Database-Assisted Spectrum Access in Dynamic Networks: A Distributed Learning Solution

YUHUA XU[1], (Member, IEEE), YITAO XU[1], AND ALAGAN ANPALAGAN[2], (Senior Member, IEEE)

[1]College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China
[2]Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada

Corresponding author: Y. Xu (yuhuaenator@gmail.com)

ABSTRACT This paper investigates the problem of database-assisted spectrum access in dynamic TV white spectrum networks, in which the active user set is varying. Since there is no central controller and information exchange, it encounters dynamic and incomplete information constraints. To solve this challenge, we formulate a state-based spectrum access game and a robust spectrum access game. It is proved that the two games are ordinal potential games with the (expected) aggregate weighted interference serving as the potential functions. A distributed learning algorithm is proposed to achieve the pure strategy Nash equilibrium (NE) of the games. It is shown that the best NE is almost the same with the optimal solution and the achievable throughput of the proposed learning algorithm is very close to the optimal one, which validates the effectiveness of the proposed game-theoretic solution.

INDEX TERMS TV white spectrum, geo-location database, ordinal potential game, learning automata.

I. INTRODUCTION

Employing TV White Spectrum (TVWS) [1]–[3] has been regarded as a promising approach to solve the spectrum shortage problem in future wireless networks, as it can effectively improve the spectrum efficiency by allowing the unlicensed users dynamically to access the idle TV channels. For the TVWS, it has been shown that obtaining the spectrum availability information by inquiring a geo-location database is more efficient than performing spectrum sensing alone [4], [5]. Also, geo-location database has been widely supported by the standards bodies and industrial organizations [6]–[9].

Currently, researchers in this field mainly focused on: i) constructing and maintaining the geo-location database [10]–[12], and ii) developing business models to analyze the revenues of the TV spectrum holders and the unlicensed users [13]–[15]. Since there is no centralized controller available, how to choose a channel for transmission in a distributed manner, aiming to avoid mutual interference among the users, remains a key challenge. However, only a few preliminary results were reported recently [16], [17], and hence it is urgent and important to study efficient database-assisted distributed spectrum access strategies.

In this paper, we consider a dynamic and distributed TVWS network. Specifically, considering practical applications of the users, they do not access the channels when there is no data to transmit. To capture this dynamics, it is assumed

that each user becomes active/inactive according to an active probability in each decision period. As a result, the active user set is varying. Furthermore, since there is no information exchange, a user does not know the chosen channels, the current state (active or inactive) and the active probabilities of other users. The *dynamic* and *incomplete* information constraints make the task of developing efficient distributed spectrum access strategies challenging.

To solve this problem, we resort to game models [18] and distributed learning technology. Specifically, we formulate a state-based spectrum access game and a robust spectrum access game, and propose a distributed learning to achieve desirable solutions. The main contributions can be summarized as follows:

1) For an arbitrary active user set, we formulate the problem of distributed spectrum access as a state-based non-cooperative game. It is proved that state-based game is an ordinal potential game with the aggregate weighted interference serving as the potential function. To address the challenges caused by the varying active user set, we formulate a robust spectrum access game, which is also proved to be a ordinal potential game. Finally, we propose a distributed learning algorithm to achieve the pure strategy Nash equilibria of the formulated games.

2) It is shown that the best Nash equilibrium is almost the same with optimal solution, which validates the

effectiveness of the formulated game models. In addition, the achievable throughput of the proposed learning algorithm is very close to the optimal one, which also validates the distributed learning algorithm in dynamic networks.

The most related work is [16], in which game-theoretic database-assisted spectrum sharing strategies were investigated. This work is differentiated in: i) all users are assumed to be always active [16], while we consider a network with varying number of active users, and iii) the spectrum access algorithms in [16] need information exchange, while the proposed learning-based spectrum access algorithm in this work is fully distributed. Also, the problem of distributed spectrum access for minimizing the aggregate weighted interference was studied in [19] and in our previous work [20], [21]. The differences in the current work are that we optimize the throughput directly and the active user set in each decision period is randomly changing. Also, some simulation results on the problem of distributed spectrum access in TVWS with changing active user set were preliminarily reported in our recent work [22].

The rest of this paper is organized as follows. In Section II, the system model and problem formulation are presented. In Section III, we formulate the state-based spectrum access game and the robust spectrum game respectively, analyze their properties, and propose a distributed learning-based spectrum access algorithm. Finally, simulation results and discussion are presented in Section IV and conclusion is drawn in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION
A. SYSTEM MODEL
We consider a distributed network with N cognitive users and M channels. Note that here each cognitive user corresponds to a communication link consisting of a transmitter and a receiver, e.g., the cognitive access point (AP) and its serving clients. Each cognitive user inquires the spectrum availability from the geo-location spectrum database, which specifies the available channel set $\mathcal{A}_n$ and the maximum allowable transmission power P_n for each user n. With the inquired information on the available channels and maximum transmission power, the users take suitable spectrum access strategies through learning. An illustration of the database-assisted spectrum access is shown in Fig. 1.

To address the user traffic in practical applications, we consider a network with a varying number of active users. Specifically, it is assumed that each user performs channel access in each slot with probability λ_n, $0 < \lambda_n \leq 1$. Such a model captures general kinds of dynamics in wireless networks [23], e.g., a user becomes active only when it has data to transmit and inactive otherwise; also, it can be regarded as a high-level abstraction of the user traffic, i.e., the active probability corresponds to the probability of non-empty buffer.

