Learning and Decision Making with Negative Externality for Opportunistic Spectrum Access

Biling Zhang[1,2], Yan Chen[1], Chih-Yu Wang[1,3], and K. J. Ray Liu[1]
[1]Department of Electrical and Computer Engineering,
University of Maryland, College Park, MD 20742, USA.
[2]School of Network Education,
Beijing University of Posts and Telecommunications, Beijing, 100876, P. R. China
[3]Graduate Institute of Communication Engineering,
National Taiwan University, Taipei, Taiwan
E-mail:{bzhang15, yan}@umd.edu, tomkywang@gmail.com, kjrliu@umd.edu,

***Abstract*—In cognitive radio networks, secondary users (SUs) are allowed to opportunistically exploit the licensed channels by sensing primary users' (PUs) activities. Once finding the spectrum holes, SUs generally need to share the available licensed channels. Therefore, one of the critical challenges for fully utilizing the spectrum resources is how the SUs obtain accurate information about the PUs' activities and make right decisions of accessing channels to avoid competition from other SUs. In this paper, we formulate SUs' learning and decision making process as a Chinese Restaurant Game by considering the scenario where SUs sense channels simultaneously and make access decisions sequentially. In the proposed game, SUs build the knowledge of the PUs' activities by their own sensing and learning the information from other SUs. They also predict their subsequent SUs' decisions to maximize their own utilities. We analyze the interactions among SUs in the proposed game and study specifically the impact of SUs' prior belief and sensing accuracy on their decisions. We also derive the theoretic results for the two-user two-channel case. Finally, we demonstrate the effectiveness and efficiency of the proposed scheme through simulations.**

***Index Terms*—Chinese Restaurant Game, opportunistic spectrum access, game theory, social learning**

I. Introduction

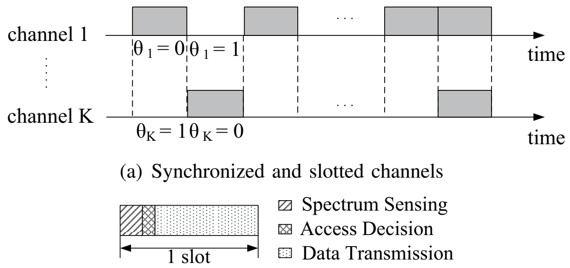
In cognitive radio networks, secondary users (SUs) as unlicensed users are allowed to use licensed spectrum bands with the constraint that they do not incur harmful interference to the primary users (PUs) who have the license of the spectrum bands. One typical cognitive radio technology is the opportunistic spectrum access, where SUs perform spectrum sensing, i.e., detect the PUs' activities, and access the spectrum once finding spectrum holes.

In the literature, many spectrum sensing approaches have been proposed to identify the spectrum holes [1] [2]. On the other hand, spectrum access aims at designing Medium Access Control (MAC) protocols to efficiently share the available spectrum resources among SUs [3] [4]. Joint spectrum sensing and access is also considered in the literature [5] [6]. Most of the aforementioned approaches assume that the utility of a specific SU is independent with the actions of other SUs. However, such an assumption is generally not true in reality, especially when we consider scenarios where SUs share or compete for certain resource. In such scenarios, the interactions among rational but selfish SUs need to be taken into account and game theory has been shown to be an effective tool to model such complex interactions [7] [8] [9] [10] [11].

Although the existing dynamic spectrum access schemes have greatly improved the spectrum utilization efficiency, due to the mobility of nodes and the dynamics of the channel variation, the accuracy of players' decisions is limited and remains a challenge to fully utilize the scarce spectrum resources [12]. Nevertheless, players in a cognitive network are generally intelligent and able to optimize their performance. They not only have the ability to recognize the changes of the surrounding environment by local observations, but also can collect global information such as signals and decisions revealed by other nodes. In such a case, the player's limited knowledge about the true system state can be expanded. The information learned by the player can be used to construct a belief on the unknown system state and improve the accuracy of the player's decision and thus the system efficiency.

In cognitive radio networks, the more SUs access the same channel, the lower rate they can achieve due to the interference among them. Such a phenomenon is known as the negative network externality [13] [14] [15]. Therefore, when making the decision of channel access, SUs should predict other SUs' decisions. Chinese Restaurant Game proposed in [16] provides a general framework for modeling strategic learning and decision processes in the social learning problem with negative network externality. The authors also illustrated three applications of Chinese Restaurant Game in wireless networking, cloud computing, and online social networking in [17]. However, since the authors in [16] and [17] mainly focus on building a general framework, the model and analysis may be too general to a specific system. Moreover, the authors only consider the homogeneous players where players have the same valuation about the resource. To better apply the Chinese Restaurant Game into cognitive radio networks, we need to carefully design the utility function of SUs by taking into account the heterogeneous characteristic of SUs, and detailedly analyze SUs' optimal actions under different conditions.

