This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

Stochastic Game-Theoretic Spectrum Access in Distributed and Dynamic Environment

Jianchao Zheng, *Student Member, IEEE*, Yueming Cai, *Senior Member, IEEE*, Ning Lu, *Student Member, IEEE*, Yuhua Xu, *Student Member, IEEE*, and Xuemin (Sherman) Shen, *Fellow, IEEE*

***Abstract*—In this paper, we investigate the problem of channel selection for interference mitigation in opportunistic spectrum access networks using a stochastic game-theoretic approach. The studied network is *distributed* and *dynamic*, where each user only has its individual information and no information exchange is available among users. Moreover, each user is considered to be dynamically active due to its specific data service requirement. Specifically, a user randomly becomes active and then competes for the wireless channel to transmit for a random duration. To capture such dynamic interactions among users, a dynamic interference graph is defined, based on which the interference mitigation problem is formulated as a graphical stochastic game. It is proved to be an exact potential game, in which the existence of the Nash equilibrium (NE) is guaranteed. Then, the performance bounds of the NE are theoretically analyzed. Furthermore, we design a fully distributed and online algorithm based on stochastic learning for the interference-mitigation channel selection, which is proved to converge to the NE of the formulated game. Finally, we conduct simulations to validate the effectiveness of the proposed algorithm for interference mitigation and throughput improvement in the distributed and dynamic environment.**

***Index Terms*—Opportunistic spectrum access, distributed channel selection, dynamic service requirement, interference mitigation, interference graph, potential game, stochastic learning.**

I. Introduction

OPPORTUNISTIC spectrum access (OSA) is a promising solution to dealing with the spectrum shortage problem in wireless communications today, since it provides flexible and efficient spectrum usage [2]–[4]. In general, there are multiple available channels in OSA systems while each user can only access a small part of the channels at a time [5], [6]. These channels are not allocated to the users in advance, and users need to choose suitable channels in an intelligent manner. Therefore, careful design of channel selection scheme is key to eliminate mutual interference among the users and thus improve the network throughput [9].

The problem of channel selection in OSA systems has been widely studied in the literature, using e.g., multi-armed bandit problem [8], partially observable Markovian decision process [6], and optimal stopping theory [7]. However, these research works mainly focus on investigating the behavior of single user in OSA systems. In order to capture the interactions among multiple users, different game-theoretic channel selection approaches have been investigated [9]–[21], which can be divided into two categories. One is *non-spatial*, in which the users are located in a small-scale area and hence any two users interfere with each other [9]–[14]. The other one is *spatial*, which is a more general case that the users are located in a large-scale area. In this case, each user's transmission only interferes with the nearby users. To capture the characteristics of local interaction, graphical games (also called local interaction games) have been proposed [15]–[21].

Paying little attention to fully distributed scenarios in the absence of information exchange, the existing game theoretic solutions mainly focus on investigating the property of the game, e.g., the existence of NE and the convergence toward NE with explicit or implicit information exchange among users [22]. Moreover, almost all the existing works study the static environment, in which the users participate in the communication process all the time. In practice, however, users may not always need to occupy the channels for communication due to their specific data service demand. A user may randomly become active and then compete for the wireless channel to transmit for a random duration. Such dynamics of user transmission behavior yields a variable set of users participating in the competition for communication, which also imposes great challenges on achieving the optimal scheme for interference mitigation. This dynamics along with the distributed property is the focus of this paper.

In this paper, we consider a distributed and dynamic multiuser spectrum access network, which presents the following three characteristics: *i)* there is no centralized controller, *ii)* there is no information exchange among users, and *iii)* the users participate in the communication dynamically and stochastically, i.e., in an opportunistic manner. In this setting, the channel selection is conducted by each user independently and autonomously; moreover, the interests of the users are conflicting, since each user selfishly minimizes its experienced interference. Therefore, we adopt game theory to analyze and solve the problem of distributed channel selection for inter-

This work is supported by the Project of Natural Science Foundations of China under Grant No. 61301163, No. 61301162 and the Jiangsu Provincial Nature Science Foundation of China under Grant No. BK20130067. This work is also supported by a research grant from the Natural Science and Engineering Research Council (NSERC), Canada.

This paper has been presented in part at the GameNets Conference [1], November 2014, Beijing, China.

Jianchao Zheng, Yueming Cai and Yuhua Xu are all with the College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China (email: longxingren.zjc.s@163.com, caiym@vip.sina.com, yuhuaenator@gmail.com).

Ning Lu and Xuemin (Sherman) Shen are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada (e-mail: n7lu@uwaterloo.ca, sshen@uwaterloo.ca).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

ference mitigation, and the stochastic learning automata [23] is incorporated into the game model to capture the dynamics of the active users. Specifically, the main contributions of this paper are:

- We investigate the distributed channel selection in a dynamic network where the active user set varies dynamically. A dynamic interference graph is first defined to capture the dynamic and local interference. Then, a graphical stochastic game is formulated based on the defined interference graph. It is worth noting that the change of the active user set leads to the change of players in the game model, which imposes a great challenge on the existing game framework.
- The graphical stochastic game is proved to be an exact potential game, in which at least one pure-strategy Nash equilibrium (NE) exists. Moreover, the upper and lower performance bounds of the NE for any network topology are analyzed. Additionally, in the case of global interference, we establish the relationship between the expected interferences perceived by any two players at NE points, and then prove that all the NE points globally minimize the aggregate network interference when the active probabilities of different users are same.
- We design a fully distributed, online adaptive, stochastic learning algorithm to find the NE solution when the active users vary dynamically. Each user independently and adaptively adjusts its channel based on the individual experienced action-reward, and thus information exchange is not needed in the distributed and dynamic environment.

The remainder of this paper is organized as follows. In Section II, we give a brief review of the related works. In Section III, we introduce our system model and problem formulation. In Section IV, we present an interference mitigation game to analyze the problem of distributed channel selection. In Section V, we propose a fully distributed, online adaptive, stochastic learning algorithm to find the NE solution. In Section VI, we validate our analysis through simulation. Finally, concluding remarks are given in Section VII.

II. Related Work

Game theoretic approaches for distributed channel selection have been extensively investigated in the literature, e.g., [9]–[21]. However, most existing schemes require information exchange in terms of chosen actions and/or received payoffs among users during the convergence towards NE points. In [24], we proposed a fully distributed algorithm for adaptive channel selection in a canonical communication network. In [22], Xu *et al.* investigated the distributed spectrum access with spatial reuse from the perspective of avoiding information exchange and proposed two uncoupled algorithms to find the NE solution. In [15], Chen *et al.* studied the spatial channel selection and strategic mobility jointly, and designed distributed algorithms based on users' local information. However, the above mentioned works only consider a static environment, in which there is no dynamics on users.

Recently, the problem of channel selection for OSA in the dynamic environment began to draw attentions. In [25], Wu *et al.* studied the problem of distributed channel selection for interference mitigation in a time-varying radio environment without information exchange. In [26]. we investigated a more general and practical system model in which the active users are dynamically variable, and designed a dynamic learning algorithm based on the no-regret procedure. However, [25], [26] only studied a small-scale case in which every pair of users are close enough to cause interference to each other, and hence the game models considered are non-spatial. Moreover, these studies are from the perspective of minimizing the experienced interference in the physical layer, without considering the multiple access control mechanisms [22], which thus limits their application in practical communications. In [27], Xu *et al.* investigated the problem of spectrum access in the dynamic spectrum environment. However, the dynamics of the active users is not considered, and the game model cannot be applied to spectrum access with spatial reuse [15].

Overall, effective game models for the distributed and dynamic environment have not been well investigated. Moreover, the task of achieving Nash equilibrium (NE) solutions in the distributed and dynamic environment is challenging. Most existing game-theoretic algorithms, e.g., best (better) response [28], fictitious play [30], spatial adaptive play [31], and no-regret learning [32] are coupled, i.e., they need indispensable information exchange for users' strategy updating in each iteration; besides, these algorithms require the environment to be static during the convergence.

For the distributed and dynamic environment, some learning technologies can be found in the literature. Specifically, reinforcement learning has been widely adopted for aggregated interference control [35], Aloha-like spectrum access [36] and opportunistic bandwidth sharing [37] in cognitive radio networks. Besides, a trial-and-error learning approach is used for joint channel selection and power control in decentralized self-organizing networks [38], and stochastic learning automata [23] is designed for discrete power control [42] and distributed channel selection [25], [27] in dynamic networks. In methodology, incorporating learning technologies into game theory [23], [39]–[43] is very interesting and promising since game theory characterizes interactions among multiple users while learning technologies address the problems of lacking information exchange in the dynamic environment. The challenge here is to investigate the convergence property of the learning algorithms when incorporated into game theory, which differs greatly in specific scenarios.

III. System Model and Problem Formulation

A. System Model

We consider a canonical wireless communication network which consists of multiple autonomous communication nodes [24]–[26], [45], as shown in Fig. 1. In this canonical network, each node is not a single communication entity but a collection of several entities with intra-node communications capability. In each collection, the entities are closely located and there is a leading entity for managing the whole collection. The leading entity chooses the operational channel and the followers share the channel by using some multiple access control schemes.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

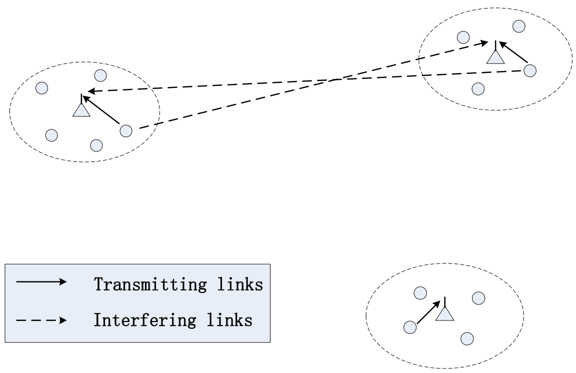


Fig. 1. Canonical network model.

[44]–[46] give some examples of the canonical communication network, e.g., a cluster head together with its members [44] and a WLAN access point along with its serving clients [45].

