A Game-Theoretic Learning Approach for Anti-Jamming Dynamic Spectrum Access in Dense Wireless Networks

Luliang Jia, Yuhua Xu, *Member, IEEE*, Youming Sun, *Student Member, IEEE*, Shuo Feng, *Student Member, IEEE*, Long Yu, and Alagan Anpalagan, *Senior Member, IEEE*

***Abstract*—In this paper, we investigate the anti-jamming channel selection problem for interference mitigation (IM) based dense wireless networks in dynamic environment, in which the active user set is variable due to their specific traffic demands. We jointly consider the mutual interference among users and external jamming in IM-based dense wireless networks, and propose a generalized maximum protocol interference and jamming model to accurately capture the mutual interference and external jamming. Then, the anti-jamming channel selection problem is formulated as an anti-jamming dynamic game, and subsequently it is proved to be an exact potential game, which has at least one pure strategy Nash equilibrium (NE). Based on the stochastic learning theory, a distributed anti-jamming channel selection algorithm (DACSA) is proposed to find the NE solution. Moreover, the simulation results are presented to demonstrate the effectiveness of the proposed DACSA algorithm.**

***Index Terms*—Distributed channel selection, exact potential game, stochastic learning, interference mitigation, anti-jamming, hypergraph.**

Manuscript received January 11, 2018; revised September 23, 2018; accepted November 30, 2018. Date of publication December 24, 2018; date of current version February 12, 2019. This work was supported in part by the Natural Science Foundation for Distinguished Young Scholars of Jiangsu Province under Grant BK20160034, in part by the National Science Foundation of China under Grants 61671473, 61631020, 61401508, and 61401505, in part by Jiangsu Provincial Natural Science Foundation of China under Grants BK20130069 and BK20151450, and in part by the Open Research Foundation of Science and Technology in the Communication Networks Laboratory. This paper was presented in part at the 17th IEEE International Conference on Communication Technology, Chengdu, China, October 2017. The review of this paper was coordinated by Prof. L. Wang. *(Corresponding author: Yuhua Xu.)*

L. Jia is with the School of Space Information, Space Engineering University, Beijing 101416, China, with the College of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China, and also with the Science and Technology on Communication Networks Laboratory, Shijiazhuang 050002, China (e-mail: jiallts@163.com).

Y. Xu, Y. Sun, and S. Feng are with the College of Communications Engineering, Army Engineering University of PLA, Nanjing 210007, China, and also with the Science and Technology on Communication Networks Laboratory, Shijiazhuang 050002, China (e-mail: yuhuaenator@gmail.com; sunyouming10@163.com; fengs13@mcmaster.ca).

L. Yu is with the 63rd Research Institute, National University of Defense Technology, Nanjing 210007, China (e-mail: elong025@163.com).

A. Anpalagan is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: alagan@ee.ryerson.ca).

Digital Object Identifier 10.1109/TVT.2018.2889336

I. Introduction

THE wireless transmission is extremely vulnerable to various security attacks or threats due to the inherent broadcast nature. The wireless communication security issue has received increasing research interest recently to fight against various attacks, such as jamming attack, eavesdropping attack, primary user emulation attack, data falsification attack, and so on [1]–[5]. In this paper, we focus on jamming attack, which can severely degrade network performance. Moreover, mutual interference among users is an important factor that restricts the network performance, and it also leads to a significant threat to the security of the spectrum availability [6], especially in the dense wireless networks with plenty of wireless devices. In our work, we focus on the mutual co-channel interference, and assume that the leakage among different channels is negligible.

Interference mitigation (IM) is an important problem, and there exist some related studies on IM. In [7], the co-channel interference among secondary users for spectrum sharing cognitive radio networks was considered, and a non-cooperative game method was adopted. In [8]–[10], the interference mitigation game was formulated for non-spatial scenario, where the users were located closely and a user would suffer interference from all other users. Considering the spatial aspect, the authors in [11] investigated the graphical game, in which mutual interference only occurs between neighboring users. However, only the strong interference relationship was considered in these graph models, and the cumulative weak interference effect was ignored [12]. In dense deployment scenarios, the traditional graph model is insufficient to characterize the mutual interference relationship. To be specific, several weak interference users may constitute a strong interference, if the cumulative interference exceeds a threshold. Therefore, it is essential to consider the cumulative interference effect in dense wireless networks. Hypergraph model [13] is a promising technology, and it can accurately model the interference relationship. It can simultaneously capture the strong interference relationship and cumulative weak interference effect, and had been employed for resource allocation in wireless communications [12], [14]–[16].

To mitigate the mutual interference among users, a hypergraph interference model was proposed, and a channel access game was formulated in [12]. Nevertheless, there exist several unsolved problems. Firstly, the work in [12] ignored the dynamics of users, and it was assumed that all users participated in communication all the time. However, the users dynamically compete for channels in practice due to their specific traffic demands. Moreover, little attention was paid to the external jammers in [12]. However, the jamming attack is a significant threat to wireless networks, and degrades the system performance in the IM-based networks. To cope with the external jammers, and several anti-jamming technologies were investigated in [17]–[25]. In [17]–[19], the anti-jamming Stackelberg game was investigated, and the optimal power strategy was derived. In [20], the anti-jamming power control game was studied, and the observation error was analyzed. In [21], the cooperative transmission game was formulated to fight against jamming attack, and the Stackelberg equilibrium was obtained. In [22], the prospect theory was employed to analyze the end-user behavior, and the interaction between the secondary user and jammer was formulated as a jamming game. The authors in [23] studied the anti-jamming defense in cognitive radio network, and an anti-jamming game was formulated. A stochastic game was formulated to analyze the interactions between the secondary users and attackers in [24]. In [25], M. K. Hanawal *et al.* formulated a Markov game in the presence of jammer, and achieved the optimal defense strategy.

Hence, it is important and timely to study the anti-jamming problem for the IM-based dense wireless networks in dynamic environment. Motivated by [12], a generalized maximum protocol interference and jamming model is proposed. Specifically, hypergraph theory is employed to model the cumulative effect of the weak interference. To accurately characterize the damage degree of the external jammers, a generalized jamming model is proposed. Then, this problem is formulated as an anti-jamming dynamic game. Furthermore, we incorporate the stochastic learning automata [26] into the formulated game to obtain the Nash equilibrium (NE) solution.

To the best of authors' knowledge, this is the first work to study the anti-jamming channel selection problem in the IM-based dense wireless networks, jointly considering the mutual interference and external jamming. The main contributions of this paper are as follows.

- The generalized maximum protocol interference and jamming (GMPIJ) model is proposed to accurately characterize the mutual interference among users and external jamming. The hypergraph interference model is adopted to capture the cumulative weak interference effect in the proposed GMPIJ model. Moreover, the proposed GMPIJ model can precisely depict the external jamming.
- We investigate the anti-jamming channel selection problem in IM-based dense wireless networks. Considering the dynamics of users, the set of active users is variable due to dynamic traffic demands. The varying number of active users brings a great technical challenge, and it will lead to the varying number of players of the game model. Subsequently, an anti-jamming dynamic game, which has variable number of players due to dynamic traffic demands, is formulated, and it is proved to be an exact potential game, which has at least one pure strategy NE.

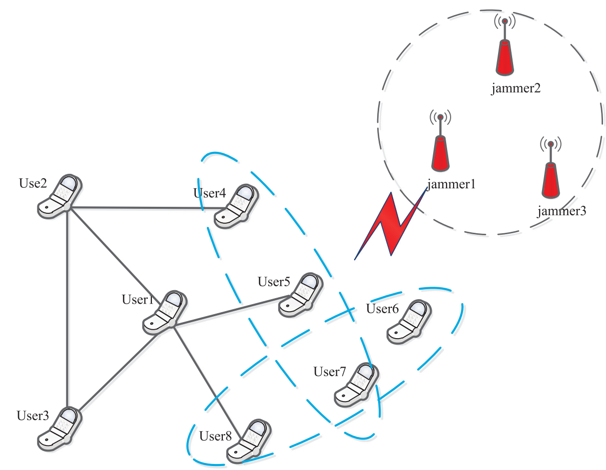


Fig. 1. System model.

- Based on the stochastic learning theory, a distributed anti-jamming channel selection algorithm (DACSA) with varying number of active users is proposed, which can asymptotically converge to the NE of the formulated anti-jamming dynamic game.