In this paper, we use Chinese Restaurant Game to model the opportunistic spectrum access problem in a cognitive radio network with multiple PUs and SUs. In our system, the SUs sense the channels simultaneously to estimate the channel state and then decide sequentially which channel to access.



(a) Synchronized and slotted channels

(b) The slot structure

Fig. 1. The system model.

We assume that each SU can only sense one channel and access one channel. Note that the channel an SU accesses may not be the channel he senses. Instead, the SU would exploit the information he knows or collects to make the optimal channel access decision. Since there exists a negative network externality in cognitive radio network, i.e., the more SUs access the same channel, the lower rate they can achieve, the SU should also predict other SUs' decisions to achieve the maximal payoff.

The rest of the paper is organized as follows. Section II describes in details our system model for the cognitive radio networks and formulates the decision making problem as a Chinese Restaurant Game. In Section III, we analyze the impact of prior belief and sensing accuracy on SUs' decisions in the two-user two-channel scenario, and derive some important theoretic results. Finally, we present the simulation results in Section IV and draw conclusions in Section V.

II. Chinese Restaurant Game Model of Cognitive Radio System

A. System Model

In this paper, we consider a primary system with K licensed channels, H_k, $k \in \mathcal{K} = \{1, 2, ..., K\}$, as shown in Fig. 1. We assume that the channels are slotted, and each channel is owned by one PU. Within each slot, according to the activity of the PU, the state of channel k is $\theta_k \in \{0, 1\}$, where "0" stands for the channel being occupied by the PU while "1" means that the channel is vacant.

Suppose that there are M secondary users (SUs), i.e., SU_m, $m \in \mathcal{M} = \{1, 2, ..., M\}$, searching vacant channels for transmission. Since SUs are not licensed users, they can only access the channel when the PUs are not present. In such a case, SUs need to perform sensing before accessing the channels. We assume that each SU will sense one of the channels and make his own decision on the PUs' activity individually. The sensing result, which represents the state of the sensed channel, is a binary signal $\{s^+, s^-\}$. The positive signal s^+ indicates that the channel is vacant while the negative signal s^- stands for the channel is occupied by the PU.

As shown in Fig. 1 (b), in our model, one slot is further divided into three sub-slots. In the first sub-slot, M SUs simultaneously perform sensing. In the second sub-slot, SUs sequentially make their accessing decisions based on the information they collected. We assume that SUs report their decisions as well as their sensing results through a dedicated common control channel which can be overheard by all other SUs. In the third sub-slot, SUs transmit their data through the channels they selected. If more than one SU choose the same channel, they will share the channel through Time Division Multiple Access (TDMA) or Code Division Multiple Access (CDMA).

B. Utility Function

Let $g = \{g_{m,k} | m \in \mathcal{M}, k \in \mathcal{K}\}$ be the channel quality of the system with $g_{m,k}$ being SU_m's channel gain in H_k. Here we assume that g is known to every SU. Given g, let $R_{m,k}(g, n)$ be the rate that SU_m can obtain when it shares channel H_k with $n - 1$ other SUs. The exact form of $R_{m,k}(g, n)$ is determined by how users share the channel. For example, if n users access a channel in a TDMA way, then $R_{m,k}(g, n) = R_{m,k}(g, 1)/n$.

Let $R_{m,k}(g)$ be the maximal rate SU_m can obtain by accessing channel H_k. Since the SU's data rate $R_{m,k}(g, n)$ is a decreasing function in terms of n, we have $R_{m,k}(g) = R_{m,k}(g, 1) = \frac{1}{2} log_2(1 + \frac{g_{m,k} P_m}{N_0^2})$, where P_m is SU_m's transmission power and N_0^2 is the variance of additive white Gaussain noise. Here we assume that all SUs use the same power to transmit and all channels have the same noise variance.

Definition (Preferential Channel): Channel H_k is the preferential channel of SU_m if $H_k = \arg \underset{H_k \in \{H_1, ..., H_K\}}{max} R_{m,k}(g)$.

We use the transmission throughput as SUs' utilities. Assuming the length of one slot is normalized to 1, the utility of SU_m accessing channel H_k can be written as

$$U_{m,k}(g, \theta_k, N_k) = R_{m,k}(g, N_k) 1(\theta_k = 1), \qquad (1)$$

where $1(\Delta)$ is the indicator function and N_k is the final number of SUs that choose to access channel H_k.

