from the perspective of minimizing generalized interference. We formulate this problem as a Stackelberg game [5]. In this game, the jammer acts as the leader and the users act as the followers. The interference mitigation problem between users is modeled as a potential game [17]. After that, we propose a hierarchical learning framework to achieve the Stackelberg Equilibrium (SE).

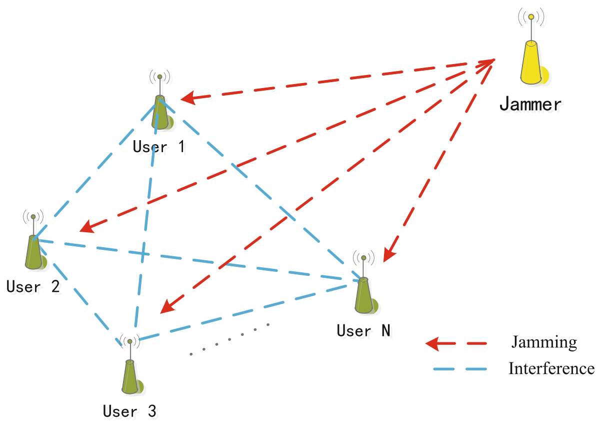

Fig. 1. Multi-user channel access model in jamming environment.

We assume that the jammer with learning ability uses Q learning to simulate the jammer's strategy. We propose the SAP-based algorithm to deal with the jamming impact as well as mutual interference between users. The contributions are summarized as follows:

(1) To study the anti-jaming problem in limited network bandwidth environment, the POC model is proposed and channel selection with anti-jamming problem is formulated as a Stackelberg game. We have proven the jammer's stationary action and users' policy consist a SE.
(2) We propose a hierarchical learning framework to deal with the proposed problem. Simulation results show that the proposed algorithm can achieve good performance in the jamming environment, and the network throughput can maintain a stable state when the jamming intensity increases.

The differences between [5] and our work are displayed as follows. Firstly, we study the channel selection problems based on POCs in limited resource environment while the channel selection problems are discussed in NOCs environment with abundant bandwidth resources. What's more, from the algorithm perspective, we formulate a SAP-based algorithm to deal with the channel selection problem while a Stochastic-learning-automata(SLA)-based algorithm is proposed in [5].

The rest of this paper is discussed as follows. In Sect. 2, the system model and the anti-jamming problem are introduced. In Sect. 3, we propose a Stackelberg game to model the anti-jamming problem. In Sect. 4, a hierarchical framework based on SAP is discussed. In Sects. 5 and 6, we present the simulation results and draw the conclusion.

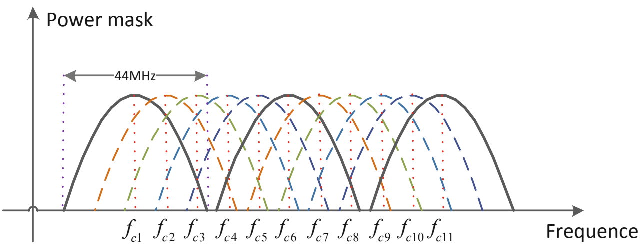

Fig. 2. The illustration of POC.

2 System Model and Problem Formulation

As shown in Fig. 1, there is a distributed mesh network consisting of N users. A smart jammer outside the network aims to block the communication of users. In this work, we take the IEEE 802.11b standard for example. Figure 2 shows the illustration of POC. In the limited spectrum environment, the bandwidth of each channel is 22 MHz and there are only 3 orthogonal channels. In POC situation, adjacent channels overlap partially and separation between two near channels is 5 MHz. As the imaginal lines show, there are 11 partially overlapping channels in the same bandwidth range of the 3 orthogonal channels mentioned above.

The user set is denoted by $\mathcal{N} = \{1, 2, ..., N\}$. Strategy profile of all users is denoted as $\mathbf{a} = \{a_1, ..., a_N\}$, where a_n is denoted as user n's action and $\mathbf{a}_{-n} = \{a_1, a_2, ..., a_{n-1}, a_{n+1}, ..., a_N\}$ is the action profile of other users except user n. The channel sets of users and the jammer are denoted as $\mathcal{M} = \{1, 2, ..., M\}$ and $\mathcal{C} = \{1, 2, ..., H\}$ and $\mathcal{M} = \mathcal{C}$. For simplicity, we assume that the jammer only selects one channel to jam each decision-making time. In this paper, as we study on the channel selection problem, we assume that the transmission power of each user is the same as each other.

The smart jammer can recognize users' actions and adjust its jamming strategy dynamically to maximize its damage. Having acknowledged the jammer's strategy, users adopt proper strategies to minimize the external jamming and interference from adjacent channels. However, as the scale of network becomes larger, network bandwidth becomes scarce in practical environment. To overcome this problem, the technology of POC is introduced in this paper.