Note that part of this work was presented in [27], and some preliminary results were reported therein. The differences and new contributions in this work are: (i) a generalized jamming model with multiple jamming ranges is formulated to accurately capture the external jamming, based on which a more practical GMPIJ model is proposed, (ii) rigorous theoretical analysis is provided, and more simulation results are presented to provide a more comprehensive evaluation.

The rest of the paper is organized as follows. The system model and problem formulation are given in Section II. In Section III, we formulate the anti-jamming dynamic game and analyze its properties. In Section IV, we propose a distributed channel selection algorithm to obtain the NE of the formulated anti-jamming dynamic game. In Section V, we conduct the simulation results. In Section VI, a discussion is presented, and a conclusion is given in Section VII.

II. System Model and Problem Formulation

A. System Model

As shown in Fig. 1, there are N users and K jammers in our model. Denote the user set as $\mathcal{N} = \{1, 2, \ldots, N\}$ and the jammer set as $\mathcal{K} = \{1, 2, \ldots, K\}$. The number of channels is M $(N > M)$, and the set of available channels for users is denoted as $\mathcal{M} = \{1, 2, \ldots, M\}$. We assume that each user autonomously competes for one of M channels. The jammers can sense the users' available channels, and each jammer chooses one channel to launch the jamming attack. For the external jammers, the set of jamming channels is denoted as $\mathcal{C} = \{1, 2, \ldots, H\}$, where H represents the number of jamming channels. For the sake of analysis, it is assumed that the set $\mathcal{M}$ of available channels for users is the same as the set $\mathcal{C}$ of jamming channels for jammers. For each user n, its received generalized

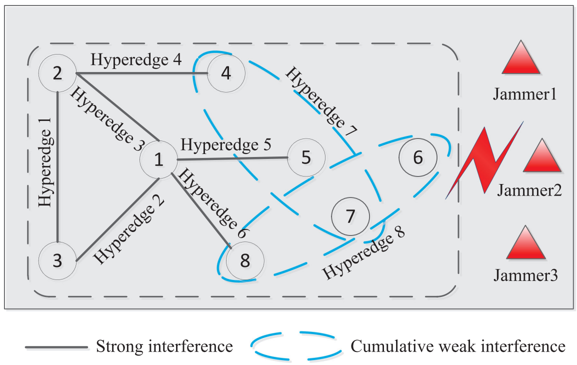


Fig. 2. Hypergraph interference model of users.

interference and jamming is expressed as:

$$D_n(a_n, a_{-n}) = I_n(a_n, a_{-n}) + \alpha J_n(a_n, \mathbf{c}), \tag{1}$$

where a_n is user n's selected channel, a_{-n} denotes the channel selection profile of all users except user n, $\mathbf{c} = [c_1, c_2, \ldots, c_K]$ denotes the strategy profile of jammers, K is the number of jammers, $c_k (k \in \{1, 2, \ldots, K\})$ is the jamming channel of jammer k, $I_n(a_n, a_{-n})$ represents the received mutual interference among users, $J_n(a_n, \mathbf{c})$ denotes the external jamming, and α is the jamming factor.

1) Precise mutual interference model

Definition 1 ([13], [14]): Denote $V = \{v_1, v_2, \ldots, v_N\}$ as a finite vertex set, a hypergraph $\Gamma = (V, E)$ is a family $E = \{e_1, e_2, \ldots, e_{|E|}\}$ of subsets of V, and it holds

$$e_i \neq \phi \, (i = 1, 2, \ldots, |E|),$$
$$\bigcup_{i=1}^{|E|} e_i = V, \tag{2}$$

where $e_i = \{v_{i_1}, v_{i_2}, \ldots, v_{i_j}\}$ denotes the hyperedges of hypergraph Γ, and it contains any subset of V. The hypergraph is the generalization of the traditional graph model, and it can effectively characterize the interference relation. In a hypergraph model, the cumulative weak interference can be captured by hyperedges, and the largest hyperedge contains Q vertices. Similar to [12], [14], we have $Q = 3$. As shown in Fig. 2, the strong interference relationship is denoted by solid lines. If the cumulative weak interference caused by multiple users reaches a threshold, it can lead to an equivalent interferer and construct a cumulative weak interference relationship, which is represented by blue circles in dashed line. To be specific, one of user 4 and user 5 cannot interfere with user 7, while the simultaneous transmission of user 4 and user 5 will interfere with user 7 if they adopt the same channel. The strong interference relationship and cumulative weak interference relationship constitute the hyperedges together. That is to say, both the solid lines and the circles form the hyperedges together in Fig. 2.

Considering the specific traffic demands of different users, it is assumed that each user is active in each time slot with probability θ_n, $0 < \theta_n \leq 1$. To characterize the dynamic

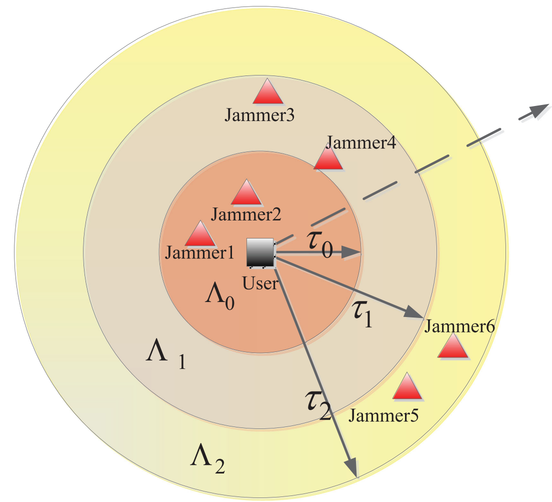


Fig. 3. The illustration of the generalized jamming model.

environment, a random vector $\mathbf{S}(v(t)) = (s_1, s_2, \ldots, s_N)$ is introduced, and $s_n \in \{0, 1\}$ (0 for inactive, and 1 for active) denotes the state of user n. Denote the active user set at time t as $\mathcal{B}(t)$, i.e., $\mathcal{B}(t) = \{n \in \mathcal{N} : s_n = 1\}$. Similar to [12], for a realization $v(t)$ at time t, the potentially maximum protocol interference of user n is defined as:

$$\hat{I}_n(a_n, a_{-n}, v(t)) = \delta(n \in \mathcal{B}(t)) \sum_{n \in N_n} \delta(e_n^w), \tag{3}$$

with

$$\delta(e_n^w) = \begin{cases} 1; & a_j = a_n, \forall j \in e_n^w \\ 0; & \text{otherwise}, \end{cases} \tag{4}$$

where N_n denotes the user n's neighboring set that contains strong interfering neighboring users and potentially weak interfering neighboring users, $\delta(n \in \mathcal{B}(t))$ specifies the state of user n, $|\cdot|$ represents the set's cardinality, $e_n^w \in E(n) = \{e_n^1, e_n^2, \ldots, e_n^{|E(n)|}\}$, where $E(n)$ is the hyperedge set that contains the vertex n.

Remark 1: In existing work [12], the dynamics of users is ignored. However, the scenario with varying user set due to specific traffic demands is practically appealing. In our work, we investigate the channel selection problem in dynamic environment, and the number of active users is variable.

2) Generalized jamming model

The damage degree of external jammers can be divided into multiple jamming levels according to the deterioration to the Signal-to-Interference-plus-Noise Ratio (SINR) value. To accurately characterize the external jamming, the jamming radius $\tau_i (i = 0, 1, 2, \cdots)$ is introduced, which measures the damage degree. The smaller the jamming radius τ_i is, the more serious damage degree it is. An illustration of the generalized jamming model is shown in Fig. 3, and the first jamming range can be defined as Λ_0, i.e., $\Lambda_0 = \{SINR \leq \tau_0\}$, and the second jamming range is expressed as Λ_1, i.e., $\Lambda_1 = \{\tau_0 < SINR \leq \tau_1\}$. By analogy, the g-th jamming range is denoted as Λ_{g-1},

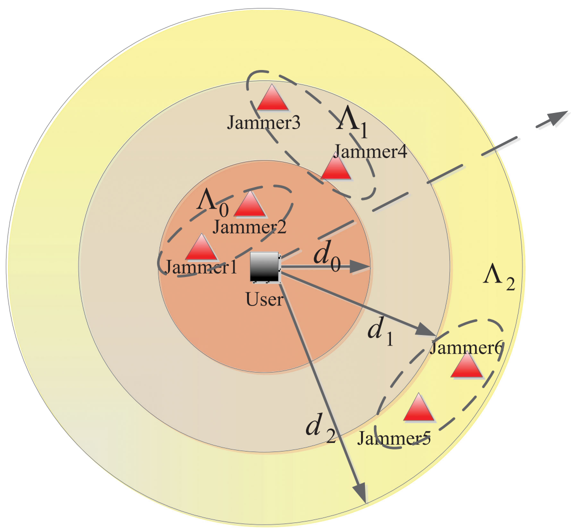


Fig. 4. Generalized jamming model.

i.e., $\Lambda_{g-1} = \{\tau_{g-1} < SINR \leq \tau_g\}$. In traditional jamming models, only the jammer located in Λ_0 is considered, whereas other jammers are ignored. However, the jammers outside the range Λ_0, especially those near the jamming radius τ_0, can still bring non-negligible jamming.

