Spatial Spectrum Access Game: Nash Equilibria and Distributed Learning

Xu Chen
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong
cx008@ie.cuhk.edu.hk

Jianwei Huang
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, Hong Kong
jwhuang@ie.cuhk.edu.hk

ABSTRACT

A key feature of wireless communications is the spatial reuse. However, the spatial aspect is not yet well understood for the purpose of designing efficient spectrum sharing mechanisms. In this paper, we propose a framework of spatial spectrum access games on directed interference graphs, which can model quite general interference relationship with spatial reuse in wireless networks. We show that a pure strategy equilibrium exists for the two classes of games: (1) any spatial spectrum access games on directed acyclic graphs, and (2) any games satisfying the congestion property on directed trees and directed forests. Under mild technical conditions, the spatial spectrum access games with random backoff and Aloha channel contention mechanisms on undirected graphs also have a pure Nash equilibrium. We then propose a distributed learning algorithm, which only utilizes users' local observations to adaptively adjust the spectrum access strategies. We show that the distributed learning algorithm can converge to an approximate mixed-strategy Nash equilibrium for any spatial spectrum access games. Numerical results demonstrate that the distributed learning algorithm achieves up to 100% performance improvement over a random access algorithm.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless communication

General Terms

Theory, Algorithms

Keywords

Cognitive Radio, Distributed Spectrum Sharing, Nash Equilibrium, Distributed Learning

1. INTRODUCTION

MobiHoc'12, June 11–14, 2012, Hilton Head Island, SC, USA.

Cognitive radio is envisioned as a promising technique to alleviate the problem of spectrum under-utilization [1]. It enables unlicensed wireless users (secondary users) to opportunistically access the licensed channels owned by legacy spectrum holders (primary users), and thus can significantly improve the spectrum efficiency [1].

A key challenge of the cognitive radio technology is how to resolve the resource competition by selfish secondary users in a decentralized fashion. If multiple secondary users transmit over the same channel simultaneously, severe interferences or collisions might occur and the data rates of all users may get reduced. Therefore, it is necessary to design efficient spectrum sharing mechanism for cognitive radio networks.

The competitions among secondary users for common spectrum have often been studied as a noncooperative game theory (e.g., [6, 9, 18, 19, 23]). Nie and Comniciu in [18] designed a self-enforcing distributed spectrum access mechanism based on potential games. Niyato and Hossain in [19] studied a price-based spectrum access mechanism for competitive secondary users. Chen and Huang in [6] investigated stable spectrum sharing mechanism design based on evolutionary game theory. Fĺęlegyhĺćzi *et al.* in [9] proposed a two-tier game framework for medium access control (MAC) mechanism design.

When not knowing spectrum information such as channel availability, secondary users need to learn the network environment and adapt the spectrum access decisions accordingly. Han *et al.* in [11] used no-regret learning to solve this problem, assuming that the users' channel selections are common information. When users' channel selections are not observable, authors in [2, 13] designed multi-agent multi-armed bandit learning algorithms to minimize the expected performance loss of distributed spectrum access.

A common assumption of the above results is that secondary users are close-by and interfere with each other when they transmit on the same channel simultaneously. However, a unique feature of wireless communication is spatial reuse. If users who transmit simultaneously are located sufficiently far away, then simultaneous transmissions over the same channel may not cause any performance degradation to any user (see Figure 1 for an illustration). Such spatial effect on spectrum sharing is less understood than many other aspects in existing literature [24].

Recently, Tekin *et al.* in [22] and Southwell *et al.* in [21] proposed a novel spatial congestion game framework to take spatial relationship into account. The key idea is to extend the classical congestion game upon an *undirected* graph, by assuming that the interferences among the players are sym-

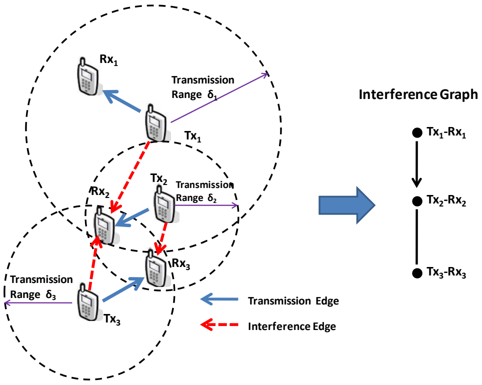

Figure 1: Illustration of distributed spectrum access with spatial reuse under the protocol interference model. Each user n is represented by a transmitter Tx_n and receiver Rx_n pair. Users 2 and 3 can not generate interference to user 1, since user 1's receiver Rx_1 is far from user 2 and 3's transmitters. On the other hand, user 1 can generate interference to user 2, since user 2's receiver Rx_2 is within the transmission range of user 1's transmitter Tx_1. Similarly, user 2 and user 3 can generate interferences to each other.

metric and a player's throughput depends on the number of players in its neighborhood that choose the same resource. As illustrated in Figure 1, however, the interference relationship among the secondary users can be asymmetric due to the heterogeneous transmission powers and locations of the users. We hence propose a more general framework of spatial spectrum access game on *directed* interference graphs, which take users' heterogeneous resource competition capabilities and asymmetric interference relationship into account. The congestion game on directed graphs has also been studied in [4], with the assumption that players have linear and homogeneous payoff functions. The game model in this paper is more generic and allows linear/nonlinear player-specific payoff functions. Moreover, we design a distributed algorithm for achieving the equilibria of the game. The main results and contributions of this paper are as follows:

- *General game formulation*: We formulate the distributed spectrum access problem as a spatial spectrum access game on directed interference graphs, with user-specific channel data rates and channel contention capabilities.
- *Existence of Nash equilibria*: We show by counterexamples that a general spatial spectrum access game may not have a pure Nash equilibrium. We then show that a pure strategy equilibrium exists for the two classes of games: (1) any spatial spectrum access games on directed acyclic graphs, and (2) any games satisfying the congestion property on directed trees and directed forests. We also show that under mild conditions the spatial spectrum access games with random backoff and Aloha channel contention mechanisms on undirected graphs are potential games and have pure Nash equilibria.
- *Distributed learning for achieving approximate Nash equilibrium*: We propose a distributed learning algorithm that can converge to an approximate mixed Nash equilibrium for any spatial spectrum access games by utilizing users' local observations only. Numerical results demonstrate that the distributed learning algorithm achieves up-to 100% performance improvement over the random access algorithm.

The rest of the paper is organized as follows. We introduce the system model and the spatial spectrum access game in Sections 2 and 3, respectively. We investigate the existence of Nash equilibria in Section 4. Then we present the distributed learning algorithm in Section 5. We illustrate the performance of the proposed algorithm through numerical results in Section 6, and finally conclude in Section 7.

2. SYSTEM MODEL

We consider a cognitive radio network with a set $\mathcal{M} = \{1, 2, ..., M\}$ of independent and stochastically heterogeneous primary channels. A set $\mathcal{N} = \{1, 2, ..., N\}$ of secondary users try to access these channels distributively when the channels are not occupied by primary (licensed) transmissions. Here we assume that each secondary user is a dedicated transmitter-receiver pair.

