

Evolutionarily Stable Opportunistic Spectrum Access in Cognitive Radio Networks

Li Xu¹, He Fang^{1,*}, Zhiwei Lin²

¹Fujian Provincial Key Laboratory of Network Security and Cryptology, School of Mathematics and Computer Science, Fujian Normal University, Fuzhou, Fujian, China, 350117.

²School of Computing and Mathematics, Ulster University, BT37 0QB UK.

*fanghe.fjnu@hotmail.com.

Abstract: In order to fully utilize limited spectrum resources of multiple channels and multiple radios in cognitive radio networks, we propose a potential game model for opportunistic spectrum access based on both accurate and inaccurate spectrum state estimation with considering the interference constraints of licensed users. Three algorithms are proposed to achieve equilibrium of the proposed game. Firstly, assuming spectrum sensing results are accurate, a joint strategy fictitious play based channel selection algorithm with incomplete information is presented, and it can achieve a pure Nash equilibrium of the proposed game. Secondly, in order to make the outcomes of game robust, an evolutionary spectrum access mechanism with complete information is introduced by using evolutionary game theory based on inaccurate spectrum state estimation so that evolutionary stable strategy (ESS) can be achieved. Finally, with incomplete network information, a distributed learning algorithm is proposed to achieve a mixed Nash equilibrium, which is proved to be an ESS. Simulation results show that these algorithms can significantly improve spectrum allocation efficiency while reducing mutual collision.

1. Introduction

The pervasive applications of mobile and wireless technologies overwhelm the limited spectrum supply. The need for smarter and better spectrum usage is becoming urgent due to the limited available spectrum. Therefore, it is necessary to exploit the existing wireless spectrum opportunistically so that the limited spectrum can be re-used more effectively. Research in the area has shown that dynamic spectrum access and cognitive radio technology are promising approaches to facilitating efficient use of fragmented and under-utilized spectrum [1]. The *dynamic spectrum access* (DSA for short), also named as *opportunistic spectrum access* (OSA for short), allows unlicensed users (or secondary users, SUs for short) to dynamically access the unoccupied bands released from licensed users (or primary users, PUs for short) on opportunistic basis [2]. In OSA, when a secondary user (SU) detects one or multiple spectrum holes, released from any PUs, it reconfigures its transmission parameters to utilize the available spectrum. After that, the SUs will have to release their spectrum belonging to PUs, and switch themselves to inactive mode whenever any PUs becomes active. In order to improve the performance of the OSA in *cognitive radio networks* (CRNs for short) and to reduce interference, *game theory* has been applied and efficient algorithms are keys to schedule the process of OSA to allow spectrum usage more efficiently.

There are many open problems and interesting approaches in this area, such as decision-theoretic

solutions for channel selection and access strategies for OSA system, spectrum assignment in CRNs with interference, and the criteria for selecting the most suitable portion of the spectrum [3, 4]. The study in [5] shows that cognitive radio can coexist with multiple parallel wireless local ad-hoc network channels by interference constraint. In [6], it is shown that the myopic sensing policy has a simple robust structure to reduce channel selection to a round-robin procedure without the need for knowing channel transition probabilities. However, most of the existing models need both complete channel statistics and state, collected by information exchange via a common control channel, which leads to significant communication overhead and energy consumption. This is very inefficient in real network. Therefore, a continuous-time Markov models is used to estimate channel statics and state for dynamic spectrum access in open spectrum wireless networks, without having to exchange those information [7]. However, this work ignores the influence of PUs in the dynamic spectrum access, and fails to improve the energy efficiency. An optimization model for both maximizing end-to-end throughput per unit of energy consumption and minimizing the delay constraint specified for a data stream is introduced, and a novel distributive and self-organized heuristic channel assignment algorithm to this optimization model is also presented in [8]. The work in [9] presents a new network selection and channel allocation mechanism, in order to increase revenue by accommodating more SUs with regard to their preferences, while at the same time, without disrespecting the primary network operator's policies. An interesting optimization problem is formulated for minimizing the accumulated interference incurred to licensed users, however, it does not take into account the interactions between SUs. The authors in [10, 11, 12] consider the access problem in a multichannel opportunistic communication system with imperfect sensing. In these methods, the state of each channel evolves as an independent and identically distributed Markov process.

Game theory is an effective tool for modeling the interactions among independent decision makers, and it has been widely applied in cognitive radio network to improve random access and energy efficiency [13, 14, 15, 16, 17, 18, 19]. With game theory, dynamic spectrum sharing has several key aspects, needed to be considered, including analysis of network users' behaviors, efficient dynamic distributed design, and optimality analysis [13]. The model in [14] characterizes Nash equilibria of random access games, studies their dynamics, and proposes distributed algorithms to achieve Nash equilibria. A further study in [17] investigates two special cases of local interaction game, which can cope with drawbacks of centralized control and local influences. However, these models did not consider the interference constraints of PUs and dynamic of a game, and therefore, these models are ineffective in CRNs.