Denote the set of nodes[1] and the set of available channels by $\mathcal{N} = \{1, 2, \ldots, N\}$ and $\mathcal{M} = \{1, 2, \ldots, M\}$, respectively. Assume that the interference exclusively comes from the nodes with the same channel. Moreover, the nodes are spatially distributed and interference only occurs among nearby nodes [17], [18]. To capture the spatial separation, we characterize the limited range of interference by an un-directional graph $G = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the vertex set corresponding to the node set and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the edge set representing the mutual interference relationship among the nodes when they transmit in the same channel. Specifically, if the distance between two nodes m and n, denoted as d_{mn}, is less than a threshold d_0, it implies that they can hear each other and hence interfere with each other when simultaneously transmitting on the same channel; thus, m and n are connected by an edge $e_{mn} \in \mathcal{E}$. Let $\mathcal{B}_n$ denote the neighboring user set of user n, i.e.,

$$\mathcal{B}_n = \{i \in \mathcal{N} : (i, n) \in \varepsilon\}. \tag{1}$$

We assume that the interference is bilateral between any two users, i.e., user m is interfered with by user n when it interferes with n. That is, $m \in \mathcal{B}_n \Leftrightarrow n \in \mathcal{B}_m$.

In consideration of the specific service requirements for different nodes, we assume that nodes are active/inactive with probability in each time slot. For a specific node, the active probability is stationary from the statistics perspective. We use θ_n to denote the active probability of node n. In general, the active probabilities of different nodes are different due to their individual service requirements, i.e., $\theta_n \neq \theta_m$ when $n \neq m$. We define a probability space as $(\Omega, \mathcal{H}, \mathbb{P})$, where Ω is a sample space, $\mathcal{H}$ is a minimal σ-algebra on subsets of Ω, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{H})$. Let ω denote an event in the sample space Ω. $\mathbf{C}(\omega) : \Omega \rightarrow 2^N$ is a random vector, where $\mathbf{C} = [c_n]_{\forall n \in \mathcal{N}}$, and $c_n \in \{0, 1\}$ denotes the state of node n (0 for silent, and 1 for active), and the notation 2^N denotes the set of all N-dimensional binary vectors. Moreover, it is emphasized here that each node does not know the active probabilities of other nodes. Define $\mathcal{C}(t)$ as the active user set at time t, $\mathcal{C}(t) = \{n \in \mathcal{N} : c_n^t = 1\}$, where c_n^t is the state of node n at time t. Moreover, define $\mathcal{B}_n(t) = \{i \in \mathcal{C}(t) : (i, n) \in \varepsilon\}$ as active neighboring user set of user n at time t. Notably, $\mathcal{B}_n(t) = \mathcal{B}_n \cap \mathcal{C}(t)$.

[1]We will use node, user and player interchangeably in this paper.

B. Problem Formulation

Since the nodes are active/inactive with probability in each time slot, the experienced interference by each node is a random variable and can vary from slot to slot. Therefore, the throughput of each node is also random in each time slot. Firstly, we analyze the state-based case. Let a_n and $\mathcal{A}_n$ denote user n's selected channel and its available channel set, respectively. Obviously, $a_n \in \mathcal{A}_n$, and $\mathcal{A}_n = \mathcal{M}, \forall n \in \mathcal{N}$. For the considered canonical network, efficient distributed approaches such as CSMA can be applied to coordinate transmissions among neighboring and interfering nodes. According to the principle of CSMA, for a specific realization $\omega[t] \in \Omega$ at time t, the individual throughput of an arbitrary active node[2] $n \in \mathcal{C}(t)$ under channel selection profile $a = (a_1, \ldots, a_N)$ is given by:

$$\hat{q}_n(a_1, \ldots, a_N, \omega[t]) = \frac{f(\hat{s}_n + 1) R_{a_n}}{\hat{s}_n + 1}, \tag{2}$$

where R_{a_n} is the transmission rate[3] of channel a_n, $f(k)$ is the throughput loss function when k nodes are competing for a single channel [27], which decreases with k and satisfies $0 < f(k) \leq 1$, and

$$\hat{s}_n = \sum_{j \in \mathcal{B}_n(t)} \delta(a_j = a_n) \tag{3}$$

is the number of neighboring nodes interfering with node n at time t. Here, $\delta(\textit{condition})$ is the indicator function and it equals 1 (resp. 0) when *condition* is true (resp. false).

As can be seen from Eq. (2), lower value of $\hat{s}_n$ is desirable from the user-side, since minimizing $\hat{s}_n$ is equivalent to maximizing its throughput. In essence, $\hat{s}_n$ represents the interference level or local congestion level for node n [16], [22]. In order to maximize network throughput, lower aggregate interference experienced by all the nodes is more preferable from the network-side [22], and the quantitative characterization of the aggregate network interference is given by:

$$\hat{I}(a_1, \ldots, a_N, \omega[t]) = \sum_{n \in \mathcal{C}(t)} \hat{s}_n. \tag{4}$$

Motivated by the previous work of addressing interference mitigation [22], [24]–[26], [45], we study this problem from the perspective of minimizing the aggregate network interference. In a dynamic case where the active nodes vary dynamically, the network-centric goal is to find an optimal channel selection $a^{\text{opt}} = (a_1^{\text{opt}}, \ldots, a_N^{\text{opt}})$ that minimizes the expected network interference defined by $I(a_1, \ldots, a_N) = \mathbb{E}_{\mathbf{C}}\left[\hat{I}(a_1, \ldots, a_N, \mathbf{C})\right]$, where $\mathbb{E}_{\mathbf{C}}[\cdot]$ is the operation of

[2]Obviously, the throughput of a non-active node is 0, since it does not perform communications.

[3]For simplicity of analysis, we assume that all channels have the same transmission rate.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

taking expectation. Thus, the optimization problem can be formally expressed as

$$(P1): \quad a^{\text{opt}} \in \arg\min_{a \in \mathcal{A}} I, \tag{5}$$

where $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \ldots \times \mathcal{A}_N$ is the joint channel allocation strategy space, and $\times$ denotes the Cartesian product.

Remark 1. *$P1$ is a combinatorial optimization problem, and finding the optimal solution is NP-hard even in a centralized manner on the condition that all the system parameters (i.e., all the users' locations and active probabilities) are apriori known. Moreover, there is no central controller in the distributed network, and obtaining the other users' active probabilities is generally unrealistic and infeasible. Therefore, designing a low-complexity, fully distributed scheme to find the optimal solution is a challenging and valuable work.*

IV. Interference Mitigation Game

In this section, the problem of distributed channel selection for interference mitigation in the dynamic environment is formulated as a stochastic game, where the communication nodes act as the game players.

A. Graph Based Stochastic Game Model

According to the defined interference graph, the experienced interference of each player only depends on its own action and the action profile of its neighboring player set $\mathcal{B}_n$. Therefore, for a specific realization $\omega[t] \in \Omega$ at time t, the interference experienced by user n can be expressed as $\hat{s}_n(a_n, a_{\mathcal{B}_n}, \omega[t])$, which is given by

$$\hat{s}_n(a_n, a_{\mathcal{B}_n}, \omega[t]) = \delta(n \in \mathcal{C}(t)) \sum_{j \in \mathcal{B}_n(t)} \delta(a_j = a_n), \tag{6}$$

where $a_{\mathcal{B}_n}$ is a channel selection profile of player n's neighbors, and $\delta(n \in \mathcal{C}(t))$ is used to specify the state of node n since it could only be interfered with when active. It should be noted that Eq. (6) and Eq. (3) are equivalent, since node n is implicitly assumed to be active in Eq. (3). Then, the state-based utility function is defined as

$$\hat{u}_n(a_n, a_{\mathcal{B}_n}, \omega[t]) = L - \hat{s}_n(a_n, a_{\mathcal{B}_n}, \omega[t]), \tag{7}$$

where $L > |\mathcal{B}_n|$ is a predefined constant for player n to guarantee the utility non-negative and compatible with the learning algorithm proposed later, where $|\mathcal{B}_n|$ denotes the cardinality of the set $\mathcal{B}_n$.

Since $\omega[t]$ is random in each time slot, we define s_n as the expected interference received by player n in the dynamic environment, specified by[4]

$$s_n(a_n, a_{\mathcal{B}_n}) = \mathbb{E}_{\mathbf{C}}[\hat{s}_n(a_n, a_{\mathcal{B}_n}, \mathbf{C})] = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \{\hat{s}_n(a_n, a_{\mathcal{B}_n}, \omega[t])\}. \tag{8}$$

[4]Eq. (8) and (9) are based on the assumption that the stochastic process is ergodic, thus, the time average is equal to the average over the whole probability space.

Accordingly, the expected utility function of player n, u_n, is defined as

$$u_n(a_n, a_{\mathcal{B}_n}) = \mathbb{E}_{\mathbf{C}}[\hat{u}_n(a_n, a_{\mathcal{B}_n}, \mathbf{C})] = L - s_n(a_n, a_{\mathcal{B}_n}). \tag{9}$$

Now, we formulate the following graphical stochastic game denoted by $\mathcal{G} = [\mathcal{N}, \mathbf{C}, \{\mathcal{A}_n\}_{n \in \mathcal{N}}, \{\mathcal{B}_n\}_{n \in \mathcal{N}}, \{u_n\}_{n \in \mathcal{N}}]$, where $\mathcal{N}$ is the set of players, $\mathcal{A}_n$ is the set of available actions (channels) for each player n. Each player independently and selfishly adjusts its strategy to maximize its individual utility, which can be expressed as:

$$(\mathcal{G}): \quad \max_{a_n \in \mathcal{A}_n} u_n(a_n, a_{\mathcal{B}_n}), \forall n \in \mathcal{N}. \tag{10}$$

B. Analysis of Nash Equilibrium (NE)

Definition 1 (Nash equilibrium): A channel selection profile $a_{NE} = (a_1^*, a_2^*, \ldots, a_N^*)$ is a pure-strategy NE if and only if no player can improve its utility function by deviating unilaterally, i.e.,

$$u_n(a_n^*, a_{\mathcal{B}_n}^*) \geq u_n(a_n, a_{\mathcal{B}_n}^*), \forall n \in \mathcal{N}, \forall a_n \in \mathcal{A}_n. \tag{11}$$

Theorem 1. *For any network topology, $\mathcal{G}$ is an exact potential game which has at least one pure-strategy NE point.*