To take users' spatial relationship into account, we denote $\boldsymbol{d}_n = (d_{Tx_n}, d_{Rx_n})$ as the **location vector** of secondary user n, where d_{Tx_n} and d_{Rx_n} denote the location of the transmitter and the receiver, respectively. Each secondary user n has a **transmission range** δ_n. Then given the location vectors of all secondary users, we can obtain the **interference graph** $G = \{\mathcal{N}, \mathcal{E}\}$ to describe the interference relationship among the users (see Figure 1 for an example). Here the vertex set $\mathcal{N}$ is the same as the secondary user set. The edge set is defined as $\mathcal{E} = \{(i,j) : ||d_{Tx_i}, d_{Rx_j}|| \leq \delta_i, \forall i, j \neq i \in \mathcal{N}\}$, where $||d_{Tx_i}, d_{Rx_j}||$ is the distance between the transmitter of user i and the receiver of user j. As illustrated in Figure 1, an interference edge can be directed or undirected. If an interference edge is directed from secondary user i to user j, then user j's data transmission will be affected by user i's transmission on the same channel, but user i will not be affected by user j. If the interference edge is undirected[1] between user i and user j, then the two users can affect each other. Note that a generic directed interference graph can consist of a mixture of directed and undirected edges. In the sequel, we call an interference graph undirected, if and only if all the edges of the graph are undirected. We also denote the set of users that can cause interference to user n as $\mathcal{N}_n = \{i : (i,n) \in \mathcal{E}, i \in \mathcal{N}\}$.

Based on the interference model above, we describe the cognitive radio network with a slotted transmission structure as follows:

- *Channel State*: the channel state for a channel m during time slot t is

$$S_m(t) = \begin{cases} 0, & \text{if channel } m \text{ is occupied} \\ & \text{by primary transmissions,} \\ 1, & \text{if channel } m \text{ is idle.} \end{cases}$$

- *Channel State Transition*: for a channel m, the channel state $S_m(t)$ is a random variable with a probability

[1]Here the edge is actually bi-directed. We follow the conventions in [22] and [21] and ignore the directions on the edge.

density function as ψ_m. In the following, we denote the channel idle probability θ_m as the mean of $S_m(t)$, i.e., $\theta_m = E_{\psi_m}[S_m(t)]$.

- *User Specific Channel Throughput*: for each secondary user n, its realized data rate $b_m^n(t)$ on an idle channel m in each time slot evolves according to a random process with a mean B_m^n, due to users' heterogeneous transmission technologies and the local environmental effects such as fading.

- *Time Slot Structure*: each secondary user n executes the following stages synchronously during each time slot:

  – *Channel Sensing*: sense one of the channels based on the channel selection decision made at the end of previous time slot.

  – *Channel Contention*: Let a_n be the channel selected by user n, and $\boldsymbol{a} = (a_1, ..., a_N)$ be the channel selection profile of all users. The probability that user n can grab the chosen idle channel a_n during a time slot is $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a})) \in (0, 1)$, which depends on the subset of user n's interfering users that choose the same channel $\mathcal{N}_n^{a_n}(\boldsymbol{a}) \triangleq \{i \in \mathcal{N}_n : a_i = a_n\}$. Here are two examples:
  1) *Random backoff mechanism*: the contention stage of a time slot is divided into $\lambda_{\max}$ mini-slots (see Figure 2). Each contending user n first counts down according to a randomly and uniformly generated integer backoff time counter (number of mini-slots) λ_n between 1 and $\lambda_{\max}$. If there is no active transmissions till the count-down timer expires, the user monitors the channel and transmits RTS/CTS messages on that channel. If multiple users choose the same backoff counter, a collision will occur and no users can grab the channel successfully. Once successfully gets the channel, the user starts to transmit its data packet. In this case, we have

$$\begin{aligned} g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a})) &= Pr\{\lambda_n < \min_{i \in \mathcal{N}_n : a_i = a_n} \{\lambda_i\}\} \\ &= \sum_{\lambda=1}^{\lambda_{\max}} Pr\{\lambda_n = \lambda\} \\ &\quad \times Pr\{\lambda_n < \min_{i \in \mathcal{N}_n : a_i = a_n} \{\lambda_i\} | \lambda_n = \lambda\} \\ &= \sum_{\lambda=1}^{\lambda_{\max}} \frac{1}{\lambda_{\max}} \left(\frac{\lambda_{\max} - \lambda}{\lambda_{\max}} \right)^{K_{a_n}^n(\boldsymbol{a})}, \end{aligned} \quad (1)$$

  where $K_{a_n}^n(\boldsymbol{a}) = |\mathcal{N}_n^{a_n}(\boldsymbol{a})| = \sum_{i \in \mathcal{N}_n} I_{\{a_i = a_n\}}$ denotes the number of user n's interfering users choosing the same channel as user n.
  2) *Aloha mechanism*: user n contends for an idle channel with a probability $p_n \in (0, 1)$ in a time slot. If multiple interfering users contend for the same channel, a collision occurs and no user can grab the channel for data transmission. In this case, we have

$$g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a})) = p_n \prod_{i \in \mathcal{N}_n^{a_n}(\boldsymbol{a})} (1 - p_i). \quad (2)$$

  Note that for the random backoff mechanism, the channel grabbing probability $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}))$ is *user homogeneous* since it only depends on the number of contending users $K_{a_n}^n(\boldsymbol{a})$. For the Aloha mechanism, the channel grabbing probability $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}))$ is *user heterogeneous* since it depends on who (instead of how many users) contend the channel.

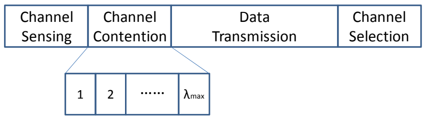

Figure 2: Time slot structure with random backoff mechanism

  – *Data Transmission*: transmit data packets if the user successfully grabs the channel.

  – *Channel Selection*: choose a channel to access during next time slot according to the distributed learning algorithm in Section 5.

Under a fixed channel selection profile $\boldsymbol{a}$, the long-run expected throughput of a secondary user n choosing channel a_n can be computed as

$$U_n(\boldsymbol{a}) = \theta_{a_n} B_{a_n}^n g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a})). \quad (3)$$

Since our analysis is from the secondary users' perspective, we will use the terms "secondary user" and "user" interchangeably. Due to page limit, the detailed proofs are given in our technical report [5].

3. SPATIAL SPECTRUM ACCESS GAME

We now consider the problem that each user tries to maximize its own throughput by choosing a proper channel distributively. Let $a_{-n} = \{a_1, ..., a_{n-1}, a_{n+1}, ..., a_N\}$ be the channels chosen by all other users except user n. Given other users' channel selections a_{-n}, the problem faced by a user n is

$$\max_{a_n \in \mathcal{M}} U(a_n, a_{-n}), \forall n \in \mathcal{N}. \quad (4)$$

The distributed nature of the channel selection problem naturally leads to a formulation based on the game theory, such that users can self organize into a mutually acceptable channel selection (**pure Nash equilibrium**) $\boldsymbol{a}^* = (a_1^*, a_2^*, ..., a_N^*)$ with

$$a_n^* = \arg \max_{a_n \in \mathcal{M}} U(a_n, a_{-n}^*), \forall n \in \mathcal{N}. \quad (5)$$

We thus formulate the distributed channel selection problem on an interference graph G as a **spatial spectrum access game** $\Gamma = (\mathcal{N}, \mathcal{M}, G, \{U_n\}_{n \in \mathcal{N}})$, where $\mathcal{N}$ is the set of players, $\mathcal{M}$ is the set of strategies, G describes the interference relationship among the players, and U_n is the payoff function of player n.