This paper steps further for those exiting problems in cognitive radio networks, aiming to evaluate the maximum achievable rate while minimizing collisions of a system under different access scenarios in CRNs. We first design a framework of opportunistic spectrum access based on game theory, which is named *OSAG*, to study the issues of how SUs compete and cooperate in multiple radios and multiple channels in CRNs based on accurate spectrum state estimation. The pure-strategy Nash equilibrium (NE) for the SUs is investigated by proposing a joint strategy fictitious play (JSFP) based channel selection algorithm. Then for the situation where spectrum sensing results are not always available, we propose a framework for analyzing spectrum access strategies by estimating inaccurate spectrum state, and by analyzing the SUs' achievable capacity and collisions with other users (including PUs and other SUs). In order to make the outcomes of game robust, we investigate an evolutionary spectrum access algorithm by using evolutionary game theory, which is a centralized learning algorithm. However, centralized learning algorithm does not fit into many real distributed environments, and therefore we turn to a new distributed learning algorithm based

on stochastic learning automata in order to achieve mixed NE. We prove that the mixed NE is an evolutionary stable strategy (ESS). We summarize our contributions as:

1. A non-cooperative potential game among the SUs is presented to fully utilize rare spectrum resources and to improve energy efficiency in CRNs. The game also considers complete and incomplete information, available in CRNs. The pure-strategy and mixed-strategy Nash equilibria for the SUs are investigated;
2. As spectrum sensing results are not always available, we propose a framework, firstly for analyzing spectrum access strategies based on both accurate and inaccurate spectrum state estimation, and secondly for analyzing the SUs' achievable capacity and collision with other users (including both all PUs and other SUs);
3. In order to maximize the performance of spectrum sharing, three learning algorithms are proposed so that SUs can achieve the Nash equilibria of the OSAG game, if the network is not on a steady state. The algorithms are evaluated by numerical experiments and the results show that the proposed scheme can better improve the performance for SUs, compared to other schemes. In the experiments, we find that our method gains higher throughput than the random scheme, and the random access game [14].

The rest of this paper is organized as follows. Section 2 introduces the system model and symbols used in the paper. In Section 3, a potential game model based on accurate spectrum state estimation, which is used to investigate the properties of Nash equilibrium, is presented, and a distributed learning algorithm for achieving the pure NE is proposed. Section 4 analyzes the proposed game based on inaccurate spectrum state estimation, and introduces two learning solutions based on inaccurate spectrum state estimation for achieving evolutionary stable strategy. Simulation results are presented in Section 5. Finally, this paper is concluded in Section 6.

2. Definitions and System Model

Table 1 Notations and Definitions

Notation	Definition
n	Number of channels in the cognitive radio network.
N	Set of the channels, where $N = \{1, 2, \dots, n\}$.
m	Number of the SUs in the cognitive radio network.
M	Set of SUs, where $M = \{1, 2, \dots, m\}$.
\wp	Transmit power of the SUs.
g	Complex channel gain of the secondary link.
\mathbf{X}^j	Channel allocation decision vector of SU j . $\mathbf{X}^j = (X_1^j, X_2^j, \dots, X_n^j)$, where X_i^j is the probability of SU j choosing channel i , $i \in N$ $j \in M$.
M_i^j	Number of SUs competing with SU j for the channel i during a decision slot, $i \in N$ $j \in M$.
O	Probability of a SU accessing a channel successfully.
Q	Probability of a SU being involved in a collision.
η	Additive white Gaussian noise at the SUs.
R	Achievable rate of SUs.
I	Interference threshold.
ω	False-alarm probability.
ϵ	Miss-detection probability.

In this paper, we assume a cognitive radio network consists of a primary network with multiple licensed users (or primary users/PUs) and a secondary network with multiple unlicensed users (or secondary users/SUs). The set of SUs is denoted as $M = \{1, 2, \dots, m\}$. Assuming that there are n available channels in cognitive radio network, the set of channels is denoted as $N = \{1, 2, \dots, n\}$. In CRNs, if the authorized spectrum is not fully utilized by PUs all the time, the PUs could lease the unused channels to SUs for monetary gains. Since the unlicensed spectrum is busy for SUs,

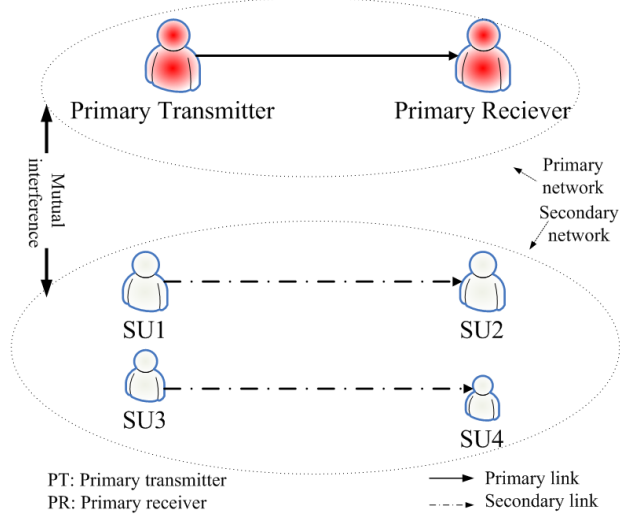


Fig. 1. The cognitive radio network consisting of a primary network and a secondary network with considering the interference between these two networks.

