Evolutionary Game for Joint Spectrum Sensing and Access in Cognitive Radio Networks

Chunxiao Jiang*†, Yan Chen*, Yang Gao*, and K. J. Ray Liu*
*Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA
†Department of Electronic Engineering, Tsinghua University, Beijing 100084, P. R. China
E-mail:{jcx, yan, yanggao, kjrliu}@umd.edu

***Abstract*—Many spectrum sensing and access algorithms have been proposed to improve secondary users' (SUs') opportunities of utilizing primary spectrums. However, most of them have separated the analysis of spectrum sensing and access. In this paper, we propose to integrate the design of spectrum sensing and access algorithms by taking into account the mutual influence of them. Due to selfish natures, SUs tend to access the primary channel without contribution to the spectrum sensing. Moreover, they may take out-of-equilibrium strategies because of the uncertainty of others' strategies. To model the complicated interactions among SUs, we formulate the joint spectrum sensing and access problem as an evolutionary game and derive the evolutionarily stable strategy (ESS) that no one will deviate from. Furthermore, we design a distributed learning algorithm for SUs to converge to the ESS. With the proposed algorithm, each SU senses and accesses the primary channel with the probabilities learned purely from its own past utility history, and finally achieves the desired ESS. Simulation results shows that our system can quickly converge to the ESS and such an ESS is robust to the sudden unfavorable deviations of selfish SUs.**

I. Introduction

Recently, cognitive radio has been proposed as an effective communication paradigm to mitigate the problem of crowded radio spectrums. Through dynamic spectrum access (DSA), the utilization efficiency of existing spectrum resources can be greatly improved [1]. In DSA, cognitive devices, called as Secondary Users (SUs), can dynamically access the licensed spectrum, under the condition that the interference to the Primary Users (PUs) is minimized.

To detect available spectrums, SUs need to perform spectrum sensing to monitor the PUs' activities. Many spectrum sensing algorithms have been proposed in the literature [2]-[5]. Spectrum sensing methods based on energy detection and waveform sensing were proposed in [2] and [3], respectively. To improve the sensing performance, Ghasemi *et al.* proposed cooperative spectrum sensing to combat shadowing/fading effects [4], while Visotsky *et al.* studied how to combine spectrum sensing results in cooperative spectrum sensing schemes [5]. After detecting available spectrums, SUs need to decide how to access the spectrum. Several spectrum access methods based on different mathematical models have been proposed, e.g., Markov decision process (MDP) based approaches [6], renewal theoretic approaches [7], and game theoretic approaches [8].

However, most existing works separated the analysis of spectrum sensing and access, i.e., either optimizing the spectrum sensing performance without considering the effect of spectrum sharing, or designing the multi-user access algorithm without considering the issue of spectrum sensing. In this paper, we will integrate the analysis of SUs' spectrum sensing and access by considering a joint spectrum sensing and access game. On one hand, when only a few SUs contribute to spectrum sensing, the false-alarm probability is relatively high, resulting in low throughput during channel access. On the other hand, when many SUs access the primary channel, the channel will be very crowded and little throughput can be obtained by each SU. Therefore, each SU should dynamically adjust its strategy accordingly through learning from its interactions with other SUs. In [9], a joint design of spectrum sensing and access was studied from a queuing theoretic view, which considered the effect of spectrum sensing errors on the performance of SUs' channel access. In this paper, we propose a game theoretic framework for joint spectrum sensing and access by considering the mutual influence of them. In [10], a coalition game theoretic approach for joint spectrum sensing and access problem was proposed, which focused on how to form coalitions among SUs to constitute a Nash-stable network partition. However, our evolutionary game approach in this paper is to derive evolutionarily stable strategies for SUs.

Since SUs are naturally selfish, they want to access the channel without contributing to the spectrum sensing. Moreover, they may take out-of-equilibrium strategies due to the uncertainty of others' strategies. Therefore, a robust Nash equilibrium (NE) is desired for each SU. To model these socio-economic interactions among SUs and find a stable NE for them, we formulate the joint spectrum sensing and access problem as an evolutionary game and derive the evolutionarily stable strategy (ESS). Moreover, we propose a distributed learning algorithm for SUs to achieve the ESS purely based on their own utility histories. Evolutionary game has been used to study users' behaviors in communication and networking [11], such as cooperative spectrum sensing [12], network selection [13], spectrum access [14] and cooperative peer-to-peer (P2P) streaming [15], which is considered as an effective approach to model users' dynamic interactions in a network.

The rest of this paper is organized as follows. Firstly, our system model is described in details in Section II. Then, we analyze the joint spectrum sensing and access problem using evolutionary game theory in Section III. Simulation results are shown in Section IV and conclusions are drawn in Section V.

978-1-4799-1353-4/13/$31.00

II. System Model

We consider a cognitive radio network with one licensed primary channel and M SUs. The primary channel is slotted and SUs are synchronous with the PU's time slots. All SUs can independently perform spectrum sensing using energy detection and report their sensing results to others. There is a narrow-band signalling channel in the secondary network for SUs to exchange sensing results [16]. In this paper, we adopt the distributed cooperative sensing architecture, where each SU independently decides whether the PU is present through combining its own sensing results and other SUs'. To achieve better sensing performance, we assume SUs report their full sensing information to others, which is known as the *soft* combination rules [17].