In our model, it is assumed that the damage degree of jammers is decided by the jamming distance. The jamming model is shown in Fig. 4. Motivated by [28], to accurately characterize the received jamming of user *n*, the generalized jamming can be defined as:

$$\begin{aligned}\hat{J}_n\left(a_n, \mathbf{c}, v(t)\right) &= \delta\left(n \in \mathcal{B}(t)\right)\Bigg[\sum_{k\in\Lambda_0} \lambda_0(d_{nk})\cdot\delta\left(a_n, c_k\right)\\ &\quad + \sum_{k\in\Lambda_1}\lambda_1(d_{nk})\cdot\delta\left(a_n, c_k\right) + \cdots\\ &\quad + \sum_{k\in\Lambda_g}\lambda_g(d_{nk})\cdot\delta\left(a_n, c_k\right)\Bigg], \end{aligned} \tag{5}$$

where d_{nk} represents the distance between user *n* and jammer *k*, $\lambda_0, \lambda_1, \cdots \lambda_g$ denotes the damage degree factor, and $\delta(\cdot)$ represents the Kronecker delta function given by:

$$\delta\left(a_n, c_k\right) = \begin{cases} 1, & a_n = c_k,\\ 0, & a_n \neq c_k. \end{cases} \tag{6}$$

The jammers located in the first jamming range are severe, and $\lambda_0(d_{nk}) = 1$. Since longer distance leads to milder jamming, and the jamming factor can be defined as $\lambda_g(d_{nk}) = (d_{nk}/d_{g-1})^{-g\beta}$, where β is the path loss exponent, d_{g-1} denotes the jamming range radius. To be specific, when a jammer is located in the second jamming range, its damage degree is almost as strong as the jammers located in the first jamming range, if $(d_{nk}/d_0)^{-\beta} \rightarrow 1$. For simplicity, we consider the jamming model with two levels in this paper, and the received jamming can be given by:

$$\begin{aligned}\hat{J}_n\left(a_n, \mathbf{c}, v(t)\right) &= \delta\left(n \in \mathcal{B}(t)\right)\Bigg[\sum_{k\in\Lambda_0}\delta\left(a_n, c_k\right)\\ &\quad + \sum_{k\in\Lambda_1}\left(\frac{d_{nk}}{d_0}\right)^{-\beta}\cdot\delta\left(a_n, c_k\right)\Bigg]. \end{aligned} \tag{7}$$

According to the above analysis, the generalized maximum protocol interference and jamming (GMPIJ) of user *n* can be expressed as:

$$\begin{aligned}\hat{D}_n\left(a_n, a_{-n}, \mathbf{c}, v(t)\right) &= \hat{I}_n + \alpha\hat{J}_n,\\ &= \delta\left(n \in \mathcal{B}(t)\right)\Bigg[\sum_{e_n^w\in N_n}\delta\left(e_n^w\right) + \alpha\sum_{k\in\Lambda_0}\delta\left(a_n, c_k\right)\\ &\quad + \alpha\sum_{k\in\Lambda_1}\lambda_1(d_{nk})\cdot\delta\left(a_n, c_k\right)\Bigg]. \end{aligned} \tag{8}$$

Remark 2: The existing work [12] only captures the mutual interference among users, whereas the impact of external jammers is ignored. The external jamming is a potential threat, which degrades the system performance. In our work, we jointly consider the mutual interference and external jamming in IM-based dense wireless networks.

For a realization $v(t)$ at time *t*, similar to [12], [29], we can define the normalized capacity of user *n* as:

$$\hat{C}_n\left(a_n, a_{-n}, \mathbf{c}, v(t)\right) = \begin{cases} 1, & \hat{D}_n\left(a_n, a_{-n}, \mathbf{c}, v(t)\right) \leq T_0,\\ 0, & \text{else}, \end{cases} \tag{9}$$

where T_0 denotes the threshold that can decode the packets correctly, and $\hat{C}_n\left(a_n, a_{-n}, v(t)\right) = 0$ if user*n* cannot decode the packets correctly. Then, the normalized network capacity can be given by:

$$\hat{C}(a_n, a_{-n}, \mathbf{c}, v(t)) = \sum_{n\in\mathcal{N}}\hat{C}_n\left(a_n, a_{-n}, \mathbf{c}, v(t)\right). \tag{10}$$

At each time slot, $v(t)$ is random, and the expected normalized capacity can be expressed as:

$$\begin{aligned}C(a_n, a_{-n}, \mathbf{c}) &= \mathrm{E}\left[\hat{C}(a_n, a_{-n}, \mathbf{c}, v(t))\right]\\ &= \lim_{T\rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T}\left\{\hat{C}(a_n, a_{-n}, \mathbf{c}, v(t))\right\}, \end{aligned} \tag{11}$$

where $E[\cdot]$ represents the statistical expectation.

B. Problem Formulation

Inspired by the definition of the network potentially maximum protocol interference (PMPI) in [12], which can characterize the interference level of the network. In our model, the network

GMPIJ can be given by:

$$\hat{D}(a_n, a_{-n}, \mathbf{c}, v(t)) = \sum_{n\in\mathcal{B}(t)} \sum_{e_n^w \in N_n} \delta\left(e_n^w\right) + \sum_{n\in\mathcal{B}(t)} \left(\alpha \sum_{k\in\Lambda_0} \delta\left(a_n, c_k\right) + \alpha \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta\left(a_n, c_k\right)\right). \quad (12)$$

The expected network GMPIJ can be defined as:

$$D(a_n, a_{-n}, \mathbf{c}) = \mathrm{E}\left[\hat{D}(a_n, a_{-n}, v(t))\right] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left\{\hat{D}(a_n, a_{-n}, \mathbf{c}, v(t))\right\}. \quad (13)$$

Motivated by the previous work, we aim to minimize the network GMPIJ, and the optimization problem can be expressed as:

$$P1 : \mathbf{a}_{opt} \in \arg\min D(a_n, a_{-n}, \mathbf{c}). \quad (14)$$

Remark 3: The above problem is a combinatorial optimization problem, and traditional optimization technologies cannot be directly applied. Besides, considering the dynamics of users, it is particularly challenging. Therefore, it is of great value to design a learning-based distributed scheme to obtain the desirable solutions.

III. Anti-Jamming Channel Selection Game

A. Game Model

Here, the GMPIJ minimization problem in dynamic environment can be formulated as an anti-jamming dynamic game. Similar to [11], [12], the user *n*'s utility at a realization $v(t)$ can be expressed as:

$$\hat{u}_n(a_n, a_{-n}, \mathbf{c}, v(t)) = L - \delta\left(n \in \mathcal{B}(t)\right) \left[I_n(a_n, a_{N_n}) + \sum_{i\in N_n} I_i(a_i, a_{N_i})\right] - \delta\left(n \in \mathcal{B}(t)\right) \alpha \left[\sum_{k\in\Lambda_0} \delta\left(a_n, c_k\right) + \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta\left(a_n, c_k\right)\right], \quad (15)$$

where *L* is a positive constant to ensure that the proposed algorithm is effective, N_n denotes user *n*'s neighboring set. The second term denotes the interference from user *n* and its neighbors, and the third term represents the jamming from the external jammers. For a realization $v(t)$ at time *t*, the user *n*'s utility is random, hence, the expected utility of user *n* is given by:

$$u_n(a_n, a_{-n}, \mathbf{c}) = E\left[\hat{u}_n(a_n, a_{-n}, \mathbf{c}, v(t))\right] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \hat{u}_n(a_n, a_{-n}, \mathbf{c}, v(t)), \quad (16)$$

Formally, the formulated anti-jamming dynamic game can be denoted as $\mathcal{G} = \{\mathbf{S}, \mathcal{N}, \mathcal{K}, \mathcal{M}, \mathcal{C}, \{u_n\}_{n\in\mathcal{N}}\}$, where $\mathbf{S} = (s_1, s_2, \ldots, s_N)$ denotes the state of users, and $s_n \in \{0, 1\}$ (0 for inactive, and 1 for active), $\mathcal{N} = \{1, 2, \ldots, N\}$ and $\mathcal{K} = \{1, 2, \ldots, K\}$ respectively represent the set of users and jammers, $\mathcal{M} = \{1, 2, \ldots, M\}$ and $\mathcal{C} = \{c_1, c_2, \ldots, c_H\}$ respectively represent the set of available channels of users and jamming channels of jammers, and $\{u_n\}_{n\in\mathcal{N}}$ denotes the user *n*'s utility. The proposed anti-jamming dynamic game can be expressed as:

$$(\mathcal{G}) : \quad \max_{a_n \in\mathcal{M}} u_n(a_n, a_{-n}, \mathbf{c}) \quad \forall n \in \mathcal{N}. \quad (17)$$

In the formulated anti-jamming dynamic game $\mathcal{G}$, each user acts as a game player. In our work, we focus on the strategies of users, and each user aims to maximize its utility and selects its optimal channel for transmission.