SUs may wish to pay for some unused channels from PUs for more communication gains. We also assume each SU is selfish, and it can sense all channels from PUs but can only access one channel after sensing part of the channels. As shown in Fig.1, where the solid and dashed lines represent the primary link (primary transmitter-primary receiver) and the secondary link (secondary transmitter-secondary receiver), respectively. We denote the transmit power of SU j as \wp^j , for $\forall j \in \{1, \dots, m\}$, and the complex channel gain of secondary link in channel i as g_i^j , which satisfies $g_i^j \in (0, 1]$, for $\forall j \in \{1, \dots, m\}$ and $\forall i \in \{1, \dots, n\}$. Note that the communication activity of the SUs creates interference at the PUs. The maximum tolerable average interference at the primary transmitter is called interference threshold [6], denoted by I . The interference threshold is introduced by the Federal Communications Commission (FCC) [20] which indicates the tolerable interference level at the primary receiver imposed by the operation of the secondary service. Technically, as long as the interference threshold constraint is satisfied, the spectrum is underutilized. Therefore, the transmit power at SUs should be adjusted so that the interference received at primary receiver is always kept below I . Different from underlay spectrum access strategy, the overlay spectrum access strategy ignores the interference threshold during inactive periods of the channels. However, the main requirement for such approach is to have an online detection of the idle and busy periods with an acceptable level of accuracy. So in this paper, we develop our spectrum access method in two different cases which are based on accurate and inaccurate spectrum state estimation. The notations and definitions are summarized in Table 1.

2.1. Spectrum Access based on Accurate Spectrum State Estimation

In a network, each SU estimates its conditional collision probability and adjusts its channel access probability. We assume that there are $M_i^j - 1$ other SUs (excluding j) competing channel i with SU j , where $i \in N$ and $j \in M$. It is well known that the probability of successfully accessing an idle channel i for SU j is

$$O_i^j = (1 - X_i^1) \cdots (1 - X_i^{j-1}) X_i^j (1 - X_i^{j+1}) \cdots (1 - X_i^{M_i^j}), \quad i \in N, j \in M, \quad (1)$$

where X_i^j is the probability of SU j choosing channel i . Then the probability of node j being involved in a collision with other SUs on channel i in a decision slot is

$$Q_i^j = 1 - O_i^j, \quad i \in N, \quad j \in M. \quad (2)$$

According to [26], the average achievable rate of SU j in channel i is given as follows

$$R_i^j = O_i^j \log\left(1 + \frac{g_i^j \wp^j}{\eta}\right), \quad i \in N, \quad j \in M, \quad (3)$$

where η is additive white Gaussian noise in SUs, and we assume all channels have identical values for η .

The main objective of this paper is to evaluate the maximum achievable capacity of SUs. Therefore, the SUs in M need to decide the optimal spectrum access actions to maximize their achievable capacity and a spectrum access and power control strategy for SU j becomes an optimization problem as

$$\begin{aligned} \max_{(X_i^j, \wp^j)} R_i^j &= \max_{(X_i^j, \wp^j)} O_i^j \log\left(1 + \frac{g_i^j \wp^j}{\eta}\right), \\ \text{s.t. } X_i^j &\in [0, 1], \quad \wp^j \in [0, \wp_{max}], \quad i \in N, \quad j \in M, \end{aligned} \quad (4)$$

where \wp_{max} is the maximum transmit power of SUs.

2.2. Spectrum Access based on Inaccurate Spectrum State Estimation

In practice, the spectrum sensing results are not always accurate. Therefore, if the spectrum access system is not designed to manage the miss-detection incidents, the collision between the primary and secondary users transmissions can potentially reduce the actual capacity of the primary network. We use ϵ to denote the probability of miss-detection, i.e., the probability of estimating the spectrum state as “idle”, while it is busy. Similarly, let ω be the probability of false-alarm, i.e., the probability of estimating the spectrum state as “busy”, while it is idle. Let $S \in \{0, 1\}$ denote the channel state, and S' denote the sensing outcome, where $S' = 0$ (1) means the channel i is sensed to be busy (idle). Using such notation, the probability of miss-detection and the probability of false-alarm can be shown as

$$\begin{aligned} \omega &= Pr\{S' = 0 | S = 1\}, \\ \epsilon &= Pr\{S' = 1 | S = 0\}. \end{aligned} \quad (5)$$

From each PU's point of view, the best case is $\epsilon = 0$, while from the SU's point of view, it should be $\omega = 0$. However, due to some practical limitations, i.e., the spectrum sensing techniques, a tolerance on miss-detection and false-alarm probabilities is acceptable. Hence, in the case of inaccurate spectrum state estimation, the communication activity of the secondary transmitter creates interference at the primary receiver. In order to express the communication activity of SUs and PUs, we show the schematic of the spectrum sharing system in the Fig.2. As shown in Fig.2, the secondary transmitter creates interference at the primary receiver through the channel with the complex channel gain of link SU j to primary receiver, g_{jp} . Besides, in the case of miss-detection, the primary transmitter also creates interference at the secondary receiver through the channel with the complex channel gain of link primary transmitter to secondary receiver, g_{pj} .