From the definition of utility we can see that an SU's utility is determined by the channel quality, the channel state and the number of SUs who share this channel. Therefore, to maximize the utilities, SUs should estimate both the channel state and the number of users who will eventually share the channel with them. Such a decision making process can be formulated as a Chinese Restaurant Game [16].

C. Chinese Restaurant Game

Let $A_s = \{1, 2, ..., K\}$ and $A_a = \{1, 2, ..., K\}$ be the sensing and access action set that SUs may choose from, respectively. Let $a_{sm} \in A_s$ and $a_{am} \in A_a$ be the sensing and access action of SU_m, and $a_s = \{a_{s1}, a_{s2}, ..., a_{sM}\}$ be all the SUs' sensing actions. We use the concept of belief to describe the SU's estimate on the channel state. Specifically, let the belief $b_{m,k}$ be the probability that channel H_k is vacant from the perspective of SU_m. Moreover, we assume that all SUs have a common prior belief on the channels as $b = \{b_{0,1}, b_{0,2}, ..., b_{0,K}\}$.

Let $s_m \in \{s^+, s^-\}$ be the signal obtained via SU_m's sensing and $\bar{s}_m \in \{s^+, s^-\} \backslash s_m$ be the complement signal of s_m. Let $S_k = \{s_m | SU_m \ \ senses \ \ H_k, \forall m\}$ and with S_k,

$$
\begin{aligned}
&Pr(v_{m,k} = X|b, p, g, a_s, n_m, h_m, s_m, a_{am}, \theta = l) \qquad (3)\\
&= \begin{cases} \sum\limits_{0\le u\le K} \int_{s\in S_{m+1,u(n_{m+1},h_{m+1})}} Pr(v_{m+1,k} = X-1|b, p, g, a_s, n_{m+1}, h_{m+1}, s_{m+1} = s, a_{am+1} = u, \theta = l) \\ f(s|\theta = l)ds, \quad a_{am} = k, \\ \sum\limits_{0\le u\le K} \int_{s\in S_{m+1,u(n_{m+1},h_{m+1})}} Pr(v_{m+1,k} = X|b, p, g, a_s, n_{m+1}, h_{m+1}, s_{m+1} = s, a_{am+1} = u, \theta = l) \\ f(s|\theta = l)ds, \quad a_{am} \neq k. \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
& a^*_{am} = BE_m(b, p, g, a_s, n_m, h_m, s_m), \\
= \quad & \arg\max_{k\in\mathcal{K}} \sum_{l\in\Theta} Pr(\theta = l|b, p, a_s, h_m, s_m) E[U_{m,k}(g, \theta_k, N_k)|b, p, g, a_s, n_m, h_m, s_m, a_{am} = k, \theta = l], \\
= \quad & \arg\max_{k\in\mathcal{K}} \sum_{l\in\Theta} \sum_{x=0}^{M-i+1} Pr(\theta = l|b, p, a_s, h_m, s_m) Pr(v_{m,k} = x|b, p, g, a_s, n_m, h_m, s_m, a_{am} = k, \theta = l) \\
& U_{m,k}(g, \theta_k, n_{m,k} + x), \qquad (4)
\end{aligned}
$$

SU_m can update its belief on H_k by following the Bayesian rule as

$$
b_{m,k}(b_{0,k}, p, S_k) = \frac{\prod\limits_{s_m\in S_k} f(s_m|\theta_k = 1) b_{0,k}}{\prod\limits_{s_m\in S_k} f(s_m|\theta_k = 1) b_{0,k} + \prod\limits_{s_m\in S_k} f(s_m|\theta_k = 0)(1 - b_{0,k})} \qquad (2)
$$

Here we assume that all the signals are independent and $f(s_m|\theta_k)$ is a predefined distribution that the signal s_m generated conditioning on the channel state θ_k. We denote $p = f(s_m = s^+|\theta_k = 1) = f(s_m = s^-|\theta_k = 0)$ be the sensing accuracy.

Besides the estimate on the channel state, an SU also needs to predict the decisions of the subsequent SUs due to the existence of negative network externality. Let $h_m = \{s_1, s_2, ..., s_{m-1}\}$ be the signals revealed by the SUs before SU_m and $n_m = \{n_{m,1}, n_{m,2}, ..., n_{m,K}\}$ be the grouping observed by SU_m when making its decision. If we denote $v_{m,k}$ be the number of SUs choosing H_k after SU_m, including SU_m itself, then through backward induction [16], we have (3) where $h_{m+1} = \{h_m, s_m\}$, $n_{m+1} = \{n_{m+1,1}, ..., n_{m+1,K}\}$, $\theta = \{\theta_1, \theta_2, ..., \theta_K\} \in \Theta$ is the system state, and $S_{m+1,u}$ is the signal space that SU_{m+1} will access H_u.