Let $\mathcal{H}_0$ and $\mathcal{H}_1$ denote the PU being absent and present, $\mathcal{D}_0$ and $\mathcal{D}_1$ denote SUs decide that the PU is absent and present, respectively. The performance of spectrum sensing is generally measured by two terms: detection probability P_d and false-alarm probability P_f. Since P_d is usually pre-required by the PU, the corresponding P_f can be calculated by [12]

$$P_f = \frac{1}{2}\text{erfc}\left(\sqrt{2\gamma+1}\cdot\text{erfc}^{-1}(2P_d)+\gamma\sqrt{\frac{M\lambda T_s}{2}}\right) \quad (1)$$

where T_s is the sensing time and λ is the sampling rate, γ is a SU's received SNR of the PU's signal under $\mathcal{H}_1$, erfc$(\cdot)$ is the complementary error function erfc$(x) = \frac{2}{\sqrt{\pi}}\int_x^{+\infty} e^{-t^2}dt$.

III. Evolutionary Game Formulation

In this section, we first define SUs' utility functions of joint spectrum sensing and access. Through analyzing the replicator dynamics, we then derive SUs' evolutionarily stable strategy (ESS). Finally, we give a distributed learning algorithm for SUs to achieve the ESS.

A. Evolutionary Game

1) Basic Concepts: In the evolutionary game, each player dynamically adjusts his/her strategy through observing the utilities under different strategies. It is an effective approach for a group of players converging to a stable equilibrium after a period of strategic interactions, and such an equilibrium is called as Evolutionarily Stable Strategy (ESS). In a distributed scheme, all players are uncertain about other players' actions and utilities. To improve his/her own utility, each player will try different strategies in different rounds and learn from the interactions using the method of understanding-by-building. During this process, the portion of players using a certain pure strategy varies with time. In the evolutionary game, replicator dynamics are used to model such a population evolution.

In our system, there are two strategy sets for SUs: one is spectrum sensing strategy set $\Lambda_1 = (s, \overline{s})$ where strategy s denotes sensing and strategy $\overline{s}$ denotes not sensing, the other is spectrum access strategy set $\Lambda_2 = (a, \overline{a})$ where strategy a denotes access and strategy $\overline{a}$ denotes not access. Let p_s denote the portion of SUs who sense the primary channel, and p_a denote the portion of SUs who access the channel if they observe that the PU is absent after cooperative spectrum sensing, i.e., $p(a)|_{\mathcal{D}_\bullet} = p_a$ and $p(a)|_{\mathcal{D}_1} = 0$. Then, the evolution dynamics of p_s and p_a are given by

$$\dot{p_s} = \eta p_s\big(\mathbb{U}_s - \mathbb{U}\big), \quad (2)$$

$$\dot{p_a} = \eta p_a\big(\mathbb{U}_a|_{\mathcal{D}_\bullet} - \mathbb{U}|_{\mathcal{D}_\bullet}\big), \quad (3)$$

where $\mathbb{U}_s$ is the average utility of SUs who participate in the cooperative spectrum sensing, $\mathbb{U}$ is the average utility of all SUs, $\mathbb{U}_a|_{\mathcal{D}_\bullet}$ is the average utility of SUs who access the primary channel given the condition of $\mathcal{D}_0$, $\mathbb{U}|_{\mathcal{D}_\bullet}$ is the average utility of all SUs given $\mathcal{D}_0$, and η is a positive scale factor. From (5), we can see that if spectrum sensing can lead to a higher utility than the average level, the portion p_s will increase and the increasing rate $\dot{p_s}/p_s$ is proportional to the difference between $\mathbb{U}_s$ and $\mathbb{U}$. Similar phenomenon can be found for the evolution of p_a.

2) Utility Functions: Since we are considering the joint spectrum sensing and access game, the utility functions are determined by both p_s and p_a. When the PU is absent, i.e., given $\mathcal{H}_0$, the utility functions of SUs with four difference actions $\{sa, \overline{s}a,\ s\overline{a},\ \overline{sa}\}$, can be written as follows:

$$U_{sa}|_{\mathcal{H}_\bullet} = \mathbb{F}(Mp_a) - \Theta_a - \Theta_s,\ \ U_{s\overline{a}}|_{\mathcal{H}_\bullet} = -\Theta_s + R, \quad (4)$$

$$U_{\overline{s}a}|_{\mathcal{H}_\bullet} = \mathbb{F}(Mp_a) - \Theta_a,\ \ U_{\overline{sa}}|_{\mathcal{H}_0} = 0, \quad (5)$$

where $\mathbb{F}(\cdot)$ is the reward of a SU obtained from channel access, $\Theta_a = T_aE_2$ is the energy consumed by data transmission, $\Theta_s = T_sE_3$ is the energy consumed by spectrum sensing, and the constant R is the reward to the SU who contributes to channel sensing but do not access the channel and $R > \Theta_s$. Here $\mathbb{F}(Mp_a)$ represents the throughput of a SU given by

$$\mathbb{F}(Mp_a) = B\log\left(1+\frac{\text{SNR}}{(Mp_a-1)\cdot\text{INR}+1}\right)\cdot T_aE_1, \quad (6)$$

where Mp_a denotes the number of SUs that choose to access the channel given $\mathcal{D}_0$, B is the bandwidth of the primary channel, INR is the interference from other SUs, and E_1 is the parameter that translates one SU's throughput reward into its energy reward.

Similar to the case $\mathcal{H}_0$, we can summarize SUs' utility functions under $\mathcal{H}_1$ as follows:

$$U_{sa}|_{\mathcal{H}_1} = -\Theta_a - \Theta_s, \quad U_{\overline{s}a}|_{\mathcal{H}_1} = -\Theta_a, \quad (7)$$

$$U_{s\overline{a}}|_{\mathcal{H}_1} = -\Theta_s + R, \quad U_{\overline{sa}}|_{\mathcal{H}_1} = 0, \quad (8)$$

where we assume that SUs cannot obtain reward by access under $\mathcal{H}_1$ due to the presence of the PU.