B. PROBLEM FORMULATION
To capture the changing number of active users, we define the system state as $\mathbf{S} = \{s_1, \ldots, s_N\}$, where $s_n = 1$ indicates that user n is active while $s_n = 0$ means it is inactive. The system state probability is then given by $\mu(s_1, \ldots, s_N) = \prod_{n=1}^{N} p_n$, where p_n is determined as follows:

$$p_n = \begin{cases} \lambda_n, & s_n = 1 \\ 1 - \lambda_n, & s_n = 0 \end{cases} \tag{1}$$

Denote an arbitrary and non-empty active user set as $\mathcal{B}$, i.e., $\mathcal{B} = \{n \in \mathcal{N} : s_n = 1\}$. For presentation, denote the set of all the active user sets as Γ. Then, the probability of an active user set can denoted by $\mu(\mathcal{B})$. We have $\sum_{\mathcal{B} \in \Gamma} \mu(\mathcal{B}) = 1 - \mu(\mathcal{B}_0)$,
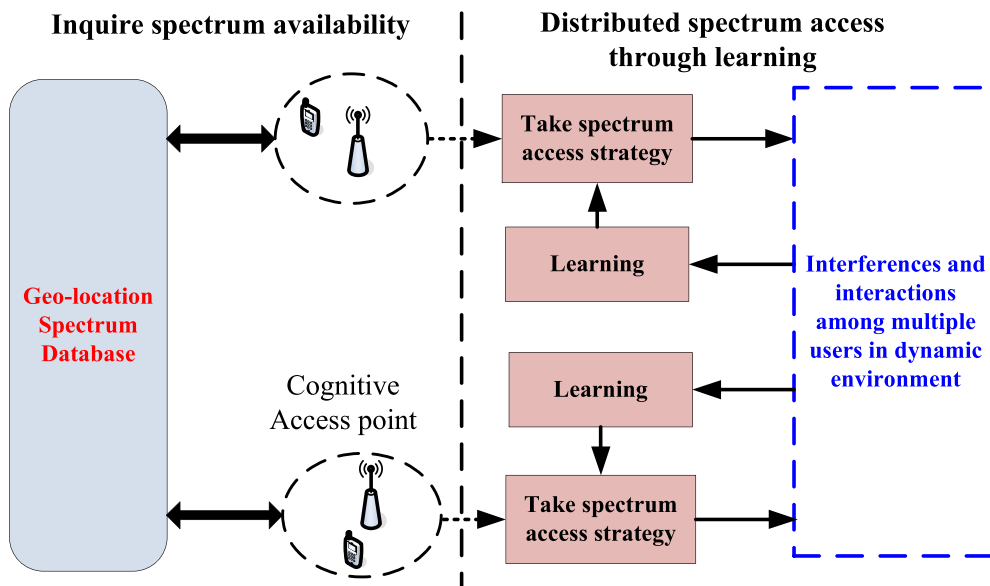

FIGURE 1. The illustrative diagram of database-assisted spectrum access.

where $\mu(\mathcal{B}_0) = \prod_{n=1}^{N}(1 - \lambda_n)$ is the probability that all the users are inactive.

With the spectrum availability information obtained from the database, user n chooses a channel $a_n \in \mathcal{A}_n$ for data transmission. For any active user set $\mathcal{B}$ and channel selection profile (a_n, a_{-n}), the received Signal-to-Interference-plus-Noise Ratio (SINR) of an active user n is determined by:

$$\eta_n(\mathcal{B}, a_n, a_{-n}) = \frac{P_n d_n^{-\alpha}}{\sum_{i \in \mathcal{B}\setminus\{n\}:a_i=a_n} P_i d_{in}^{-\alpha} + \sigma}, \qquad (2)$$

where α is the path loss factor, d_n is the distance between user n and its dedicated receiver, d_{in} is the distance between user i and n, $\sum_{i \in \mathcal{B}\setminus\{n\}:a_i=a_n} P_i d_{in}^{-\alpha} + \sigma$ is the aggregate interference from all other active users also choosing channel a_n, and σ is the background noise. Then, the achievable throughput of user n is given by:

$$R_n(\mathcal{B}, a_n, a_{-n}) = B \log\left(1 + \eta_n(\mathcal{B}, a_n, a_{-n})\right), \qquad (3)$$

where B is the channel bandwidth. Therefore, there are two possible optimization goals for each user, i.e.,

$$\mathbf{P1}: \quad \max_{a_n} \ R_n(\mathcal{B}, a_n, a_{-n}), \quad \forall \mathcal{B} \in \Gamma, \qquad (4)$$

or

$$\mathbf{P2}: \quad \max_{a_n} \ \mathrm{E}_{\mathcal{B}}[R_n(\mathcal{B}, a_n, a_{-n})] = \sum_{\mathcal{B} \in \Gamma} \mu(\mathcal{B}) R_n(\mathcal{B}, a_n, a_{-n}) \qquad (5)$$

Since the network is always changing, it is not feasible to optimize the achievable throughput for each active user set. Thus, we pay attention to solving **P2**. However, the task of solving **P2** is challenging due to the following imperfect information constraints:

- *Dynamic:* The active user set in the system is always changing; in particular, it may change randomly in each decision period.
- *Incomplete:* There is no information exchange among the users, which leads to: i) a user does not know the active probabilities and chosen channels of other users, and ii) the state distribution $\mu(\mathcal{B})$ is unknown to the all users.

Based on the above analysis, it is seen that centralized approaches are not available and we need to develop a distributed and learning-based approach for solving problem **P2**.

III. SPECTRUM ACCESS GAMES AND DISTRIBUTED LEARNING ALGORITHM

Since no centralized controller is available in the considered distributed network and all the users take their spectrum access strategies distributively and autonomously, we formulate this problem as a non-cooperative game. In the following, we present the formulated game models, analyze its properties, and propose a distributed learning algorithm to converge to stable solutions in dynamic environment.

A. GAME FORMULATION AND ANALYSIS

In this subsection, we first present a state-based spectrum access game, in which an inherent system state specifies the active user set. Based on the state-based game, we then present a robust game, in which the expectations over all possible system states are considered. Note that the state-based game corresponds to problem **P1** while the robust game corresponds to problem **P2**.