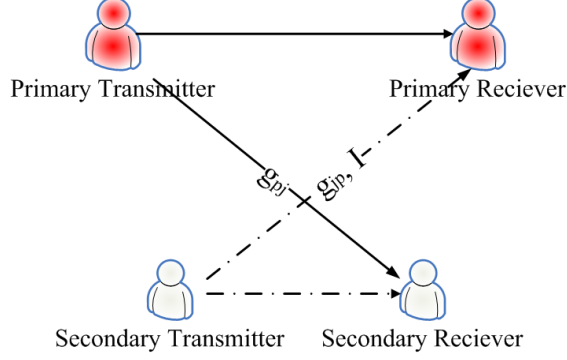


Fig. 2. The schematic of the spectrum sharing system based on inaccurate spectrum state estimation.

Table 2 shows the interference at the PUs imposed by the SUs based on inaccurate spectrum state estimation. As shown in Table 2, only in the case of $S = 0$ and $S' = 1$, the SU j causes interference at primary receiver. According to the Fig.2 and Table 2, we can calculate the average interference at primary receiver imposed by the SU j as

$$\epsilon g_{jp} \wp^j. \quad (6)$$

Table 2 The interference imposed by the SUs based on inaccurate spectrum state estimation

Sensing result S Channel state S'	0 0	0 1	1 0	1 1
Probability	$1 - \epsilon$	ω	ϵ	$1 - \omega$
Transmit power of SU j	0	0	\wp^j	\wp^j
Complex gain of the channel	0	0	g_{jp}	g_{jp}
Does SU j cause interference at primary receiver?	No	No	Yes	No

We denote the interference threshold as I , hence, the transmitted power at secondary transmitters should be adjusted so that the interference received at primary receiver is lower than I . The constraint of the interference can be defined as

$$\epsilon g_{jp} \wp^j \leq I. \quad (7)$$

In the case of spectrum access based on inaccurate spectrum state estimation, the average achievable rate of SU j in channel i is

$$\widehat{R}_i^j = \widehat{O}_i^j \log\left(1 + \frac{g_i^j \wp^j}{\eta + \epsilon g_{pj} \wp_p}\right), \quad i \in N, j \in M, \quad (8)$$

where \wp_p is transmit power of PU, g_{pj} represents the channel with the complex channel gain of link primary transmitter to secondary receiver, and \widehat{O}_i^j is the probability of successfully accessing an idle channel i for SU j , and

$$\widehat{O}_i^j = (1 - \epsilon)(1 - X_i^1) \cdots (1 - X_i^{j-1}) X_i^j (1 - X_i^{j+1}) \cdots (1 - X_i^{M_i^j}), \quad i \in N, j \in M. \quad (9)$$

In the case of miss-detection, the successful access probability of SUs is decreased by ϵ . Similarly, in the case of false-alarm, the access probability of the SUs decreases by ω . Then the

optimization problem in (4) can be rewritten as

$$\begin{aligned} \max_{(X_i^j, \wp^j)} \widehat{R}_i^j &= \widehat{O}_i^j \log\left(1 + \frac{g_i^j \wp^j}{\eta + \epsilon g_{pj} \wp_p}\right), \\ \text{s.t. } \epsilon g_{jp} \wp^j &\leq I, X_i^j \in [0, 1], \wp^j \in [0, \wp_{max}], i \in N, j \in M. \end{aligned} \quad (10)$$

Remark 1: The channel allocation and power control problems formulated in (4) and (10) are difficult to solve since the same spectrum access strategy of multiple SUs may cause collisions in the secondary network. Therefore, an unrational access strategy may cause transmission failure. To overcome this, in the rest of this paper, a potential game model is proposed to allocate SU's power and spectrum access strategy among all SUs to maximize system's revenue in both distributed and centralized fashion.

3. Game Theoretic Spectrum Access

In order to cope with the problem of scarce spectrum resources with considering interference constrain, we study opportunistic spectrum access joint with power control based on game theory. In this section, we first design a framework of opportunistic spectrum access based on game theory (named OSAG) to demonstrate competition and cooperation among multiple radios and multiple channels in CRNs with accurate spectrum state estimation. Then a distributed algorithm is proposed to achieve pure NE. Finally, we develop our analysis based on inaccurate spectrum sensing results.

3.1. OSAG Model based on Accurate Spectrum State Estimation

In order to achieve high efficiency (high throughput and low collision) and better fairness, we propose an opportunistic spectrum accessing strategy to stabilize the network into a steady state. We first show the definition of OSAG model with pure strategy as follows:

Definition 1: The opportunistic spectrum access based on game theory (OSAG) model consists:

Player set: The set of secondary users in the cognitive radio network M ;

Strategy set: $N \cup [0, \wp_{max}]$, $(x^j, \wp^j) \in N \cup [0, \wp_{max}]$ is denoted as node j 's action, $j \in M$, where $N = \{1, 2, \dots, n\}$ is the set of channels, and $[0, \wp_{max}]$ represents the transmit power set of SUs;

Utility set: $\{u^j, j = 1, \dots, m\}$, where u^j represents the utility function of the SU j .