In POC situation, interference between user n and m is affected by the physical distance d_{mn} and channel distance $\delta_{mn} = |a_m - a_n|$ [12]. For example, when the physical distance between two users are large, they could select the same channel without interfering with each other. To the opposite, when the physical distance is small, users have to choose different channels to decrease the interference. In this work, we have to consider two aspects. Firstly, due to the property of partially overlapping, interference from adjacent channels inevitably degrades the network performance. Secondly, the smart jammer could sense and recognize the users' channel selection strategy and destroy the information transmission.

As two users select channels independently, the overlapping power mask is expressed as [12]:

$$H(f) = \int p_m(f) \cdot p_n(f) df, \tag{1}$$

$p_m(f)$ and $p_n(f)$ are the power mask selected independently by user m and n. The central frequency of one channel is denoted as f_c and the power mask can be expressed as:

$$p(f) = \begin{cases} 0\text{dB}\,, & |f - f_c| \leq 11\text{MHz} \\ -30\text{dB}\,, & 11\text{MHz} \leq |f - f_c| \leq 22\text{MHz}. \\ -50\text{dB}\,, & |f - f_c| \geq 22\text{MHz} \end{cases} \tag{2}$$

As is verified in [14], the overlapping power mask between user m and user n can be expressed as:

$$H(\delta_{mn}) = \begin{cases} 1, & \delta_{mn} = 0 \\ 0.605, & \delta_{mn} = 1 \\ 0.305, & \delta_{mn} = 2 \\ 0.108, & \delta_{mn} = 3 \\ 0.012, & \delta_{mn} = 4 \\ 0, & \delta_{mn} \geq 5 \end{cases}. \tag{3}$$

where δ_{mn} denotes the channel distance between user m and n, $\delta_{mn} = |a_m - a_n|$. Users within the interference range of user n is denoted as neighbour set. It can be expressed as:

$$\xi_n = \{m \in \mathcal{N} : d_{mn} < d_\tau\}, \forall n \in \mathcal{N}, \tag{4}$$

where the interference range is denoted by d_τ. When the physical distance between user m and n is larger than d_τ, interference can be ignored and user m is not user n's neighbour. Oppositely, when d_{mn} is smaller than d_τ, user m is the neighbour of user n and interference can not be ignored. As a result, the interference which user n suffers is equal to the sum of interference which comes from all its neighbours. Mathematically, it is expressed as:

$$I_n = \sum_{m \in \xi_n} p_m H(\delta_{mn})(d_{mn})^{-\alpha_1} \varepsilon_{mn}^{a_n}, \forall m, n \in \mathcal{N}, \tag{5}$$

where α_1 and $\varepsilon_{mn}^{a_n}$ denote the pass-loss exponent and instantaneous fading coefficient between users, respectively. Similarly, the jamming caused by malicious jammer can be denoted by J_n. d_{jn} and $\delta_{jn} = |c_j - a_n|$ denote as the physical distance and channel distance between user n and the jammer. c_j and a_n are the selected channels of the jammer and user n, respectively. The overlapping power mask between the smart jammer and user n can be expressed as [14]:

$$H(\delta_{jn}) = \begin{cases} 1, & \delta_{jn} = 0 \\ 0.605, & \delta_{jn} = 1 \\ 0.305, & \delta_{jn} = 2 \\ 0.108, & \delta_{jn} = 3 \\ 0.012, & \delta_{jn} = 4 \\ 0, & \delta_{jn} \geq 5 \end{cases}. \tag{6}$$

Then, the jamming is denoted as:

$$J_n = p_j H(\delta_{jn})(d_{jn})^{-\alpha_2}\varepsilon_{jn}^{c_j}, \forall n \in \mathcal{N}, \tag{7}$$

where d_{jn}, α_2 and $\varepsilon_{jn}^{c_j}$ are physical distance, pass-loss exponent and instantaneous fading coefficient, respectively. As a result, the user n's throughput should be expressed as:

$$R_n(a_n, \mathbf{a}_{-n}, c_j) = B\log(1 + \frac{p_n(d_{nn})^{-\alpha_1}\varepsilon_{nn}^{a_n}}{N_0 + I_n + J_n}), \forall n \in \mathcal{N}, \tag{8}$$

where B and N_0 are channel's bandwidth and background noise, respectively. I_n and J_n are denoted by the adjacent interference and external jamming. p_n is the transmission power of user n and d_{nn} is the physical distance between user n's transmitter and receiver. In a whole, the optimization goal is the maximization of network throughput. Mathematically, the network throughput is denoted as:

$$R = \sum_{n\in\mathcal{N}} R_n(a_n, \mathbf{a}_{-n}, c_j) = \sum_{n\in\mathcal{N}} B\log(1 + \frac{p_n(d_{nn})^{-\alpha_1}\varepsilon_{nn}^{a_n}}{N_0 + I_n + J_n}), \forall n \in \mathcal{N}. \tag{9}$$