B. Replicator Dynamics of Spectrum Sensing

The replicator dynamics of spectrum sensing are given in (2), where we need to derive the average utility of SUs who perform channel sensing $\mathbb{U}_s$ and the average utility of all SUs $\mathbb{U}$. According to utility functions (4)-(8), we can calculate the average utility of performing sensing and not performing sensing given $\mathcal{H}_0$ or $\mathcal{H}_1$ as follows:

$$\mathbb{U}_s|_{\mathcal{H}_\bullet} = p(a)|_{\mathcal{H}_\bullet}\cdot U_{sa}|_{\mathcal{H}_\bullet} + p(\overline{a})|_{\mathcal{H}_\bullet}\cdot U_{s\overline{a}}|_{\mathcal{H}_\bullet}, \quad (9)$$

$$\mathbb{U}_s|_{\mathcal{H}_1} = p(a)|_{\mathcal{H}_1}\cdot U_{sa}|_{\mathcal{H}_1} + p(\overline{a})|_{\mathcal{H}_1}\cdot U_{s\overline{a}}|_{\mathcal{H}_1}, \quad (10)$$

$$\mathbb{U}_{\overline{s}}|_{\mathcal{H}_\bullet} = p(a)|_{\mathcal{H}_\bullet} \cdot U_{\overline{s}a}|_{\mathcal{H}_\bullet} + p(\overline{a})|_{\mathcal{H}_\bullet} \cdot U_{\overline{sa}}|_{\mathcal{H}_\bullet}, \quad (11)$$

$$\mathbb{U}_{\overline{s}}|_{\mathcal{H}_1} = p(a)|_{\mathcal{H}_1} \cdot U_{\overline{s}a}|_{\mathcal{H}_1} + p(\overline{a})|_{\mathcal{H}_1} \cdot U_{\overline{sa}}|_{\mathcal{H}_1}, \quad (12)$$

where $p(a)|_{\mathcal{H}_\bullet}$ and $p(a)|_{\mathcal{H}_1}$ denote the portion of SUs who access the primary channel given channel condition $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, which can be calculated by

$$p(a)|_{\mathcal{H}_\bullet} = p(a)|_{(\mathcal{D}_\bullet,\mathcal{H}_\bullet)} p(\mathcal{D}_0)|_{\mathcal{H}_\bullet} = p_a\Big(1 - P_f(Mp_s)\Big), \quad (13)$$

$$p(a)|_{\mathcal{H}_1} = p(a)|_{(\mathcal{D}_\bullet,\mathcal{H}_1)} p(\mathcal{D}_0)|_{\mathcal{H}_1} = p_a\Big(1 - P_d\Big), \quad (14)$$

where $P_f(Mp_s)$ is the false-alarm probability when Mp_s SUs cooperatively sense the primary channel.

Thus, we can derive the average utility of SUs who sense the primary channel, $\mathbb{U}_s$, the average utility of SUs who do not sense, $\mathbb{U}_{\overline{s}}$, and the average utility of all SUs, $\mathbb{U}$, as follows:

$$\mathbb{U}_s = p_0 \cdot \mathbb{U}_s|_{\mathcal{H}_\bullet} + p_1 \cdot \mathbb{U}_s|_{\mathcal{H}_1}, \quad (15)$$

$$\mathbb{U}_{\overline{s}} = p_0 \cdot \mathbb{U}_{\overline{s}}|_{\mathcal{H}_\bullet} + p_1 \cdot \mathbb{U}_{\overline{s}}|_{\mathcal{H}_1}, \quad (16)$$

$$\mathbb{U} = p_s \cdot \mathbb{U}_s + (1 - p_s) \cdot \mathbb{U}_{\overline{s}}, \quad (17)$$

where p_0 is the probability that the PU is absent, i.e., the probability of $\mathcal{H}_0$, and $p_1 = 1 - p_0$ is the probability of $\mathcal{H}_1$. Combing (2) and (4)-(17), we can re-write the replicator dynamics of spectrum sensing by (18) below.

C. Replicator Dynamics of Spectrum Access

Similar to the analysis of replicator dynamics of spectrum sensing, we should first calculate the average utility of SUs accessing the primary channel $\mathbb{U}_a|_{\mathcal{D}_\bullet}$ and the average utility of all SUs given $\mathcal{D}_0$, $\mathbb{U}|_{\mathcal{D}_\bullet}$. The SUs' utilities with four pure strategies $\{sa, \overline{s}a, s\overline{a}, \overline{sa}\}$ given $\mathcal{D}_0$ can be written by (19)-(22), where $p(\mathcal{H}_0)|_{(a,\mathcal{D}_\bullet)}$ and $p(\mathcal{H}_0)|_{(\overline{a},\mathcal{D}_\bullet)}$ are the probabilities that the PU is absent when SUs decide to access and not access, respectively, $U_{\cdot\cdot}|_{(\mathcal{D}_\bullet,\mathcal{H}_\bullet)}$ and $U_{\cdot\cdot}|_{(\mathcal{D}_\bullet,\mathcal{H}_1)}$ are SUs' utility functions when given $\{\mathcal{D}_0, \mathcal{H}_0\}$ and $\{\mathcal{D}_0, \mathcal{H}_1\}$, respectively. Since given $\mathcal{D}_0$, the accessing behavior is independent with the status of the PU, we have $p(\mathcal{H}_0)|_{(a,\mathcal{D}_\bullet)} = p(\mathcal{H}_0)|_{(\overline{a},\mathcal{D}_\bullet)} = p(\mathcal{H}_0)|_{\mathcal{D}_\bullet}$, where $p(\mathcal{H}_0)|_{\mathcal{D}_\bullet}$ can be calculated by the Bayes' rule given as

$$p(\mathcal{H}_0)|_{\mathcal{D}_\bullet} = \frac{p_0 \cdot p(\mathcal{D}_0)|_{\mathcal{H}_\bullet}}{p_0 \cdot p(\mathcal{D}_0)|_{\mathcal{H}_\bullet} + p_1 \cdot p(\mathcal{D}_0)|_{\mathcal{H}_1}} = \frac{p_0\big(1 - P_f(Mp_s)\big)}{1 - p_0 P_f(Mp_s) - p_1 P_d}. \quad (23)$$