Obviously, $\mathbf{X}^j = (X_1^j, X_2^j, \dots, X_n^j)^\dagger$, where $X_k^j \in \{0, 1\}$ for $1 \leq k \leq n$ denotes a channel is chosen or not, is a mixed strategy of OSAG game. If node j chooses channel x^j , $\mathbf{X}^j = (X_1^j, X_2^j, \dots, X_n^j)^\dagger = (0, \dots, 0, 1, 0, \dots, 0)^\dagger$, where the corresponding entry $x^j = 1$. For each SU, a utility function is defined as the gain of signal transmission minus the cost of collision, hence, the instantaneous utility of node j in a decision slot is designed as

$$u^j(x^j, \wp^j) = aR^j(x^j, \wp^j) - bQ_i^j = O_{x^j}^j \left(a \log\left(1 + \frac{g_{x^j}^j \wp^j}{\eta}\right) + b \right) - b, \quad (11)$$

where a is the gain per unit of rate, and b is the cost per unit collision. In Equation (11), $aR^j(x^j, \wp^j)$ represents the income of SU based on the achievable rate, and bQ_i^j is the cost of contention. For each SU, it wants to improve its capacity of communication and reduce its collision with other

SUs. Hence, it aims at maximizing its utility in (11). The utility of system, which is defined as the minimum utility of all the SUs:

$$u'(\mathbf{x}, \wp) = \min_{j \in M} u^j(x^j, \wp^j), \quad (12)$$

where $\mathbf{x} = (x^1, \dots, x^m)^\dagger$ and $\wp = (\wp^1, \dots, \wp^m)^\dagger$ are decision vectors of SUs in M . Then the system's objective based on accurate spectrum state estimation is to find the optimal channel selection \mathbf{x} and transmit power \wp of SUs such that the system utility is maximized. Formally,

$$\begin{aligned} (\mathbf{x}, \wp)_{opt} &= \arg \max u'(\mathbf{x}, \wp), \\ \text{s.t. } x^j &\in N, \wp^j \in [0, \wp_{max}], j = 1, \dots, m. \end{aligned} \quad (13)$$

The property of OSAG model with pure strategy is characterized by the following theorem:

Theorem 1. The OSAG model is an ordinal potential game, and the system utility maximization problem in (13) constitutes a pure strategy Nash equilibrium (NE) of the proposed game.

Proof. We construct the potential function as follows

$$\phi(x^j, \wp^j) = u^j(x^j, \wp^j) + b = O_{x^j}^j(a \log(1 + \frac{g_{x^j}^j \wp^j}{\eta}) + b). \quad (14)$$

For every channel, we don't consider the differences between different channels in this paper, then we have

$$u^j(x^j, \wp^j) - u^{j'}(x^{j'}, \wp^{j'}) = \phi(x^j, \wp^j) - \phi(x^{j'}, \wp^{j'}). \quad (15)$$

Finally we get $\phi(x^j, \wp^j) - \phi(x^{j'}, \wp^{j'}) \geq 0 \Leftrightarrow u^j(x^j, \wp^j) - u^{j'}(x^{j'}, \wp^{j'}) \geq 0$.

Besides, ordinal potential game belongs to potential games, which has been widely applied to wireless communication systems. Potential game exhibits several good properties and the most important two aspects are as follows: 1) Every potential game has at least one pure strategy NE; 2) Any global or local maximization of the potential function constitutes a pure strategy NE. Hence, the theorem holds based on the above properties. \square

Remark 2: It is difficult to obtain optimal solution for Equation (13), since 1) there is no control center for SUs, 2) the number of SUs is unknown, 3) the spectrum sensing mechanisms may be imperfect and 4) there is no information exchange between SUs.

3.2. A Distributed Learning Algorithm based on Joint Strategy Fictitious Play for Achieving the Pure NE

In this subsection, we assume that the spectrum sensing results are accurate. In this case, the interference at PUs caused by SUs equals to zero. Therefore, the optimization problem for Equation (13) turns into

$$\begin{aligned} \mathbf{x}_{opt} &= \arg \max u'(\mathbf{x}, \wp_{max}), \\ \text{s.t. } \mathbf{x} &= (x^1, \dots, x^m)^\dagger, \text{ where } x^j \in N, j = 1, \dots, m. \end{aligned} \quad (16)$$

The dynamics of OSAG studies how interacting players could converge to a Nash equilibrium, it is a difficult problem in general. In the setting of random access, SUs can observe the outcome

(collision or throughput) of the actions of others, but do not have direct knowledge of other SUs' actions and payoffs. Motivated by the underlying idea of joint strategy fictitious play (JSFP), we propose a JSFP based channel selection (JCS) algorithm, with which the SUs can finally adjust their behaviors towards a pure NE point with incomplete information. In JCS algorithm, each player tracks the empirical frequencies of the joint actions of all other players.