Since it is hard to obtain the global information of the open and dense mesh network, we could not optimize this problem from the perspective of maximizing the throughput. Therefore, we view this problem from the perspective of minimizing generalized interference. The generalized jamming is denoted as D_n. It means the sum of adjacent interference and external jamming, that is:

$$D_n = I_n + J_n, \tag{10}$$

and the network generalized jamming is expressed as:

$$D = \sum_{n\in\mathcal{N}} D_n. \tag{11}$$

According to [12], the expected throughput of user n decreases monotonously while the generalized interference increases. Their relationship is displayed as followed:

$$R_n = f(D_n), \tag{12}$$

where the function $f(\cdot)$ is a monotone decreasing function equals to the minimization of the generalized interference of the network. The network throughput can be expressed as:

$$R = \sum_{n\in\mathcal{N}} R_n. \tag{13}$$

The optimization goal of the channel selection with anti-jamming problem is the maximization of the network which can be formulated as follows:

$$P1 : (a_1, a_2, ..., a_n, c_j) = \arg\max R. \tag{14}$$

3 Hierarchical Channel Selection Stackelberg Game

In this paper, we model the multi-users channel selection problem as a Stackelberg game. Firstly, the jammer learns the users' channel selection strategy and chooses the optimal strategy to degrade the network performance. After sensing the jammer's strategy, users adjust their strategies to minimize the generalized interference, so as to maximize the network throughput. Mathematically, the hierarchical channel selection Stackelberg game is denoted as $\mathcal{G} = \{\mathcal{N}, \mathcal{J}, \mathcal{M}, \mathcal{C}, u_n, u_j\}$. $\mathcal{N}$ and $\mathcal{J}$ denote user set and the smart jammer. $\mathcal{M}$ and $\mathcal{C}$ denote the channel sets of users and the smart jammer.We denote u_n and u_j as the utility function of user n and the jammer. Based on the game $\mathcal{G}$, there exists two sub-games, which is defined as leader sub-game $\mathcal{G}_l$ and follower sub-game $\mathcal{G}_f$. The leader sub-game can be expressed as:

$$\mathcal{G}_l = \{\mathcal{J}, \mathcal{C}, u_j(a_n, \mathbf{a}_{-n}, c_j)\}, \tag{15}$$

where $\mathcal{J}$ and $\mathcal{C}$ denote the smart jammer and available channel for the jammer. $u_j(a_n, a_{-n}, c_j)$ is the utility function of the smart jammer. From the perspective of jammer, it aims to select proper channel in order to maximize the damage to the network. Therefore, the jammer's utility function is:

$$u_j(a_n, \mathbf{a}_{-n}, c_j) = \sum_{n \in \mathcal{N}} J_n = \sum_{n \in \mathcal{N}} p_j H(\delta_{jn})(d_{jn})^{-\beta} \varepsilon_{jn}^{c_j}, \forall n \in \mathcal{N}. \tag{16}$$

The follower sub-game can be expressed as:

$$\mathcal{G}_f = \{\mathcal{N}, \mathcal{M}, u_n(a_n, \mathbf{a}_{-n}, c_j)\}, \tag{17}$$

where $\mathcal{N}$ and $\mathcal{M}$ denote the user set and available channel for these users. $u_n(a_n, a_{-n}, c_j)$ is user n's utility function. From the perspective of users, the optimization goal is to minimize the generalized interference. Mathematically, user n's utility function can be given as:

$$\begin{aligned} &u_n(a_n, \mathbf{a}_{-n}, c_j) = -D_n = -(I_n + J_n) \\ &= -(\sum_{m \in \mathcal{N}} p_m H(\delta_{mn})(d_{mn})^{-\alpha} \varepsilon_{mn}^{a_n} + p_j H(\delta_{jn})(d_{jn})^{-\beta} \varepsilon_{jn}^{c_j}), \forall m, n \in \mathcal{N}. \end{aligned} \tag{18}$$

Definition 1 (Exact Potential Game (EPG) [17]). *$\mathcal{G}_f$ is an exact potential game as long as there exists a potential function φ whose variation is equal to variation in the utility function caused by any player's unilateral deviation. Mathematically,*

$$\begin{aligned} &\varphi(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \varphi(a_n, \mathbf{a}_{-n}, c_j) = u_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - u_n(a_n, \mathbf{a}_{-n}, c_j), \\ &\forall n \in \mathcal{N}, \overline{a}_n, a_n \subseteq \mathcal{M}, \overline{a}_n \neq a_n, \end{aligned} \tag{19}$$

where $\overline{a}_n$ denotes the changed action of user n after unilateral deviation.