Proof: Following the similar idea of proof in [26], we first construct a potential function as

$$\Phi(a_n, a_{-n}) = -\frac{1}{2} I(a_1, \ldots, a_N) = -\frac{1}{2} \sum_{n \in \mathcal{N}} \mathbb{E}_{\mathbf{C}}[s_n(a_n, a_{\mathcal{B}_n}, \mathbf{C})]$$
$$= -\lim_{T \to \infty} \frac{1}{2T} \sum_{t=1}^{T} \sum_{n \in \mathcal{C}(t)} \sum_{m \in \mathcal{B}_n(t)} \delta(a_m = a_n), \tag{12}$$

where $a_{-n} \in \mathcal{A}_1 \times \mathcal{A}_{n-1} \times \mathcal{A}_{n+1} \ldots \times \mathcal{A}_N$ denotes a channel selection profile of all the players excluding n. Then, we have

$$\begin{aligned}
&\Phi(a_n, a_{-n}) \\
&= -\lim_{T \to \infty} \frac{1}{2T} \sum_{t=1}^{T} \sum_{n \in \mathcal{C}(t)} \sum_{m \in \mathcal{B}_n(t)} \delta(a_n = a_m) \\
&= -\lim_{T \to \infty} \frac{1}{2T} \sum_{t=1}^{T} \Bigg(\delta(n \in \mathcal{C}(t)) \sum_{m \in \mathcal{B}_n(t)} \delta(a_n = a_m) \\
&\quad + \sum_{i \in \mathcal{C}(t), i \neq n} \sum_{m \in \mathcal{B}_i(t)} \delta(a_i = a_m) \Bigg) \\
&= -\lim_{T \to \infty} \frac{1}{2T} \sum_{t=1}^{T} \Bigg(\delta(n \in \mathcal{C}(t)) \sum_{m \in \mathcal{B}_n(t)} \delta(a_n = a_m) \\
&\quad + \sum_{i \in \mathcal{C}(t), i \neq n} \sum_{m \in \mathcal{B}_i(t), m \neq n} \delta(a_i = a_m) \\
&\quad + \sum_{i \in \mathcal{C}(t), i \neq n} \delta(a_i = a_n)\, \delta(n \in \mathcal{B}_i(t)) \Bigg).
\end{aligned} \tag{13}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

According to the definition of $\mathcal{B}_n(t)$, we have

$$\begin{aligned}&\sum_{i\in\mathcal{C}(t),i\neq n}\delta(a_i=a_n)\,\delta(n\in\mathcal{B}_i(t))\\&=\sum_{i\neq n}\delta(a_i=a_n)\,\delta(n\in\mathcal{C}(t))\,\delta(i\in\mathcal{B}_n(t))\\&=\delta(n\in\mathcal{C}(t))\sum_{i\in\mathcal{B}_n(t)}\delta(a_i=a_n)\\&=\delta(n\in\mathcal{C}(t))\sum_{i\in\mathcal{B}_n(t)}\delta(a_n=a_i).\end{aligned}\tag{14}$$

Therefore,

$$\begin{aligned}&\Phi(a_n,a_{-n})\\&=-\lim_{T\to\infty}\frac{1}{2T}\sum_{t=1}^{T}\left(\delta(n\in\mathcal{C}(t))\sum_{m\in\mathcal{B}_n(t)}\delta(a_n=a_m)\right.\\&\quad\left.+\Psi_{-n}+\delta(n\in\mathcal{C}(t))\sum_{i\in\mathcal{B}_n(t)}\delta(a_n=a_i)\right)\\&=-\lim_{T\to\infty}\frac{1}{2T}\sum_{t=1}^{T}\left(2\delta(n\in\mathcal{C}(t))\sum_{m\in\mathcal{B}_n(t)}\delta(a_n=a_m)+\Psi_{-n}\right)\\&=-\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left(\hat{s}_n(a_n,a_{\mathcal{B}_n},\omega[t])\right)-\lim_{T\to\infty}\frac{1}{2T}\sum_{t=1}^{T}\Psi_{-n}\\&=-\mathbb{E}_{\mathbf{C}}\left[s_n(a_n,a_{\mathcal{B}_n},\mathbf{C})\right]-\lim_{T\to\infty}\frac{1}{2T}\sum_{t=1}^{T}\Psi_{-n},\end{aligned}\tag{15}$$

where $\Psi_{-n}=\sum_{i\in\mathcal{C}(t),i\neq n}\sum_{m\in\mathcal{B}_i(t),m\neq n}\delta(a_i=a_m)$. Because Ψ_{-n} is independent of player n's strategy, we can derive that

$$\begin{aligned}&\Phi(a'_n,a_{-n})-\Phi(a_n,a_{-n})\\&=\mathbb{E}_{\mathbf{C}}\left[\hat{s}_n(a_n,a_{\mathcal{B}_n},\mathbf{C})\right]-\mathbb{E}_{\mathbf{C}}\left[\hat{s}_n(a'_n,a_{\mathcal{B}_n},\mathbf{C})\right]\\&=u_n(a'_n,a_{-n})-u_n(a_n,a_{-n}),\end{aligned}\tag{16}$$

which indicates that the change in individual utility function caused by any player's unilateral deviation of channel selection strategy is equal to the change in the potential function. According to the definition given in [28], it is known that $\mathcal{G}$ is an exact potential game with Φ serving as the potential function. Moreover, the most important property of the exact potential game is that it has at least one pure-strategy NE point [28], [31], [33]. Therefore, Theorem 1 is proved. ∎

To evaluate the performance of the proposed game $\mathcal{G}$, the aggregate network interference of a pure-strategy NE point $a_{NE}=(a_1^*,a_2^*,\ldots,a_N^*)$ is given by:

$$I(a_{NE})=\sum_{n\in\mathcal{N}}s_n\left(a_n^*,a_{\mathcal{B}_n}^*\right),\tag{17}$$

and thus the best pure-strategy NE is defined to be the one that minimizes the network interference, i.e.,

$$a_{NE}^{\text{opt}}\in\arg\min_{a\in\mathcal{A}_{NE}}I,\tag{18}$$

where $\mathcal{A}_{NE}$ is the set of the NE points.

It can be seen from (10) that the players in the game are selfish, which may lead to inefficiency and dilemma, known as the *tragedy of commons* [29]. Analyzing the achievable performance of NE points is interesting but challenging. We will address it in the following theorems.

Theorem 2. *For any network topology, the best pure-strategy NE point of $\mathcal{G}$ is a global minimum of problem P1.*

Proof: According to the properties of the potential game, the best pure-strategy NE point, as specified by Eq. (18), coincides with the global maximizer of the potential function [28]. That is,

$$a_{NE}^{\text{opt}}\in\arg\max_{a\in\mathcal{A}}\Phi(a_n,a_{-n}).\tag{19}$$

Referring to Eq. (12), the relationship between the defined potential function and the aggregate network interference is characterized by $\Phi(a_n,a_{-n})=-\frac{1}{2}I(a_1,\ldots,a_N)$. Therefore, we can derive

$$a_{NE}^{\text{opt}}\in\arg\min_{a\in\mathcal{A}}I(a_1,\ldots,a_N).\tag{20}$$

Applying Eq. (5), we can conclude that the best pure-strategy NE point is a global minimum of $P1$. This concludes the proof. ∎

Theorem 2 indicates that the best NE point of $\mathcal{G}$ lies at the global minimum of the formulated network interference minimization problem $P1$. This result is interesting and promising, since the competitive and selfish decisions lead to a global optimality. The above analysis gives its best achievable performance, and we further investigate the lower bound of the performance following the similar idea of proof in [22].

Theorem 3. *For any network topology, the aggregate network interference at any NE point is bounded by $\frac{1}{M}\sum_{n\in\mathcal{N}}\sum_{j\in\mathcal{B}_n}\theta_n\theta_j$.*

Proof: According to the definition of utility function (9) and the definition of NE (11), for any pure-strategy NE point $a_{NE}=(a_1^*,\ldots,a_N^*)$, the following inequality holds:

$$s_n\left(a_n^*,a_{\mathcal{B}_n}^*\right)\le s_n\left(a_n,a_{\mathcal{B}_n}^*\right),\forall n\in\mathcal{N},a_n\in\mathcal{A}_n.\tag{21}$$

Summing the two-sides of (21) yields the following:

$$|\mathcal{A}_n|\cdot s_n\left(a_n^*,a_{\mathcal{B}_n}^*\right)\le\sum_{a_n\in\mathcal{A}_n}s_n\left(a_n,a_{\mathcal{B}_n}^*\right),\forall n\in\mathcal{N},\tag{22}$$

where $|\mathcal{A}_n|$ is the number of the available channels of user n. We can rewrite (22) as follows:

$$s_n\left(a_n^*,a_{\mathcal{B}_n}^*\right)\le\frac{\sum_{a_n\in\mathcal{A}_n}s_n\left(a_n,a_{\mathcal{B}_n}^*\right)}{|\mathcal{A}_n|},\forall n\in\mathcal{N}.\tag{23}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

According to Eq. (8), we have

$$\begin{aligned}
&\textstyle\sum_{a_n \in \mathcal{A}_n} s_n\left(a_n, a^*_{\mathcal{B}_n}\right) \\
&= \sum_{a_n \in \mathcal{A}_n} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \hat{s}_n\left(a_n, a^*_{\mathcal{B}_n}, \omega[t]\right) \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{a_n \in \mathcal{A}_n} \hat{s}_n\left(a_n, a^*_{\mathcal{B}_n}, \omega[t]\right) \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{a_n \in \mathcal{A}_n} \delta(n \in \mathcal{C}(t)) \sum_{j \in \mathcal{B}_n(t)} \delta\left(a_n = a_j\right) \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \delta(n \in \mathcal{C}(t)) \sum_{j \in \mathcal{B}_n(t)} \sum_{a_n \in \mathcal{A}_n} \delta\left(a_n = a_j\right).
\end{aligned} \tag{24}$$

Because $a_j \in \mathcal{A}_j = \mathcal{M} = \mathcal{A}_n$, the following equation holds

$$\begin{aligned}
&\sum_{a_n \in \mathcal{A}_n} \delta\left(a_n = a_j\right) \\
&= \delta\left(a_j = a_j\right) + \sum_{a_n \in \mathcal{A}_n, a_n \neq a_j} \delta\left(a_n = a_j\right) \\
&= 1.
\end{aligned} \tag{25}$$