Define the empirical frequency, $v_{x^j}^j[t]$, as the percentage of stages at which player j has chosen the action $x^j[t]$ up to time $t - 1$, i.e.,

$$v_{x^j}^j[t] = \frac{1}{t} \sum_{k=0}^{t-1} u^j[t](x^j[k]), \quad (17)$$

where $\mathbf{v}^j[t] = (v_1^j[t], \dots, v_n^j[t])^\dagger$ represents the empirical frequency vector. Note that the dimension of $\mathbf{v}^j[t]$ is n . $v_{x^{-j}}^{-j}[t]$ is the percentage of stages at which other players have chosen the joint action profile $\mathbf{x}^{-j}[t]$ up to time $t - 1$, i.e.,

$$v_{x^{-j}}^{-j}[t] = \frac{1}{t} \sum_{k=0}^{t-1} u^j(\mathbf{x}^{-j}[k]), \quad (18)$$

where $\mathbf{x}^{-j}[k] = (x^1[k], \dots, x^{j-1}[k], x^{j+1}[k], \dots, x^m[k])^\dagger$. $\mathbf{v}^{-j}[t]$ is denoted as the empirical frequency vector formed by the components $\{v_{x^{-j}}^{-j}[t]\}_{x^{-j}}$. Note that the dimension of $\mathbf{v}^{-j}[t]$ is the cardinality $|\times_{j \neq k} N^k|$.

JCS requires that each agent maintains a hypothetical payoff for each action $x^j[t]$ such that

$$V^j(x^j[t], \mathbf{v}^{-j}[t]) = \sum_{x^{-j}} V^j(x^j[t], \mathbf{x}^{-j}[t]) v_{x^{-j}}^{-j}[t]. \quad (19)$$

Note that this hypothetical payoff can be computed recursively and a player only needs to access the payoff for alternative actions at each time step. For the game of spectrum access, the pure Nash equilibrium $\{x_{opt}^j\}$, $j \in M$ satisfies

$$x_{opt}^j = \operatorname{argmax} V^j(x^j[t], \mathbf{v}^{-j}[t]). \quad (20)$$

Substituting (18) into (19) results in

$$V^j(x^j[t], \mathbf{v}^{-j}[t]) = \frac{1}{t} \sum_{\eta=0}^{t-1} V^j(x^j[t], x_k^{-j}). \quad (21)$$

We denote $V^j(x^j[t], \mathbf{v}^{-j}[t])$ as $V_{x^j}^j[t]$, then we have

$$V_{x^j}^j[t+1] = \frac{t}{t+1} V_{x^j}^j[t] + \frac{1}{t+1} V^j(x^j, \mathbf{x}^{-j}[t]). \quad (22)$$

Lemma 1. In a finite n -person game, if at any time $t > 0$ the joint action $\mathbf{x}[t]$ generated by a JSFP is: 1) A pure Nash equilibrium; 2) The action $x^j[t] \in BR^j[t] := \operatorname{argmax} V^j(x^j[t], \mathbf{v}^{-j}[t])$ for all players, then $x^j[t+k] = x^j[t]$ for all $k > 0$ and $j \in M$ [21].

Remark 3: The OSAG model provides a general analytical framework to model a large set of system-wide QoS models via the specification of local utility functions. Besides, system-wide fairness or service differentiation can be achieved in a distributed manner as long as each node executes JCS algorithm.

Algorithm 1 JSFP based Channel Selection (JCS) Algorithm

Given the number of channels n , and the complex channel gain of secondary link g .

1. Initialize : $t = 0$

- each SU j randomly take action $x^j[0]$ from N , $j \in M$;

2. Learn :

- calculate the utility values of SU j and the system according to (11) and (12);
 - update the recursion (22);
 - if $x^j[t-1] = x_{opt}^j$, then node j chooses action $x^j[t]$ satisfies $x^j[t] = x^j[t-1]$;
 - otherwise, j chooses an action $x^j[t]$ at time t according to the recursion (22) and (20), until $x^j[t] = x^j[t-1] = x_{opt}^j$.
-

4. Learning Solutions for Achieving Evolutionary Stable based on Inaccurate Spectrum State Estimation

In this section, we study the case when inaccurate spectrum sensing happens in CRNs. We first conduct an analysis of the proposed game based on inaccurate spectrum state estimation. Then a learning algorithm based on evolutionary game theory with complete information is proposed to achieve the evolutionary stable strategy (ESS) of proposed game. However, this algorithm may lead to significant communication overhead, and then it may be impossible in some distributed systems. We thus propose an algorithm based on the stochastic learning automaton with incomplete information to achieve the mixed NE of OSAG model for distributed spectrum access.

4.1. Analysis of the Proposed Game based on Inaccurate Spectrum State Estimation

According to the constraint of the interference in (7), the optimization problem in Eq. (13) becomes

$$\begin{aligned} (\mathbf{x}, \boldsymbol{\wp})_{opt} &= \arg \max u'(\mathbf{x}, \boldsymbol{\wp}), \\ s.t. \quad \epsilon g_{jp} \wp^j &\leq I, \quad x^j \in N, \quad \wp^j \in [0, \wp_{max}], \quad j \in M, \end{aligned} \quad (23)$$

and the utility function of SU j is represented by

$$u^j(x^j, \wp^j) = a\widehat{R}^j(x^j, \wp^j) - b\widehat{Q}_{x^j}^j = \widehat{O}_{x^j}^j \left(a \log \left(1 + \frac{g_{x^j}^j \wp^j}{\eta + \epsilon g_{pj} \wp_p} \right) + b \right) - b. \quad (24)$$

In the case of miss-detection, the successfully access probability of SUs is decreased by ϵ . Similarly, in the case of false-alarm, the access probability of the SUs decreases by ω . In order to maximize its own utility with the interference constrain, each SU adjusts its individual power-allocation and spectrum access strategy. We find that:

Theorem 2. The proposed game based on inaccurate spectrum state estimation has an optimal strategy (x_{opt}^j, \wp_{opt}^j) such that $u^j(x^j, \wp^j)$ is maximized.