Then given b, p, a_s, h_m, n_m and s_m, SU_m's best response for maximizing its expected utility can be written as (4) where $Pr(\theta = l|b, p, a_s, h_m, s_m) = f(b, p, a_s, h_m, s_m)$, a function of b, p, a_s, h_m and s_m, is the probability that the system state θ is l.

III. Analysis of the Game for the Two-User Two-Channel Scenario

In this section, we analyze the interactions among SUs for the two-user two-channel scenario, i.e., $K = 2$ and $M = 2$. We first derive SUs' optimal access actions under different b and p by assuming the channel quality g, the sensing action a_s, and the corresponding sensing results are given. Then, we discuss the SUs' expected actions before knowing the sensing results. Due to page limitation, all the proofs for the Lemmas and Theorems are shown in the supplementary information [18].

A. Optimal Actions and Action Regions with Sensing Results

To give more insight of the proposed approach, we first assume that SUs' prior belief on both channels are the same, i.e., $b_{0,1} = b_{0,2} = b_0$. As discussed in the previous section, we use backward induction to derive SUs' optimal action. In the following, we first analyze SU_2's strategies and obtain the corresponding optimal action regions as described in Theorem 1.

Theorem 1: Suppose H_i is the preferential channel of SU_2. When $a_{s1} \neq a_{s2}$ and $s_1 \neq s_2$, or $a_{s1} = a_{s2}$ and $s_1 = s_2$, there are three possible action regions for SU_2 on the plane of b_0 and p as follows.

- $\Psi_1 = \{(b_0, p)|\frac{b_{2,i}(b_0,p,a_s,s_1,s_2)}{b_{2,-i}(b_0,p,a_s,s_1,s_2)} > \frac{R_{2,-i}(g,1)}{R_{2,i}(g,2)}\}$ with the optimal action $a^*_{a2} = i$,
- $\Psi_2 = \{(b_0, p)|\frac{R_{2,-i}(g,2)}{R_{2,i}(g,1)} < \frac{b_{2,i}(b_0,p,a_s,s_1,s_2)}{b_{2,-i}(b_0,p,a_s,s_1,s_2)} < \frac{R_{2,-i}(g,1)}{R_{2,i}(g,2)}\}$ with the optimal action $a^*_{a2} = -a_{a1}$,
- $\Psi_3 = \{(b_0, p)|\frac{b_{2,i}(b_0,p,a_s,s_1,s_2)}{b_{2,-i}(b_0,p,a_s,s_1,s_2)} < \frac{R_{2,-i}(g,2)}{R_{2,i}(g,1)}\}$ with the optimal action $a^*_{a2} = -i$,

where $-i \in \mathcal{K}\backslash i$, and $b_{2,i}(b_0, p, a_s, s_1, s_2)$ and $b_{2,-i}(b_0, p, a_s, s_1, s_2)$ are given by (2).

On the other hand, when $a_{s1} = a_{s2}$ and $s_1 \neq s_2$, or $a_{s1} \neq a_{s2}$ and $s_1 = s_2$, there will be only one possible optimal action on the whole plane of b_0 and p .

Based on SU_2's optimal action regions, we can analyze SU_1's strategies and derive the corresponding optimal action regions as follows.

Theorem 2: Suppose H_j is the preferential channel of SU_1. Then, SU_1's optimal actions and the corresponding action regions can be written as follows.

- $\Phi_1 = \underset{d}{\cup}\phi_d$ with the optimal action $a^*_{a1} = j$,
- $\Phi_2 = \underset{d}{\cup}\bar{\phi}_d$ with the optimal action $a^*_{a1} = -j$,

where $-j \in \mathcal{K}\backslash j$, $d \in \mathcal{D} = \{1, 2, 3\}$, ϕ_d and $\bar{\phi}_d$ are defined in (5) and (6), respectively.

$$\phi_d = \{(b0, p)|\{\frac{\bar{b}_{1,j}(b_0, p, a_s, s_1, \Psi_d)}{\bar{b}_{1,-j}(b_0, p, a_s, s_1, \Psi_d)} > \frac{R_{1,-j}(g)}{R_{1,j}(g)}\} \cap \Psi_d\}, \tag{5}$$

$$\bar{\phi}_d = \Psi_d \backslash \phi_d. \tag{6}$$

From the analysis of SUs' optimal strategies and the corresponding action regions in Theorem 1 and 2, we have the following observations.