Given the actions, SUs' utilities are independent with $\mathcal{D}_0$, i.e., $U_{\cdot\cdot}|_{(\mathcal{D}_\bullet,\mathcal{H}_\bullet)} = U_{\cdot\cdot}|_{\mathcal{H}_\bullet}$ and $U_{\cdot\cdot}|_{(\mathcal{D}_\bullet,\mathcal{H}_1)} = U_{\cdot\cdot}|_{\mathcal{H}_1}$. In such a case, we can derive the average utilities of SUs who access and who do not access the primary channel $\mathbb{U}_a|_{\mathcal{D}_\bullet}$ and $\mathbb{U}_{\overline{a}}|_{\mathcal{D}_\bullet}$, and the average utility of all SUs $\mathbb{U}|_{\mathcal{D}_\bullet}$ as follows:

$$\mathbb{U}_a|_{\mathcal{D}_\bullet} = p_s \cdot U_{sa}|_{\mathcal{D}_\bullet} + (1 - p_s) \cdot U_{\overline{s}a}|_{\mathcal{D}_\bullet}, \quad (24)$$

$$\mathbb{U}_{\overline{a}}|_{\mathcal{D}_\bullet} = p_s \cdot U_{s\overline{a}}|_{\mathcal{D}_\bullet} + (1 - p_s) \cdot U_{\overline{sa}}|_{\mathcal{D}_0}, \quad (25)$$

$$\mathbb{U}|_{\mathcal{D}_\bullet} = p_a \cdot \mathbb{U}_a|_{\mathcal{D}_\bullet} + (1 - p_a) \cdot \mathbb{U}_{\overline{a}}|_{\mathcal{D}_\bullet}. \quad (26)$$

Combining (2) and (19)-(26), we can re-write the replicator dynamics of spectrum access by (27) below.

D. Analysis of Evolutionarily Stable Strategy

At equilibrium, we have $\dot{p_s} = 0$ and $\dot{p_a} = 0$. According to (18) and (27), we can get seven possible equilibria: $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(p_{s_1}, 1)$, $(1, p_{a_1})$, (p_{s_2}, p_{a_2}), where p_{s_1} satisfies $P_f(Mp_{s_1}) = \big(\frac{\Theta_s}{R} - p_1 P_d\big)/p_0$, p_{a_1} satisfies $\mathbb{F}(Mp_{a_1}) = \frac{(\Theta_\bullet+R)\big(1-p_\bullet P_f(M)-p_1 P_\bullet\big)}{p_\bullet\big(1-P_f(M)\big)}$ and (p_{s_2}, p_{a_2}) is the solution to the following equations

$$\begin{cases} -\Theta_s + \Big(1-p_a+p_a\Big(p_0 P_f(Mp_s)+p_1 P_d\Big)\Big)R = 0, \\ \frac{p_0\big(1-P_f(Mp_s)\big)}{1-p_0 P_f(Mp_s)-p_1 P_d}\mathbb{F}(Mp_a) - \Theta_a - p_s R = 0. \end{cases} \quad (28)$$

According to the evolutionary game theory [11], an equilibrium of the replicator dynamics equation is an ESS if and if only it is a locally asymptotically stable point in a dynamic system. In the following *Lemma 1* and *Theorem 1*, we will check which equilibria are the ESSs.

Lemma 1: The false-alarm probability P_f is a decreasing function in terms of p_s, and the reward from channel access $\mathbb{F}$ is a decreasing function in terms of p_a, i.e., $\frac{dP_f(Mp_s)}{dp_s} < 0$ and $\frac{d\mathbb{F}(Mp_a)}{dp_a} < 0$.

Proof: This can be easily proved by taking derivatives on (1) and (6). ■

Theorem 1: For the joint spectrum sensing and access evolutionary game, there are three ESSs: $(p_s^*, p_a^*) = (1, 0)$, $(1, p_{a_1})$ and (p_{s_2}, p_{a_2}), under different conditions of the rewards R listed as follows:

$$\dot{p_s} = \eta p_s(1-p_s)(\mathbb{U}_s - \mathbb{U}_{\overline{s}}) = \eta p_s(1-p_s)\Bigg(-\Theta_s + \Big(1-p_a+p_a\Big(p_0 P_f(Mp_s)+p_1 P_d\Big)\Big)R\Bigg). \quad (18)$$

$$U_{sa}|_{\mathcal{D}_\bullet} = p(\mathcal{H}_0)|_{(a,\mathcal{D}_\bullet)} \cdot U_{sa}|_{(\mathcal{D}_\bullet,\mathcal{H}_\bullet)} + \Big(1 - p(\mathcal{H}_0)|_{(a,\mathcal{D}_\bullet)}\Big) \cdot U_{sa}|_{(\mathcal{D}_\bullet,\mathcal{H}_1)}, \quad (19)$$