It is known that not every finite strategic game possesses a pure Nash equilibrium [17]. We then introduce a more general concept of mixed Nash equilibrium. Let $\boldsymbol{\sigma}_n \triangleq (\sigma_1^n, ..., \sigma_M^n)$ denote the mixed strategy of user n, where $0 \leq \sigma_m^n \leq 1$ is the probability of user n choosing channel m, and $\sum_{m=1}^M \sigma_m^n = 1$. For simplicity, we use the same payoff notation $U_n(\boldsymbol{\sigma}_1, ..., \boldsymbol{\sigma}_N)$ to denote the expected throughput of user n under the mixed strategy profile $(\boldsymbol{\sigma}_1, ..., \boldsymbol{\sigma}_N)$, and it can be computed as

$$U_n(\boldsymbol{\sigma}_1, ..., \boldsymbol{\sigma}_N) = \sum_{a_1=1}^{M} \sigma^1_{a_1} ... \sum_{a_N=1}^{M} \sigma^N_{a_N} U_n(a_1, ..., a_N). \quad (6)$$

Similarly to the pure Nash equilibrium, the mixed Nash equilibrium is defined as:

DEFINITION 1 (**Mixed Nash Equilibrium** [17]). *The mixed strategy profile* $\boldsymbol{\sigma}^* = (\boldsymbol{\sigma}^*_1, ..., \boldsymbol{\sigma}^*_N)$ *is a mixed Nash equilibrium, if for every user* $n \in \mathcal{N}$*, we have*

$$U_n(\boldsymbol{\sigma}^*_n, \boldsymbol{\sigma}^*_{-n}) \geq U_n(\boldsymbol{\sigma}_n, \boldsymbol{\sigma}^*_{-n}), \forall \boldsymbol{\sigma}_n \neq \boldsymbol{\sigma}^*_n,$$

where $\boldsymbol{\sigma}^*_{-n}$ *denote the mixed strategy choices of all other users except user* n.

Note that the pure Nash equilibrium is a special case of the mixed Nash equilibrium, wherein every user chooses a single channel with probability one. One critical issue in game theory is the existence of both mixed and pure Nash equilibria, which motivates the study in the following Section 4.

4. EXISTENCE OF NASH EQUILIBRIA

In this part, we study the existence of Nash equilibria in a spatial spectrum access game. Since a spatial spectrum access game is a finite strategic game (i.e., with finite number of players and finite number of channels), we know that it always admits a mixed Nash equilibrium according to [17].

On the other hand, not every finite strategic game possesses a pure Nash equilibrium [17]. A pure Nash equilibrium is much preferable than a general mixed strategy Nash equilibrium, as in a pure strategy equilibrium users can achieve mutually acceptable channel selections without randomly picking and switching channels all the time. This motivates us to further investigate the existence of pure Nash Equilibria of the spatial spectrum access games.

4.1 Existence of Pure Nash Equilibria on Directed Interference Graphs

We first study the existence of pure Nash Equilibria on directed interference graphs.

First of all, we can construct a game which does not have a pure Nash equilibrium.

THEOREM 1. *There exists a spatial spectrum access game on a directed interference graph not admitting any pure Nash equilibrium.*

Figure 3 shows such an example. It is easy to verify that for all 8 possible channel selection profiles, there always exists one user (out of these three users) having an incentive to change its channel selection unilaterally to improve its throughput.

We then focus on identifying the conditions under which the game admits a pure Nash equilibrium. To proceed, we first introduce the following lemma.

LEMMA 1. *Consider any spatial spectrum access game on a given directed interference graph G that has a pure Nash equilibrium. Then we can construct a new spatial spectrum access game by adding a new player, who can not generate interference to any player in the original game and may receive interference from one or multiple players in the original game. The new game also has a pure Nash equilibrium.*

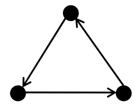

Figure 3: An example of spatial spectrum access game without pure Nash equilibria. There are two channels available and the throughput of a user n is given as $U_n(\boldsymbol{a}) = p\prod_{i\in\mathcal{N}_n^{a_n}(\boldsymbol{a})}(1-p)$. If all three players (nodes) choose channel 1, then each player has the incentive of choosing channel 2 to improve its throughput assuming that the other two players do not change their channel choices. We can show that such derivation will happen for all 8 possible strategy profiles $\boldsymbol{a} = (a_1, a_2, a_3)$, where $a_i \in \{1, 2\}$ for $i \in \{1, 2, 3\}$.

We know that any directed acyclic graph (i.e., a directed graph contains no directed cycles) can be given a topological sort (i.e., an ordering of the nodes), such that if node $i < j$ then there are no edges directed from the node j to node i in the ordering [3]. From Lemma 1, we know that

COROLLARY 1. *Any spatial spectrum access game on a directed acyclic graph has a pure Nash equilibrium.*

To obtain more insightful results, we next impose the following property on the spatial spectrum access games:

DEFINITION 2 (**Congestion Property**). *User* n*'s channel grabbing probability* $g_n\left(\mathcal{N}_n^{a_n}(\boldsymbol{a})\right)$ *satisfies the congestion property if for any* $\tilde{\mathcal{N}}_n^{a_n}(\boldsymbol{a}) \subseteq \mathcal{N}_n^{a_n}(\boldsymbol{a})$*, we have*

$$g_n(\tilde{\mathcal{N}}_n^{a_n}(\boldsymbol{a})) \geq g_n\left(\mathcal{N}_n^{a_n}(\boldsymbol{a})\right). \quad (7)$$

Furthermore, a spatial spectrum access game satisfies the congestion property if (7) holds for all users $n \in \mathcal{N}$.

The congestion property means that the more contending users exist, the less chance a user can grab the channel. Such a property is natural for practical wireless systems such as the random backoff and Aloha systems. We can show that

LEMMA 2. *Consider any spatial spectrum access game satisfying the congestion property on a given directed interference graph G that has a pure Nash equilibrium. Then we can construct a new spatial spectrum access game by adding a new player, whose channel grabbing probability satisfies the congestion property and who can have interference relationship with at most one player* $n \in \mathcal{N}$ *in the original game. The new game also has a pure Nash equilibrium.*

DEFINITION 3 (**Directed Tree** [3]). *A directed graph is called a directed tree if the corresponding undirected graph obtained by ignoring the directions on the edges of the original directed graph is a tree.*

Note that a (undirected) tree is a special case of directed trees. Since any spatial spectrum access game over a single node always has a pure Nash equilibrium, we can then construct the directed tree recursively by introducing a new node and adding an (directed or undirected) edge between this node and one existing node. From Lemma 2, we obtain that

COROLLARY 2. *Any spatial spectrum access game satisfying the congestion property on a directed tree has a pure Nash equilibrium.*

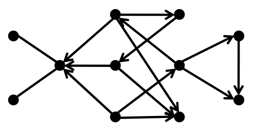

Figure 4: An interference graph that consists of directed acyclic graphs and directed trees

DEFINITION 4 (**Directed Forest** [3]). *A directed graph is called a directed forest if it consists of a disjoint union of directed trees.*

Similarly, we can obtain from Lemma 2 that

COROLLARY 3. *Any spatial spectrum access game satisfying the congestion property on a directed forest has a pure Nash equilibrium.*

As illustrated in Figure 4, according to Lemmas 1 and 2, we can construct more complicated directed interference graphs over which a spatial spectrum access game satisfying the congestion property has a pure Nash equilibrium.

4.2 Existence of Pure Nash Equilibria on Undirected Interference Graphs

We now study the case that the interference graph is undirected. This is a good approximation of reality if the transmitter of each user is close to its receiver, and all users' transmit powers are roughly the same.

When an undirected interference graph is a tree, according to Corollary 2, any spatial spectrum access game satisfying the congestion property has a pure Nash equilibrium. However, for those non-tree undirected graphs without a topological sort, the existence of pure Nash equilibrium can not be proved following the results in previous Section 4.1. This motivates us to further study the existence of pure Nash equilibria on generic undirected interference graphs.