B. Analysis of NE

Definition 2 (NE [30]): If and only if no player can derive more profit by deviating unilaterally, a strategy profile $a^* = (a_1^*, a_2^*, \ldots, a_N^*)$ is a pure strategy NE. Mathematically,

$$u_n(a_n^*, a_{-n}^*, \mathbf{c}) \geq u_n(a_n, a_{-n}^*, \mathbf{c}) \quad \forall n \in \mathcal{N}, a_n \in \mathcal{M}. \quad (18)$$

Definition 3 (Exact potential game [31], [32]): $\mathcal{G}$ is an exact potential game, if there exists a potential function $\Phi(a_n, a_{-n}, \mathbf{c})$ that satisfies the following condition.

$$u_n(a_n, a_{-n}, \mathbf{c}) - u_n(\bar{a}_n, a_{-n}, \mathbf{c}) = \Phi(a_n, a_{-n}, \mathbf{c}) - \Phi(\bar{a}_n, a_{-n}, \mathbf{c}), \quad (19)$$

where $\bar{a}_n$ is the player *n*'s action after deviation.

Theorem 1: The proposed anti-jamming dynamic game $\mathcal{G}$ is an exact potential game, which has at least one pure strategy NE.

Proof: Inspired by the potential function proposed in [10]–[12], we construct the potential function as:

$$\begin{aligned}\Phi(a_n, a_{-n}, \mathbf{c}) &= -D(a_n, a_{-n}, \mathbf{c}) \\ &= \Phi_1(a_n, a_{-n}, \mathbf{c}) + \Phi_2(a_n, a_{-n}, \mathbf{c}) \\ &= -\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{n\in\mathcal{B}(t)} I_n(a_n, a_{N_n}) \\ &\quad - \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{n\in\mathcal{B}(t)} \alpha \left[\sum_{k\in\Lambda_0} \delta\left(a_n, c_k\right) + \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta\left(a_n, c_k\right)\right]. \end{aligned} \quad (20)$$

Now, it is assumed that an arbitrary player *n* unilaterally changes its strategy selection from a_n to $\bar{a}_n$, then the corresponding change of the player *n*'s utility can be expressed as:

$$\begin{aligned}
&u_n(a_n, a_{-n}, \mathbf{c}) - u_n(\bar{a}_n, a_{-n}, \mathbf{c}) \\
&= \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \hat{u}_n(a_n, a_{-n}, \mathbf{c}, v(t)) \\
&\quad - \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \hat{u}_n(\bar{a}_n, a_{-n}, \mathbf{c}, v(t)) \\
&= \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} [\delta(n \in \mathcal{B}(t))(I_n(\bar{a}_n, a_{N_n}) - I_n(a_n, a_{N_n}))] \\
&\quad + \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \left[\sum_{i\in\mathcal{B}(t), i\neq n, i\in N_i} (I_i(a_i, a_{N_i}) - I_i(a_i, \bar{a}_{N_i}))\right] \\
&\quad + \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \left\{\delta(n \in \mathcal{B}(t))\,\alpha \left[\sum_{k\in\Lambda_0} \delta(\bar{a}_n, c_k)\right.\right. \\
&\quad + \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta(\bar{a}_n, c_k) - \sum_{k\in\Lambda_0} \delta(a_n, c_k) \\
&\quad \left.\left. - \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta(a_n, c_k)\right]\right\}. \qquad (21)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\Phi_1(a_n, a_{-n}, \mathbf{c}) - \Phi_1(\bar{a}_n, a_{-n}, \mathbf{c}) \\
&= -\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sum_{n\in\mathcal{B}(t)} I_n(a_n, a_{N_n}) \\
&\quad + \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sum_{n\in\mathcal{B}(t)} I_n(\bar{a}_n, a_{N_n}) \\
&= \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} [\delta(n \in \mathcal{B}(t))(I_n(\bar{a}_n, a_{N_n}) \\
&\quad - I_n(a_n, a_{N_n}))] \\
&\quad + \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sum_{i\in\mathcal{B}(t), i\neq n, i\in N_i} (I_i(a_i, \bar{a}_{N_i}) - I_i(a_i, a_{N_i})) \\
&\quad + \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sum_{l\in\mathcal{B}(t), l\neq n} (I_i(a_i, \bar{a}_{N_i}) - I_i(a_i, a_{N_i})),
\end{aligned} \qquad (22)$$

Only the neighboring users affect the player *n*'s utility, and hence we can obtain

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sum_{l\in\mathcal{B}(t), l\neq n} (I_i(a_i, \bar{a}_{N_i}) - I_i(a_i, a_{N_i})) = 0. \qquad (23)$$

In addition,

$$\begin{aligned}
&\Phi_2(a_n, a_{-n}, \mathbf{c}) = -\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \sum_{n\in\mathcal{B}(t)} \alpha \left[\sum_{k\in\Lambda_0} \delta(a_n, c_k)\right. \\
&\quad \left. + \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta(a_n, c_k)\right] \\
&= -\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \left\{\delta(n \in \mathcal{B}(t))\,\alpha \left[\sum_{k\in\Lambda_0} \delta(a_n, c_k)\right.\right. \\
&\quad \left. + \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta(a_n, c_k)\right] + \sum_{i\in\mathcal{B}(t)/n} \alpha \left[\sum_{k\in\Lambda_0} \delta(a_i, c_k)\right. \\
&\quad \left.\left. + \sum_{k\in\Lambda_1} \lambda_1(d_{ik}) \cdot \delta(a_i, c_k)\right]\right\} \\
&= -\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \left\{\delta(n \in \mathcal{B}(t))\,\alpha \left[\sum_{k\in\Lambda_0} \delta(a_n, c_k)\right.\right. \\
&\quad \left.\left. + \sum_{k\in\Lambda_1} \lambda_1(d_{nk}) \cdot \delta(a_n, c_k)\right]\right\} + \Gamma(a_{-n}), \qquad (24)
\end{aligned}$$

where $\Gamma(a_{-n})$ is independent of a_n.

Applying (21), (22) and (24), we have

$$\begin{aligned}
&\Phi(a_n, a_{-n}, \mathbf{c}) - \Phi(\bar{a}_n, a_{-n}, \mathbf{c}) \\
&\quad = u_n(a_n, a_{-n}, \mathbf{c}) - u_n(\bar{a}_n, a_{-n}, \mathbf{c}). \qquad (25)
\end{aligned}$$

Therefore, the proposed anti-jamming dynamic game $\mathcal{G}$ is an exact potential game, and it has at least one pure strategy NE. ■

Theorem 2: The optimal pure strategy NE constitutes the optimal solution of the optimization problem P1.

Proof: Following the properties of the potential game, the optimal pure strategy NE point is consistent with the global or local maxima of the potential function [31], [32]. Therefore, we can obtain

$$\mathbf{a}_{opt} \in \arg\max \Phi(a_n, a_{-n}, \mathbf{c}). \qquad (26)$$

According to (20), the relationship between the network GMPIJ and the defined potential function can be expressed as $\Phi(a_n, a_{-n}, \mathbf{c}) = -D(a_n, a_{-n}, \mathbf{c})$. Consequently, we achieve

$$\mathbf{a}_{opt} \in \arg\min D(a_n, a_{-n}, \mathbf{c}). \qquad (27)$$

Applying (14), the optimal pure strategy NE is a global minimum of P1. ■

IV. Learning Algorithm for Anti-Jamming Dynamic Game

The formulated anti-jamming dynamic game $\mathcal{G}$ is an exact potential game according to Theorem 1, and a large number of learning algorithms can be employed to achieve the NE, e.g., best response [31], spatial adaptive play [11], [12], and multi-agent learning [33]. However, these approaches are suitable for

Algorithm 1: Distributed Anti-Jamming Channel Selection Algorithm (DACSA).