- When SUs have the same preferential channel, they will share the preferential channel in region ϕ_1 and share the non-preferential channel in region $\bar{\phi}_3$.
- When SUs have their own preferential channel, respectively, they will share SU_1's preferential channel in region ϕ_3 and share SU_2's preferential channel in region $\bar{\phi}_1$.
- Given a_s and s_1, SU_1's action will be independent from the actual signal SU_2 receives.

B. Expected Actions without Sensing Results

In the previous subsection, we derive SUs' optimal strategies and the corresponding action regions given the sensing results. In this subsection, we will analyze the symmetric property of SUs' expected actions without the sensing results. Note that the expected action can be served as the SUs' prior information about their optimal actions before actually performing sensing.

For any $(b_x, p_y) \in \{(b_0, p)\}$, the expected action of SU_i, $i \in \{1, 2\}$, is defined as

$$\varphi_i(b_x, p_y) = \sum_{s \in \{h_i, s_i\}} Pr(s|b_x, p_y) * a_{ai}(s, b_x, p_y), \tag{7}$$

where s is the signal(s) SU_i collected, $Pr(s|b_x, p_y)$ is the probability of receiving s under b_x and p_y, and $a_{ai}(s, b_x, p_y)$ is SU_i's action when he receives s.

To show the symmetric property of the expected actions, we first characterize, in Lemma 1 and Lemma 2, the symmetric property of SUs' optimal actions and action regions when receiving opposite sensing results.

Lemma 1: Given a_s and g, SU_2 will choose the same optimal strategy in the action region $\Phi_d(b_0, p)$ with sensing results (s_1, s_2) and the action region $\Phi_d(b_0, 1-p)$ with sensing results $(\bar{s}_1, \bar{s}_2)$.

Lemma 2: Given a_s and g, SU_1 will choose the same optimal strategy in the action region $\phi_d(b_0, p)$ with sensing results s_1 and the action region $\phi_d(b_0, 1-p)$ with sensing results $\bar{s}_1$.

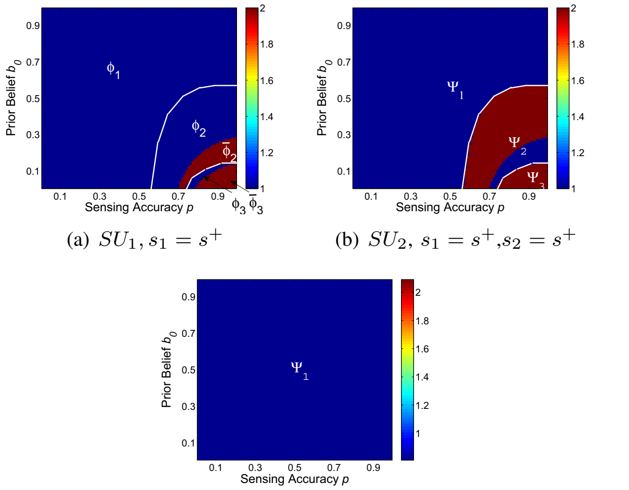
With the Lemmas above, we are ready to show the symmetric property of SUs' expected actions.

Theorem 3: Given a_s, the expected actions of SU_2 are symmetrical to p=0.5.

Theorem 4: Given a_s, the expected actions of SU_1 are symmetrical to p=0.5.

IV. Simulation Results

In this section, we evaluate the proposed game theoretic approach in terms of optimal action, action region, the expected action, and the system performance.



(a) $SU_1, s_1 = s^+$

(b) SU_2, $s_1 = s^+$,$s_2 = s^+$

(c) SU_2, $s_1 = s^+$, $s_2 = s^-$

Fig. 2. Action regions of SU_1 and SU_2 with $a_{s1} = 2$, $a_{s2} = 2$, $g = [10, 1; 10, 1]$ and $b_{01} = b_{02} = b_0$.

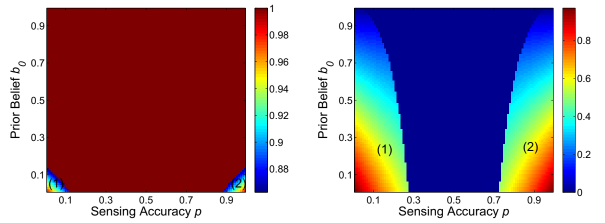
A. Actions with Sensing Results

In the first simulation, we evaluate SUs' strategies and the corresponding action regions by assuming that channel H_1 is the preferential channel for both SUs with channel gain 10, and channel H_2 is the non-preferential channel for both SUs with channel gain 1. Fig. 2 shows the optimal action regions when both SUs sense channel H_2.