1) STATE-BASED SPECTRUM ACCESS GAME

Formally, the state-based spectrum access game model is denoted as $\mathcal{F}_1 = [\mathcal{N}, \mathcal{B}, \{\mathcal{A}_n\}_{n\in\mathcal{N}}, \{u_n(\mathcal{B})\}_{n\in\mathcal{N}}]$, where $\mathcal{N} = \{1, \ldots, N\}$ is a set of players (users), $\mathcal{B}$ is active user set, $\mathcal{A}_n$ is a set of available actions (channels) for user n, and $u_n(\mathcal{B})$ is the utility function of player n. The utility function is defined as the available transmission, i.e.,

$$u_n(\mathcal{B}, a_n, a_{-n}) = R_n(\mathcal{B}, a_n, a_{-n}), \quad \forall n \in \mathcal{N}, \quad \forall \mathcal{B} \in \Gamma \qquad (6)$$

Each user in the game intends to maximize its individual utility, which means that the state-based spectrum access game can be expressed as:

$$(\mathcal{F}_1): \quad \max_{a_n \in \mathcal{A}_n} u_n(\mathcal{B}, a_n, a_{-n}), \quad \forall n \in \mathcal{B}. \qquad (7)$$

In order to investigate the properties of $\mathcal{F}_1$, we first present the following definitions, which is directly drawn from [24].

Definition 1 (Nash Equilibrium): For an arbitrary active user set $\mathcal{B}$, a spectrum access profile $a^* = (a_1^*, \ldots, a_{|\mathcal{B}|}^*)$ is a pure strategy NE if and only if no player can improve its utility by deviating unilaterally, i.e.,

$$u_n(\mathcal{B}, a_n^*, a_{-n}^*) \geq u_n(\mathcal{B}, a_n, a_{-n}^*), \quad \forall n \in \mathcal{B}, \quad \forall a_n \in \mathcal{A}_n,$$
$$a_n \neq a_n^* \qquad (8)$$

Definition 2 (Ordinal Potential Game): A game is an ordinal potential game (OPG) if there exists an ordinal potential function $\phi : \mathcal{A}_1 \times \cdots \times \mathcal{A}_N \to R$ such that for all $n \in \mathcal{N}$, all $a_n \in \mathcal{A}_n$, and $a_n' \in \mathcal{A}_n$, the following holds:

$$u_n(\mathcal{B}, a_n, a_{-n}) - u_n(\mathcal{B}, a_n', a_{-n}) > 0$$
$$\Leftrightarrow \phi(\mathcal{B}, a_n, a_{-n}) - \phi(\mathcal{B}, a_n', a_{-n}) > 0 \qquad (9)$$

That is, the change in the utility function caused by the unilateral action change of an arbitrary each user has the same trend with that in the ordinal potential function. Following the similar methodology presented in [16], we have the following theorem.

Theorem 1: For any active user set $\mathcal{B}$, the state-based spectrum access game $\mathcal{F}_1$ is an ordinal potential game.

Proof: For presentation, for any active user set $\mathcal{B}$ and an arbitrary user $\forall n \in \mathcal{B}$, denote

$$v_n(\mathcal{B}, a_n, a_{-n}) = -\sum_{i \in \mathcal{B}\setminus\{n\}:a_i=a_n} P_i P_n d_{in}^{-\alpha}, \qquad (10)$$

which can be regarded as the weighted interference [19] experienced by user n. Define $\phi : \mathcal{A}_1 \times \cdots \times \mathcal{A}_{|\mathcal{B}|} \to R$

as

$$\phi(\mathcal{B}, a_n, a_{-n}) = \sum_{n\in\mathcal{B}} v_n(\mathcal{B}, a_n, a_{-n})$$
$$= -\sum_{n\in\mathcal{B}} \sum_{i\in\mathcal{B}\setminus\{n\}: a_i = a_n} P_i P_n d_{in}^{-\alpha}, \quad (11)$$

which is the aggregate weighted interference experienced by all the active users.

If an arbitrary player n unilaterally changes its channel selection from a_n to a_n^*, then the change in v_n caused by this unilateral change is as follows:

$$v_n(\mathcal{B}, a_n^*, a_{-n}) - v_n(\mathcal{B}, a_n, a_{-n})$$
$$= \sum_{i\in\mathcal{B}\setminus\{n\}: a_i = a_n} P_i P_n d_{in}^{-\alpha} - \sum_{i\in\mathcal{B}\setminus\{n\}: a_i = a_n^*} P_i P_n d_{in}^{-\alpha} \quad (12)$$

For analysis convenience, denote the users choosing the same channel with user n as $\mathcal{I}_n(a_n, \mathcal{B}) = \{i \in \mathcal{B}\setminus n : a_i = a_n\}$. Then, the change in potential function ϕ caused by the unilateral change of user n can be expressed as follows:

$$\phi(\mathcal{B}, a_n^*, a_{-n}) - \phi(\mathcal{B}, a_n, a_{-n})$$
$$= v_n(\mathcal{B}, a_n^*, a_{-n}) - v_n(\mathcal{B}, a_n, a_{-n})$$
$$+ \sum_{i\in\mathcal{I}_n(a_n^*, \mathcal{B})} P_n P_i d_{ni}^{-\alpha} - \sum_{i\in\mathcal{I}_n(a_n, \mathcal{B})} P_n P_i d_{ni}^{-\alpha}$$
$$= 2\Big(v_n(\mathcal{B}, a_n^*, a_{-n}) - v_n(\mathcal{B}, a_n, a_{-n})\Big) \quad (13)$$