First of all, [15] showed that a 3-players and 3-resources congestion game with user-specific congestion weights may not have a pure Nash equilibrium. Such a congestion game can be considered as a spatial spectrum access game on a complete undirected interference graph (by regarding the resources as channels). When all users have homogeneous channel contention capabilities and all channels have the same mean data rates, [22] showed that the spatial spectrum access game on any undirected interference graphs has a pure Nash equilibrium. Clearly, the applicability of such a channel-homogeneous model is quite limited, since the channel throughputs in practical wireless networks are often heterogeneous. We hence next focus on exploring the random backoff and Aloha systems with user-specific data rates, which provide useful insights for the user-homogeneous and user-heterogeneous channel contention mechanisms, respectively.

Here we resort to a useful tool of potential game[2], which is defined as

DEFINITION 5 (**Potential Game** [16]). *A game is called a potential game if it admits a potential function $\Phi(\boldsymbol{a})$ such that for every $n \in \mathcal{N}$ and $a_{-n} \in \mathcal{M}^{N-1}$,*

$$\text{sgn}\left(\Phi(a_n^{'}, a_{-n}) - \Phi(a_n, a_{-n})\right) = \text{sgn}\left(U_n(a_n^{'}, a_{-n}) - U_n(a_n, a_{-n})\right),$$

where $\text{sgn}(\cdot)$ is the sign function.

[2] Note that it is much more difficult to find a proper potential function to take into account users' asymmetric relationships (i.e., directions of edges on graph) when the interference graph is directed. Hence in this study we only apply the tool of potential game in the undirected case.

DEFINITION 6 (**Better Response Update** [16]). *The event where a player n changes to an action $a_n^{'}$ from the action a_n is a better response update if and only if $U_n(a_n^{'}, a_{-n}) > U_n(a_n, a_{-n})$.*

An appealing property of the potential game is that it always admits a pure Nash equilibrium and the finite improvement property, which is defined as

DEFINITION 7 (**Finite Improvement Property** [16]). *A game has the finite improvement property if any asynchronous better response update process (i.e., no more than one player updates the strategy at any given time) terminates at a pure Nash equilibrium within a finite number of updates.*

Based on the potential game theory, we first study the random backoff mechanism. We show in Theorem 2 that when the undirected interference graph is complete, there exists indeed a pure Nash equilibrium.

THEOREM 2. *Any spatial spectrum access game on a complete undirected interference graph with the random backoff mechanism is a potential game with the potential function*

$$\Phi(\boldsymbol{a}) = \prod_{n=1}^{N} \theta_{a_n} B_{a_n}^n \prod_{m=1}^{M} \prod_{c=1}^{K_m(\boldsymbol{a})} g_n(c), \tag{8}$$

where $K_m(\boldsymbol{a})$ is the number of users choosing channel m under the strategy profile $\boldsymbol{a}$, and hence has a pure Nash equilibrium.

We then consider the random backoff mechanism in the asymptotic case that $\lambda_{\max}$ goes to infinity. This can be a good approximation of reality when the number of backoff mini-slots is much greater than the number of interfering users, and collision rarely occurs. In this case, we have

$$\begin{aligned} g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a})) &= \lim_{\lambda_{\max}\to\infty} \sum_{\lambda=1}^{\lambda_{\max}} \frac{1}{\lambda_{\max}} \left(\frac{\lambda_{\max}-\lambda}{\lambda_{\max}}\right)^{K_{a_n}^n(\boldsymbol{a})} \\ &= \int_0^1 x^{K_{a_n}^n(\boldsymbol{a})} dx = \frac{1}{1+K_{a_n}^n(\boldsymbol{a})}, \end{aligned} \tag{9}$$

here $K_{a_n}^n(\boldsymbol{a})$ denotes the number of users that choose channel a_n and can interfere with user n. Equation (9) implies that the channel opportunity is equally shared among $1+K_{a_n}^n(\boldsymbol{a})$ contending users (including user n). We consider the user specific throughput as

$$U_n(a_n, a_{-n}) = h_n \theta_{a_n} B_{a_n} \frac{1}{1+K_{a_n}^n(\boldsymbol{a})}, \tag{10}$$

where h_n is regarded as a user-specific transmission gain. For example, a user n can possess a higher transmission gain if the distance between its transmitter and receiver is shorter than other users given that all the users transmit with the same power level. We show that

THEOREM 3. *Any spatial spectrum access game on any undirected interference graph with user-specific transmission gains and the random backoff mechanism in the asymptotic case is a potential game with the potential function*

$$\Phi(\boldsymbol{a}) = -\sum_{n=1}^{N} \left(\frac{1 + \frac{1}{2} K_{a_n}^n(\boldsymbol{a})}{\theta_{a_n} B_{a_n}} \right), \quad (11)$$

and hence has a pure Nash equilibrium.

We now consider the Aloha mechanism. According to (2), we have the throughput function as

$$U_n(\boldsymbol{a}) = \theta_{a_n} B_{a_n}^n p_n \prod_{i \in \mathcal{N}_n^{a_n}(\boldsymbol{a})} (1 - p_i). \quad (12)$$

We can show that

THEOREM 4. *Any spatial spectrum access game on any undirected interference graph with the Aloha mechanism is a potential game with the potential function*

$$\Phi(\boldsymbol{a}) = \sum_{i=1}^{N} -\log(1-p_i) \times \left(\frac{1}{2} \sum_{j \in \mathcal{N}_i^{a_i}(\boldsymbol{a})} \log(1-p_j) + \log\left(\theta_{a_i} B_{a_i}^i p_i\right) \right), \quad (13)$$

and hence has a pure Nash equilibrium.

When a spatial spectrum access game is a potential game, we can design a distributed algorithm such that each user asynchronously updates the channel selection myopically to increase its throughput. According to the finite improvement property of the potential game, such an algorithm can achieve a pure Nash equilibrium within finite number of iterations. However, asynchronous better response updates require each user to have the complete information of other users' channel selections. This can only be achieved with extensive information exchange among the users, which may not be always feasible. It will be very nice to design a distributed algorithm that achieves the equilibrium without information exchange.

5. DISTRIBUTED LEARNING FOR SPATIAL SPECTRUM ACCESS

In this part, we discuss how to achieve an equilibrium for the spatial spectrum access games. As shown in Section 4, a generic spatial spectrum access game does not necessarily have a pure Nash equilibrium, and thus it is impossible to design a mechanism achieving pure Nash equilibria in general. We hence target on approaching the mixed Nash equilibria. Govindan and Wilson in [10] proposed a global Newton method to compute the mixed Nash equilibria for any finite strategic games. This method hence can be applied to find the mixed Nash equilibria for the spatial spectrum access games. However, such an approach is a centralized optimization, which requires that each user has the complete information of other users and compute the solution accordingly. This is often infeasible in a cognitive radio network, since acquiring complete information requires heavy information exchange among the users, and setting up and maintaining a common control channel for message broadcasting demands high system overheads [1]. Moreover, this approach is not incentive compatible since some users may not be willing to share their local information due to the energy consumption of information broadcasting. We thus propose a distributed learning algorithm for any spatial spectrum access games, and the algorithm does not require any information exchange among users. Each user only estimates its expected throughput locally, and learns to adjust its channel selection strategy adaptively. We show that the distributed learning algorithm can converge to a mixed Nash equilibrium approximately.