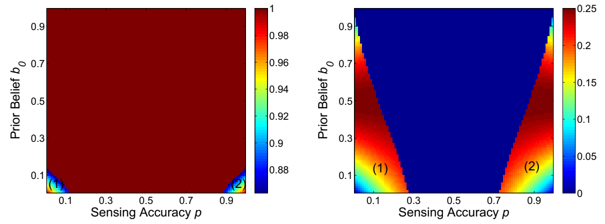
Since both SUs sense channel H_2, SUs' believes on channel H_1 remain unchanged while SUs' believes on channel H_2 will be updated according to the sensing results. From Fig. 2 (b) and (c), we can see that there are three action regions when both sensing results are positive, while there is only one action region when one of the sensing results is positive and the other is negative. Such phenomenon verifies the theoretical results in Theorem 1.

As shown in Fig. 2 (b), Ψ_1 is the action region where SU_2 accesses its preferential channel H_1. Such a phenomenon can be explained as follows. When $p < 0.5$, SU_2's belief on H_1 is larger than its belief on H_2 and accessing H_1 can bring a larger payoff due to the higher channel gain. Therefore, SU_2 chooses H_1 when $p < 0.5$. When $p > 0.5$, although SU_2's belief on H_1 is smaller than its belief on H_2, the larger payoff of accessing H_1 in action region Ψ_1 can make up the loss caused by the low belief even considering SU_1 may also access the same channel. Nevertheless, when (b_0, p) shifts from region Ψ_1 to Ψ_2, the gain of accessing H_1 can no longer compensate the loss of low belief and sharing channel with SU_1. Therefore, the best strategy for SU_2 in region Ψ_2 is to access the different channel from SU_1. In the region Ψ_3, SU_2's belief on H_1 is so low that the payoff of accessing H_1 is smaller than that of accessing H_2 even though H_2 may be shared by SU_1.

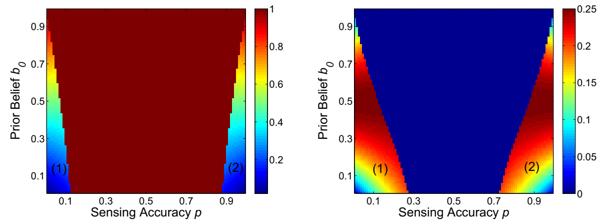
The action regions of SU_1 are shown in Fig. 2 (a). We can see that there are two possible action regions for SU_1 in each of SU_2's action region Ψ_d, which verifies the results in Theorem 2. The reason that there is only one action region ϕ_1 in Ψ_1 is that no (b_0, p) in Ψ_1 satisfies the condition defined



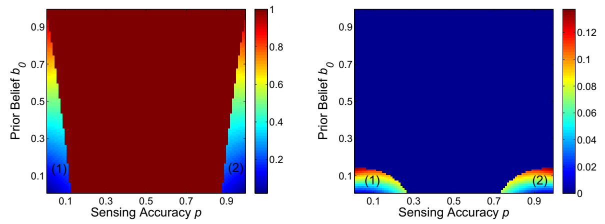
(a) $a_{s1} = 2$, $a_{s2} = 2$



(b) $a_{s1} = 2$, $a_{s2} = 1$



(c) $a_{s1} = 1$, $a_{s2} = 2$



(d) $a_{s1} = 1$, $a_{s2} = 1$

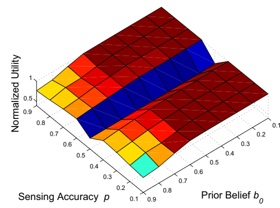
Fig. 3. Expected actions of SU_1 and SU_2 with g=[10,1;1,10].

in (6).

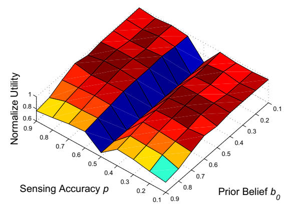
B. Expected Actions without Sensing Results

In this subsection, we evaluate SUs' expected actions without the sensing results and the outcomes are shown in Fig. 3. From Fig. 3, we can see that the expected actions of both SU_1 and SU_2 are symmetrical to $p = 0.5$, which verify Theorem 3 and Theorem 4.

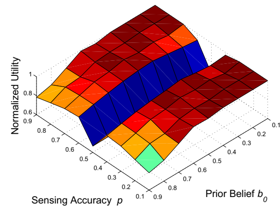
From Fig. 3, we can also see both SUs deviate from their preferential channels when (b_0, p) lies in the regions marked with (1) and (2). This is because in these regions, SUs' belief on the non-preferential channel can make up the loss of payoff when switching from the preferential channel. Moreover, when p becomes larger and b_0 becomes smaller, the probability of deviating becomes larger if they sense the preferential channels and becomes smaller if they sense the non-preferential channels. This is because SUs' expected actions depend on the signals they received, which is determined by p and b_0.