Step 1: Set $t = 0$, and initialize the mixed strategy $\theta_{nm}(t) = 1/|\mathcal{M}|, \forall n \in \mathcal{N}, \forall m \in \mathcal{M}$.
Step 2: In the *t*th slot, each active user $n \in \mathcal{B}(t)$ selects channel $a_n(t)$ according to its current strategy $q_n(t)$.
Step 3: Active users achieve its received utility $\hat{u}_n(t)$.
Step 4: Active users update their strategies according to the following rules:

$$\begin{aligned} q_{nm}(t+1) &= q_{nm}(t) + b\tilde{u}_n(t)(1 - q_{nm}(t)), m = a_n(t), \\ q_{nm}(t+1) &= q_{nm}(t) - b\tilde{u}_n(t) q_{nm}(t), m \neq a_n(t), \end{aligned} \tag{28}$$

where $0 < b < 1$ denotes the learning step size, and $\tilde{u}_n(t) = \hat{u}_n(t)/L$ is the normalized utility. The inactive users remain their strategies unchanged, and we have

$$q_n(t) = q_n(t+1), \quad n \in \mathcal{B}(t)/\mathcal{N}. \tag{29}$$

Step 5: Update $t = t + 1$, until the stopping criterion holds.

the game models with fixed players, and cannot be applied in dynamic environment with varying number of players.

In this section, to find the NE of the proposed anti-jamming dynamic game $\mathcal{G}$, a distributed anti-jamming channel selection algorithm (DACSA) based on stochastic learning automata [26] is proposed in dynamic environment. Each user updates its channel selection, and therefore it does not need a centralized controller. The information exchange can be realized by a common control channel. To characterize the proposed DACSA algorithm, the anti-jamming dynamic game $\mathcal{G}$ is extended to a mixed strategy form. The user *n*'s channel selection probability vector at time slot *t* can be defined as $q_n(t) = (q_{n1}(t), \ldots, q_{nm}(t), \ldots, q_{nM}(t))$, where $q_{nm}(t)$ represents the probability that user *n* selects channel *m*. The proposed algorithm is shown in Algorithm 1. Similar to [8], [18], the stop criterion is either the maximum iteration number is achieved, or the probability vector holds $q_n(t) = q_n(t+1)$.

Theorem 3: As the learning step size $b \to 0$, the proposed DACSA algorithm asymptotically converges to a pure strategy NE point of the formulated game $\mathcal{G}$.

Proof: By Theorem 1, the proposed anti-jamming dynamic game $\mathcal{G}$ is an exact potential game. The stochastic learning automata can converge to a pure strategy NE for any exact potential game according to [35, Th. 6]. Therefore, this theorem is proved. The proof is similar to [34]–[37]. Motivated by [34], as $b \to 0$, the sequence $\mathbf{q}(t)$ converges to the solution of the following equation.

$$\frac{d\mathbf{q}}{dt} = f(\mathbf{q}), \mathbf{q}(0) = \mathbf{q}^0, \tag{30}$$

where $f(\mathbf{q}) = E[G(\mathbf{q}(t), \mathbf{a}(t), \mathbf{u}(t)) | \mathbf{q}(t)]$, $\mathbf{q}(0)$ denotes the initial condition, and $G(\mathbf{q}(t), \mathbf{a}(t), \mathbf{u}(t)) = \mathbf{q}(t+1)$ represents the update rule in eq. (28). Referring to [34], when user *n* selects channel *s*, and other users employ the mixed strategy $\mathbf{q}_{-n}$, we define user *n*'s expected utility as:

$$h_{ns}(\mathbf{q}) = \sum_{a_t, t \neq n} u_n(a_1, \ldots, a_{n-1}, s, a_{n+1}, \ldots, a_N) \prod_{t \neq n} q_{ta_t}. \tag{31}$$

Referring to [8], [34], [37], we have $H(\mathbf{q}) = E[\Phi(\mathbf{q})]$, where $\Phi(\mathbf{q})$ represents the potential function. Thus, we can obtain

$$\begin{aligned} H(m, \mathbf{q}_{-n}) = &\sum_{a_t, t \neq n} \Phi_n(a_1, \ldots, a_{n-1}, m, a_{n+1}, \ldots, a_N) \\ &\times \prod_{t \neq n} q_{ta_t}. \end{aligned} \tag{32}$$

Applying (31), (32), and (25), we have

$$H(m_1, \mathbf{q}_{-n}) - H(m_2, \mathbf{q}_{-n}) = h_{nm_1}(\mathbf{q}) - h_{nm_2}(\mathbf{q}). \tag{33}$$

Similar to [26, Th. 3.2] and [35, Th. 5], Theorem 3 is obtained. ■

Motivated by [38], the update rules of the proposed Algorithm 1 has two vector-vector sums and one scalar-vector product. Besides, it involves the normalized operations of user *n*'s utility. Thus, the complexity of the procedure of updating the channel selection can be expressed as $\mathcal{O}(2M + 2)$, in which*M* represents the number of available channels. For the proposed Algorithm 1, the complexity for estimating the number of the jamming and interfering users can be given by $\mathcal{O}(F)$, where *F* represents a constant determined by the estimation period and the employed methods. Moreover, the number of the convergence iterations of the proposed Algorithm 1 is denoted as *K*. Therefore, the total complexity can be given by:

$$\mathcal{C}_{algorithm} = K(\mathcal{O}(F) + \mathcal{O}(2M + 2)). \tag{34}$$

V. Numerical Results and Discussions

In this section, we present the simulation results to evaluate the performance of the proposed DACSA algorithm in dynamic environment. Similar to [12], [14], [29], under the condition of fixed number of users, the hypergraph is randomly constructed with $Q = 3$. In the simulation, the number of hyperedges is $|E|$, and $|E_s|$ and $|E_c|$ respectively represent the number of hyperedges of the strong interference relationship and cumulative interference relationship, and $|E_s| + |E_c| = |E|$. We assume that there exists one jammer located in the first jamming range, and there are two jammers located in the second jamming range. An illustration of the simulation settings is shown in Fig. 5. The simulation parameters are given as follows: the path loss exponent $\beta = 3$, the first jamming range radius $d_0 = 500$ m, the second jamming range radius $d_1 = 1000$ m, the positive constant $L = 20$, the learning step size $b = 0.08$, the number of available channels $M = 4$, and the number of users $N = 15$. For simplicity, it is assumed that all the users have the same active probabilities, i.e., $\theta_n = \theta, n \in \mathcal{N}$.

A. Convergence Behavior

The convergence behavior of the proposed DACSA algorithm is shown in Fig. 6, and we take an arbitrarily selected user 1 as an example. The user 1 randomly selects the channel with equal

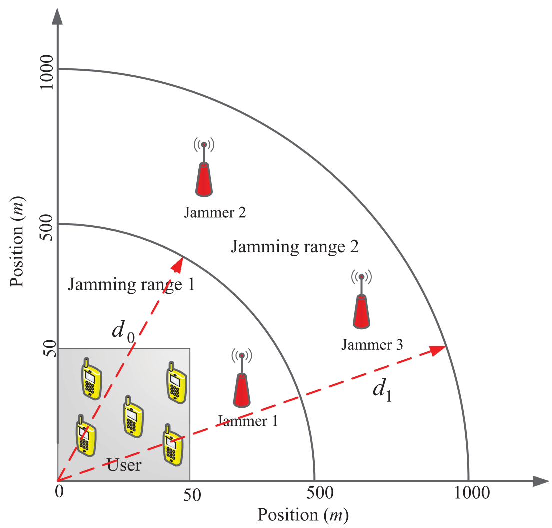


Fig. 5. An illustration of the simulation settings.