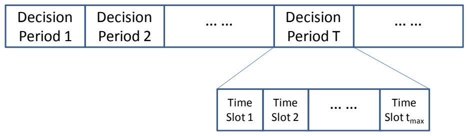

Figure 5: Time structure of a decision period

5.1 Expected Throughput Estimation

We first introduce the estimation of user's expected throughput based on local observations. To achieve an accurate estimation, a user needs to gather a large number of local observation samples. This motivates us to divide the spectrum access time into a sequence of *decision periods* indexed by $T(=1,2,...)$, where each decision period consists of $t_{\max}$ time slots (see Figure 5 as an illustration). During a single decision period, a user accesses the *same* channel in all $t_{\max}$ time slots. Thus the total number of users accessing each channel does not change within a decision period, which allows users to better learn the environment.

Suppose user n chooses channel m to access at decision period T. According to (3), a user's expected throughput during period T depends on the probability of grabbing the channel $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$ on that period, the channel idle probability θ_m, and the mean data rate B_m^n. Similarly to our work in [7] on the expected throughput estimation for imitation-based spectrum sharing mechanism design, we will apply the maximum likelihood estimation (MLE) to get accurate estimations of there parameters for the distributed learning mechanism design, due to MLE's efficiency and ease of implementation.

5.1.1 Maximum Likelihood Estimation

At the beginning of each time slot $t(=1,...,t_{\max})$ of a decision period T, a user n will sense the same channel $a_n(T) = m$. If the channel is idle, the user will compete to grab the channel according to a specified channel contention mechanism. At the end of each time slot t, a user observes $S_m^n(T,t)$, $I_m^n(T,t)$, and $b_m^n(T,t)$. Here $S_m^n(T,t)$ denotes the state of the chosen channel m (i.e., whether occupied by the primary traffic), $I_m^n(T,t)$ indicates whether the user has successfully grabs the channel, i.e.,

$$I_m^n(T,t) = \begin{cases} 1, & \text{if user } n \text{ successfully} \\ & \text{grabs the channel } m, \\ 0, & \text{otherwise,} \end{cases}$$

and $b_m^n(T,t)$ is the received data rate on the chosen channel m by user n at time slot t. At the end of each decision period T, each user n can collect a set of local observations $\Omega_n(T) = \{S_m^n(T,t), I_m^n(T,t), b_m^n(T,t)\}_{t=1}^{t_{\max}}$. Note that if $S_m^n(T,t) = 0$

(i.e., the channel is occupied by the primary traffic), we also set $I_m^n(T,t)$ and $b_m^n(T,t)$ to be 0.

When the channel m is idle (i.e., no primary traffic), user n will grab the channel with probability $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$. Since there are a total of $\sum_{t=1}^{t_{\max}} S_m^n(T,t)$ rounds of channel contentions in the period T and each round is independent and identically distributed (i.i.d.), the total number of successful channel captures $\sum_{t=1}^{t_{\max}} I_m^n(T,t)$ follows the Binomial distribution. A user n can then compute the likelihood of $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$, i.e., the probability of the realized observations $\Omega_n(T)$ given the parameter $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$ as

$$\begin{aligned}&\mathcal{L}[\Omega_n(T)|g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))]\\&=\binom{\sum_{t=1}^{t_{\max}} S_m^n(T,t)}{\sum_{t=1}^{t_{\max}} I_m^n(T,t)} g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))^{\sum_{t=1}^{t_{\max}} I_m^n(T,t)}\\&\quad\times(1-g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T))))^{\sum_{t=1}^{t_{\max}} S_m^n(T,t)-\sum_{t=1}^{t_{\max}} I_m^n(T,t)}.\end{aligned}$$

Then MLE of $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$ can be computed by maximizing the log-likelihood function $\ln\mathcal{L}[\Omega_n(T)|g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))]$, i.e., $\max_{g(k(T))}\ln\mathcal{L}[\Omega_n(T)|g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))]$. By the first order condition, we obtain the optimal solution as $\tilde{g}_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T))) = \frac{\sum_{t=1}^{t_{\max}} I_m^n(T,t)}{\sum_{t=1}^{t_{\max}} S_m^n(T,t)}$, which is the sample averaging estimation. When length of decision period $t_{\max}$ is large, by the central limit theorem, we know that

$$\begin{aligned}&\tilde{g}_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))\\&\sim\mathcal{N}\left(g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T))),\frac{g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))(1-g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T))))}{\sum_{t=1}^{t_{\max}} S_m^n(T,t)}\right),\end{aligned}$$

where $\mathcal{N}(\cdot)$ denotes the normal distribution.

Similarly, we can apply the MLE to estimate the channel idle probability θ_m and the mean channel data rate B_m^n. More specifically, when the channel state $S_m(t)$ and the realized data rate $b_m^n(t)$ are i.i.d. random variables, we can easily obtain the closed-form estimations as $\hat{\theta}_m = \frac{\sum_{t=1}^{t_{\max}} S_m^n(T,t)}{t_{\max}}$ and $\hat{B}_m^n = \frac{\sum_{t=1}^{t_{\max}} b_m^n(T,t)}{\sum_{t=1}^{t_{\max}} I_m^n(T,t)}$, respectively[3]. By the MLE, we can obtain the estimation of $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$, θ_m and B_m^n as $\tilde{g}_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$, $\tilde{\theta}_m$ and $\tilde{B}_m^n$, respectively, and then estimate the true expected throughput $U_n(\boldsymbol{a}(T))$ as $\tilde{U}_n(T) = \tilde{\theta}_m\tilde{B}_m^n\tilde{g}_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$. Since according to the central limit theorem $\tilde{g}_n$, $\tilde{\theta}_m$, and $\tilde{B}_m^n$ follow independent normal distributions with the mean $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$, θ_m, and B_m^n, respectively, we thus have

$$E[\tilde{U}_n(\boldsymbol{a}(T))] = E[\tilde{\theta}_m\tilde{B}_m\tilde{g}_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))] = U_n(\boldsymbol{a}(T)),$$

i.e., the estimation of expected throughput $U_n(\boldsymbol{a}(T))$ is unbiased.

5.2 Distributed Learning Algorithm

Based on the expected throughput estimation, we now propose the distributed learning algorithm for spatial spectrum access games. The idea is to extend the principle of single-agent reinforcement learning to a multi-agent setting. Such multi-agent reinforcement learning algorithm has also been applied to the classical congestion games on complete graphs [14,20]. Here we apply the learning algorithm to the generalized spatial congestion games on any generic graphs and derive the convergence conditions accordingly. The algorithm works as follows.

At the beginning of each period T, a user $n\in\mathcal{N}$ chooses a channel $a_n(T)\in\mathcal{M}$ to access according to its mixed strategy $\boldsymbol{\sigma}_n(T) = (\sigma_m^n(T),\forall m\in\mathcal{M})$, where $\sigma_m^n(T)$ is the probability of choosing channel m. The mixed strategy is generated according to $\boldsymbol{P}_n(T) = (P_m^n(T),\forall m\in\mathcal{M})$, which represents its *perceptions* of the payoff performance of choosing different channels based on local estimations. Perceptions are based on local observations in the past and may not accurately reflect the expected payoff. For example, if a user n has not accessed a channel m for many decision intervals, then perception $P_m^n(T)$ can be out of date. The key challenge for the learning algorithm is to update the perceptions with proper parameters such that perceptions equal to expected payoffs at the equilibrium.