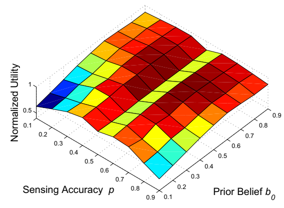
(a) myopic/CRG



(b) learning/CRG



(c) signal/CRG



(d) random/CRG

Fig. 4. Normalized utility of SU_2 with g=[10,10;10,10;10,10;10,10].

C. System Performance

In this subsection, we evaluate the proposed approach in terms of system performance. Since the simulation results are similar for different channel sharing models, here we only show those with the TDMA model where the SU's utility is defined as

$$U_{m,k}(g, \theta_k, N_k) = R_{m,k}(g)1(\theta_k = 1)/N_k. \tag{8}$$

We compare our approach with four other strategies: random, signal, learning, and myopic strategies. In the random strategy, SUs randomly and uniformly choose to access one of the channels. In the signal strategy, SUs make their decisions purely based on their own signal and the goal is to choose the channel that can maximize their expected utility as follows

$$a_{am}^{signal} = \arg \underset{k \in \mathcal{K}}{max} \sum_{l \in \Theta} Pr(\theta = l | b_0, p, a_s, s_m) U_{m,k}(g, \theta_k, 1), \tag{9}$$

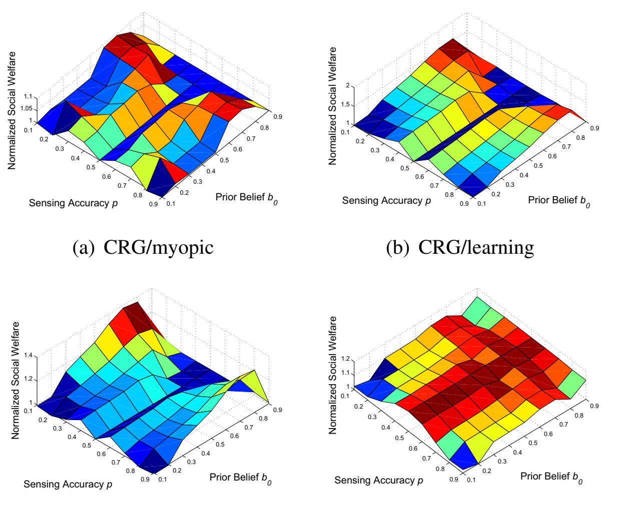
The learning strategy is an extension of the signal strategy. Under this strategy, the SU learns the channel state not only by his own signal but also by the signals revealed by the previous SUs. Therefore, the learning strategy can be obtained as

$$a_{am}^{learn} = \arg \underset{k \in \mathcal{K}}{max} \ b_{m,k} U_{m,k}(g, \theta_k = 1, 1), \tag{10}$$

In the myopic strategy, a myopic SU makes the decision according to his own signal, all signals revealed by the previous SUs, and the current grouping. The objective of the SU under myopic strategy is maximizing his current expected utility given by

$$a_{am}^{myopic} = \arg \underset{k \in \mathcal{K}}{max} \ b_{m,k} U_{m,k}(g, \theta_k = 1, n_{m,k} + 1), \tag{11}$$

We first verify that the proposed approach leads to the Nash equilibrium, i.e., any deviation to other strategies will lead to a utility loss. We assume that among the SUs, SU_2 may adopt one of the following five strategies: the proposed strategy denoted as CRG, random, signal, learning, and myopic. The rest of SUs all use the proposed strategy. We measure the ratio



(a) CRG/myopic (b) CRG/learning

(c) CRG/signal (d) CRG/random

Fig. 5. Normalized social welfare with g=[10,7;7,10;10,7;7,10].

between the utility generated by any four other strategies and the utility generated by CRG, and the results are shown in Fig. 4. From Fig. 4, we can see that the ratio is smaller than or equal to 1 for any b_0, p, and g, which means that the proposed strategy is indeed a Nash equilibrium.

In the following, we study the system performance in term of social welfare, i.e., the sum of all SUs' utilities in the system. In this simulation, all SUs in the system will adopt the same strategy. The results are presented in form of normalized social welfare, i.e., the ratio between the social welfare generated by CRG and the social welfare generated by any four other strategies.