$$U_{\overline{s}a}|_{\mathcal{D}_\bullet} = p(\mathcal{H}_0)|_{(a,\mathcal{D}_\bullet)} \cdot U_{\overline{s}a}|_{(\mathcal{D}_\bullet,\mathcal{H}_\bullet)} + \Big(1 - p(\mathcal{H}_0)|_{(a,\mathcal{D}_\bullet)}\Big) \cdot U_{\overline{s}a}|_{(\mathcal{D}_\bullet,\mathcal{H}_1)}, \quad (20)$$

$$U_{s\overline{a}}|_{\mathcal{D}_\bullet} = p(\mathcal{H}_0)|_{(\overline{a},\mathcal{D}_\bullet)} \cdot U_{s\overline{a}}|_{(\mathcal{D}_\bullet,\mathcal{H}_\bullet)} + \Big(1 - p(\mathcal{H}_0)|_{(\overline{a},\mathcal{D}_\bullet)}\Big) \cdot U_{s\overline{a}}|_{(\mathcal{D}_\bullet,\mathcal{H}_1)}, \quad (21)$$

$$U_{\overline{sa}}|_{\mathcal{D}_0} = p(\mathcal{H}_0)|_{(\overline{a},\mathcal{D}_\bullet)} \cdot U_{\overline{sa}}|_{(\mathcal{D}_0,\mathcal{H}_0)} + \Big(1 - p(\mathcal{H}_0)|_{(\overline{a},\mathcal{D}_\bullet)}\Big) \cdot U_{\overline{sa}}|_{(\mathcal{D}_0,\mathcal{H}_1)}. \quad (22)$$

$$\dot{p_a} = \eta p_a(1 - p_a)(\mathbb{U}_a|_{\mathcal{D}_\bullet} - \mathbb{U}_{\overline{a}}|_{\mathcal{D}_\bullet}) = \eta p_a(1 - p_a)\Bigg(\frac{p_0\big(1 - P_f(Mp_s)\big)}{1 - p_0 P_f(Mp_s) - p_1 P_d}\mathbb{F}(Mp_a) - \Theta_a - p_s R\Bigg). \quad (27)$$

$$(p_s^*, p_a^*) = \begin{cases} (1,0), & R > \frac{p_0\left(1-P_f(M)\right)}{1-p_0P_f(M)-p_1P_d}\mathbb{F}(1) - \Theta_a, \\ (1, p_{a_1}), & R > \frac{\Theta_s}{1-p_{a_1}\left(1-p_0P_f(M)-p_1P_d\right)}, \\ (p_{s_2}, p_{a_2}), & R < R', \end{cases} \quad (29)$$

where

$$R' = \frac{p_0^2 p_{s_2}\left(1-P_f(Mp_{s_2})\right)\frac{d\mathbb{F}(Mp_{a_2})}{dp_{a_2}} - \frac{p_0p_1(1-P_d)\mathbb{F}(Mp_{a_2})}{1-p_0P_f(Mp_{s_2})-p_1P_d}}{\left(1-p_0P_f(Mp_{s_2})-p_1P_d\right)^2} \frac{dP_f(Mp_{s_2})}{dp_{s_2}}.$$

Proof: Due to page limitation, we show the proof in the supplementary information [18]. ■

E. A Distributed Learning Algorithm for ESS

In the above joint spectrum sensing and access evolutionary games, we have obtained the ESSs for SUs. Thus, a group of SUs can achieve the ESS using the replicator dynamics equations (18) (27). We can see that solving these equations requires the exchange of utilities among all SUs to find the average utilities such as $\mathbb{U}_s$ and $\mathbb{U}_a|_{\mathcal{D}_0}$. However, in a distributed network, it is generally difficult to make each SU reveal such private information. In this section, we will present a distributed learning algorithm that can gradually converge to ESS without private utility information exchange.

In the evolutionary biology, the Wright-Fisher model has been widely used to study the population reproduction dynamics under natural selection [19]. The model is based on the assumption that the probability of an individual adopting a certain strategy is proportional to the expected utility of the population using that strategy. Let $\tilde{p}_s$ and $\tilde{p}_a$ be the probabilities of a SU sensing and accessing the primary channel, respectively. According to the Wright-Fisher model, $\tilde{p}_s$ is proportional to the total utility of SUs sensing the channel. In such a case, SUs' strategy of spectrum sensing at time slot $t+1$, $\tilde{p}_s(t+1)$, can be calculated by

$$\tilde{p}_s(t+1) = \frac{p_s(t)\widetilde{\mathbb{U}}_s(t)}{p_s(t)\widetilde{\mathbb{U}}_s(t) + \Big(1-p_s(t)\Big)\widetilde{\mathbb{U}}_{\overline{s}}(t)}, \quad (30)$$

where $\widetilde{\mathbb{U}}_s(t)$ and $\widetilde{\mathbb{U}}_{\overline{s}}(t)$ are the average utilities of SUs who have sensed and not sensed the channel at the tth time slot, respectively, the denominator is the average utility of all SUs, which is the normalization term that ensures $\tilde{p}_s + \tilde{p}_{\overline{s}} = 1$. If SUs observe that the PU is present after cooperative spectrum sensing at the beginning of the time slot, they will always not access the primary channel within this slot. Otherwise, each SU will access the channel with probability $\tilde{p}_a$. With the Wright-Fisher model, $\tilde{p}_a$ is proportional to the total utility of SUs who choose to access the channel. In such a case, one SU's strategy of channel access, $\tilde{p}_a$, can be computed by

$$\tilde{p}_a(t'+1) = \frac{p_a(t')\widetilde{\mathbb{U}}_a(t')}{p_a(t')\widetilde{\mathbb{U}}_a(t') + \Big(1-p_a(t')\Big)\widetilde{\mathbb{U}}_{\overline{a}}(t')}, \quad (31)$$

where t' and $t'+1$ represent the time slots when SUs observe that the PU is absent, $\widetilde{\mathbb{U}}_a(t')$ and $\widetilde{\mathbb{U}}_{\overline{a}}(t')$ are the average utilities of SUs who have accessed and not accessed the channel at the t'th time slot, respectively.