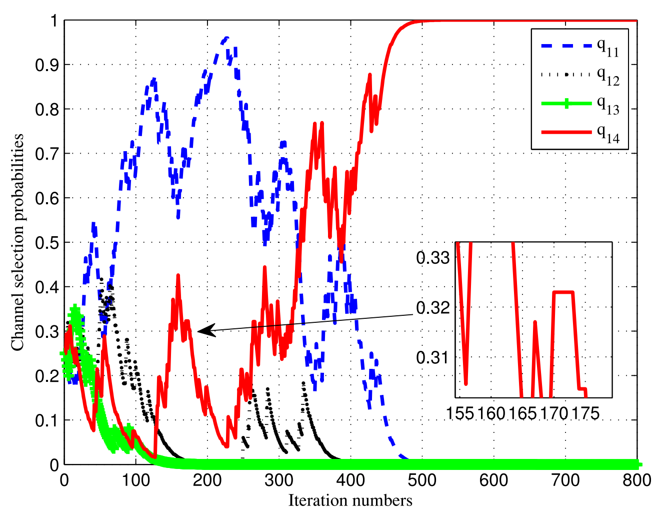


Fig. 6. Evolution of channel selection probabilities of user 1 ($\theta = 0.8, \alpha = 1$).

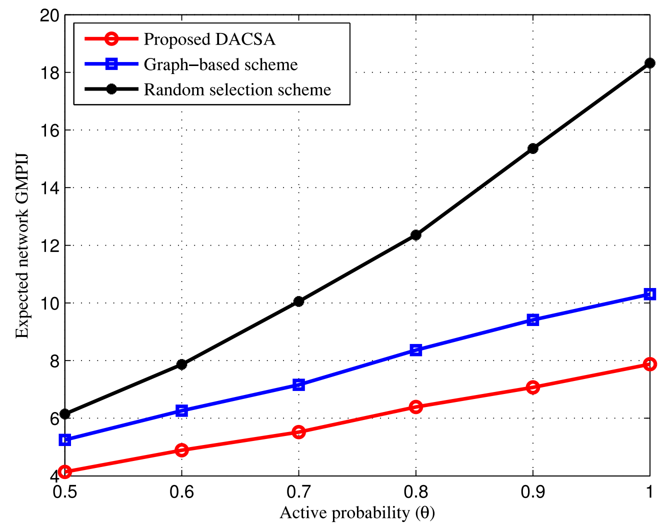


Fig. 7. The expected network GMPIJ versus the active probabilities θ ($\alpha = 1$, $M = 4$, $|E| = 30$, $|E_c| = 10$).

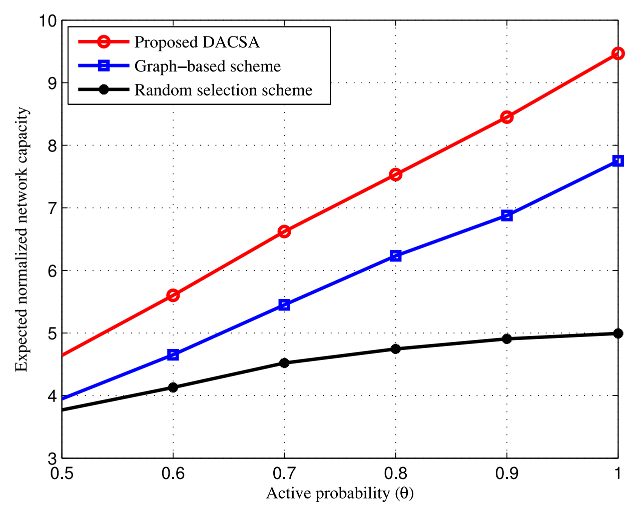


Fig. 8. The expected normalized network capacity versus the active probabilities θ ($\alpha = 1$, $M = 4$, $|E| = 30$, $|E_c| = 10$).

probabilities at time slot $t = 0$, and it is noted that the channel selection probabilities converge in about 500 iterations. Moreover, at certain time slots, the channel selection probabilities remain unchanged, which means that the user 1 is inactive during these time slots.

B. Performance Evaluation

To demonstrate the effectiveness of the proposed DACSA algorithm, we compare the simulation results of the graph-based scheme and random selection scheme. Each user randomly selects one channel at each time slot in the random selection scheme. We obtain the simulation results by conducting 1000 independent trials and then taking the mean value.

To evaluate the impact of the active probability θ, Fig. 7 and Fig. 8 respectively show the performance of the expected network GMPIJ and normalized network capacity versus the active probability θ. Fig. 7 indicates that the expected network GMPIJ increases with growing the active probability θ. The reason is that higher active probability leads to more serious mutual interference. As can be seen from Fig. 7, the solution achieved by the random selection scheme is the worst, and it brings the highest expected network GMPIJ. The solution to the proposed DACSA algorithm is the best, and it causes the lowest expected network GMPIJ. The reason is that the proposed DACSA algorithm converges to a desirable solution, and the random selection scheme is an instinctive approach. Since the cumulative interference effect is ignored, the performance of the graph-based scheme is worse than the proposed DACSA algorithm.

As indicated in Fig. 8, the expected normalized network capacity increases with the active probability θ. The reason is that the larger active probability θ leads to more frequent communication. Compared to the graph-based scheme and random selection scheme, the proposed DACSA algorithm achieves the higher normalized network capacity.

In Fig. 9, the statistical distribution for the 1000 trials is shown to analyze the uncertainty. Fig. 9 shows that the simulation data

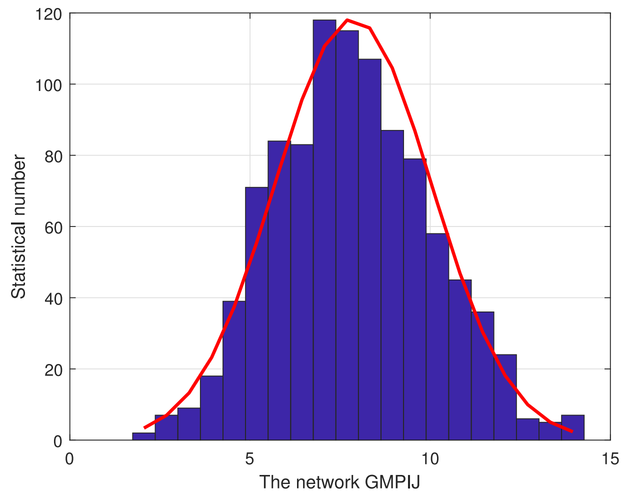


Fig. 9. The statistical distribution for 1000 trials ($\alpha = 1$, $M = 4$, $\theta = 1$, $|E| = 30$, $|E_s| = 10$).

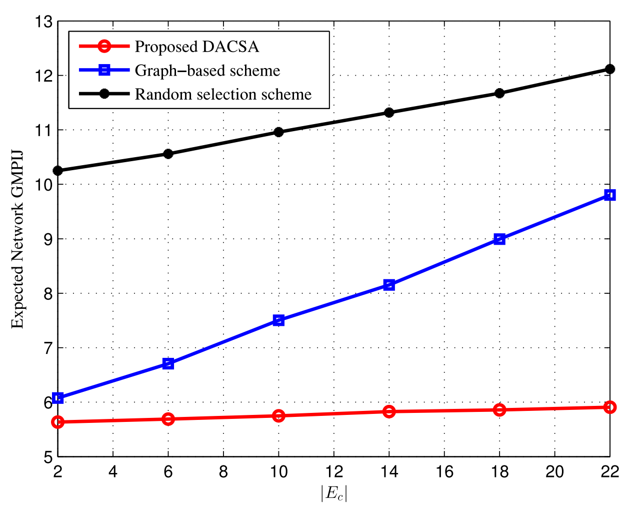


Fig. 10. The expected network GMPIJ versus the number of $|E_c|$ ($\alpha = 1$, $M = 4$, $\theta = 0.8$, $|E_s| = 15$).

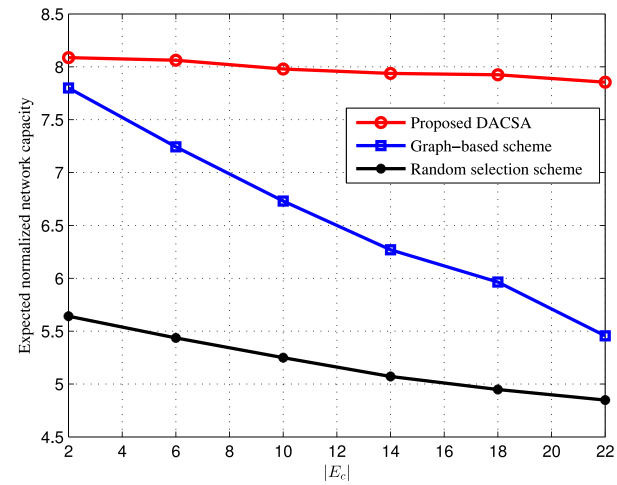


Fig. 11. The expected normalized network capacity versus the number of $|E_c|$ ($\alpha = 1$, $M = 4$, $\theta = 0.8$, $|E_s| = 15$).