Similarly to the single-agent learning, we choose the Boltzmann distribution as the mapping from perceptions to mixed strategies, i.e.,

$$\sigma_m^n(T) = \frac{e^{\gamma P_m^n(T)}}{\sum_{i=1}^M e^{\gamma P_i^n(T)}},\forall m\in\mathcal{M}, \tag{14}$$

where γ is the temperature that controls the randomness of channel selections. When $\gamma\to 0$, each user will choose to access channels uniformly at random. When $\gamma\to\infty$, user n always chooses the channel with the largest perception value $P_m^n(T)$ among all channel $m\in\mathcal{M}$. We will show later on that the choice of γ trades off convergence and performance of the learning algorithm.

At the end of a decision period T, a user n computes its estimated expected payoff $\tilde{U}_n(\boldsymbol{a}(T))$ as in Section 5.1 (i.e., by using the MLE method based on the set of local observations $\Omega_n(T)$ during the period), and adjusts its perceptions as

$$P_m^n(T+1) = \begin{cases}(1-\mu_T)P_m^n(T)+\mu_T\tilde{U}_n(\boldsymbol{a}(T)), & \text{if } a_n(T)=m,\\ P_m^n(T), & \text{otherwise,}\end{cases} \tag{15}$$

where $(\mu_T\in(0,1),\forall T)$ are the smoothing factors. A user only changes the perception of the channel just accessed in the current decision period, and keeps the perceptions of other channels unchanged.

Algorithm 1 summarizes the distributed learning algorithm. Next we study the convergence of the learning algorithm based on the theory of stochastic approximation [12].

5.3 Convergence of Distributed Learning Algorithm

We now study the convergence of the proposed distributed learning algorithm.

First, the perception value update in (15) can be written in the following equivalent form,

$$P_m^n(T+1)-P_m^n(T) = \mu_T[Z_m^n(T)-P_m^n(T)],\forall n\in\mathcal{N},m\in\mathcal{M}, \tag{16}$$

where $Z_m^n(T)$ is the update value defined as

$$Z_m^n(T) = \begin{cases}\tilde{U}_n(\boldsymbol{a}(T)), & \text{if } a_n(T)=m,\\ P_m^n(T), & \text{otherwise.}\end{cases} \tag{17}$$

For the sake of brevity, we denote the perception values, update values, and mixed strategies of all the users as

[3]When $S_m(t)$ and $b_m^n(t)$ are non-i.i.d. random variables, the MLE can also be derived based on the specific probability distribution functions by following the similar procedure as introduced in Section 5.1.1.

Algorithm 1 Distributed Learning Algorithm For Spatial Spectrum Access Game

initialization:
  set the temperature γ.
  set the initial perception values $P_m^n(0) = \frac{1}{M}$ for each user $n \in \mathcal{N}$.
  set the period index $T = 0$.
end initialization

loop for each decision period T and each user $n \in \mathcal{N}$ in parallel:
  select a channel $m \in \mathcal{M}$ according to (14).
  for each time slot t in the period T **do**
    sense and contend to access the channel m.
    record the observations $S_m^n(T,t)$, $I_m^n(T,t)$ and $b_m^n(T,t)$.
  end for
  estimate $g_n(\mathcal{N}_n^{a_n}(\boldsymbol{a}(T)))$, θ_m, and B_m^n by the maximum likelihood estimation.
  compute the estimated expected payoff $\tilde{U}_n(\boldsymbol{a}(T))$.
  update the perceptions value $\boldsymbol{P}_n(T)$ according to (15).
  set the period index $T = T+1$.
end loop

$\boldsymbol{P}(T) \triangleq (P_m^n(T), \forall m \in \mathcal{M}, n \in \mathcal{N})$, $\boldsymbol{Z}(T) \triangleq (Z_m^n(T), \forall m \in \mathcal{M}, n \in \mathcal{N})$, and $\boldsymbol{\sigma}(T) \triangleq (\sigma_m^n(T), \forall m \in \mathcal{M}, n \in \mathcal{N})$, respectively.

Let $Pr\{\mathcal{N}_n^m(\boldsymbol{a}(T))|\boldsymbol{P}(T), a_n(T) = m\}$ denote the conditional probability that, given that the users' perceptions are $\boldsymbol{P}(T)$ and user n chooses channel m, the set of users that choose the same channel m in user n's neighborhood $\mathcal{N}_n$ is $\mathcal{N}_n^m(\boldsymbol{a}(T)) \subseteq \mathcal{N}_n$. Since each user independently chooses a channel according to its mixed strategy $\boldsymbol{\sigma}_n(T)$, then the random set $\mathcal{N}_n^m(\boldsymbol{a}(T))$ follows the Binomial distribution of $|\mathcal{N}_n|$ independent non-homogeneous Bernoulli tries with the probability mass function as

$$\begin{aligned} & Pr\{\mathcal{N}_n^m(\boldsymbol{a}(T))|\boldsymbol{P}(T), a_n(T) = m\} \\ = & \prod_{i \in \mathcal{N}_n^m(\boldsymbol{a}(T))} (\sigma_m^i(T)) \prod_{i \in \mathcal{N}_n \setminus \mathcal{N}_n^m(\boldsymbol{a}(T))} (1 - \sigma_m^i(T)) \\ = & \prod_{i \in \mathcal{N}_n} (\sigma_m^i(T))^{I_{\{a_i(T)=m\}}} (1 - \sigma_m^i(T))^{1 - I_{\{a_i(T)=m\}}}, \end{aligned} \quad (18)$$

where $I_{\{a_i(T)=m\}} = 1$ if user i chooses channel m, and $I_{\{a_i(T)=m\}} = 0$ otherwise.

Since the update value $Z_m^n(T)$ depends on user n's estimated payoff $\tilde{U}_n(\boldsymbol{a}(T))$ (which in turn dependents on $\mathcal{N}_n^m(\boldsymbol{a}(T))$), thus $Z_m^n(T)$ is also a random variable. The equations in (16) are hence stochastic difference equations, which are difficult to analyze directly. We thus focus on the analysis of its *mean dynamics* [12]. To proceed, we define the mapping from the perceptions $\boldsymbol{P}(T)$ to the expected payoff of user n choosing channel m as $Q_m^n(\boldsymbol{P}(T)) \triangleq E[U_n(\boldsymbol{a}(T))|\boldsymbol{P}(T), a_n(T) = m]$. Here the expectation $E[\cdot]$ is taken with respective to the mixed strategies $\boldsymbol{\sigma}(T)$ of all users (i.e., the perceptions $\boldsymbol{P}(T)$ of all users due to (14)). We show that

Lemma 3. *For the distributed learning algorithm, if the temperature satisfies*

$$\gamma < \frac{1}{2 \max_{m \in \mathcal{M}, n \in \mathcal{N}} \{\theta_m B_m^n\} \max_{n \in \mathcal{N}} \{|\mathcal{N}_n|\}}, \quad (19)$$

the mapping from the perceptions to the expected payoff $Q(\boldsymbol{P}(T)) \triangleq (Q_m^n(\boldsymbol{P}(T)), m \in \mathcal{M}, n \in \mathcal{N})$ forms a maximum-norm contraction.

Note that the condition (19) is a sufficient condition to form a contraction mapping, which is in turn is a sufficient condition for convergence. Simulation results show that a slightly larger γ may also lead to the convergence of the mapping. Based on the property of contraction mapping, there exists a fixed point $\boldsymbol{P}^*$ such that $Q(\boldsymbol{P}^*) = \boldsymbol{P}^*$. By the theory of stochastic approximations [12], we show that the distributed learning algorithm also converges to the same limit point $\boldsymbol{P}^*$.