Theorem 1. *When c_j is determined, the users' sub-game is an EPG and there exists at least one pure strategy NE.*

Proof: Motivated by [19], when the strategy of jammer is given as c_j, the potential function of the follower sub-game $\mathcal{G}_f$ is:

$$\begin{aligned}&\varphi(a_n, \mathbf{a}_{-n}, c_j) = \varphi_1(a_n, \mathbf{a}_{-n}, c_j) + \varphi_2(a_n, \mathbf{a}_{-n}, c_j)\\&= -\tfrac{1}{2}\sum_{n\in\mathcal{N}} I_n(a_n, \mathbf{a}_{-n}, c_j) - \sum_{n\in\mathcal{N}} J_n(a_n, \mathbf{a}_{-n}, c_j)\end{aligned} \tag{20}$$

We denote the function $\varphi_1(a_n, \mathbf{a}_{-n}, c_j)$ and $\varphi_2(a_n, \mathbf{a}_{-n}, c_j)$ with the following rules:

$$\begin{aligned}\varphi_1(a_n, \mathbf{a}_{-n}, c_j) &= -\tfrac{1}{2}\sum_{n\in\mathcal{N}} I_n(a_n, \mathbf{a}_{-n}, c_j),\\ \varphi_2(a_n, \mathbf{a}_{-n}, c_j) &= -\sum_{n\in\mathcal{N}} J_n(a_n, \mathbf{a}_{-n}, c_j).\end{aligned} \tag{21}$$

In the POC situation, when the physical distance between two users is out of the interference range, they are not neighbour and there is no interference between them. For the function $\varphi_1(a_n, a_{-n}, c_j)$, as user n's action changes from a_n to $\overline{a}_n$, the interference of the non-neighbors does not change at all. Therefore, the variation of the function φ_1 can be expressed as:

$$\begin{aligned}&\varphi_1(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \varphi_1(a_n, \mathbf{a}_{-n}, c_j)\\&= -\tfrac{1}{2}\Big(\sum_{n\in\mathcal{N}} I_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \sum_{n\in\mathcal{N}} I_n(a_n, \mathbf{a}_{-n}, c_j)\\&= -\tfrac{1}{2}\Big(\sum_{m\in\zeta_n} I_m(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \sum_{m\in\zeta_n} I_m(a_n, \mathbf{a}_{-n}, c_j)\\&\quad + I_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - I_n(a_n, \mathbf{a}_{-n}, c_j)\Big)\end{aligned} \tag{22}$$

where ζ_n denotes the user n's neighbour set.

For user n's neighbor m, the variation of the interference before and after user n unilaterally changes its action is shown as:

$$I_m(\overline{a}_n, \mathbf{a}_{-n}, c_j) - I_m(a_n, \mathbf{a}_{-n}, c_j) = p_n(d_{mn})^{-\alpha}\varepsilon_{mn}^{a_n}[H(\overline{\delta}_{mn}) - H(\delta_{mn})]. \tag{23}$$

And for user n, the change of interference can be expressed as:

$$I_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - I_n(a_n, \mathbf{a}_{-n}, c_j) = \sum_{m\in\zeta_n} p_m(d_{mn})^{-\alpha}\varepsilon_{mn}^{a_n}[H(\overline{\delta}_{mn}) - H(\delta_{mn})]. \tag{24}$$

In this work, we assume the transmission power of each user is the same as each other. Therefore, combining Eqs. (23) and (24), Eq. (22) can be expressed as:

$$\begin{aligned}&\varphi_1(\overline{a}_n, a_{-n}, c_j) - \varphi_1(a_n, a_{-n}, c_j)\\&= -\tfrac{1}{2}\Big(\sum_{m\in\zeta_n} I_m(\overline{a}_n, a_{-n}, c_j) - \sum_{m\in\zeta_n} I_m(a_n, a_{-n}, c_j) + I_n(\overline{a}_n, a_{-n}, c_j) - I_n(a_n, a_{-n}, c_j)\Big)\\&= -\tfrac{1}{2}\Big(\sum_{m\in\zeta_n} p_n(d_{mn})^{-\alpha}\varepsilon_{mn}^{a_n}[H(\overline{\delta}_{mn}) - H(\delta_{mn})] + \sum_{m\in\zeta_n} p_m(d_{mn})^{-\alpha}\varepsilon_{mn}^{a_n}[H(\overline{\delta}_{mn}) - H(\delta_{mn})]\Big)\\&= -\sum_{m\in\zeta_n} p_m(d_{mn})^{-\alpha}\varepsilon_{mn}^{a_n}[H(\overline{\delta}_{mn}) - H(\delta_{mn})]\\&= -(I_n(\overline{a}_n, a_{-n}, c_j) - I_n(a_n, a_{-n}, c_j)).\end{aligned} \tag{25}$$