Then,

$$\begin{aligned}
&\textstyle\sum_{a_n \in \mathcal{A}_n} s_n\left(a_n, a^*_{\mathcal{B}_n}\right) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \delta(n \in \mathcal{C}(t)) \sum_{j \in \mathcal{B}_n(t)} 1 \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} c^t_n \sum_{j \in \mathcal{B}_n} c^t_j = \sum_{j \in \mathcal{B}_n} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} c^t_n c^t_j \\
&= \sum_{j \in \mathcal{B}_n} \mathbb{E}_{\mathbf{C}}\left[c_n c_j\right] = \sum_{j \in \mathcal{B}_n} \mathbb{E}_{\mathbf{C}}\left[c_n\right] \mathbb{E}_{\mathbf{C}}\left[c_j\right],
\end{aligned} \tag{26}$$

where the last equation holds because activities of any two players (say n and j) are independent events. Then,

$$\textstyle\sum_{a_n \in \mathcal{A}_n} s_n\left(a_n, a^*_{\mathcal{B}_n}\right) = \sum_{j \in \mathcal{B}_n} \theta_n \theta_j, \tag{27}$$

and

$$s_n\left(a^*_n, a^*_{\mathcal{B}_n}\right) \leq \frac{\sum_{j \in \mathcal{B}_n} \theta_n \theta_j}{|\mathcal{A}_n|}, \forall n \in \mathcal{N}. \tag{28}$$

Accordingly, it follows that

$$I\left(a_{NE}\right) = \textstyle\sum_{n \in \mathcal{N}} s_n\left(a^*_n, a^*_{\mathcal{B}_n}\right) \leq \frac{\sum_{n \in \mathcal{N}} \sum_{j \in \mathcal{B}_n} \theta_n \theta_j}{|\mathcal{A}_n|}, \tag{29}$$

where $|\mathcal{A}_n|$ is the number of available channels, $|\mathcal{A}_n| \equiv M$. This concludes the proof.

■

Theorem 3 characterizes the upper bound of aggregate network interference at any NE point. It is shown that in order to obtain less aggregate interference, a larger number of available channels ($|\mathcal{A}_n|$) is preferable. The reason is that as the number of channels increases, the users can choose different channels to avoid mutual interference. Secondly, lower active probabilities of nodes can achieve lower interference. It is like the time-division multiplexing that each node shares a small portion of time slots for channel access. In addition, Theorem 3 demonstrates that smaller number of neighboring users ($|\mathcal{B}_n|$) can result in lower aggregate interference. As shown in Fig. 2, $|\mathcal{B}_n|$ equals to 2 and $N-1$ for the two interference graphs, respectively. If the active probabilities of all nodes are assumed to be equal to θ, the upper bounds of the aggregate interference at NE are $2N\theta^2/M$ and $N(N-1)\theta^2/M$, respectively. In particular, if the active probabilities of all nodes are set to 1, all the results in [22] can be achieved.

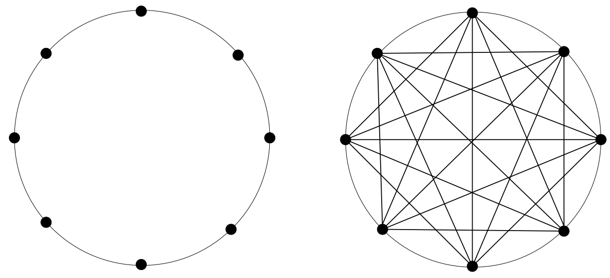


Fig. 2. Examples of interference graph (In Fig. (a), each node is interfered with only by the neighboring two nodes, while in Fig. (b) each node is interfered with by all the other nodes).

C. The Case of Global Interference

As for the interference graph in Fig. 2(b), any two nodes interfere with each other, i.e., all nodes locate in a small-scale mutually-interfering area, which generates the global interference. It is a common and general application case. In the following, we will investigate its special property.

Lemma 1. *In the case of global interference, at any NE point $a_{NE} = (a^*_1, a^*_2, \ldots, a^*_N)$, $\forall i, n \in \mathcal{N}, i \neq n$,*
if $a^*_i \neq a^*_n$, $\frac{\theta_n}{\theta_i} s_i\left(a^*_i, a^*_{\mathcal{B}_i}\right) - \theta^2_n \leq s_n\left(a^*_n, a^*_{\mathcal{B}_n}\right) \leq \frac{\theta_n}{\theta_i} s_i\left(a^*_i, a^*_{\mathcal{B}_i}\right) + \theta_n\theta_i$;
if $a^*_i = a^*_n$, $s_n\left(a^*_n, a^*_{\mathcal{B}_n}\right) = \frac{\theta_n}{\theta_i} s_i\left(a^*_i, a^*_{\mathcal{B}_i}\right) + \theta_n\theta_i - \theta^2_n$.

Proof: According to the definition of utility function (9) and the definition of NE (11), for any pure-strategy NE point $a_{NE} = (a^*_1, \ldots, a^*_N)$, we have

$$s_n\left(a^*_n, a^*_{\mathcal{B}_n}\right) \leq s_n\left(a_n, a^*_{\mathcal{B}_n}\right), \forall n \in \mathcal{N}, a_n \in \mathcal{A}_n. \tag{30}$$

Because $a^*_i \in \mathcal{A}_i = \mathcal{M} = \mathcal{A}_n$, $\forall i \in \mathcal{N}$, we can get

$$s_n\left(a^*_n, a^*_{\mathcal{B}_n}\right) \leq s_n\left(a^*_i, a^*_{\mathcal{B}_n}\right), \forall i, n \in \mathcal{N}, i \neq n. \tag{31}$$

Due to the definition of interference, we have

$$\begin{aligned}
&s_n\left(a^*_i, a^*_{\mathcal{B}_n}\right) \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \hat{s}_n\left(a^*_i, a^*_{\mathcal{B}_n}, \omega[t]\right) \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \delta(n \in \mathcal{C}(t)) \sum_{j \in \mathcal{B}_n(t)} \delta\left(a^*_j = a^*_i\right) \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} c^t_n \sum_{j \in \mathcal{B}_n} c^t_j \delta\left(a^*_j = a^*_i\right) \\
&= \theta_n \sum_{j \in \mathcal{B}_n} \theta_j \delta\left(a^*_j = a^*_i\right).
\end{aligned} \tag{32}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

Since any two nodes interfere with each other, $\mathcal{B}_n = \mathcal{N} \setminus \{n\}$. Thus,

$$s_n\left(a_i^*, a_{\mathcal{B}_n}^*\right) = \theta_n \sum_{j \in \mathcal{N} \setminus \{n\}} \theta_j \delta\left(a_j^* = a_i^*\right). \tag{33}$$

Similarly,

$$s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) = \theta_i \sum_{j \in \mathcal{N} \setminus \{i\}} \theta_j \delta\left(a_j^* = a_i^*\right). \tag{34}$$

1) If $a_i^* \neq a_n^*$, $\delta\left(a_n^* = a_i^*\right) = 0$. Therefore,

$$\begin{aligned} s_n\left(a_i^*, a_{\mathcal{B}_n}^*\right) &= \theta_n \sum_{j \in \mathcal{N}} \theta_j \delta\left(a_j^* = a_i^*\right) \\ &= \theta_n \sum_{j \in \mathcal{N} \setminus \{i\}} \theta_j \delta\left(a_j^* = a_i^*\right) + \theta_n \theta_i \\ &= \frac{\theta_n}{\theta_i} s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) + \theta_n \theta_i, \end{aligned} \tag{35}$$

which along with Eq. (31) yields

$$s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) \leq \frac{\theta_n}{\theta_i} s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) + \theta_n \theta_i. \tag{36}$$

Due to the symmetrical property, we have

$$s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) \leq \frac{\theta_i}{\theta_n} s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) + \theta_i \theta_n, \tag{37}$$

which is equivalent to

$$s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) \geq \frac{\theta_n}{\theta_i} s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) - \theta_n^2. \tag{38}$$

2) If $a_i^* = a_n^*$, $\delta\left(a_n^* = a_i^*\right) = 1$, then,

$$\begin{aligned} s_n\left(a_i^*, a_{\mathcal{B}_n}^*\right) &= \theta_n \sum_{j \in \mathcal{N}} \theta_j \delta\left(a_j^* = a_i^*\right) - \theta_n^2 \\ &= \theta_n \sum_{j \in \mathcal{N} \setminus \{i\}} \theta_j \delta\left(a_j^* = a_i^*\right) + \theta_n \theta_i - \theta_n^2 \\ &= \frac{\theta_n}{\theta_i} s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) + \theta_n \theta_i - \theta_n^2, \end{aligned} \tag{39}$$

which along with Eq. (31) yields

$$s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) \leq \frac{\theta_n}{\theta_i} s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) + \theta_n \theta_i - \theta_n^2. \tag{40}$$

Due to the symmetrical property, we have

$$s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) \leq \frac{\theta_i}{\theta_n} s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) + \theta_i \theta_n - \theta_i^2, \tag{41}$$

which is equivalent to

$$s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) \geq \frac{\theta_n}{\theta_i} s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) + \theta_n \theta_i - \theta_n^2. \tag{42}$$

Based on Eqs. (40) and (42), we can derive

$$s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) = \frac{\theta_n}{\theta_i} s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) + \theta_n \theta_i - \theta_n^2. \tag{43}$$

This concludes the proof. ■

Lemma 1 illustrates the relationship between the expected interferences (i.e., $s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right)$, $s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right)$) perceived by any two players at NE points. With this lemma, the following theorems can be achieved.