Proof. Since $u^j(x^j, \wp^j)$ is a continuous function on the compact set $[0, \wp_{max}]$, it can achieve its maximum value at some $\wp \in [0, \wp_{max}]$ [23]. The first order partial derivative of $u^j(x^j, \wp^j)$ with respect to \wp^j , $j \in M$ is shown as

$$\frac{\partial u^j(x^j, \wp^j)}{\partial \wp^j} = \frac{a\widehat{O}_{x^j}^j g_{x^j}^j (\eta + \epsilon g_{pj} \wp_p) \ln 2}{\eta + \epsilon g_{pj} \wp_p + g_{x^j}^j \wp^j} \geq 0. \quad (25)$$

It is easy to obtain that $u^j(x^j, \wp^j)$ is an increasing function on the compact set $[0, \wp_{max}]$. Then we can get the optimal power strategy based on the interference constrain in Eq. (7) as follows

$$\wp_{opt}^j = \frac{I}{\epsilon g_{jp}}, j \in M. \quad (26)$$

Therefore, there is a unique optimal power-allocation strategy of SU j to the optimization problem. Then for the given \wp_{opt}^j , probability of miss-detection ϵ and probability of false-alarm ω , we can easily prove that the proposed game with inaccurate spectrum sensing results is an ordinal potential game based on Theorem 1. More important, there is a unique x_{opt}^j for SU j , $j \in M$, to the optimal problem (23). □

Remark 4: Eq. (26) shows that for a given interference threshold I , a SU is allowed to transmit signals with a lower power if the probability of miss-detection ϵ is higher. Besides, from the SUs viewpoint, the ideal spectrum sensing procedure has false-alarm probability equal to zero. Hence, it is more difficult for SUs to handle the spectrum access strategies in the CRNs.

4.2. Evolutionary Learning Algorithm with Complete Information

In OSAG, the population state $\mathbf{y}[t] = (y_1[t], y_2[t], \dots, y_n[t])^\dagger$ describes the dynamics of the reproduction process, where for $i \in N$, $y_i[t]$ is the proportion of SUs in secondary network selecting channel i at time t . Specially, $y_{x^j}[t] = M_{x^j}^j/m$ represents the proportion of node j selecting a pure strategy x_j at time t . Then we design the replicator dynamics as follows

$$\widetilde{y}_{x^j}[t] = \beta \left(\frac{u^j(x^j[t], \mathbf{y}[t])}{\bar{u}(\mathbf{y}[t])} - 1 \right), \forall j \in M, \quad (27)$$

where $\bar{u}(\mathbf{y}[t])$ is the average utility of the population, $u^j(x^j, \mathbf{y}[t])$ is the utility of SU j , choosing strategy $x^j[t]$ based on the population $\mathbf{y}[t]$ at time t , and $\beta > 0$ is the rate of strategy adaptation. The replicator dynamics means that the strategy works better than the average will be promoted. The average utility of the network is defined as follows

$$\bar{u}(\mathbf{y}[t]) = \frac{1}{m} \sum_{j=1}^m u^j(x^j, \mathbf{y}[t]). \quad (28)$$

In order to make the outcomes of OSAG robust and to have good stability properties under an important sort of dynamic, an ESS is investigated as follows:

Definition 2: A pure strategy \mathbf{y}^{opt} is an ESS, if for all $\mathbf{y} \neq \mathbf{y}^{opt}$ in some vicinity of \mathbf{y}^{opt} , there is an expected payoff function $\pi(\mathbf{y}, \mathbf{y}^{opt})$ satisfying

- 1) $\pi(\mathbf{y}, \mathbf{y}^{opt}) \leq \pi(\mathbf{y}^{opt}, \mathbf{y}^{opt})$;
- 2) if $\pi(\mathbf{y}, \mathbf{y}^{opt}) = \pi(\mathbf{y}^{opt}, \mathbf{y}^{opt})$, then $\pi(\mathbf{y}, \mathbf{y}) \leq \pi(\mathbf{y}^{opt}, \mathbf{y})$.

To be specific, in condition 1) \mathbf{y}^{opt} is the best response strategy of the game, hence it is a NE; Condition 2) is interpreted as a stability condition. Considering the spectrum sensing results may be wrong, for the given miss-detection probability ϵ and false-alarm probability ω , we can easily obtain the optimal power strategy of SU j according to Eq. (26). Then system's utility $u'(\mathbf{x}, \wp)$ turns into a function of \mathbf{y} , and we redenote it as $u'(\mathbf{y})$. Hence the optimal problem Eq. (23) turns into

$$\mathbf{y}_{opt} = \arg \max u'(\mathbf{y}), \text{ s.t. } \epsilon g_p \wp_{opt} \leq I, \quad (29)$$

where $\wp_{opt} = \max_{j \in M} \wp_{opt}^j$ represents the maximum optimal power strategy of SUs in the secondary network.