On the other hand, the strategy unilaterally changed by user n does not influence the jamming of user n's neighbor. Therefore, the variation of the function φ_2 is shown as:

$$\begin{aligned}
&\varphi_2(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \varphi_2(a_n, \mathbf{a}_{-n}, c_j) \\
&= -(\sum_{m \in \{\mathcal{N}/n\}} J_m(\overline{a}_n, \mathbf{a}_{-n}, c_j) + J_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) \\
&\quad - \sum_{i \in \{\mathcal{N}/n\}} J_i(a_n, \mathbf{a}_{-n}, c_j) - J_n(a_n, \mathbf{a}_{-n}, c_j)) \\
&= -(\sum_{m \in \{\mathcal{N}/n\}} J_m(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \sum_{i \in \{\mathcal{N}/n\}} J_i(a_n, \mathbf{a}_{-n}, c_j)) \\
&\quad -(J_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - J_n(a_n, \mathbf{a}_{-n}, c_j)) \\
&= -(J_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - J_n(a_n, \mathbf{a}_{-n}, c_j)).
\end{aligned} \tag{26}$$

As a result, the variation of the potential function φ is:

$$\begin{aligned}
&\varphi(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \varphi(a_n, \mathbf{a}_{-n}, c_j) \\
&= \varphi_1(\overline{a}_n, \mathbf{a}_{-n}, c_j) + \varphi_2(\overline{a}_n, \mathbf{a}_{-n}, c_j) - (\varphi_1(a_n, \mathbf{a}_{-n}, c_j) + \varphi_2(a_n, \mathbf{a}_{-n}, c_j)) \\
&= -(I_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - I_n(a_n, \mathbf{a}_{-n}, c_j)) + (-(J_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - J_n(a_n, \mathbf{a}_{-n}, c_j))) \\
&= -(I_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) + J_n(\overline{a}_n, \mathbf{a}_{-n}, c_j)) - (-I_n(a_n, \mathbf{a}_{-n}, c_j) + J_n(a_n, \mathbf{a}_{-n}, c_j)) \\
&= u_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - u_n(a_n, \mathbf{a}_{-n}, c_j).
\end{aligned} \tag{27}$$

According to the proof above, when the jamming strategy is determined and user n changes its actions unilaterally, the variation of φ is equal to the variation of the user n's utility function. Therefore, the follower sub-game is an EPG and there exacts at least a pure strategy NE. As user n's throughput is in monotone decreasing relationship with generalized interference, $u'_n = f(D_n)$ [19], we formulate the throughput maximization game as:

$$\mathcal{G}'_f : \max_{a_n \in \mathcal{M}} u'_n(a_n, \mathbf{a}_{-n}, c_j), \forall n \in \mathcal{N}. \tag{28}$$

Definition 2 (Ordinary Potential Game(OPG) [16]). *If there is a potential game which satisfies the following equation, this game is an OPG:*

$$\begin{aligned}
&\mathrm{sgn}[u'_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - u'_n(a_n, \mathbf{a}_{-n}, c_j)] \\
&= \mathrm{sgn}[\varphi'(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \varphi'(a_n, \mathbf{a}_{-n}, c_j)] \\
&\forall n \in \mathcal{N}, \overline{a}_n, a_n \subseteq \mathcal{M}, \overline{a}_n \neq a_n,
\end{aligned} \tag{29}$$

where $\mathrm{sgn}(\cdot)$ is expressed as:

$$\mathrm{sgn}(n) = \begin{cases} 1, & n > 0 \\ 0, & n = 0 \\ -1, & n < 0. \end{cases} \tag{30}$$

Theorem 2. *The throughput maximization game is an OPG and there exists at least a pure strategy NE.*

Proof. The following proof follows the lines given in [19]. The utility function of the throughput maximization game is denoted as $u'_n(a_n, \mathbf{a}_{-n}, c_j) = R_n$ and the potential function is denoted as:

$$\varphi'(a_n, \mathbf{a}_{-n}, c_j) = -f[\varphi(a_n, \mathbf{a}_{-n}, c_j)]. \tag{31}$$

Assuming user n changes its action from a_n to $\overline{a}_n$, the variation of the utility function is expressed as:

$$\begin{aligned} &\Delta u'_n(a_n, \mathbf{a}_{-n}, c_j) = u'_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - u'_n(a_n, \mathbf{a}_{-n}, c_j) \\ &= R_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - R_n(a_n, \mathbf{a}_{-n}, c_j) \\ &\approx f[D_n(\overline{a}_n, \mathbf{a}_{-n}, c_j)] - f[D_n(a_n, \mathbf{a}_{-n}, c_j)], \end{aligned} \tag{32}$$

and the changes of φ is:

$$\begin{aligned} &\Delta\varphi'(a_n, \mathbf{a}_{-n}, c_j) = \varphi'(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \varphi'(a_n, \mathbf{a}_{-n}, c_j) \\ &= f[\varphi(a_n, \mathbf{a}_{-n}, c_j)] - f[\varphi(\overline{a}_n, \mathbf{a}_{-n}, c_j)]. \end{aligned} \tag{33}$$