Fig. 5 show the results of the scenario where the first and the third SUs have the same preferential channel and the second and the fourth SUs have the same preferential channel. In the preferential channel their channel gain is 10 while in the non-preferential channel their channel gain is 7. From Fig. 5, we can see that the social welfare with CRG has been increased 3%, 21%, 10% and 11% compared to that with myopic, learning, signal and random, respectively. That's because with high quality signals, an SU can get accurate information of the channel state and avoid the conflict with the PU. What's more, by observing the actions of previous SUs and estimate the actions of subsequent SUs, the SU can also avoid sharing the channel with too many other SUs. Such a mechanism finally contributes to the SU's right decision making and better payoff.

V. Conclusion

In this paper, we formulate SUs' decision making process problem in opportunistic spectrum access as a Chinese Restaurant Game. With the proposed game theoretic approach, SUs can make better decisions and achieve better performance through learning from others and estimating others' decisions. We theoretically derive SUs' optimal access actions and the corresponding action regions under different initial conditions. We also study some general properties such as symmetric property of SUs' expected action under different channel qualities for the two-user two-channel scenario. Simulation results verify our theoretic results and demonstrate the effectiveness and efficiency of the proposed scheme.

Acknowledgement

This work is partially supported by National Key Technologies R&D Program of China under Grant $2012ZX03005008$.

References

[1] G. Ganesan, and Y. Li, "Cooperative spectrum sensing in cognitive radio, part I: Two user networks", *IEEE Trans. Wireless. Commun*, vol. 6, no. 6, pp. 2204-2213, 2007.

[2] S. Haykin, D. J. Thomson and J. H. Reed, "Spectrum sensing for cognitive radio", *Proceedings of the IEEE*, vol. 97, no.5, pp. 849-877, 2009.

[3] L. Ma, X. Han and C. C. Shen, "Dynamic open spectrum sharing MAC protocol for wireless ad hoc networks", *New Frontiers in Dynamic Spectrum Access Networks*, 2005.

[4] R. Urgaonkar, and M. J. Neely, "Opportunistic scheduling with reliability guarantees in cognitive radio networks", *IEEE Trans. Mobile Computing*, vol. 8, no.6, pp. 766-777, 2009.

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

[6] K. W. Choi and E. Hossain, "Opportunistic access to spectrum holes between packet bursts: A learning-based approach", *IEEE Trans. Wireless. Commun*, vol. 10, no. 8, pp. 2497-2509, 2011.

[7] K. J. R. Liu and B. Wang, Cognitive radio networking and security: A game theoretical view, Cambridge University Press, 2010.

[8] B. Wang, Y. Wu, and K.J.R. Liu, "Game theory for cognitive radio networks: An overview", *Computer Networks*, vol 54, no 14, pp.2537-2561, 2010.

[9] Z. Ji and K. J. R. Liu, "Dynamic spectrum sharing: A game theoretical overview", *IEEE Communications Magazine*, vol 45, no 5, pp.88-94, 2007.

[10] B. Wang, K. J. R. Liu, and T. C. Clancy, "Evolutionary cooperative spectrum sensing game: How to collaborate?", *IEEE Trans. Commun.*, vol 58, no 3, pp.890-900, 2010.

[11] Y. Chen and K. J. R. Liu,"Indirect reciprocity game modelling for cooperation stimulation in cognitive networks", *IEEE Trans. Commun.*, vol 59, no 1, pp.159-168, 2011.

[12] V. Kone, L. Yang, X. Yang, B. Y. Zhao and H. Zheng, "On the feasibility of effective opportunistic spectrum access", *Proceedings of the 10th Internet Measurement Conference (IMC)*, 2010.

[13] M. L. Katz and C. Shapiro, "Technology adoption in the presence of network externalities", *Journal of Political Economy*, pp. 822C841, 1986.

[14] W. H. Sandholm, "Negative externalities and evolutionary implementation", *Review of Economic Studies*, vol 72, no 3, pp. 885C915, 2005.

[15] G. Fagiolo, "Endogenous neighborhood formation in a local coordination model with negative network externalities", *Journal of Economic Dynamics and Control*, vol. 29, no. 1-2, pp. 297C319, 2005.

[16] C. Y. Wang, Y. Chen, and K. J. R. Liu, "Chinese Restaurant Game - Part I: Theory of learning with negative network Externality". Arxiv preprint arXiv:1112.2188, 2011.

[17] C. Y. Wang, Y. Chen, and K. J. R. Liu, "Chinese Restaurant Game - Part II: Applications to wireless networking, cloud computing, and online social networking", Arxiv preprint arXiv:1112.2188, 2011.

[18] Supplementary information: http://www.ece.umd.edu/∼yan/CRG.pdf.