Based on the assumption that the number of SUs M is sufficiently large, the portion of SUs who sense the primary channel is equal to the probability of one individual SU choosing to sense the channel, i.e., $p_s = \tilde{p}_s$. Similarly, $p_a = \tilde{p}_a$ if SUs observe the PU is absent. In such a case, we have

$$\tilde{p}_s(t+1) = \frac{\tilde{p}_s(t)\widetilde{\mathbb{U}}_s(t)}{\tilde{p}_s(t)\widetilde{\mathbb{U}}_s(t) + \Big(1-\tilde{p}_s(t)\Big)\widetilde{\mathbb{U}}_{\overline{s}}(t)}, \quad (32)$$

$$\tilde{p}_a(t'+1) = \frac{\tilde{p}_a(t')\widetilde{\mathbb{U}}_a(t')}{\tilde{p}_a(t')\widetilde{\mathbb{U}}_a(t') + \Big(1-\tilde{p}_a(t')\Big)\widetilde{\mathbb{U}}_{\overline{a}}(t')}. \quad (33)$$

From (32), we can see that when $\widetilde{\mathbb{U}}_s = \widetilde{\mathbb{U}}_{\overline{s}}$ or $\tilde{p}_s = 0$ or 1, $\dot{\tilde{p}}_s = 0$, i.e., the equilibrium is achieved. According to the replicator dynamics equation of spectrum sensing in (18), $\widetilde{\mathbb{U}}_s = \widetilde{\mathbb{U}}_{\overline{s}}$ and $\tilde{p}_s = 0$ or 1 are the solutions to the replicator dynamics. The same argument can be applied to $\tilde{p}_a$. Therefore, the Wright-Fisher model is equivalent to the replicator dynamics equations when M is sufficiently large. Note that although the total number of SUs M may be large, the ESS, e.g., (p_{s_2}, p_{a_2}), can ensure that only portions of SUs will sense and access the channel. In such a case, the actual number of SUs that need to exchange sensing data or share the channel simultaneously can be small, which can be seen from the simulation results.

From (32), we can see that in order to update $\tilde{p}_s(t+1)$, each SU needs to learn about the average utilities $\widetilde{\mathbb{U}}_s(t)$ and $\widetilde{\mathbb{U}}_{\overline{s}}(t)$. Here, we assume that each slot can be further divided into L subslots and each SU uses the same strategy of channel sensing and access during all L subslots, i.e., same $\tilde{p}_s$ and $\tilde{p}_a$. During each time slot t or t', the SU computes the approximated average utilities $\widetilde{\mathbb{U}}_s(t)$, $\widetilde{\mathbb{U}}_{\overline{s}}(t)$, $\widetilde{\mathbb{U}}_a(t')$ and $\widetilde{\mathbb{U}}_{\overline{a}}(t')$ by

$$\widetilde{\mathbb{U}}_s(t) = \frac{\sum_{g=1}^{n} U_{sa}(t,g) + \sum_{h=1}^{m} U_{s\overline{a}}(t,h)}{n+m}, \quad (34)$$

$$\widetilde{\mathbb{U}}_{\overline{s}}(t) = \frac{\sum_{g=1}^{n} U_{sa}(t,g) + \sum_{h=1}^{m} U_{s\overline{a}}(t,h)}{n+m}, \quad (35)$$

$$\widetilde{\mathbb{U}}_a(t') = \frac{\sum_{g'=1}^{n} U_{sa}(t',g') + \sum_{h'=1}^{m} U_{\overline{s}a}(t',h')}{n+m}, \quad (36)$$

$$\widetilde{\mathbb{U}}_{\overline{a}}(t') = \frac{\sum_{g'=1}^{n} U_{s\overline{a}}(t',g') + \sum_{h'=1}^{m} U_{\overline{s}\overline{a}}(t',h')}{n+m}, \quad (37)$$

where g denotes the subslots when the SU chooses sensing and accessing, h denotes the subslots when the SU chooses sensing but not accessing, while g' and h' denote the subslots when SUs observe that the PU is absent. At the end of each slot t, the SU adjust its strategies according to (32) and (33). Thus, each SU can gradually learn the ESS. Based on (32)-(37), each SU can gradually learn the ESS. In Algorithm 1, we summarize the detailed procedures of the proposed distributed learning algorithm. In the following, we will examine the effectiveness of the learning algorithm through simulations.

Algorithm 1 A Distributed Learning Algorithm for ESS.

```
• Given the time slot index t = 0, the SU initializes its
  strategy p̃_s and p̃_a with p̃_s(0) and p̃_a(0).
for each time slot t do
  for n = 1 : L do
    • Sense the primary channel with probability p̃_s(t).
    • Exchange its sensing data with others on the
      signalling channel.
    if The SU observes that the PU is absent then
      • Access the sensed channel with probability p̃_a(t).
      • Estimate the average utilities of sensing and not
        sensing using (34) and (35).
      • Estimate the average utilities of accessing and
        not accessing using (36) and (37).
    else
      • Do not access the sensed channel.
      • Estimate the average utilities of sensing and not
        sensing using (34) and (35).
    end if
  end for
  • Update the probability of sensing and accessing,
    p̃_s(t + 1) and p̃_a(t + 1), using (32) and (33).
end for
```

IV. Simulation Results

In this section, we conduct simulations to verify the effectiveness of our analysis. All the parameters used in the simulation are listed in Table I. We simulate the proposed learning algorithm with 20 SUs, and adjust the value of the reward R to see which ESS the system will converge to.