Theorem 5. *For the distributed learning algorithm, if the temperature γ satisfies (19), $\sum_T \mu_T = \infty$ and $\sum_T \mu_T^2 < \infty$, then the sequence $\{\boldsymbol{P}(T), \forall T \geq 0\}$ converges to the unique limit point $\boldsymbol{P}^* \triangleq (P_m^{n*}, \forall m \in \mathcal{M}, n \in \mathcal{N})$ of the differential equations ($\forall m \in \mathcal{M}, n \in \mathcal{N}$)*

$$\frac{dP_m^n(T)}{dT} = \sigma_m^n(T) \left(Q_m^n(\boldsymbol{P}(T)) - P_m^n(T) \right), \quad (20)$$

with probability one. Further, the limit point $\boldsymbol{P}^*$ satisfies

$$Q_m^n(\boldsymbol{P}^*) = P_m^{n*}, \forall m \in \mathcal{M}, n \in \mathcal{N}. \quad (21)$$

We next explore the property of the equilibrium $\boldsymbol{P}^*$ of the distributed learning algorithm. From Theorem 5, we see that

$$Q_m^n(\boldsymbol{P}^*) = E[U_n(\boldsymbol{a}(T))|\boldsymbol{P}^*, a_n(T) = m] = P_m^{n*}. \quad (22)$$

It means that the perception value P_m^{n*} is an accurate estimation of the expected payoff in the equilibrium. Moreover, we show that the mixed strategy $\boldsymbol{\sigma}^*$ is an approximate Nash equilibrium.

Definition 8 (**Approximate Nash Equilibrium** [8]). *A mixed strategy profile $\bar{\boldsymbol{\sigma}} = (\bar{\boldsymbol{\sigma}}_1, ..., \bar{\boldsymbol{\sigma}}_N)$ is a ξ- approximate Nash equilibrium if*

$$U_n(\bar{\boldsymbol{\sigma}}_n, \bar{\boldsymbol{\sigma}}_{-n}) \geq \max_{\boldsymbol{\sigma}_n} U_n(\boldsymbol{\sigma}_n, \bar{\boldsymbol{\sigma}}_{-n}) - \xi, \forall n \in \mathcal{N},$$

where $U_n(\bar{\boldsymbol{\sigma}}_n, \bar{\boldsymbol{\sigma}}_{-n})$ denotes the expected payoff of player n under mixed strategy $\bar{\boldsymbol{\sigma}}$, and $\bar{\boldsymbol{\sigma}}_{-n}$ denotes the mixed strategy profile of other users except player n.

Here $\xi \geq 0$ is the gap from a (precise) mixed Nash equilibrium. For the distributed learning algorithm, we show that

Theorem 6. *For the distributed learning algorithm, the mixed strategy $\boldsymbol{\sigma}^*$ in the equilibrium $\boldsymbol{P}^*$ is a ξ-approximate Nash equilibrium, with $\xi = \max_{n \in \mathcal{N}} \{-\frac{1}{\gamma} \sum_{m=1}^M \sigma_m^{n*} \ln \sigma_m^{n*}\}$.*

The gap ξ can be interpreted as the *weighted entropy*, which describes the randomness of the learning exploration. A larger ξ means worse learning performance. When each user adopts the uniformly random access, the gap ξ reaches the maximum value and results in the worst learning performance. Theorems 5 and 6 together illustrate the tradeoff between the convergence and performance through the choice of γ. A small enough γ is required to explore the environment (so that users are not getting stuck in channels with the *current* best payoffs) and guarantee the convergence of distributed learning. If γ is too small, however, then the performance gap ξ is large due to over-exploration.

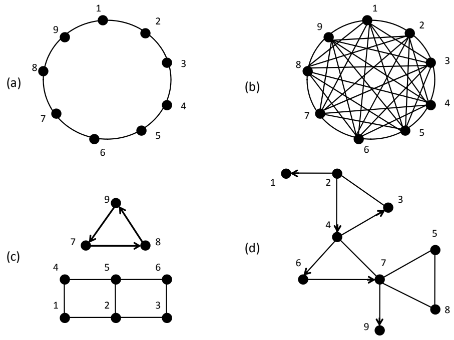

Figure 6: Interference Graphs

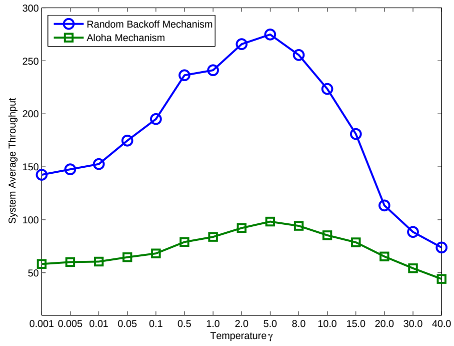

Figure 7: The system performance of the distributed learning algorithm with different temperature γ

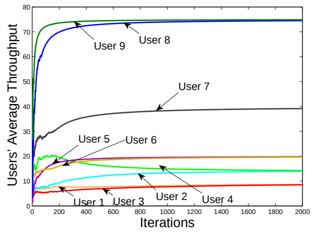

Figure 8: Users' average throughput on interference graph (d) with random backoff mechanism

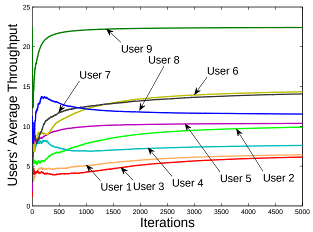

Figure 9: Users' average throughput on interference graph (d) with Aloha mechanism

6. NUMERICAL RESULTS

We now evaluate the proposed distributed learning algorithm by simulations. We consider a Rayleigh fading channel environment. The data rate of user n on an idle channel m is given according to the Shannon capacity, i.e., $b_m^n(t) = V_m \log_2\left(1 + \frac{\zeta_n g_m^n(t)}{N_0}\right)$, where V_m is the bandwidth of channel m, ζ_n is the power adopted by user n, N_0 is the noise power, and $g_m^n(t)$ is the channel gain, which is a random variable that follows the exponential distribution with the mean $\bar{g}_m^n$. In the following simulations, we set $V_m = 10$ MHz, $N_0 = -100$ dBm, and $\zeta_n = 100$ mW. By choosing different mean channel gains $\bar{g}_m^n$, we have different mean data rates $B_m^n = E[b_m^n(t)]$ for different channels and users. For simplicity, we set channel state $S_m(t)$ as an i.i.d. Bernoulli random variable with the idle probability $\theta_m = 0.5$.

We consider a network of $M = 5$ channels and $N = 9$ users with four different interference graphs (see Figure 6). Graphs (a) and (b) are undirected, and Graphs (c) and (d) are directed. Let $\vec{B}_n = \{B_1^n, ..., B_M^n\}$ be the mean data rate vector of user n. We set $\vec{B}_1 = \vec{B}_2 = \vec{B}_3 = \{2, 6, 16, 20, 30\}$ Mbps, $\vec{B}_4 = \vec{B}_5 = \vec{B}_6 = \{4, 12, 32, 40, 60\}$ Mbps, and $\vec{B}_7 = \vec{B}_8 = \vec{B}_9 = \{10, 30, 80, 100, 150\}$ Mbps. We implement both the random backoff and Aloha mechanisms for channel contention. For the random backoff mechanism, we set the number of backoff mini-slots in a time slot $\lambda_{\max} = 10$. For the Aloha mechanism, the channel contention probabilities of the users are $\{0.7, 0.7, 0.7, 0.5, 0.5, 0.5, 0.3, 0.3, 0.3\}$, respectively. Notice that in this study we focus on channel choices instead of the adjustment of contention probabilities.