Observing the equation above, we find that the variation of user n's utility function equals to that of the potential function, that is:

$$D_n(\overline{a}_n, \mathbf{a}_{-n}, c_j) - D_n(a_n, \mathbf{a}_{-n}, c_j) = \varphi(\overline{a}_n, \mathbf{a}_{-n}, c_j) - \varphi(a_n, \mathbf{a}_{-n}, c_j). \tag{34}$$

Due to the fact that $f(\cdot)$ is a monotone decreasing function, we can prove that:

$$\mathrm{sgn}[\Delta u'_n] = \mathrm{sgn}[\Delta\varphi'_n]. \tag{35}$$

Note that OPG has the same properties with EPG.

Theorem 3. *The jammer's strategy and users' NE policy constitute a SE in the hierarchical channel selection game.*

Proof. When jammer's strategy is given, the Stackelberg game becomes a non-cooperative game. It has been proven the follower sub-game is an EPG and there exacts at least a pure strategy NE. The stationary strategy of the jammer can be expressed as:

$$\vartheta_0^* = \arg\max_{\vartheta_0} u_j(\vartheta_0, NE(\vartheta_0)), \tag{39}$$

every finite strategic game has a mixed strategy equilibrium. As a result, $(\vartheta_0^*, NE(\vartheta_0^*))$ constitutes a SE in the sense of stationary strategy.

4 Hierarchical Learning Algorithm for Channel Selection Problem

To overcome the anti-jamming channel selection problem, we propose a hierarchical learning framework. We propose a Spatial Adaptive Play (SAP) [22] algorithm to select the channels in the follower sub-game and a Q-learning

Algorithm 1. Hierarchical learning algorithm for anti-jamming channel selection problem

1: **Initialzation:** Set k=0, $t = 0$ and initialize the mixed strategy of both users and smart jammer as $q_{nm}(t) = 1/|\mathcal{M}|, q_{oh}(k) = 1/|\mathcal{C}|, \forall m \in \mathcal{M}, \forall h \in \mathcal{C}$.
2: **loop** $k = 0, 1, 2, ..., K_{\max}$ ($K_{\max}$ is the preestablished iteration number of the smart jammer). In the kth epoch, the smart jammer randomly selects a channel $c_j(k)$ and records this channel.
3: For the epoch k, users update their strategies based on SAP.
(1) **loop** $t = 0, 1, 2, ..., T_{\max}$ ($T_{\max}$ is the preestablished iteration number of users). In the time slot t, every user randomly selects one channel and calculates the reward value $u_n(t)$ based on equation (36).
(2) Among all users, randomly select one user n to adjust its action while keeping other users' strategies unchanged.
(3) For user n, exchange the information with neighbourhood and the surrounding environment. Traverse all available channels and calculate the corresponding utility functions.
(4) Updating the channel selection probabilities of all channels based on the Boltzmann distribution:

$$q_n(t+1) = \frac{\exp\{\omega u_n(a_n(t), c_j(k))\}}{\sum\limits_{a_n \in \mathcal{M}} \exp\{\omega u_n(a_n(t), c_j(k))\}}, \tag{36}$$

where ω is the temperature parameter which controls the tradeoff of exploration-exploitation.
(5) Based on the channel selection probability, the user n selects one of the channels.
(6) **End loop**
4: The jammer calculates $u_j(k)$ and updates its Q values based on the following rule:

$$Q_0^{k+1}(c_j(k+1)) = (1-\alpha)Q_0^k(c_j(k)) + \alpha u_j(k), \tag{37}$$

where $\alpha = \lambda^k \in [0, 1)$ is the learning rate, and the jammer updates its strategy based on the same rule as users.

$$q_0(k) = \frac{\exp\{Q_0^k(c_j(k))/\tau_0\}}{\sum\limits_{c_j \in \mathcal{C}} \exp\{Q_0^k(c_j(k))/\tau_0\}}, \tag{38}$$

where τ_0 is the temperature coefficient with the same function as ω.
5: If epoch number reaches the stopping criterion, cycle stops running, otherwise updates $k = k + 1$ and switches to step 2.
6: **End loop**

algorithm for the jammer to update the jamming strategy in the leader subgame. Both two algorithms update strategies based on the probabilities of the channels selected by players. We define the mixed strategy of user n as $q_n(t) = (q_{n1}(t), ..., q_{nm}(t), ..., q_{nM}(t))$, where $q_{nm}(t)$ is the probability of channel m selected by user n and $\sum\limits_{m \in \mathcal{M}} q_{nm}(t) = 1$. Similarly, the mixed strategy of the smart jammer is denoted as $q_0(k) = (q_{01}(k), ..., q_{0h}(k), ..., q_{0H}(k))$, where $q_{oh}(k)$

means the probability of channel h selected by the smart jammer at epoch k and $\sum_{h \in \mathcal{C}} q_{oh}(k) = 1$. They update their strategies based on different time scales.