Theorem 4. *In the case of global interference, if the active probabilities of different players are equal, all the NE points of the game $\mathcal{G}$ globally minimize the expected aggregate interference.*

Proof: Assume $\theta_i = \theta$, $\forall i \in \mathcal{N}$. According to Lemma 1, at any NE point $a_{NE} = (a_1^*, a_2^*, \ldots, a_N^*)$, the following inequality holds:

$$s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) \leq s_i\left(a_i^*, a_{\mathcal{B}_i}^*\right) + \theta^2, \forall i, n \in \mathcal{N}, i \neq n. \tag{44}$$

Letting W_m denote the number of users on channel m, it is easy to get

$$s_n\left(a_n^*, a_{\mathcal{B}_n}^*\right) = \theta^2\left(W_{a_n^*} - 1\right). \tag{45}$$

Substituting Eq. (45) into Eq. (44), we can derive

$$W_{a_n^*} \leq W_{a_i^*} + 1, \forall i, n \in \mathcal{N}, i \neq n. \tag{46}$$

Besides, when the number of users is larger than the number of channels, no channel would be idle at the NE point[5]. That is, $\bigcup_{i \in \mathcal{N}} \{a_i^*\} = \mathcal{M}$, where $\{a_i^*\}$ denotes the set containing element a_i^*. Thus, according to Eq. (46), we have

$$W_{m_1} \leq W_{m_2} + 1, \forall m_1, m_2 \in \mathcal{M}. \tag{47}$$

Moreover, when the number of users is smaller than the number of channels, it is easy to know that each channel is occupied by at most one user at the NE point, which is compatible with Eq. (47). Therefore, we can conclude that Eq. (47) always holds at any NE point. According to Eq. (47), we can derive the unique channel profile[6] at the NE points as

$$\begin{aligned} &(W_1, W_2, \ldots, W_M) \\ &= \left(\left\lfloor \frac{N}{M} \right\rfloor + 1, \ldots, \left\lfloor \frac{N}{M} \right\rfloor + 1, \left\lfloor \frac{N}{M} \right\rfloor, \ldots, \left\lfloor \frac{N}{M} \right\rfloor\right), \end{aligned} \tag{48}$$

where $\left\lfloor \frac{N}{M} \right\rfloor$ denotes the largest integer but not bigger than $\frac{N}{M}$. Moreover, we can obtain that the number of $\left\lfloor \frac{N}{M} \right\rfloor + 1$ in the channel profile is $N - M\left\lfloor \frac{N}{M} \right\rfloor$, and the number of $\left\lfloor \frac{N}{M} \right\rfloor$ is $M - N + M\left\lfloor \frac{N}{M} \right\rfloor$.

Since all the NE points correspond to the unique channel profile, and the unique channel profile decides the unique value of the aggregate interference, we can conclude that all the NE points have the same performance. In addition, we have proved in Theorem 2 that one of the NE points is a global optimum. Therefore, all the NE points globally minimize the expected aggregate interference. The proof is completed. ■

V. Dynamic Stochastic Learning with Active Players Changing

According to the above analysis, the NE points exhibit very desirable and attractive properties. Thus, it is another important work to develop an efficient algorithm to achieve the NE in the distributed and dynamic environment.

[5]If there is an idle channel, users will definitely deviate their current strategies to access the idle channel, which contradicts with the definition of NE.

[6]Here we ignore the difference between channels, since it does not impact our conclusions.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

Since the distributed channel selection problem is formulated as an exact potential game, there are a number of learning algorithms available in the literature to obtain the pure-strategy NE, e.g., best response [28], fictitious play [30], spatial adaptive play [31], and log-linear learning [34]. However, these algorithms cannot be applied in the studied dynamic network since they are originally designed for static game models with information exchange among the players. Recently, two learning algorithms were designed for dynamic channel selection in canonical networks with block-fading channels [25] and opportunistic spectrum access in cognitive radio networks with dynamic channel availability [27], respectively. Nevertheless, they are designed for game models with fixed players, and thus cannot be applied to the studied dynamic networks with variable active players.

A. SLA Based Dynamic Channel Selection Algorithm

In the following, based on the stochastic learning automata (SLA) [23], we propose a fully distributed and online adaptive channel selection algorithm when the active players dynamically vary. Specifically, each game player (i.e., the communication node) is regarded as a learning automaton, which selects the channel according to a probability vector over the available channel set, and updates the probability vector based on the received reward from the unknown environment independently. We denote the channel selection probability vector (a.k.a. mixed strategy) for an arbitrary player n as $\mathbf{p}_n = (p_{n1}, \ldots, p_{nM})$, where p_{nm} represents the probability to select channel m. The detailed algorithm process is illustrated as follows.

1) **Initialization:** At the initial time $t = 0$, set the initial channel selection probabilities of each player to be $p^t_{nm} = \frac{1}{M}$, $\forall n \in \mathcal{N}$, $m \in \mathcal{M}$.
2) **Updating channel selection strategy:** At time t, each active player, say n, stochastically selects a channel a^t_n according to its current channel selection probability vector $\mathbf{p}^t_n$. The non-active players keep their channel selection strategies unchanged.
3) **Measuring received utility:** All the active players adhere to their channel selections in an estimation period, in which the active players access the spectrum for data transmission, and measure the received utilities $\hat{u}^t_n$ using Eqs. (3) and (7). At the same time, the non-active players perform no actions.
4) **Updating channel selection probability:** All the active players update their channel selection probabilities according to the following rules:
$$\begin{array}{ll} p^{t+1}_{nm} = p^t_{nm} + b\, r^t_n (1 - p^t_{nm}), & m = a^t_n \\ p^{t+1}_{nm} = p^t_{nm} - b\, r^t_n\, p^t_{nm}, & m \neq a^t_n, \end{array} \tag{49}$$
where $0 < b < 1$ is the learning step size, and $r^t_n = \hat{u}^t_n / L$ is the normalized received utility. The non-active players keep their channel selection probabilities unchanged.
5) **Stopping criterion:** If the channel selection probability vector of $t + 1$ is equal to that of t, stop the algorithm. Otherwise, t pluses one, and go to 2).

In the step 3 of the algorithm, the interference perceived by a user can be estimated by many approaches in the literature. Here, we adopt the method proposed in [22] for this issue. Specifically, assume that all the users have the same transmission probability and each estimation period consists of H sub-slots. Let T_n denote the number of sub-slots in which user n successfully accesses the channel, then the maximum likelihood estimation of the interference experienced by user n is given by $\hat{s}_n = \frac{H}{T_n} - 1$. Thus, the received utility is $\hat{u}_n = L - \hat{s}_n = L - \frac{H}{T_n} + 1$.

Remark 2. *The proposed learning solution is fully distributed, since the updating rule specified by Eq. (49) only depends on the individual experienced action-reward. Also, there is no need of some coordination mechanisms for monitoring the actions taken by different players, and each player independently and automatically updates its action without any information exchange. Moreover, the proposed algorithm is online adaptive, since players access the spectrum for data transmission in each iteration of the algorithm process.*

B. Convergence Analysis

Convergence analysis is important and indispensable for the design of distributed algorithms. In the literature, the convergence of the SLA based algorithms has been proved for coordination games where all the players have the same utility function [23] and potential games [27]. However, the active player set is assumed to be fixed in their game models. Therefore, our formulated channel selection game $\mathcal{G}$ where the active players dynamically vary is beyond the scope of existing work, and thus the convergence of the proposed learning algorithm needs to be evaluated. By utilizing the ordinary differential equation (ODE) and stochastic approximation theory, we can achieve the following theorem.

Theorem 5. *With a sufficiently small step size b, the proposed SLA based dynamic channel selection algorithm converges to a pure NE point of $\mathcal{G}$.*

Proof: Firstly, we re-write the updating rule in Eq. (49) as
$$\mathbf{p}^{t+1}_n = \mathbf{p}^t_n + b\, c^t_n\, r^t_n (\mathbf{I}_{a^t_n} - \mathbf{p}^t_n), \tag{50}$$
where c^t_n denotes the active/inactive state of player n at time t, and $\mathbf{I}_{a^t_n}$ is a unit vector with the a^t_n-th element being one. Let $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$ denote the mixed strategy profile of all the players, and thus we can achieve the evolution of the mixed strategy profile of the game $\mathcal{G}$ as:
$$\mathbf{P}^{t+1} = \mathbf{P}^t + b\, G(\mathbf{P}^t, \mathbf{c}^t, \mathbf{r}^t, \mathbf{a}^t), \tag{51}$$
where $\mathbf{c}^t = (c^t_1, c^t_2, \ldots, c^t_N)$, $\mathbf{r}^t = (r^t_1, r^t_2 \ldots, r^t_N)$, and $\mathbf{a}^t = (a^t_1, a^t_2, \ldots, a^t_N)$. Then, according to Theorem 3.1 in [23], we know that when the step size b is sufficiently small, i.e., $b \to 0$, the sequence $\{\mathbf{P}^t\}$ will converge weakly to the solution of the ODE,
$$\frac{d\mathbf{P}}{dt} = f(\mathbf{P}), \mathbf{P}(0) = \mathbf{P}^0, \tag{52}$$
where $f(\mathbf{P}) = \mathbb{E}\left[G(\mathbf{P}^t, \mathbf{c}^t, \mathbf{r}^t, \mathbf{a}^t) | \mathbf{P}^t\right]$, and $\mathbf{P}(0)$ is the initial value of the ODE which is equal to the initial channel selection probability matrix $\mathbf{P}^0$.

Besides, let $h_n(m, \mathbf{P}_{\mathcal{B}_n})$ denote player n's expected reward function when it adopts pure strategy m and its neighbors adopt mixed strategy $\mathbf{P}_{\mathcal{B}_n}$. Then, we have

$$h_n(m,\mathbf{P}_{\mathcal{B}_n}) = \sum_{a_k, k\in\mathcal{B}_n} u_n\left(m, a_{\mathcal{B}_n}\right) \prod_{k\in\mathcal{B}_n} p_{ka_k}, \quad (53)$$

where p_{ka_k} is the probability of player k to choose pure strategy a_k, u_n is the expected utility function specified by Eq. (9) when the active players dynamically vary. In addition, take $H(\mathbf{P}) = \mathbb{E}\left[\Phi(\mathbf{P})\right]$, and thus

$$H_n(m,\mathbf{P}_{-n}) = \sum_{a_k, k\neq n} \Phi_n(a_1, \ldots, a_{n-1}, m, a_{n+1}, \ldots, a_N) \prod_{k\neq n} p_{ka_k}. \quad (54)$$

Applying Eqs. (16), (53) and (54), we can derive

$$H_n(m_1,\mathbf{P}_{-n}) - H_n(m_2,\mathbf{P}_{-n}) = h_n(m_1,\mathbf{P}_{\mathcal{B}_n}) - h_n(m_2,\mathbf{P}_{\mathcal{B}_n}). \quad (55)$$

Then, following similar proof of Theorem 5 in [27], we can get that the proposed dynamic channel selection algorithm converges to a stationary point of the ODE (52). Moreover, Theorem 3.2 in [23] has demonstrated that all stationary points of the ODE that are not NE are unstable. Therefore, Theorem 5 is proved. ∎

VI. Simulation Results and Analysis

In this section, we conduct simulations to evaluate the performance of the proposed channel selection algorithm in the distributed and dynamic environment.