For the distributed learning algorithm initialization, we set the length of each decision period $t_{\max} = 200$, which can achieve a good estimation of the mean data rate. We set the smooth factor $\mu_T = \frac{1}{200+T}$, which satisfies the condition $\sum_T \mu_T = \infty$ and $\sum_T \mu_T^2 < \infty$.

We first evaluate the distributed learning algorithm with different choices of temperature γ on the interference graph (d) in Figure 6. We run the learning algorithm sufficiently long until the time average system throughput does not change. The result in Figure 7 verifies the trade-off between the convergence and performance, and demonstrates that a proper temperature γ can offer the best performance. When is γ small, the gap ξ in Theorem 6 can be large. When γ is very large, the algorithm may get stuck in local optimum and the performance is again negatively affected. We set $\gamma = 5.0$ in the following simulations since it achieves good system performance in both random backoff and Aloha mechanisms as in Figure 7.

We then look at the learning dynamics. Figures 8 and 9 show the dynamics of users' time average throughputs with random backoff and Aloha mechanisms, respectively. These results demonstrate the convergence of the distributed learning algorithm in both mechanisms.

To benchmark the performance of the distributed learning algorithm, we compare it with the solutions obtained by the following two algorithms:

- Random Access: each user chooses a channel to access purely randomly.
- Centralized Optimization: the solution obtained by solving the centralized global optimization of $\max_{\boldsymbol{a}} \sum_{n\in\mathcal{N}} U_n(\boldsymbol{a})$.

We implement these algorithms together with the distributed learning algorithm on the four types of interference graphs in Figure 6. The results are shown in Figures 10 and 11. For the random backoff (Aloha, respectively) mechanism, we see that the distributed learning algorithm achieves up-to 100% (65%, respectively) performance improvement over the random access algorithm. Compared with the centralized optimal solution, the performance loss of the distributed learning in the full-interference graph (b) is 28% (34%, respectively). Such performance loss is not due to the algorithm design; instead it is due to the selfish nature of the users. In the partial-interference graphs (a), (c), and (d), the performance loss can be further reduced to less than 10% (17%, respectively). This shows that the negative

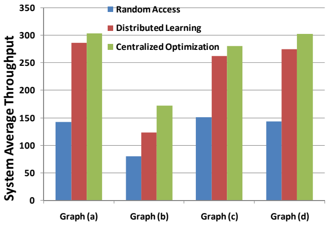

Figure 10: Comparison of distributed learning, random access, and centralized optimization with the random backoff mechanism

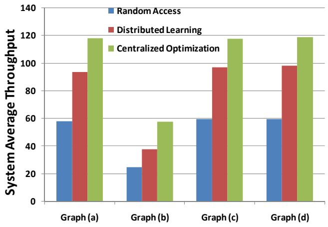

Figure 11: Comparison of distributed learning, random access, and centralized optimization with the Aloha mechanism

impact of users' selfish behavior is smaller when users can share the spectrum more efficiently through spatial reuse.

7. CONCLUSION

In this paper, we explore the spatial aspect of distributed spectrum sharing, and propose a framework of spatial spectrum access game on directed interference graphs. We investigate the critical issue of the existence of pure Nash equilibria, and develop a distributed learning algorithm converging to an approximate mixed Nash equilibrium for any spatial spectrum access games. Numerical results show that the algorithm is efficient and has significant performance gain over a random access algorithm that does not take the spatial effect into consideration.

For the future work, we are going to design distributed algorithms that maximize the system-wide throughput.

ACKNOWLEDGMENTS

This work is supported by the General Research Funds (Project Number 412509, 412710, and 412511) established under the University Grant Committee of the Hong Kong Special Administrative Region, China.

8. REFERENCES

[1] I. Akyildiz, W. Lee, M. Vuran, and S. Mohanty. Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Computer Networks*, 50(13):2127–2159, 2006.

[2] A. Anandkumar, N. Michael, and A. Tang. Opportunistic spectrum access with multiple users: learning under competition. In *Proc. of IEEE INFOCOM*, San Diego, USA, March 2010.

[3] N. Biggs, E. Lloyd, and R. Wilson. *Graph Theory*. Oxford University Press, 1986.

[4] V. Bilò, A. Fanelli, M. Flammini, and L. Moscardelli. Graphical congestion games with linear latencies. In *Proc. of ACM SPAA*, Munich, Germany, June 2008.

[5] X. Chen and J. Huang. Spatial spectrum access game: Nash equilibria and distributed learning. Technical report, The Chinese University of Hong Kong, 2012. Available at http://jianwei.ie.cuhk.edu.hk/publication/MobihocTech.pdf.

[6] X. Chen and J. Huang. Evolutionarily stable open spectrum access in a many-users regime. In *Proc. of IEEE GLOBECOM*, Houston, USA, Dec. 2011.

[7] X. Chen and J. Huang. Imitative spectrum access. In *Proc. of IEEE WiOpt*, Paderborn, Germany, May 2012.

[8] C. Daskalakisa, A. Mehtab, and C. Papadimitriou. A note on approximate Nash equilibria. *Theoretical Computer Science*, 410(17):1581–1588, 2009.

[9] M. Fĺęlegyhĺćzi, M. Cagalj, and J.-P. Hubaux. Efficient mac in cognitive radio systems: A game-theoretic approach. *IEEE Transactions on wireless Communications*, 8:1984–1995, 2009.

[10] S. Govindan and R. Wilson. A global newton method to compute nash equilibria. *Journal of Economic Theory*, 110:65–86, 2003.

[11] Z. Han, C. Pandana, and K. J. R. Liu. Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning. In *Proc. of IEEE WCNC*, Hong Kong, March 2007.

[12] H. Kushner and G. Yin. *Stochastic Approximation and Recursive: Algorithms and Applications*. Springer-Verlag, New York, 2003.

[13] K. Liu and Q. Zhao. Decentralized multi-armed bandit with multiple distributed players. In *Proc. of IEEE IAT*, San Diego, USA, Jan. 2010.

[14] E. Melo. Congestion pricing and learning in traffic network games. *Journal of Public Economic Theory*, 13(3):351–367, 2011.

[15] I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13:111–124, 1996.

[16] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[17] J. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36:48–49, 1950.

[18] N. Nie and C. Comniciu. Adaptive channel allocation spectrum etiquette for cognitive radio networks. In *Proc. of IEEE DySPAN*, Baltimore, USA, Nov. 2005.

[19] D. Niyato and E. Hossain. Competitive spectrum sharing in cognitive radio networks: a dynamic game approach. *IEEE Transactions on Wireless Communications*, 7:2651–2660, 2008.

[20] D. Shah and J. Shin. Dynamics in congestion games. *ACM SIGMETRICS Performance Evaluation Review*, 38(1):107–118, 2010.

[21] R. Southwell and J. Huang. Convergence dynamics of resource-homogeneous congestion game. In *Proc. of ICST GameNets*, Shanghai, China, April 2011.

[22] C. Tekin, M. Liu, R. Southwell, J. Huang, and S. Ahmad. Atomic congestion games on graphs and their applications in networking. *to appear in IEEE/ACM Transactions on Networking*, 2012.

[23] B. Wang, Y. Wua, and K. R. Liu. Game theory for cognitive radio networks: An overview. *Computer Networks*, 54:2537–2561, 2010.

[24] M. Weiss, M. Al-Tamaimi, and L. Cui. Dynamic geospatial spectrum modelling: taxonomy, options and consequences. In *Proc. of TPRC Conference*, Fairfax, USA, Sep. 2010.