TABLE I
Parameters used in the simulation.

Parameter	Value	Parameter	Value
p_0	0.9	P_d	0.9
γ	-15dB	T_s	10ms
T_a	100ms	λ	1MHz
B	8MHz	SNR	-10dB
INR	-20dB	E_1	0.03mw/bit
E_2	0.5mw/s	E_3	2mw/s

A. Convergence of ESS

In Fig. 1, we show the convergence of the population states p_s and p_a, where the reward to SUs who only contribute to sense but do not access the channel, R, is set as 40, 100 and 400, respectively. In the simulation, we set $p_s = 0.1$ and $p_a = 0.9$ at the beginning, which means a large portion of SUs access the primary channel without contributing to channel sensing. From Fig. 1, it can be seen that with our scheme, SUs will quickly give up such an undesired strategy, and the system finally converges to different ESSs under different R.

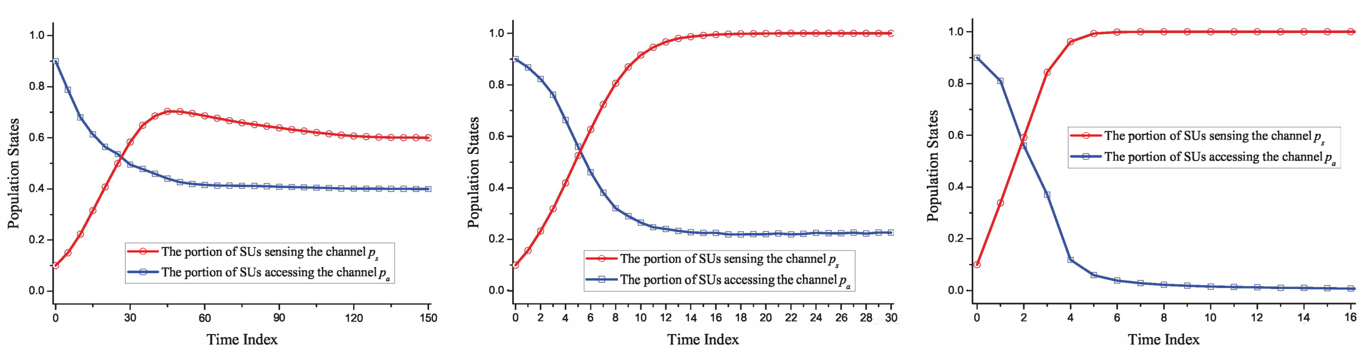
In Fig. 1-(a), when the reward $R = 40$, the ESS is $(p_s^*, p_a^*) = (0.6, 0.4)$ which is corresponding to the case that part of SUs sense and access the primary channel. In such a case, p_s converging to 0.6 is because 60 percents of SUs cooperatively sensing the channel can already achieve a relatively low false-alarm probability, and p_a converging to 0.4 is because more than 40 percents of SUs simultaneously accessing the channel will severely impair the throughput of each other.

In Fig. 1-(b), when the reward $R = 100$, the ESS is $(p_s^*, p_a^*) = (1, 0.25)$ which is corresponding to the case that all SUs sense but part of them access the primary channel. Although the false-alarm probability is already low enough when $p_s = 0.6$, the increasing of the reward R enhances the utility of sensing but not accessing, which attracts more SUs to sense but fewer SUs to access the primary channel.

In Fig. 1-(c), when the reward $R = 400$, the ESS is $(p_s^*, p_a^*) = (1, 0)$ which is corresponding to the case that all SUs sense but no one accesses the primary channel. In such a case, the reward R is too high that each SU can already obtain relatively high utility from just sensing without accessing the channel. Therefore, we can see that R should be properly set so that the network converges to a desired ESS.

B. Stability of ESS

In order to verify the stability of the ESS, we let SUs deviate from the equilibrium when the system is at ESS. As shown in Fig. 2, we first let SUs deviate from cooperative sensing by setting $p_s = 0.1$ at $t = 200$. It can be seen that both p_s and p_a return back to the ESS quickly after the perturbation. We can also see that p_a increases a little when p_s falls to 0.1. This is because a huge reduction of p_s leads to the decrease of both $\mathbb{U}_a$ and $\mathbb{U}_{\overline{a}}$. If the reduction $\Delta\mathbb{U}_a < \Delta\mathbb{U}_{\overline{a}}$, $\dot{p}_a$ will be larger



(a) $R = 40$. (b) $R = 100$. (c) $R = 400$.

Fig. 1. Population states of the joint spectrum sensing and access evolutionary game.

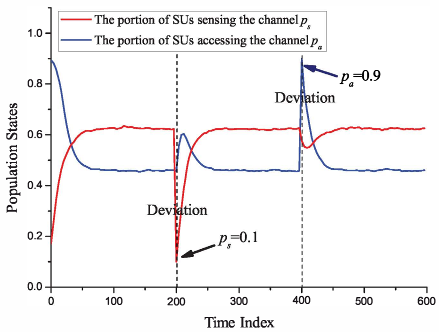


Fig. 2. Stability of ESS with $R = 50$.