We consider a canonical network where communication nodes are randomly scattered in a d-by-d square area. Large scale network ($d = 1000$m [47]) and small scale network ($d = 100$m [25], [26]) are both investigated. All the communication nodes are assumed to use the IEEE 802.11b standard with 2Mb/s data rate, and 3 non-overlapping channels are available. The interference range of co-channel communications is set to be $d_0 = 200$m [47]. The individual throughput of a user is calculated by $\frac{R}{\hat{s}_n+1}$, which is approximately obtained according to the principle of perfect CSMA/CA [22], i.e., setting the throughput loss function as $f(\hat{s}_n+1) \approx 1$ in Eq. (2). In addition, the positive constant in Eq. (7) is set to be $L = 2$, and the step size of the proposed algorithm is set to be $b = 0.1$. For convenience of analysis, the active probabilities of all the nodes are assumed to be the same, i.e., $\theta_n = \theta$, $\forall n \in \mathcal{N}$, except for special investigation in small-scale networks (as shown in Fig. 11).

A. Convergence Behavior

For convergence analysis of the proposed stochastic learning algorithm, we generate a random network topology involving 60 nodes, as shown in Fig. 3. The convergence behavior of one arbitrarily selected node (marked as node 1) is presented in Fig. 4. At the beginning, it randomly selects the channels with equal probabilities. As the algorithm iterates, its channel selection probabilities evolve with the time and converge in about 560 iterations. At the convergence time, we can see $P_{12} = 1$, and $P_{11} = P_{13} = 0$, which demonstrates that the node finally chooses channel 2 for communication. It is seen that the channel selection probabilities keep unchanged in multiple successive slots (e.g., p_{11} from slot 343 to slot

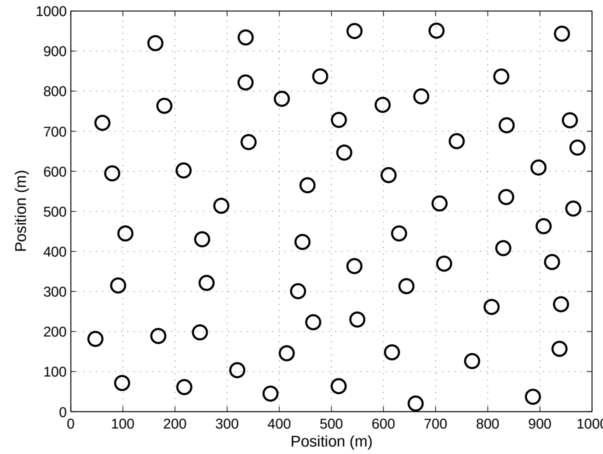


Fig. 3. A random network topology with 60 nodes.

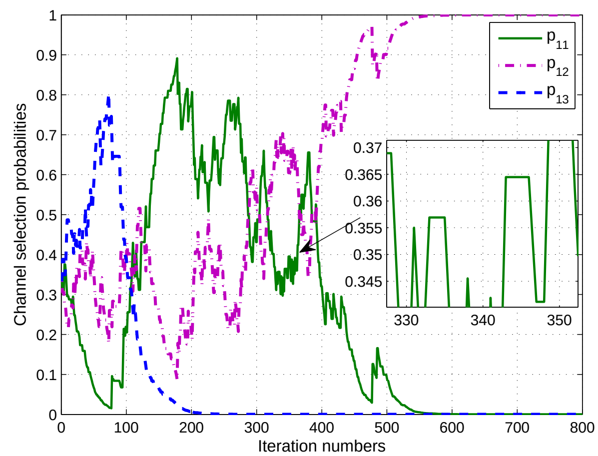


Fig. 4. Evolution of channel selection probabilities of one arbitrarily selected node (θ=0.6).

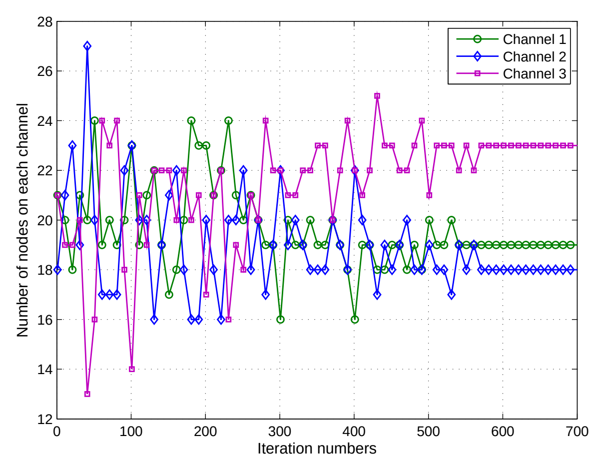


Fig. 5. Evolution of the number of nodes selecting different channels (θ=0.6).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

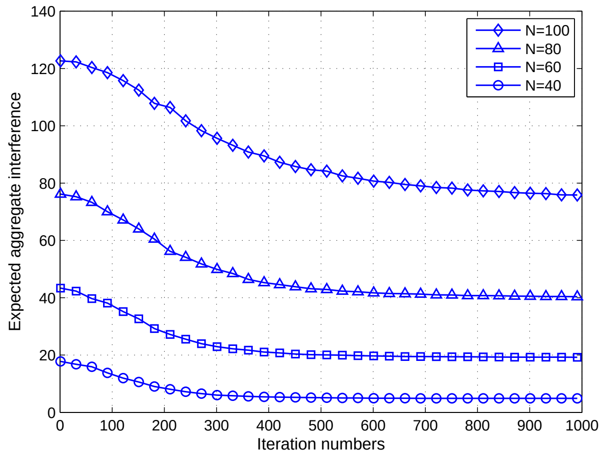


Fig. 6. Convergence comparison for different numbers of nodes (θ=0.6).

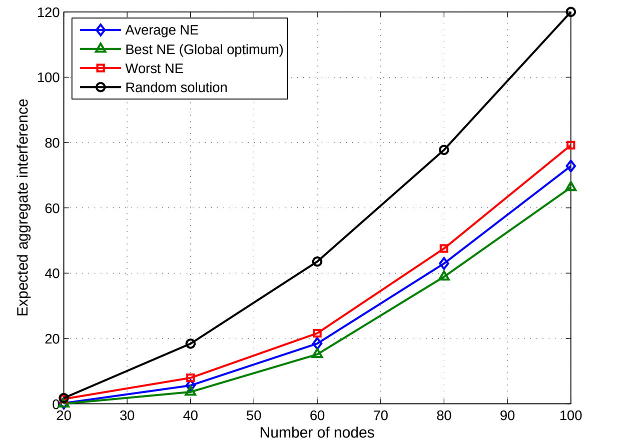


Fig. 7. Performance comparison for different solutions in terms of the expected aggregate interference in large networks ($\theta = 0.6$).

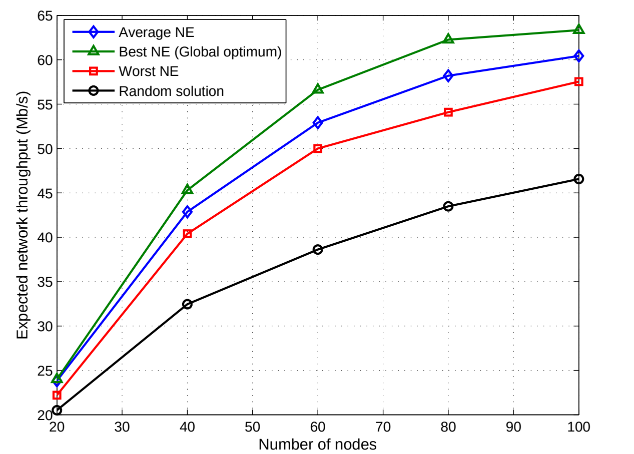


Fig. 8. Performance comparison for different solutions in terms of the achieved network throughput in large networks ($\theta = 0.6$).

347), which indicates the node keeps inactive in these slots.

Moreover, the evolution of the number of nodes on different channels is plotted in Fig. 5. Since the nodes adjust their channel strategies continuously based on their channel selection probabilities before the convergence time, the number of nodes on the channels varies accordingly. After the convergence, all the nodes remain their current channel strategies, and thus the number of nodes on the channels keeps unchanged. The simulation results validate the convergence of the proposed algorithm when the number of the active users dynamically changes.

Besides, Fig. 6 plots the convergence of network utility (i.e., expected aggregate interference) for different numbers of nodes N. As shown in Fig. 6, the convergence time increases with the number of nodes. When the number of nodes is no more than 60, the convergence can be achieved within 500 iterations. However, when the number of nodes is beyond 80, (even more than) 1000 iterations are needed for convergence. Nevertheless, the longer convergence time will not impact its practical application, since the proposed algorithm is online adaptive, which means nodes can access the spectrum for data transmission in each iteration of the algorithm process.

B. Performance Evaluation

In order to evaluate the performance of the proposed algorithm, we present the performance of the random selection scheme, the best NE, and the worst NE for comparison. In the random selection scheme, each node randomly chooses a channel in each slot. Since the active player set varies randomly and there is no information exchange, random channel selection seems to be an instinctive method. It is noted the exhaustive enumeration approach cannot be available to search the globally optimal channel selection due to the huge computation complexity when the number of nodes is large. Thus, we achieve the globally optimal solution by running the proposed algorithm for multiple times and taking the best solution, since Theorem 2 demonstrates that the best NE is the global optimum. Because the scenario of small networks presents many special properties (Section IV.C), we evaluate the performance in large and small networks separately.

1) Scenario of large networks: Fig. 7 plots performance comparison of the expected aggregate interference achieved by different solutions. The presented results are obtained by simulating 1000 independent trials and then taking the expected value. Intuitively, the solution to the random channel selection scheme is the worst which causes the most severe interference. The equilibrium solution achieved by our proposed stochastic learning algorithm is much better, which approaches the globally optimal solution. It is because that the learning equilibrium solution may converge to a locally/globally optimal channel selection profile as characterized by Theorem 1, and hence achieves near-optimal performance on average. Moreover, the gap between the worst NE and the best NE is small; while the gap between the random solution and the best NE increases with the number of nodes. In addition, the interference values achieved by all the solutions increase with the number of nodes, and the increasing speed gets faster.