Since u_n and v_n is related by:

$$u_n(\mathcal{B}, a_n, a_{-n}) = B\log\Big(1 + \frac{P_n d_n^{-\alpha}}{-v_n(\mathcal{B}, a_n, a_{-n})/P_n + \sigma}\Big),$$
$$(14)$$

and $\log\Big(1 + \frac{P_n d_n^{-\alpha}}{-x/P_n + \sigma}\Big)$ is increasing with respect to x, it follows that:

$$\Big(u_n(\mathcal{B}, a_n^*, a_{-n}) - u_n(\mathcal{B}, a_n, a_{-n})\Big)$$
$$\times \Big(v_n(\mathcal{B}, a_n^*, a_{-n}) - v_n(\mathcal{B}, a_n, a_{-n})\Big) \geq 0, \quad \forall a_n, a_n^* \in \mathcal{A}_n$$
$$(15)$$

For any active user set $\mathcal{B}$, combining (13) and (15) yields the following:

$$\Big(u_n(\mathcal{B}, a_n^*, a_{-n}) - u_n(\mathcal{B}, a_n, a_{-n})\Big)$$
$$\times \Big(\phi(\mathcal{B}, a_n^*, a_{-n}) - \phi(\mathcal{B}, a_n, a_{-n})\Big) \geq 0, \quad \forall a_n, a_n^* \in \mathcal{A}_n$$
$$(16)$$

which satisfies the definition of OPG, as characterized by (9). Thus, the state-based spectrum access game $\mathcal{F}_1$ is an ordinal potential game with ϕ serving as the potential function, which proves Theorem 1. ∎

2) ROBUST SPECTRUM ACCESS GAME

As discussed above, it is not feasible to perform optimization for each active user set since the network is always changing. Thus, based on the state-based spectrum access game $\mathcal{F}_1$, we formulate a robust spectrum game below. Specifically, the robust spectrum access game is denoted as $\mathcal{F}_2 = [\mathcal{N}, \{\mathcal{A}_n\}_{n\in\mathcal{N}}, \{\omega_n\}_{n\in\mathcal{N}}]$, where $\mathcal{N} = \{1, \ldots, N\}$ is a set of players (users), $\mathcal{A}_n$ is a set of available actions (channels) for user n, and ω_n is the utility function of player n. The utility function in robust spectrum access game is defined as the expected available transmission rate over all possible active user sets [22], [23], [25], i.e.,

$$\omega_n(a_n, a_{-n}) = \mathrm{E}_{\mathcal{B}}[u_n(\mathcal{B}, a_n, a_{-n})]$$
$$= \sum_{\mathcal{B}\in\Gamma} \mu(\mathcal{B}) u_n(\mathcal{B}, a_n, a_{-n}) \quad (17)$$

Similarly, the robust spectrum access game can be expressed as:

$$(\mathcal{F}_2): \quad \max_{a_n\in A_n} \omega_n(a_n, a_{-n}), \quad \forall n \in \mathcal{N}. \quad (18)$$

Theorem 2: The robust spectrum access game $\mathcal{F}_2$ is also an ordinal potential game.

Proof: We define the potential function as:

$$\Phi(a_n, a_{-n}) = \mathrm{E}_{\mathcal{B}}\big[\phi(\mathcal{B}, a_n, a_{-n})\big], \quad (19)$$

where ϕ is specified by (11).

If an arbitrary player n unilaterally changes its channel selection from a_n to a_n^*, then the change in w_n caused by this unilateral change is as follows:

$$\omega_n(a_n^*, a_{-n}) - \omega_n(a_n, a_{-n})$$
$$= \mathrm{E}_{\mathcal{B}}[u_n(\mathcal{B}, a_n^*, a_{-n}) - u_n(\mathcal{B}, a_n, a_{-n})] \quad (20)$$

Similarly, the challenge in the Φ is determined by:

$$\Phi(a_n^*, a_{-n}) - \Phi(a_n, a_{-n})$$
$$= \mathrm{E}_{\mathcal{B}}[\phi(\mathcal{B}, a_n^*, a_{-n}) - \phi(\mathcal{B}, a_n, a_{-n})] \quad (21)$$

Using the result obtained in (16), the following always holds:

$$\Big(\omega_n(a_n^*, a_{-n}) - \omega_n(a_n, a_{-n})\Big)$$
$$\times \Big(\Phi(a_n^*, a_{-n}) - \Phi(a_n, a_{-n})\Big) \geq 0, \quad \forall a_n, a_n^* \in \mathcal{A}_n$$
$$(22)$$

Thus, it is proved that the robust spectrum access game $\mathcal{F}_2$ is also an ordinal potential game with $\Phi(a_n, a_{-n})$ serving as the potential function. ∎

3) DISCUSSION ON THE GAME MODELS

Ordinal potential game (OPG) admits the following two promising features [24]: (i) every OPG has at least one pure strategy Nash equilibrium, and (ii) an action profile that maximizes the ordinal potential function is also a pure strategy Nash equilibrium. Some further discussions on the two spectrum game models are listed below:

- Both the state-based and robust spectrum access games have at least one pure strategy NE.

- For the state-based spectrum access game, the aggregate weighted interference serves as the potential function, as specified by (11). For the robust spectrum access game, the expected aggregate weighted interference serve as the potential game, as specified by (19). Thus, it is known all the NEs of the games minimize the (expected) aggregate weighted interference respectively. Furthermore, it has been shown that minimizing lower weighted interference leads to higher throughput [19]–[21]. Thus, it can be expected that the two games would also achieve high throughput.

B. DISTRIBUTED LEARNING FOR ACHIEVING NASH EQUILIBRIA

As the expected aggregate interference serves as the potential function for the robust spectrum access game, it is desirable to develop distributed algorithms to achieve the Nash equilibria. Conventional algorithms in the game community, e.g., best response dynamic [24], fictitious play [26] and spatial adaptive play [27], can not be applied since they need information exchange among the players. To eliminate the requirement of information exchange, some distributed algorithms have been applied in wireless applications, e.g., MAX-logit [28] and Q-learning [29]. However, MAX-logit is only suitable for static environment; although Q-learning can be applied in dynamic networks, its convergence in multiuser environment can not be guaranteed.