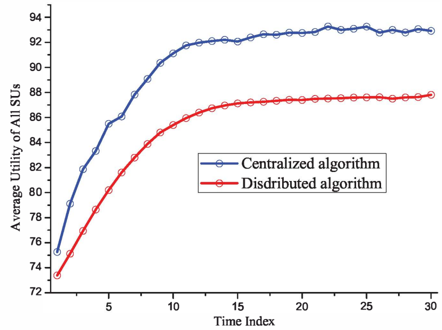


Fig. 3. Utility comparison.

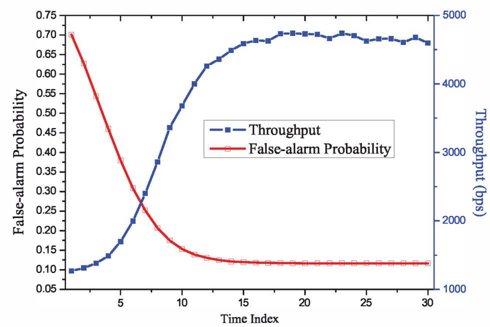


Fig. 4. Sensing and access performances.

than 0 according to (2), which results in the increasing of p_a. When $t = 400$, we let SUs deviate from the equilibrium again by setting $p_a = 0.9$. In such a case, the utility from channel access is extremely low and SUs will not sense and access the channel. That is why p_s begins to drop down when p_a is set to be 0.9.

C. Performance Evaluation

We first compare the performance of our distributed learning algorithm with that of the centralized algorithm. In the centralized model, there is a data center in charge of collecting each SU's utility information in each slot and globally adjusting SUs' strategies p_s and p_a in the next time slot. Fig. 3 shows the comparison results in terms of the average utility of all SUs, from which we can see that the gap between our distributed algorithm and the centralized one is about 6%. Nevertheless, the centralized algorithm requires all SUs to truthfully report their private utility information, while our distributed algorithm does not.

We further conduct simulation to evaluate the performance of our joint channel sensing and access algorithm. Fig. 4 shows the performances of SUs' false-alarm probability and throughput during the ESS convergence process. We can see that along with the system converging to the ESS, the false-alarm probability gradually tends to the lowest limit, while the throughput gradually achieve the highest limit. Moreover, from Fig. 1 and Fig. 4, we can see that the false-alarm probability is a decreasing function in terms of p_s and SUs' throughput is a decreasing function in terms of p_a, which is consistent with the results proved in *Lemma 1*.

V. Conclusion

In this paper, we analyzed how SUs should cooperate with each other in the joint spectrum sensing and access problem using evolutionary game theory. Through solving the joint replicator dynamics equations of channel sensing and accessing, we derived different ESSs under different conditions. Based on the nature selection theory, we proposed a distributed learning algorithm that enable the SUs to achieve the ESSs purely based on their own utility histories. From simulation results, we can see that by adjusting the reward to the contributors, the population states of the network will converge to the desired ESS.

References

[1] B. Wang and K. J. R. Liu, "Advances in cognitive radios: A survey," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 5–23, 2011.

[2] S. Shankar, C. Cordeiro, and K. Challapali, "Spectrum agile radios: utilization and sensing architectures," in *Proc. IEEE DySPAN*, 2005, pp. 160–169.

[3] H. Tang, "Some physical layer issues of wide-band cognitive radio systems," in *Proc. IEEE DySPAN*, 2005, pp. 151–159.

[4] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. IEEE DySPAN*, 2005, pp. 131–136.

[5] E. Visotsky, S. Kuffner, and R. Peterson, "On collaborative detection of TV transmissions in support of dynamic spectrum sharing," in *Proc. IEEE DySPAN*, 2005, pp. 338–345.

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

[7] C. Jiang, Y. Chen, K. J. R. Liu, and Y. Ren, "Renewal-theoretical dynamic spectrum access in cognitive radio networks with unknown primary behavior," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 406–416, 2013.

[8] Z. Ji and K. J. R. Liu, "Dynamic spectrum sharing: a game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88–94, 2010.

[9] A. A. El-Sherif and K. J. R. Liu, "Joint design of spectrum sensing and channel access in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1743–1753, 2011.

[10] W. Saad, Z. Han, R. Zheng, A. Hjorungnes, T. Basar, and H. V. Poor, "Coalitional games in partition form for joint spectrum sensing and access in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 195–209, 2012.

[11] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*. Cambridge, MA: MIT Press, 2003.

[12] 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.

[13] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. Veh. Tech.*, vol. 58, no. 4, pp. 2008–2017, 2009.

[14] X. Chen and J. Huang, "Evolutionarily stable open spectrum access in a many-users regime," in *Proc. IEEE GLOBECOM*, 2011, pp. 1–5.

[15] Y. Chen, B. Wang, W. S. Lin, Y. Wu, and K. J. R. Liu, "Cooperative peer-to-peer streaming: an evolutionary game-theoretic approach," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 20, no. 10, pp. 1346–1357, 2010.

[16] P. Baronti and etc, "Wireless sensor networks: A survey on the state of the art and the 802.15.4 and zigbee standards," *Computer Communication*, vol. 30, no. 7, pp. 1655–1695, 2007.

[17] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communication*, vol. 4, no. 3, pp. 40–62, 2011.

[18] C. Jiang, Y. Chen, Y. Gao, and K. J. R. Liu, "Supplementary information for evolutionary game for joint spectrum sensing and access in cognitive radio networks," [Online]. Available: http://www.sig.umd.edu/jcx/GLOBECOM2013SuppInfo.pdf.

[19] R. Fisher, *The Genetical Theory of Natural Selection*. Clarendon Press, 1930.