For further evaluation, we present the performance compar-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

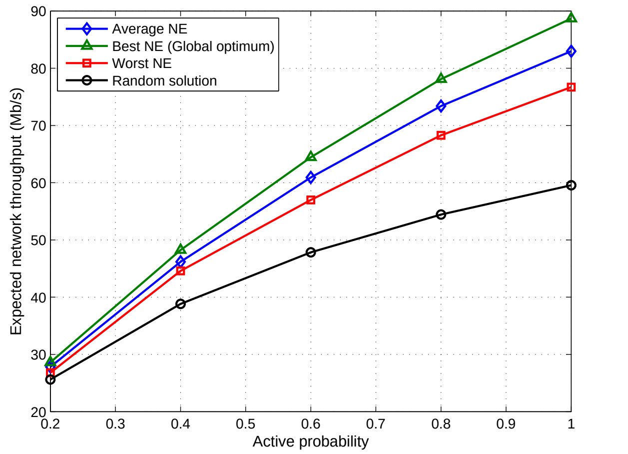


Fig. 9. Performance comparison for different solutions in terms of the achieved network throughput in large networks (N=100).

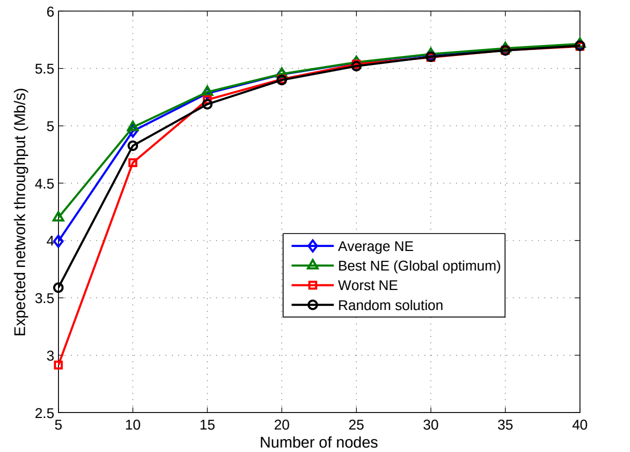


Fig. 10. Performance comparison for different solutions in terms of the achieved network throughput in small networks (θ=0.6).

ison in terms of the achieved expected network throughput in Fig. 8. As shown in the figure, the expected network throughput obtained by different solutions increases with the number of nodes in the network. It is clear since the network throughput is the sum of individual throughput achieved by each node. Moreover, it can be seen that the increasing speed of the curves gets slower when the number of nodes in the network becomes larger, which is due to the impact of more severe interference. As for the performance comparison, the random channel selection scheme achieves the worst solution, and the NE solutions obtained by our proposed stochastic learning are much better. In addition, we can see that the performance gaps between different algorithms all increase with the number of nodes, which is because the solution space gets larger when the number of nodes increases.

To analyze the impact of the active probability of the nodes, we plot the achieved expected network throughput versus the active probability θ in Fig. 9. It can be seen from the figure that the expected network throughput obtained by different solutions increases with value of active probability, since larger active probability means more frequent communication. Also, the figure shows that the increasing speed of the curves gets slower when the active probability becomes larger, because larger active probability leads to more aggressive communication and more severe interference. In terms of the performance comparison, the NE solutions obtained by our proposed stochastic learning are much better than the random solution. Moreover, the performance gaps between different algorithms all get larger when the active probability increases.

2) Scenario of small networks: In this part, we will evaluate the performance of the proposed algorithm in small-scale networks (d = 100m). Since the distance of any two nodes is smaller than the interference range d_0 = 200m, global interference is generated.

In Fig. 10, we present the achieved expected network throughput achieved by different solutions in small networks. As shown in the figure, the achieved expected network throughput increases with the number of nodes in the network,

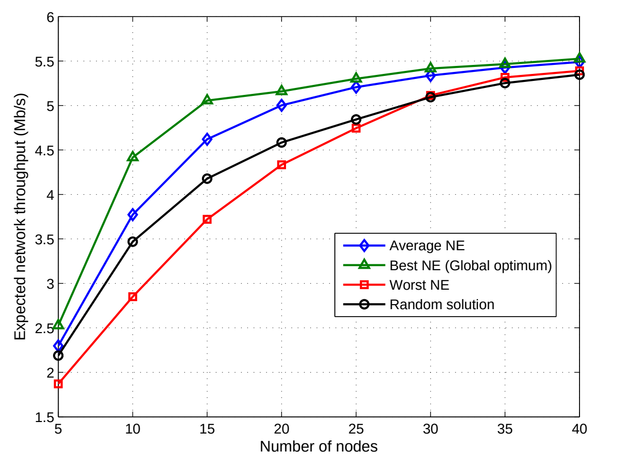


Fig. 11. Performance comparison for different solutions in terms of the achieved network throughput in small networks (θ varies for different nodes).

but the increasing speed gets much slower when the number of nodes in the network is larger than 20. Since any two nodes interfere with each other in a small network, the increasing of the number of nodes will cause more severe interference than in a large network. As for the performance comparison, we can see that all NE solutions present similar performance, which corresponds to Theorem 4. Moreover, we have given the unique channel profile of the NE as Eq. (48), which in essence is the uniform allocation. Thus, the random channel selection scheme can achieve quite close performance as the NE solutions (especially when the number of nodes is large), since each node randomly selects the channels with equal probabilities.

Although we have proved in Theorem 4 that all the NE solutions have the same performance in the case of global interference, it can be seen from Fig. 10 that performance gap exists when the number of nodes is small. It is because of the parameter setting (i.e., step size b) of the proposed stochastic learning algorithm, since it is proved in Theorem 5 that the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

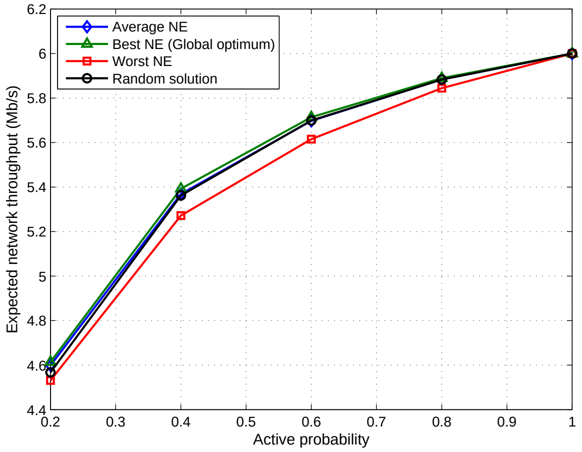


Fig. 12. Performance comparison for different solutions in terms of the achieved network throughput in small networks (N=40).

NE can be converged to only when b is sufficiently small. In other words, the solutions found by the proposed algorithm may include non-equilibrium points. However, b cannot be set to be sufficiently small in practical application since smaller b leads to slower convergence speed. In addition, the predefined positive constant L also has impact on the convergence and performance of the proposed algorithm. The readers can refer to [25] for more simulations and analysis on the selection of the parameters.

Since the conclusion of Theorem 4 holds only when the active probabilities of nodes are all equal, we further evaluate the performance of the proposed algorithm when θ randomly varies for different nodes in Fig. 11. In this case, the performance gap between the best NE and the worst NE is significant, as shown in the figure. However, the average performance of the NE points is better than that of the random solution, while the latter is better than the worst NE. In addition, the variation rule of the curves versus the number of nodes is similar to that in Fig. 10, which is omitted here for brevity.

Besides, Fig. 12 plots the achieved expected network throughput versus the active probability θ. Because the active probabilities of different nodes are equal, the performance of all the NE are theoretically proved to be equal. Therefore, they (i.e., average NE, best NE, and worst NE) present similar performance in the simulation figure. The small gaps between them are also due to the parameter setting of the proposed stochastic learning algorithm, as analyzed in Fig. 10. Moreover, the random channel selection scheme can achieve quite close performance as the NE solutions. In addition, it can be seen that the expected network throughput increases with the active probability, since larger active probability indicates more frequent communication. Also, the increasing speed of the curves get slower when the active probability increases, since larger active probability leads to more aggressive communication and more severe interference.

VII. Conclusion

In this paper, we have investigated the problem of channel selection for interference mitigation in a distributed and dynamic canonical communication network. The problem has been formulated as a graphical stochastic game and we have proved that the formulated game is an exact potential game which has at least one Nash equilibrium (NE). Furthermore, the performance bounds of the NE have been theoretically analyzed. Accordingly, we have designed a fully distributed, online adaptive, stochastic learning algorithm to converge to the NE of the formulated game. Simulation results validate the effectiveness of our proposed algorithm. For the future work, we will further improve the algorithm convergence speed.

References

[1] J. Zheng, Y. Cai, and Y. Xu, “MAC-layer interference mitigation in dynamic and distributed environment: Dynamic graphic game with stochastic learning,” *GameNets*, 2014.

[2] Y. Zhao, S. Mao, J. Neel, and J. Reed, “Performance evaluation of cognitive radios: Metrics, utility functions, and methodology,” *Proc. IEEE*, vol. 97, no. 4, pp. 642-659, Apr. 2009.

[3] S. Haykin, “Cognitive radio: Brain-empowered wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[4] M. Masonta, M. Mzyece, and N. Ntlatlapa, “Spectrum decision in cognitive radio networks: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1088-1107, 2013.

[5] M. Bkassiny, Y. Li, and S. K. Jayaweera, “A survey on machine-learning techniques in cognitive radios,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1136-1159, 2013.

[6] Q. Zhao, L. Tong, A. Swami, and Y. Chen, “Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework,” *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589-600, Apr. 2007.

[7] H. Jiang, L. Lai, R. Fan, and H. V. Poor, “Optimal selection of channel sensing order in cognitive radio,” *IEEE Trans. Wireless Commun.,* vol. 8, no. 1, pp. 297-307, Jan. 2009.

[8] K. Liu and Q. Zhao, “Distributed learning in multi-armed bandit with multiple players,” *IEEE Trans. Signal Process.,* vol. 58, no. 11, pp. 5667-5681, Nov. 2010.

[9] N. Nie and C. Comaniciu, “Adaptive channel allocation spectrum etiquette for cognitive radio networks,” *Mobile Networks & Applications*, vol. 11, no. 6, pp.779-797, 2006.

[10] M. van der Schaar and F. Fu, “Spectrum access games and strategic learning in cognitive radio networks for delay-critical applications,”*Proc. IEEE*, vol. 97, no. 4, pp. 720-740, Apr. 2009.