In this paper, we propose a learning-based distributed spectrum access algorithm, which is mainly based the stochastic learning automata [30], [31]. To begin with, denote $\mathbf{q}_n(k) = \{q_{n1}(k), \ldots, q_{n|\mathcal{A}_n|}(k)\}$ as the mixed strategy of player n in the kth slot, in which q_{nm} is the probability of choosing channel m. The key ideas of the proposed distributed learning algorithm are: i) the active users choose the channels according to their mixed strategies, and then update their mixed strategies based on the received payoffs, and iii) an inactive user does nothing. Specifically, the learning procedure is as follows:

Initialization: set $k = 1$ and set the initial mixed strategy of each user as $q_{nm}(k) = \frac{1}{|\mathcal{A}_n|}, \forall n \in \mathcal{N}, \forall m \in \mathcal{A}_n$.
Loop for $k = 1, 2, \ldots$
Denote $\mathcal{B}(k)$ as the active user set in the current slot.
1). Channel selection: according to its current mixed strategy $\mathbf{q}_n(k)$, any active user n in $\mathcal{B}(k)$ randomly chooses a channel $a_n(k)$ from its available channel set $\mathcal{A}_n$ in slot k.
2). Channel access and transmit: All the active users transmit on the chosen channels, and they receive the instantaneous transmission throughput $R_n(k)$, which is determined by (3).
3). Update mixed strategy: All the active users $n \in \mathcal{B}(k)$ update their mixed strategies according to the following rules:

$$q_{nm}(k+1) = q_{nm}(k) + br_n(k)(1 - q_{nm}(k)), \quad m = a_n(k)$$
$$q_{nm}(k+1) = q_{nm}(k) - br_n(k)q_{nm}(k), \quad m \neq a_n(k),$$
$$(23)$$

where b is the learning step size, and $r_n(k)$ is the normalized received payoff defined as follows:

$$r_n(k) = \frac{R_n(k)}{R_n^{max}}, \tag{24}$$

where R_n^{max} is the interference-free transmission throughput of user n, i.e., $R_n^{max} = B \log \left(1 + \frac{P_n d_n^{-\alpha}}{\sigma}\right)$.

All the inactive users keep their mixed strategies unchanged, i.e.,

$$\mathbf{q}_n(k+1) = \mathbf{q}_n(k), \quad \forall n \in \mathcal{N} \backslash \mathcal{B}(k) \tag{25}$$

End loop

The rationale behind the update rule (23) is as follows: when a channel is selected and a positive payoff is received, the probability of choosing the channel in the next step increases while the probabilities of choosing other channels decrease accordingly. When all the players adhere to this rule, the system will finally converge to a stable state. Note that the above learning algorithm is fully distributed since an active user only needs its individual action-payoff information. Furthermore, its asymptotical convergence property is characterized by the following theorem.

Theorem 3: When the learning step size goes sufficiently small, i.e., $b \to 0$, the proposed distributed learning algorithm asymptotically converges to a pure strategy NE point of robust spectrum access game $\mathcal{F}_2$.

Proof: It has been rigorously proved that the stochastic learning automata converges to pure strategy Nash equilibria of exact potential games in [32]. In methodology, the differences in this work are summarized as: i) all the users are always active in [32], while they are randomly active or inactive in this work, and ii) for exact potential games, the change in the utility function cased by the unilateral action change of an arbitrary user is the same with that in the potential function, i.e.,

$$u_n(a_n, a_{-n}) - u_n(a_n', a_{-n}) = \phi(a_n, a_{-n}) - \phi(a_n', a_{-n}) \tag{26}$$

When proving the convergence for exact potential games, the following inequality is vital (see [32, eq. (C.40)]):

$$\left(u_n(a_n, a_{-n}) - u_n(a_n', a_{-n})\right)\left(\phi(a_n, a_{-n}) - \phi(a_n', a_{-n})\right) \geq 0 \tag{27}$$

Note that the above inequality also holds for ordinal potential games (See equation (16)). Thus, following similar lines for the proof given in [32] (Theorem 5), and with some additional modifications for processing the user active probability λ_n, as similar lines proposed in [23], this theorem can be proved. However, to avoid unnecessary repetition, the detailed proof is omitted. ∎

IV. SIMULATION RESULTS AND DISCUSSION
The simulation scenario follows the setting given in [16]. The cognitive APs are randomly located in a 500m×500m

square area. There are $M = 5$ channels with bandwidth $B = 6$MHz, the noise power is $\sigma = -100$ dBm, and the path loss factor is $\alpha = 4$. The distance between AP n and its associated boundary user is $d_n = 20$ m. By inquiring the geo-location spectrum database, each AP operates with a specific transmission power P_n and has a different set of available channels. The step size for the learning algorithm is $b = 0.1$.
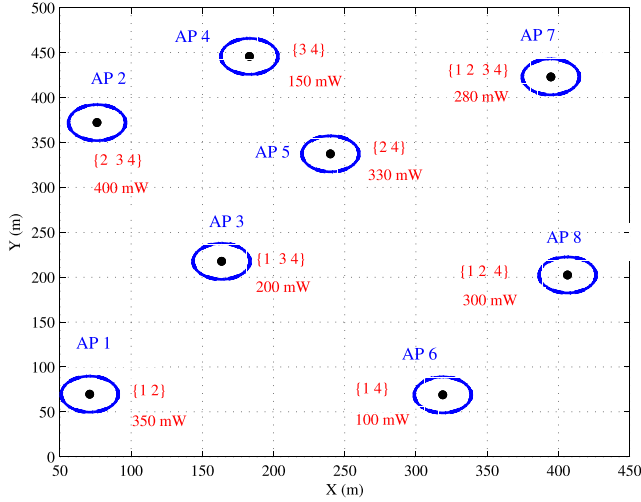

FIGURE 2. A network consisting of eight cognitive APs. By inquiring the geo-location spectrum database, each AP knows its available channel set and the transmitting transmission power, e.g, the available channel set and transmission power of AP 1 are {1, 2} and 350 mW, respectively.