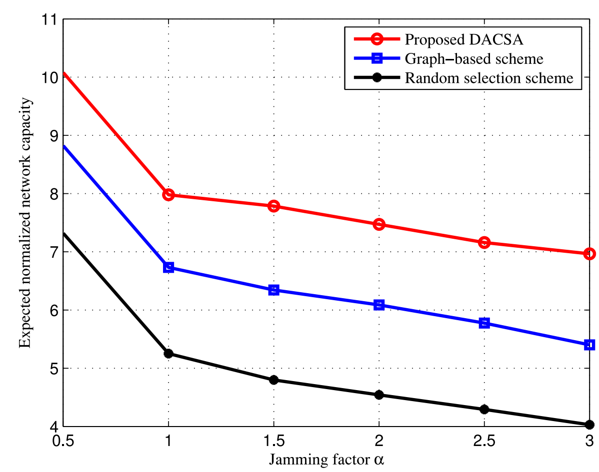


Fig. 12. The expected normalized network capacity versus jamming factor α ($M = 4$, $\theta = 0.8$, $|E| = 25$, $|E_c| = 10$).

is mainly concentrated in the vicinity of the mean value. When the confidence level is 95%, the mean value is 7.8632, and the confidence interval is [7.7282, 7.9983].

Fig. 10 and Fig. 11 show the impact of the number of $|E_c|$, when $|E_s|$ is fixed. As indicated in Fig. 10, the proposed DACSA algorithm performs much better than the graph-based scheme. It is because that the proposed DACSA algorithm has considered the effect of the cumulative interference constraint. Furthermore, the improvement for the proposed DACSA algorithm is more significant in comparison to the graph-based scheme with increasing $|E_c|$. The reason is that the interference relationship becomes more complicated with the increase of $|E_c|$. In such cases, the network is denser, and one user is exposed to more sources that will potentially cause interference. For further evaluation, Fig. 11 plots the performance comparison in terms of expected normalized network capacity by different solutions. It can be seen that the proposed DACSA algorithm is superior to the graph-based scheme and yields higher expected normalized network capacity.

In Fig. 12, we evaluate the impact of the jamming factor α. It is noted from the figure that the expected normalized network capacity decreases with the increase of jamming factor α. The reason is that larger jamming factor α means more serious damage. Moreover, the performance of the proposed DACSA algorithm outperforms the graph-based scheme and random selection scheme for different jamming factors.

VI. Discussion

In IM-based dense wireless networks, mutual interference significantly restricts the system performance, and external jamming is a potential threat to degrade channel utilization. The most important characteristic in this paper is that we jointly consider the mutual interference among users and external jamming for IM-based dense wireless networks, and propose a distributed anti-jamming channel selection algorithm. To accurately characterize the mutual interference and external jamming, a GMPIJ model is proposed. Specifically, to characterize the damage of external jammmers, we introduce the jamming factor α, which is related to the jamming level and signal processing capability.

In addition, it is also noted that some studies are needed in the near future.

- In this paper, the proposed GMPIJ model in this paper is our preliminary framework, and other mathematical descriptions will be investigated to characterize the mutual interference and external jamming more accurately in our future work.
- We assume that the jamming detection is perfect in this paper. However, imperfect detection will have an impact on system performance in practice. In the near future, we will investigate the impact of imperfect jamming detection, and design the anti-jamming channel selection scheme with imperfect jamming detection for IM-based dense wireless networks.
- The convergence speed of the proposed DACSA algorithm is slow, and we will investigate the efficient learning algorithm with faster convergence speed in future work.

VII. Conclusion

In this paper, we have investigated the anti-jamming channel selection problem for interference mitigation (IM)-based dense wireless networks. The dynamics of users is considered, and the active user set is varying. A generalized maximum protocol interference and jamming (GMPIJ) model is proposed to accurately characterize the mutual interference among users and external jamming. Moreover, an anti-jamming dynamic game has been formulated, and its properties are analyzed. In addition, a distributed anti-jamming channel selection algorithm (DACSA) based on stochastic learning automata is proposed to converge to the Nash equilibrium of the formulated game in dynamic environment. Finally, simulation results show that the performance improvement gain of the proposed DACSA algorithm is significant when compared with the graph-based scheme and random selection scheme.

References

[1] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.

[2] Y. E. Sagduyu, R. A. Berry, and A. Ephremides, "Jamming games in wireless networks with incomplete information," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 112–118, Aug. 2011.

[3] L. Jia, Y. Xu, Y. Sun, S. Feng, and A. Anpalagan, "Stackelberg game approaches for anti-jamming defence in wireless networks," *IEEE Wireless Commun.*, vol. 25, no. 6, pp. 120–128, Dec. 2018, doi: 10.1109/MWC.2017.1700363.

[4] X. Liu, Y. Xu, L. Jia, Q. Wu, and A. Anpalagan, "Anti-jamming communications using spectrum waterfall: A deep reinforcement learning approach," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 998–1001, May 2018.

[5] L. Zhang, G. Ding, Q. Wu, Y. Zou, Z. Han, and J. Wang, "Byzantine attack and defense in cognitive radio networks: A survey," *IEEE Commun. Surv. Tut.*, vol. 17, no. 3, pp. 1342–1363, Jul.–Sep. 2015.

[6] *Technical and Operational Principles for HF Sky-Wave Communication Stations to Improve the Man-Made Noise HF Environment*, ITU-R 258/5, International Telecommunication Union, Geneva, Switzerland, 2015.

[7] Y. Wu and D. H. K. Tsang, "Distributed power allocation algorithm for spectrum sharing cognitive radio networks with QoS guarantee," in *Proc. IEEE INFOCOM*, 2009, pp. 981–989.

[8] Q. Wu, Y. Xu, J. Wang, L. Shen, J. Zheng, and A. Anpalagan, "Distributed channel selection in time-varying radio environment: Interference mitigation game with uncoupled stochastic learning," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4524–4538, Nov. 2013.

[9] J. Zheng, Y. Cai, W. Yang, Y. Wei, and W. Yang, "A fully distributed algorithm for dynamic channel adaptation in canonical communication networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 491–494, Oct. 2013.

[10] J. Zheng, Y. Cai, Y. Xu, and A. Anpalagan, "Distributed channel selection for interference mitigation in dynamic environment: A game theoretic stochastic learning solution," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4757–4762, Nov. 2014.

[11] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE J. Sel. Topics Signal Process.*, vol. 6 , no. 2, pp. 180–194, Apr. 2012.

[12] Y. Sun, Q. Wu, Y. Xu, Y. Zhang, F. Sun, and J. Wang, "Distributed channel access for device-to-device communications: A hypergraph-based learning solution," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 180–183, Jan. 2017.

[13] A. Bretto, *Hypergraph Theory: An Introduction.* New York, NY, USA: Springer, 2013.

[14] H. Zhang, L. Song, and Z. Han, "Radio resource allocation for device-to-device underlay communication using hypergraph theory," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4852–4861, Jul. 2016.

[15] H. Zhang, Y. Ji, L. Song, and H. Zhu, "Hypergraph based resource allocation for cross-cell device-to-device communications," in *Proc. IEEE Int. Conf. Commun.*, 2016, pp. 1–6.

[16] Q. Li and R. Negi, "Maximal scheduling in wireless ad hoc networks with hypergraph interference models," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 297–310, Jan. 2012.

[17] L. Jia, F. Yao, Y. Sun, Y. Niu, and Y. Zhu, "Bayesian Stackelberg game for anti-jamming with incomplete information," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 1991–1994, Oct. 2016.

[18] L. Jia, F. Yao, Y. Sun, Y. Xu, S. Feng, and A. Anpalagan, "A hierarchical learning solution for anti-jamming Stackelberg game with discrete power strategies," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 818–821, Dec. 2017.

[19] D. Yang, G. Xue, J. Zhang, A. Richa, and X. Fang, "Coping with a smart jammer in wireless networks: A Stackelberg game approach," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4038–4047, Aug. 2013.

[20] L. Xiao, T. Chen, J. Liu, and H. Dai, "Anti-jamming transmission Stackelberg game with observation errors," *IEEE Commun. Lett.*, vol. 19, no. 6, pp. 949–952, Jun. 2015.