[11] A. Attar, M. Nakhai, and A. Aghvami, “Cognitive radio game for secondary spectrum access problem,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2121-2131, Apr. 2009.

[12] D. Niyato and E. Hossain, “Competitive spectrum sharing in cognitive radio networks: A dynamic game approach,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2651-2660, July 2008.

[13] M. Felegyhazi, M. Cagalj, and J. P. Hubaux, “Efficient MAC in cognitive radio systems: A game-theoretic approach,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1984-1995, Apr. 2009.

[14] M. Maskery, et al., “Decentralized dynamic spectrum access for cognitive radios: Cooperative design of a non-cooperative game,” *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 459-469, Feb. 2009.

[15] X. Chen and J. Huang, “Distributed spectrum access with spatial reuse,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 593-603, Mar. 2013.

[16] R. Southwell, X. Chen, and J. Huang, “Quality of service games for spectrum sharing,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 3, pp. 589-600, Mar. 2014.

[17] H. Li and Z. Han, “Competitive spectrum access in cognitive radio networks: Graphical game and learning,” in *Proc. IEEE WCNC*, pp. 1-6, 2010.

[18] M. Azarafrooz and R. Chandramouli, “Distributed learning in secondary spectrum sharing graphical game,” in *Proc. IEEE GLOBECOM*, pp. 1-6, 2011.

[19] M. Liu, et al., “Congestion games with resource reuse and applications in spectrum sharing,” *GameNets*, pp. 171-179, 2009.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

[20] C. Peng, H. Zheng, and B. Zhao, “Utilization and fairness in spectrum assignemnt for opportunistic spectrum access,” *Mobile Netw. Applications*, vol. 11, no. 4, pp. 555-576, 2006.

[21] N. Cheng, N. Zhang, N. Lu, X. Shen, J. W. Mark, and F. Liu, “Opportunistic spectrum access for CR-VANETs: A game-theoretic approach,” *IEEE Trans. Veh. Tech.*, vol. 63, no. 1, pp. 237-251, Jan. 2014.

[22] Y. Xu, Q. Wu, L. Shen, J. Wang, and A. Anpalgan, “Opportunistic spectrum access with spatial reuse: Graphical game and uncoupled learning solutions,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4814-4826, Oct. 2013.

[23] 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.

[24] 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.

[25] 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. Tech.*, vol. 62, no. 9, pp. 4524-4538, Nov. 2013.

[26] 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. Tech.*, in press, 2014. DOI: 10.1109/TVT.2014.2311496.

[27] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y. 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.

[28] D. Monderer. and L. S. Shapley, “Potential games,” *Games and Economic Behavior*, vol. 14, pp. 124-143, 1996.

[29] H. Kameda and E. Altman, “Inefficient noncooperation in networking games of common-pool resources,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1260-1268, Sep. 2008.

[30] J. Marden, G. Arslan, and J. Shamma, “Joint strategy fictitious play with inertia for potential games,” *IEEE Trans. Automat. Control*, vol. 54, no. 2, pp. 208-220, Feb. 2009.

[31] J. Zheng, Y. Cai, Y. Liu, Y. Xu, B. Duan, and X. Shen, “Optimal power allocation and user scheduling in multicell networks: Base station cooperation using a game-theoretic approach,” *IEEE Trans. Wireless Commun.*, in press, 2014. DOI: 10.1109/TWC.2014.2334673.

[32] J. Zheng, Y. Cai, and D. Wu, “Subcarrier allocation based on correlated equilibrium in multi-cell OFDMA systems,” *EURASIP J. Wireless Comm. and Netw.*, vol. 2012, pp. 233-245, 2012.

[33] Y. Cai, J. Zheng, Y. Wei, Y. Xu, and A. Anpalagan, “A joint game-theoretic interference coordination approach in uplink multi-cell OFDMA networks,” *Wireless Pers. Commun.*, Oct. 2014. DOI: 10.1007/s11277-014-2081-8.

[34] Y. Xu, Q. Wu, J. Wang, L. Shen, and A. Anpalgan, “Opportunistic spectrum access using partially overlapping channels: Graphical game and uncoupled learning,” *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3906-3918, Sep. 2013.

[35] A.-G. Serrano and L. Giupponi, “Distributed Q-Learning for aggregated interference control in cognitive radio networks,” *IEEE Trans. Veh. Tech.*, vol. 59, no. 4, pp.1823-1834, May 2010.

[36] H. Li, ”Multi-agent Q-learning for Aloha-like spectrum access in cognitive radio systems,” *EURASIP Journal on Wireless Communications and Networking,* vol. 2010, pp. 1-15.

[37] P. Venkatraman, B. Hamdaoui, and M. Guizani, “Opportunistic bandwidth sharing through reinforcement learning”, *IEEE Trans. Veh. Tech.*, pp. 3148-3153, July 2010.

[38] L. Rose, S. Perlaza, C. Martret, and M. Debbah, “Self-organization in decentralized networks: A trial and error learning approach,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 268-279, Jan. 2014.

[39] Y. Shoham, R. Powers, and T. Grenager, “If multi-agent learning is the answer, what is the question?” *Artificial Intelligence*, pp. 365-377, 2007.

[40] P. Stone, “Multiagent learning is not the answer. it is the question,” *Artificial Intelligence*, pp. 402-405, 2007.

[41] R. Sharma and M. Gopal, “Synergizing reinforcement learning and game theory-A new direction for control,” *Applied Soft Computing*, no. 10, pp. 675-688, 2010.

[42] Y. Xing and R. Chandramouli, “Stochastic learning solution for distributed discrete power control game in wireless data networks,” *IEEE/ACM Trans. Netw.*, vol. 16, no. 4, pp. 932-944, Aug. 2008.

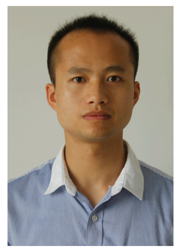
[43] P. Zhou, Y. Chang, and J. A. Copeland, “Reinforcement learning for repeated power control game in cognitive radio networks,” *IEEE/ACM Trans. Netw.*, vol. 30, no. 1, pp. 54-69, Jan. 2012.

[44] L. Cao and H. Zheng, “Distributed rule-regulated spectrum sharing,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 130-145, Jan. 2008.

[45] B. Babadi and V. Tarokh, “GADIA: A greedy asynchronous distributed interference avoidance algorithm,” *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6228–6252, Dec. 2010.

[46] N. Bambos, “Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects,” *IEEE Personal Commun.*, vol. 5, no. 3, pp. 50-59, Jun. 1998.

[47] Z. Feng and Y. Yang, “How much improvement can we get from partially overlapped channels?” in *Proc. IEEE WCNC*, pp. 2957-2962, 2008.



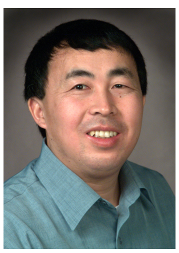
Jianchao Zheng (S’12) received the B.S. degree in electronic engineering from College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2010. He is currently pursuing the Ph.D. degree in communications and information system in College of Communications Engineering, PLA University of Science and Technology. His research interests focus on Interference mitigation techniques, learning theory, game theory, and optimization techniques.

Yueming Cai (M’05-SM’12) received the B.S. degree in Physics from Xiamen University, Xiamen, China in 1982, the M.S. degree in Micro-electronics Engineering and the Ph.D. degree in Communications and Information Systems both from Southeast University, Nanjing, China in 1988 and 1996 respectively. His current research interest includes cooperative communications, signal processing in communications, wireless sensor networks, and physical layer security.

Ning Lu (S’12) received the B.Sc. and M.Sc. degrees from Tongji University, Shanghai, China, in 2007 and 2010, respectively, both in electrical engineering. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His current research interests include capacity and delay analysis, real-time scheduling, and cross-layer design for vehicular networks. Mr. Lu served as a Technical Program Committee Member for IEEE PIMRC’12, WCSP’13, WCSP’14, and ICNC’15.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2014.2366559, IEEE Transactions on Vehicular Technology

Yuhua Xu (S'08) received his B.S. degree in Communications Engineering, and Ph.D. degree in Communications and Information Systems from College of Communications Engineering, PLA University of Science and Technology, Nanjing, China, in 2006 and 2014 respectively. He is currently an assistant professor in College of Communications Engineering, PLA University of Science and Technology. His research interests focus on opportunistic spectrum access, learning theory, game theory, and distributed optimization techniques for wireless communications. He has published several papers in international conferences and reputed journals in his research area. In 2011 and 2012, he was awarded Certificate of Appreciation as Exemplary Reviewer for the IEEE Communications Letters.



Xuemin (Sherman) Shen (M'97-SM'02-F'09) received the B.Sc.(1982) degree from Dalian Maritime University (China) and the M.Sc. (1987) and Ph.D. degrees (1990) from Rutgers University, New Jersey (USA), all in electrical engineering. He is a Professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo, Canada. He was the Associate Chair for Graduate Studies from 2004 to 2008. Dr. Shen's research focuses on resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He is a co-author/editor of six books, and has published more than 600 papers and book chapters in wireless communications and networks, control and filtering. Dr. Shen is an elected member of IEEE ComSoc Board of Governor, and the Chair of Distinguished Lecturers Selection Committee. Dr. Shen served as the Technical Program Committee Chair/Co-Chair for IEEE Infocom'14, IEEE VTC'10 Fall, the Symposia Chair for IEEE ICC'10, the Tutorial Chair for IEEE VTC'11 Spring and IEEE ICC'08, the Technical Program Committee Chair for IEEE Globecom'07, the General Co-Chair for Chinacom'07 and QShine'06, the Chair for IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as the Editor-in-Chief for IEEE Network, Peer-to-Peer Networking and Application, and IET Communications; a Founding Area Editor for IEEE Transactions on Wireless Communications; an Associate Editor for IEEE Transactions on Vehicular Technology, Computer Networks, and ACM/Wireless Networks, etc.; and the Guest Editor for IEEE JSAC, IEEE Wireless Communications, IEEE Communications Magazine, and ACM Mobile Networks and Applications, etc. Dr. Shen received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007 and 2010 from the University of Waterloo, the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. Dr. Shen is a registered Professional Engineer of Ontario, Canada, an IEEE Fellow, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, and a Distinguished Lecturer of IEEE Vehicular Technology Society and Communications Society.