A. CONVERGENCE BEHAVIOR

We consider a specific network which is shown in Fig. 2. Due to the different locations, the action (available channel) sets of the users are different. For example, the action set of users 1, 2, 3 and 4 are $\mathcal{A}_1 = \{1, 2\}$, $\mathcal{A}_2 = \{2, 3, 4\}$, $\mathcal{A}_3 = \{1, 3, 4\}$ and $\mathcal{A}_4 = \{3, 4\}$ respectively. For presentation, it is assumed that all the users have the same active probability $\lambda = 0.8$. The convergence behavior of the proposed distributed spectrum access algorithm is shown in Fig. 3. Specifically, the channel selection probabilities of users 2, 3 and 4 are presented. It is noted that the channel selection probabilities of user 2 converges $\{0, 0, 1\}$ over its available channel set $\{2, 3, 4\}$, which means that it finally chooses channel 4 for data transmission. Similarly, the selection probabilities of users 3 and 4 also finally converge. The results validate the convergence of the proposed distributed learning algorithm.

B. THROUGHPUT PERFORMANCE

For comparison, we consider four approaches: optimal, best NE, worst NE and the distributed learning algorithm. 1) Optimal: problem **P2** is solved directly in a centralized manner. Due to the fact it is a combinatorial optimization problem, we apply the exhaustive search approach to obtain the optimal solution. 2) Best NE and worst NE: by assuming that information exchange among the users is available, the best response algorithm [24] is applied to achieve pure
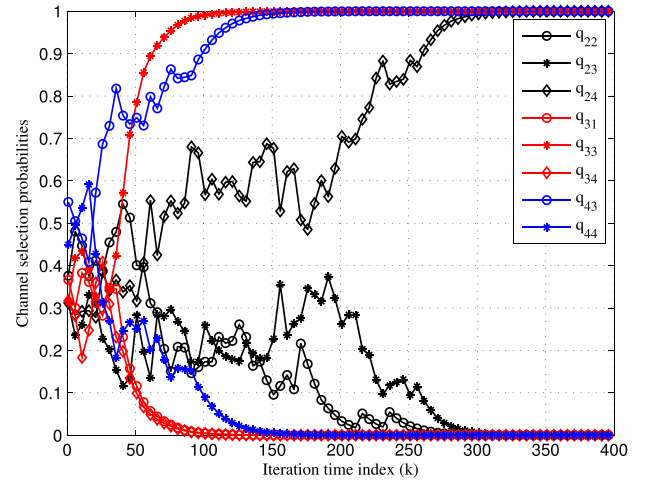

FIGURE 3. The convergence of the proposed distributed learning algorithm.

strategy NE of the robust spectrum access game $\mathcal{F}_2$ in a distributed manner. We carried out 500 independent trails and then take the best one and worst one respectively. Note that the best NE and worst NE serve as the upper and lower bounds of the game. 3) The proposed distributed learning: in the absence of information exchange and centralized controller, all the users adhere to the proposed distributed learning algorithm.
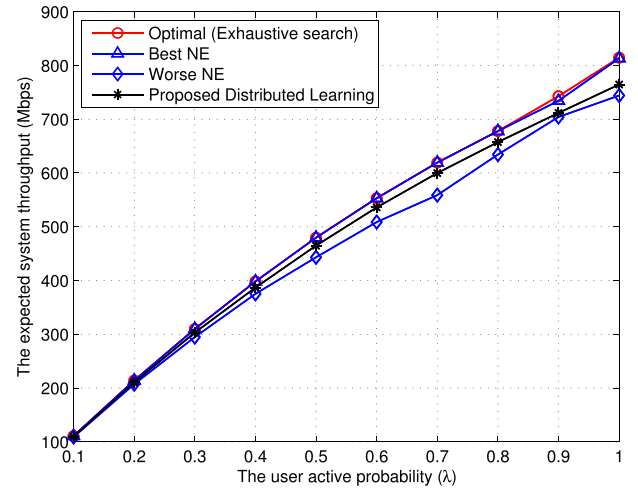

FIGURE 4. The expected throughput versus the active probabilities of the users.

First, we evaluate the throughput performance for the specific network shown in Fig. 2. For presentation, all the users have the same active probability. The expected throughput when varying the user active probability is shown in Fig. 4. The result of the proposed learning is by simulating 200 independent trials and then taking the expected results. Some important results can be observed from the figure: i) the best NE is almost the same with the optimal solution, while the throughput gap between the worse NE and the optimal solution is trivial, which validate the effectiveness of the

proposed robust spectrum game. ii) the proposed distributed learning is very close to the optimal one. Note that some similar numerical results on the throughput comparison were reported in [22].