[21] L. Xiao, Y. Li, J. Liu, and Y. Zhao, "Power control with reinforcement learning in cooperative cognitive radio networks against jamming," *J. Supercomput.*, vol. 71, no. 9, pp. 1–21, Apr. 2015.

[22] L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor, "User-centric view of jamming games in cognitive radio networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 12, pp. 2578–2590, Dec. 2015.

[23] Y. Wu, B. Wang, K. Liu, and T. Clancy, "Anti-jamming games in multichannel cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 4–15, Jan. 2012.

[24] B. Wang, Y. Wu, K. Liu, and T. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 877–889, Apr. 2011.

[25] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Joint adaptation of frequency hopping and transmission rate for anti-jamming wireless systems," *IEEE Trans. Mobile Comput.*, vol. 15, no. 9, pp. 2247–2259, Sep. 2016.

[26] P. Sastry, V. Phansalkar, and M. Thathachar, "Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 5, pp. 769–777, May 1994.

[27] L. Jia, Y. Xu, Y. Zhang, Y. Sun, Y. Zhu, and X. Dai, "A distributed anti-jamming channel selection algorithm for interference mitigation-based wireless networks," in *Proc. IEEE 17th Int. Conf. Commun. Technol.*, Chengdu, China, 2017, pp. 151–155.

[28] K. Yao, Q. Wu, Y. Xu, and J. Jing, "Distributed ABS-slot access in dense heterogeneous networks: A potential game approach with generalized interference model," *IEEE Access*, vol. 5, pp. 94–104, 2017.

[29] Y. Sun, Z. Du, Y. Xu, Y. Zhang, L. Jia, and A. Anpalagan, "Directed-hypergraph-based channel allocation for ultradense cloud D2D communications with asymmetric interference," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7712–7718, Aug. 2018.

[30] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks.* Cambridge, U.K.: Cambridge Univ. Press, 2012.

[31] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, May 1996.

[32] K. Yamamoto, "A comprehensive survey of potential game approaches to wireless networks," *IEICE Trans. Commun.*, vol. E98.B, no. 9, pp. 1804–1823, Jan. 2015.

[33] Y. Xu, J. Wang, Q. Wu, J. Zheng, L. Shen, and A. Anpalagan, "Dynamic spectrum access in time-varying environment: Distributed learning beyond expectation optimization," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5305–5318, Dec. 2017.

[34] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Wireless Commun.*, vol. 11 , no. 4, pp. 1380–1391, Apr. 2012.

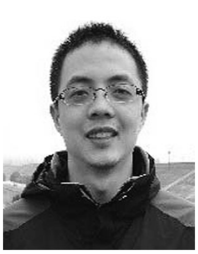
[35] J. Zheng, Y. Cai, N. Lu, Y. Xu, and X. Shen, "Stochastic game-theoretic spectrum access in distributed and dynamic environment," *IEEE Trans. Veh. Tech.*, vol. 64, no. 10, pp. 4807–4820, Oct. 2015.

[36] W. Zhong, Y. Xu, M. Tao, and Y. Cai, "Game theoretic multimode precoding strategy selection for MIMO multiple access channels," *IEEE Signal Process. Lett.*, vol. 17, no. 6, pp. 563–566, Jun. 2010.

[37] F. Yao, L. Jia, Y. Sun, Y. Xu, S. Feng, and Y. Zhu, "A hierarchical learning approach to anti-jamming channel selection strategies," *Wireless Netw.*, vol. 25, no. 1, pp. 201–213, Jan. 2019, doi: 10.1007/s11276-017-1551-9.

[38] L. Yu, Y. Xu, Q. Wu, and L. Jia, "Self-Organizing hit avoidance in distributed frequency hopping multiple access networks," *IEEE Access*, vol. 5, pp. 26614–26622, 2017.

Luliang Jia received the M.S. and Ph.D. degrees in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2014 and 2018, respectively. He is currently an Assistant Professor with the School of Space Information, Space Engineering University, Beijing, China. His research interests include game theory, learning theory, and communication anti-jamming technology.

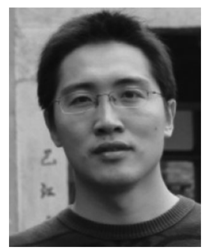


Yuhua Xu received the B.S. degree in communications engineering and the Ph.D. degree in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2006 and 2014, respectively. He is currently with the College of Communications Engineering, Army Engineering University of PLA, Nanjing, China, as an Associate Professor. He has authored/coauthored several papers in international conferences and reputed journals in his research area. His research interests focus on UAV communication networks, opportunistic spectrum access, learning theory, and distributed optimization techniques for wireless communications. He was the recipient of the Certificate of Appreciation as an Exemplary Reviewer for the IEEE COMMUNICATIONS LETTERS in 2011 and 2012, the IEEE Signal Processing Society's 2015 Young Author Best Paper Award, and the Funds for Distinguished Young Scholars of Jiangsu Province in 2016.

Youming Sun received the B.S. degree in electronic and information engineering from Yanshan University, Qinhuangdao, China, in 2010 and the M.S. degree in 2013 from the National Digital Switching System Engineering and Technological Research Center (NDSC), Zhengzhou, China, where he is currently working toward the Ph.D. degree in communications and information system.

His research interests include resource allocation in small cell networks, cognitive radio networks, game theory, and statistical learning. He currently serves as a regular reviewer for many technical journals, including the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE SYSTEMS JOURNAL, IEEE ACCESS, WIRELESS NETWORKS, *IET Communications*, and *KSII Transactions on Internet and Information Systems*. He was a Technical Program Committees member for the IEEE International Conference on Wireless Communications and Signal Processing 2015.

Shuo Feng (S'15) received the B.Sc. (Hons.) degree in electrical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2011 and the M.Sc. degree in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2014. He is currently working toward the Ph.D. degree in computational science and engineering with McMaster University, Hamilton, ON, Canada. His research interests include cognitive radio networks, machine learning, cognitive dynamic system, and information geometry. He is an invited reviewer for several journals, such as the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, etc. He served as the Technical Program Committee member for the IEEE Wireless Communications and Signal Processing 2015 and WCCN 2015. He was a co-recipient of the Best Paper Award from the IEEE VTC 2014-Fall.



Long Yu received the B.S. degree in mobile communications and the M.S. degree in communications and information system from the Institute of Communications Engineering, Nanjing, China, in 2003 and 2006, respectively. He is currently working toward the Ph.D. degree in communications and information system with the Institute of Communications Engineering, PLA University of Science and Technology, Nanjing, China. His research interests focus on wireless security, communication anti-jamming techniques, and game theory.

Alagan Anpalagan received the BASc., MASc., and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada. He is currently a Professor with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON, Canada, where he directs a research group working on radio resource management and radio access and networking areas within the WINCORE Laboratory. He co-authored three edited books, *Design and Deployment of Small Cell Networks* (Cambridge University Press, 2014), *Routing in Opportunistic Networks* (Springer, 2013), and *Handbook on Green Information and Communication Systems* (Academic Press, 2012).

Dr. Anpalagan served as an Editor for the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS from 2012 to 2014, IEEE COMMUNICATIONS LETTERS from 2010 to 2013, *Springer Wireless Personal Communications* from 2011 to 2013, and *EURASIP Journal of Wireless Communications and Networking* from 2004 to 2009. He currently serves as the TPC Vice-Chair, IEEE VTC Fall-2017, and served as the TPC Co-Chair, IEEE Globecom'15: SAC Green Communication and Computing, IEEE WPMC'12 Wireless Networks, IEEE PIMRC'11 Cognitive Radio and Spectrum Management, and IEEE CCECE'04/08. He served as the Vice-Chair, IEEE SIG on Green and Sustainable Networking and Computing with Cognition and Cooperation in 2015, IEEE Canada Central Area Chair from 2012 to 2014, IEEE Toronto Section Chair from 2006 to 2007, ComSoc Toronto Chapter Chair from 2004 to 2005, and IEEE Canada Professional Activities Committee Chair from 2009 to 2011. He was the recipient of the Dean's Teaching Award in 2011; Faculty Scholastic, Research and Creativity Award in 2010, 2014, 2017; Faculty Service Award in 2011, 2013 in Ryerson University. He was also the recipient of the Exemplary Editor Award from the IEEE ComSoc in 2013 and the Editor-in-Chief Top10 Choice Award in the Transactions on Emerging Telecommunications Technology in 2012. He is a Registered Professional Engineer in the province of Ontario, Canada, and the Fellow of the Institution of Engineering and Technology.