As shown in Algorithm 1, a SAP algorithm is proposed for the legitimate users to overcome the channel selection problem. In each learning process, only one user can be selected to update actions while other users keep their actions unchanged. This process is repeated running once the probability achieve the convergence. The smart jammer achieves its optimal action based on Q-learning [23]. Q-learning is widely used in dynamic and unknown environment with great performance. In this algorithm, the jammer repeatedly interacts with the environment with error trial in order to achieve the higher reward value. In the process of reinforcing the reward value, the jammer updates its Q function and makes a new action.

In the hierarchical learning framework, each epoch consists of T time slots and when the smart jammer updates one time, the users update their strategies T times. For both jammer and users, the iteration stops when the channel section probabilities vector achieves the convergence or the iteration numbers reach the preestablished number.

5 Simulation Results and Discussion

In this section, we consider a $120\,\text{m} \times 120\,\text{m}$ scale network. The users exist in a region of $100\,\text{m} \times 100\,\text{m}$ and the smart jammer is located in the region of [100,120 m]. Parameters in this simulation are set as follows: the users' transmitting power $p_n = 2\,\text{W}$ and the jammer's transmitting power $p_j = 15\,\text{W}$, the path loss exponent $\alpha = \beta = 3$, the background noise $N_0 = -130\,\text{dBm}$, the instantaneous fading coefficient $\varepsilon_{jn}^{c_j} = \varepsilon_{mn}^{a_m} = \varepsilon_{nn}^{a_n} = 3$, the physical distance between user n's transmitter and receiver $d_{nn} = 30\,\text{m}$.

Figure 3 shows the influence of users' number on the network throughput. Assuming there are 11 available channels and the power of jammer is set as 15 W. When the number of users is 25, the network throughput is 110.4 Mbps, and when the number of user is 45, the network throughput decreases to 71.2 Mbps. We can find that the network throughput decreases with the number of users increasing. In the dense network with large number of users, influence between users would degrade the network throughput.

Figure 4 shows the effect of jammer's power on the network performance. We set the number of available channels and users as 11 and 35. Assuming jammer's power as 15 W, 25 W and 35 W. According to the figure, we can find that the network throughput would decrease with the jammer's power enhancing. The simulation result shows the network throughput can maintain a stable state when the jamming intensity increases, which coincides with the practical situations.

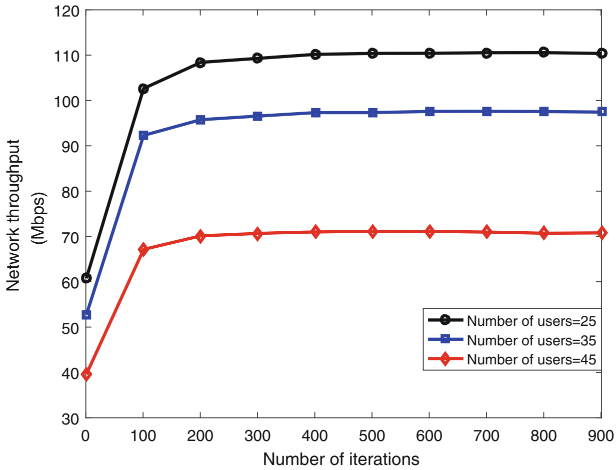

Fig. 3. The influence of users' number on the network throughput.

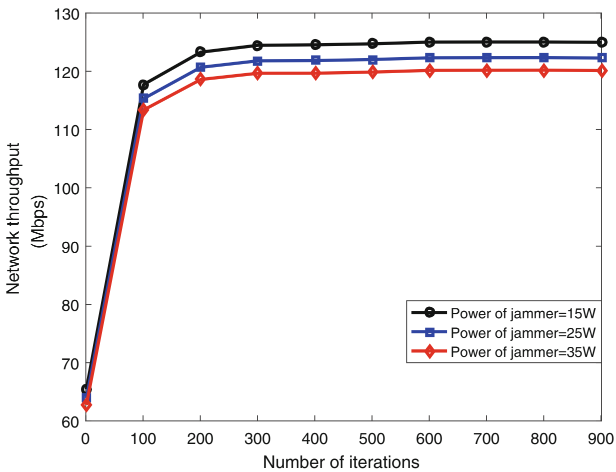

Fig. 4. The influence of jammer's power on the network throughput.