Secondly, we also consider the specific network shown in Fig. 2, but the active probabilities of the users are different. Specifically, the active probabilities of the users are set to {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8} (tagged as Scenario 1), {0.3, 0.3, 0.3, 0.6, 0.6, 0.9, 0.9, 0.9} (tagged as Scenario 2), {0.3, 0.4, 0.5, 0.5, 0.5, 0.6, 0.7, 0.8} (tagged as Scenario 3), {0.4, 0.4, 0.4, 0.6, 0.6, 0.6, 0.6, 0.6} (tagged as Scenario 4), {0.3, 0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.7} (tagged as Scenario 5), {0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.7} (tagged as Scenario 6), respectively. The throughput comparison results are shown in Fig. 5. It also noted from the figure that for the scenarios with heterogeneous active probabilities, the best NE is almost the same with the optimal solution and the proposed distributed learning is close to the optimal one. These results validate the effectiveness of the formulated spectrum access game as well as the proposed distributed learning algorithm, in both homogeneous and heterogeneous scenarios.
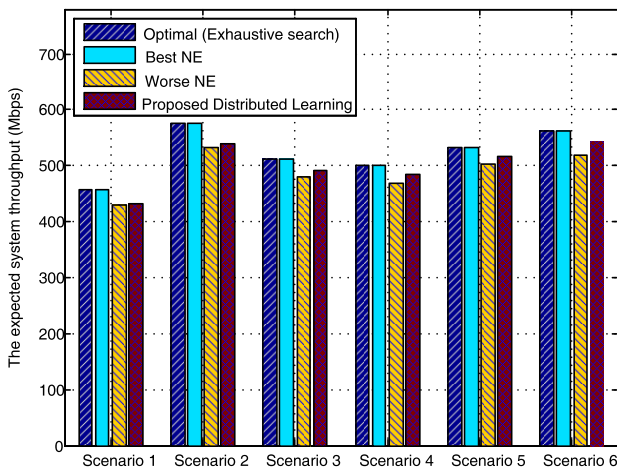

FIGURE 5. The throughput performance comparison for six scenarios with heterogeneous active probabilities.

Thirdly, we evaluate the throughput performance for general networks. Specifically, the cognitive APs are randomly located in the square region. Each channel is independently vacant with probability $\theta = 0.7$ for each AP, and the transmission power of each AP is randomly chosen from the set {100, 200, 250, 300, 350, 280, 400}mW. The throughput comparison when varying the number of cognitive APs is shown in Fig. 6. For each number of APs, e.g., $N = 10$, we simulate 100 independent topologies and take the average result. When the network scales up, exhaustive search is not feasible due to the heavy computational complexity. However, it is believed that the best NE would be very close to the optimal one. It is shown from the figure that the proposed distributed learning is close to the best NE, which again validates the proposed game-theoretic solution. Also, as the network scales up, the normalized throughput
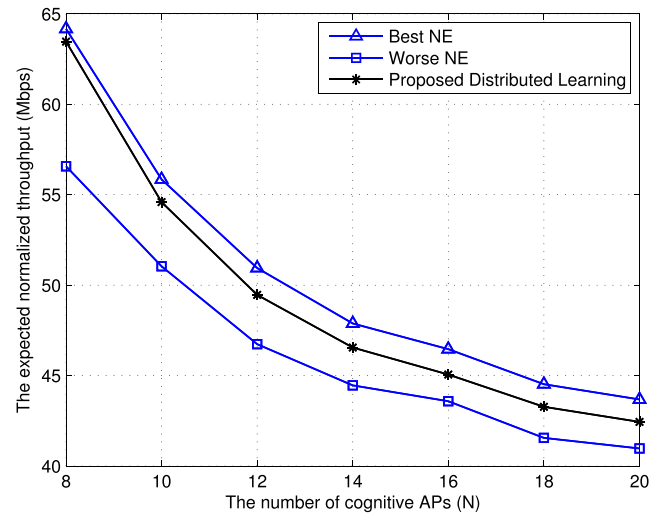

FIGURE 6. The throughput performance comparison for general networks (the active probabilities of all the users are $\lambda = 0.8$).

decreases due to the increase in the mutual interference, as can be expected.

To summarize, the simulation results show that the best NE is almost the same with optimal solution, and the proposed distributed learning algorithm is very close to the optimal one. Recalling the dynamic and incomplete information constraints in the considered system, we believe that the proposed game-theoretic learning solution is desirable for practical applications.

V. CONCLUSION

This paper investigated the problem of database-assisted spectrum access in dynamic networks, in which the active user set is varying. Since there is no central controller, it encounters dynamic and incomplete information constraints. To solve this challenge, we formulated a state-based spectrum access game and a robust spectrum access game. We proved that the two games are ordinal potential games with the (expected) aggregate interference serving as the potential functions, and proposed a distributed learning algorithm without information exchange to achieve the pure strategy Nash equilibrium (NE) of the games. Simulation results show that the best NE is almost the same with the optimal solution and the achievable throughput of the proposed learning algorithm is very close to the optimal one, which validates the effectiveness of the proposed game-theoretic solution. Note that there are still some new challenges and open issues to be studied. For example, users can access more than one channel when equipped with the multiple radio technology. A game-theoretic carrier aggregation in unlicensed spectrum bands is ongoing and will be reported soon.

REFERENCES

[1] *Second Memorandum Opinion and Order*, document FCC 10-174, 2010.
[2] *TV White Spaces—A Consultation on White Space Device Requirements*, Ofcom, London, U.K., Nov. 2012.

[3] *Implementing Geolocation*, Ofcom, London, U.K., Nov. 2010.

[4] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, "SenseLess: A database-driven white spaces network," in *Proc. IEEE Symp. New Frontiers DySPAN*, May 2011, pp. 10–21.

[5] S. Deb, V. Srinivasan, and R. Maheshwari, "Dynamic spectrum access in DTV whitespaces: Design rules, architecture and algorithms," in *Proc. ACM 15th Annu. Int. Conf. MobiCom*, 2009, pp. 1–12.

[6] J. Stine, D. Swain-Walsh, and M. Sherman, "IEEE 1900.5 enabled whitespace database architecture evolution," in *Proc. IEEE Int. Symp. DySPAN*, Apr. 2014, pp. 103–112.

[7] *Google Spectrum Database*. [Online]. Available: https://www.google.com/get/spectrumdatabase/, accessed May 8, 2015.

[8] *WhiteSpace Alliance*. [Online]. [Online]. Available: http://www.whitespacealliance.org, accessed May 8, 2015.

[9] *Microsoft Research WhiteFiService*. [Online]. Available: http://whitespaces.msresearch.us/, accessed May 8, 2015.