6 Conclusion

In this paper, we had investigated the intelligent anti-jamming problems with limited network bandwidth. Different from most papers focusing on orthogonal channels, this paper studied the channel selection problems in POC situation, where channels overlapped partially. The adversarial relationship between users and the samrt jammer was formulated as a hierarchical channel selection Stackelberg game. A hierarchical learning algorithm was proposed to reach the SE. Finally, simulation results showed that the network throughput can maintain a stable state when the jamming intensity increases, and it coincides with the practical situations.

References

1. Yang, D., Zhang, J., Fang, X., et al.: Optimal transmission power control in the presence of a smart jammer. In: Global Communications Conference, pp. 5506–5511 (2015)
2. Yang, D., Xue, G., Zhang, J., et al.: Coping with a smart jammer in wireless networks: a Stackelberg game approach. IEEE Trans. Wirel. Commun. **12**(8), 4038–4047 (2013)
3. Xiao, L., Chen, T., Liu, J., et al.: Anti-jamming transmission Stackelberg game with observation errors. IEEE Commun. Lett. **19**(6), 949–952 (2015)
4. Lv, S., Xiao, L., Hu, Q., Wang, X., Hu, C., Sun, L.: Anti-jamming power control game in unmanned aerial vehicle networks. In: IEEE GLOBECOM, pp. 1–6 (2017)
5. Yao, F., Jia, L., et al.: A hierarchical learning approach to anti-jamming channel selection strategies. Wirel. Netw. **25**, 201–213 (2017). https://doi.org/10.1007/s11276-017-1551-9
6. Wu, Y., Wang, B., Liu, K., et al.: Anti-jamming games in multi-channel cognitive radio networks. IEEE J. Sel. Areas Commun. **30**, 4–15 (2012)
7. Jia, L., et al.: A multi-domain anti-jamming defence scheme in heterogeneous wireless networks. IEEE Trans. Veh. Technol. (2018)
8. Hanawal, M.K., Abdel-Rahman, M.J., Krunz, M.: Joint adaptation of frequency hopping and transmission rate for anti-jamming wireless system. IEEE Trans. Mob. Comput. **15**(9), 2247–2259 (2016)
9. Hanawal, M.K., Abdel-Rahman, M.J., Krunz, M.: Game theoretic anti-jamming dynamic frequency hopping and rate adaption in wireless system. In: IEEE Wiopt, pp. 247–254 (2014)
10. Mishra, A., Rozner, E., Banerjee, S., et al.: Exploiting partially overlapping channels in wireless networks: turning a peril into an advantage. In: Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement, Berkeley, United States, pp. 311–316 (2005)
11. Mishra, A., Rozner, E., Banerjee, S., et al.: Exploiting partially overlapping channels not considered harmful. SIGMetrics Performance, Saint Malo, France (2006)
12. Yao, K., et al.: Oriented partially overlapping channel access in wireless networks: a game-theoretic learning approach. Wirel. Netw. (to appear)
13. Ding, Y., Huang, Y., Zeng, G., et al.: Using partially overlapping channels to improve throughput in wireless mesh networks. IEEE Trans. Mob. Comput. **11**(11), 1720–1733 (2012)
14. Tandjaoui, A., Kaddour, M., et al.: Refining the impact of partially overlapping channels in wireless mesh networks through a cross-layer optimization model. In: IEEE STWiMob, pp. 1–8 (2016)
15. Xu, Y., Wu, Q., Wang, J., et al.: Opportunistic spectrum access using partially overlapping channels: graphical game and uncoupled learning. IEEE Trans. Commun. **61**(9), 3906–3918 (2013)
16. Myerson, R.: Game Theory: Analysis of Conflict. Harvard University Press, Cambridge (1991)
17. Monderer, D., Shapley, L.S.: Potential games. Games Econ. Behav. **14**, 124–143 (1996)

18. Han, Z., Niyato, D., Saad, W., Basar, T., et al.: Game Theory in Wireless and Communication Networks. Cambridge University Press, Cambridge (2012)
19. Yao, K., Wu, Q., Xu, Y., Jing, J.: Distributed ABS-slot access in dense heterogeneous networks: a potential game approach with generalized interference model. IEEE Access **5**, 94–104 (2017)
20. Rad, A.H.M., Wong, V.W.S.: Joint optimal channel assignment and congestion control for multi-channel wireless mesh network. In: IEEE International Conference on Communications, Istanbul, Turkey, pp. 1984–1989 (2006)
21. Mohsenian, R.A.D.A.H., Wong, V.W.S.: Congestion-aware channel assignment for multi-cahnnel wireless mesh networks. Comput. Netw. **53**(14), 2502–2516 (2009)
22. Young, H.P.: Individual Strategy and Social Structure: An Evolutionary Theory of Institutions. Princeton University Press, Princeton (2001)
23. Watkins, C.J.C.H., Danyan, P.: Q-learning. Mach. Learn. **8**, 279–292 (1992)