[10] J. Ojaniemi, J. Poikonen, and R. Wichman, "Effect of geolocation database update algorithms to the use of TV white spaces," in *Proc. 7th Int. ICST Conf. Cognit. Radio Oriented Wireless Netw. Commun. (CROWNCOM)*, Jun. 2012, pp. 18–23.

[11] H. B. Yilmaz and T. Tugcu, "Location estimation-based radio environment map construction in fading channels," *Wireless Commun. Mobile Comput.*, vol. 15, no. 3, pp. 561–570, 2015.

[12] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chen, and R. Chandra, "Exploring indoor white spaces in metropolises," in *Proc. ACM 19th Annu. Int. Conf. MobiCom*, 2013, pp. 255–266.

[13] X. Feng, Q. Zhang, and J. Zhang, "A hybrid pricing framework for TV white space database," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2626–2635, May 2014.

[14] Y. Luo, L. Gao, and J. Huang, "Price and inventory competition in oligopoly TV white space markets," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 5, pp. 1002–1013, May 2015.

[15] Y. Luo, L. Gao, and J. Huang, "Trade information, not spectrum: A novel TV white space information market model," in *Proc. IEEE 12th Int. Symp. WiOpt*, May 2014, pp. 405–412.

[16] X. Chen and J. Huang, "Database-assisted distributed spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2349–2361, Nov. 2013.

[17] Y. Liu, R. Yu, M. Pan, and Y. Zhang, "Adaptive channel access in spectrum database-driven cognitive radio networks," in *Proc. IEEE ICC*, Jun. 2014, pp. 4933–4938.

[18] R. B. Myerson, *Game Theory: Analysis of Conflict*. Cambridge, MA, USA: Harvard Univ. Press, 1991.

[19] 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.

[20] 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.

[21] J. Wang, Y. Xu, A. Anpalagan, Q. Wu, and Z. Gao, "Optimal distributed interference avoidance: Potential game and learning," *Trans. Emerg. Telecommun. Technol.*, vol. 23, no. 4, pp. 317–326, 2012.

[22] Y. Xu, J. Wang, Q. Wu, Z. Du, L. Shen, and A. Anpalagan, "A game theoretic perspective on self-organizing optimization for cognitive small cells," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 100–108, Jul. 2015.

[23] Y. Xu, Q. Wu, J. Wang, L. Shen, and A. Anpalagan, "Robust multiuser sequential channel sensing and access in dynamic cognitive radio networks: Potential games and stochastic learning," *IEEE Trans. Veh. Technol.*, doi:10.1109/TVT.2014.2356554.

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

[25] H. Tembine, *Distributed Strategic Learning for Wireless Engineers*. Boca Raton, FL, USA: CRC Press, 2012.

[26] J. R. Marden, G. Arslan, and J. S. Shamma, "Joint strategy fictitious play with inertia for potential games," *IEEE Trans. Autom. Control*, vol. 54, no. 2, pp. 208–220, Feb. 2009.

[27] 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.

[28] Y. Song, S. H. Y. Wong, and K.-W. Lee, "Optimal gateway selection in multi-domain wireless networks: A potential game perspective," in *Proc. ACM 17th Annu. Int. Conf. MobiCom*, 2011, pp. 325–336.

[29] H. Li, "Multiagent *Q*-learning for Aloha-like spectrum access in cognitive radio systems," *EURASIP J. Wireless Commun. Netw.*, vol. 2010, pp. 876216-1–876216-15, May 2010.

[30] K. Verbeeck and A. Nowé, "Colonies of learning automata," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 32, no. 6, pp. 772–780, Dec. 2002.

[31] P. S. Sastry, V. 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.

[32] 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.

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, in 2006 and 2014, respectively. He has been with the College of Communications Engineering, PLA University of Science and Technology, since 2012, where he is currently as an Assistant Professor. He has authored several papers in international conferences and reputed journals in his research area. His research interests focus on opportunistic spectrum access, learning theory, game theory, and distributed optimization techniques for wireless communications. He is an Editor of *KSII Transactions on Internet and Information Systems*. In 2011 and 2012, he received the Certificate of Appreciation as an Exemplary Reviewer for the IEEE COMMUNICATIONS LETTERS.

YITAO XU received the B.S. degree in optical communications and the M.S. and Ph.D. degrees in communications and information systems from the Institute of Communications Engineering, Nanjing, China, in 1994, 2000, and 2004, respectively. He is currently a Professor with the PLA University of Science and Technology, China. His current research interests are soft-defined radio and 5G and signal processing for wireless communications.

ALAGAN ANPALAGAN (SM'05) received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical engineering from the University of Toronto, Canada, in 1995, 1997, and 2001, respectively. Since 2001, he has been with Ryerson University, Toronto, Canada, where he co-founded the WINCORE Laboratory in 2002 and leads the Radio Resource Management Group and the Wireless Access and Networking R&D Group. He is currently an Associate Professor and the Program Director of Graduate Studies with the Department of Electrical and Computer Engineering, Ryerson University.

He has authored extensively in international conferences and journals in his research area. His research interests are in general, wireless communication, mobile networks, and system performance analysis, and in particular, QoS-aware radio resource management, joint study of wireless physical/link layer characteristics, cooperative communications, cognitive radios, cross-layer resource optimization, and wireless sensor networking.

Dr. Anpalagan's editorial duties include a Guest Editor of the Special Issues on Radio Resource Management in 3G+ Wireless Systems (2005–2006) of the *Fairness in Radio Resource Management for Wireless Networks*, and an Associate Editor of the *EURASIP Journal of Wireless Communications and Networking*. He served as an IEEE Toronto Section Chair (2006–2007), the Communications Chapter Chair (2004–2005), and the Technical Program Co-Chair of the IEEE Canadian Conference on Electrical and Computer Engineering (2008 and 2004). He is a Registered Professional Engineer in the province of Ontario, Canada.

• • •