Based on the evolutionary game theory and the concept of ESS, we propose a learning solution for achieving ESS in channel selection shown in Algorithm 2, which is based on the complete information.

Algorithm 2 Evolutionary Learning for Achieving ESS in Channel Selection (ELES) Algorithm

Given the number of channels n , the number of SUs m , interference threshold I , the complex channel gain of secondary link g , miss-detection probability ϵ and false-alarm probability ω , and the rate of strategy adaptation β .

1. Initialize : $t = 0$

- each SU $j \in M$ randomly take action $x^j[0]$;
- obtain the population state $\mathbf{y}[0]$;
- estimate the expected utility $u^j[0]$ and $u'[0]$ according to (11) and (12), respectively;

2. Learn : loop for each t and each SU

- broadcast the chosen channel to the other SUs through a common control channel;
 - receive the information of other SUs' channel selection and calculate the population state $\mathbf{y}[t]$;
 - calculate the utility $u^j(x^j[y], \mathbf{y}[t])$ according to (11) and the average utility of network $\bar{u}(\mathbf{y}[t])$ by (28);
 - if $u^j(x^j[y], \mathbf{y}[t]) < \bar{u}(\mathbf{y}[t])$, select another channel $x^j[t]$ with a probability: $\max\{\frac{u^j(x^j[t+1], \mathbf{y}[t])}{\bar{u}(\mathbf{y}[t])} - 1, 0\}$.
-

In general, the equilibrium of the replicator dynamics may not be an ESS [25]. The following theorem shows evolutionary stability of the equilibrium.

Theorem 3. The ELES algorithm converges to a equilibrium \mathbf{y}^{opt} such that SUs on different channels achieve the same expected utility shown as follow:

$$\begin{aligned} u^j(x^j, \mathbf{y}^{opt}) &= u^{j'}(x^{j'}, \mathbf{y}^{opt}), \\ s.t. \forall j, j' \in M, \text{ and } x^j, x^{j'} \in N. \end{aligned} \quad (30)$$

Specially, this equilibrium is an ESS.

Proof. Since $u^j(x^j, \mathbf{y}^{opt}) = u^{j'}(x^{j'}, \mathbf{y}^{opt})$, for $\forall j, j' \in M$, it follows that

$$\bar{u}(\mathbf{y}^{opt}[t]) = \frac{1}{m} \sum_{j=1}^m u^j(x^j, \mathbf{y}^{opt}[t]) = u^j(x^j, \mathbf{y}^{opt}), \text{ so } \widetilde{y}_{x^j}^{opt} = 0.$$

It means that the \mathbf{y}^{opt} in (29) is a equilibrium of the proposed game.

Then we suppose a SU k , $k \in M$, makes an unilateral deviation to another channel $x^{k*} \neq x^k$, the population state turns into

$$\mathbf{y}^* = (y_1, \dots, y_{x^k-1}, y_{x^k} - \frac{1}{m}, y_{x^k+1}, \dots, y_{x^{k*}-1}, y_{x^{k*}} + \frac{1}{m}, y_{x^{k*}+1}, \dots, y_n)^\dagger. \quad (31)$$

Hence, the system's utility becomes

$$\begin{aligned} u'(\mathbf{y}^*) &= \min_{j \in M} u^j(x^j) = \min_{j \in M} \widehat{O}_{x^j}^j (a \log(1 + \frac{g_{x^j}^j \wp^j}{\eta + \epsilon g_{pj} \wp_p}) + b) - b, \\ &= \min_{j=k*} u^j(x^{k*}) < u^j(x^j), \quad j \neq k*, k. \end{aligned} \quad (32)$$

For the equilibrium \mathbf{y}^{opt} , we have $u^k(x^k, \mathbf{y}^{opt}) = u^{k^*}(x^{k^*}, \mathbf{y}^{opt}) = u'(\mathbf{y}^{opt})$. Hence, $u'(\mathbf{y}^*) < u'(\mathbf{y}^{opt})$ is satisfied, for $\forall x^{k^*} \neq x^k, k, k^* \in M$.

In conclusion, based on the definition of ESS in Definition 2, we can get that \mathbf{y}^{opt} is a strict NE, and also an ESS. \square

Remark 5: In algorithm 2, it is necessary for each SU to know the number of SUs at every time slots and the system average payoff, decisions on channel selection from other SUs and the population state. However, it is unlikely that all these information will be available. Therefore, we propose a new distributed algorithm in the next subsection without the need for information exchange.

4.3. Dynamic and Distributed Learning Algorithm based on the Stochastic Learning Automaton with Incomplete Information for Issues

As we already mentioned, x^{opt} is a pure NE of OSAG. However, during the process of players learning from the strategies interactions, the percentage of players using a certain pure strategy may change. The population evolution is characterized by replicator dynamics in evolutionary game theory. A strategy is ESS if and only if it is asymptotically stable to the replicator dynamics [24]. It is important to achieve an ESS with SUs' local information based on the inaccurate spectrum state estimation. Here in Algorithm 3, we present a distributed learning solution, called DLES for short, by using the inaccurate spectrum sensing results to achieve the ESS.

Algorithm 3 Distributed Learning for Achieving ESS in Channel Selection (DLES) Algorithm

Given the number of channels n , maximum transmit power of SUs ρ_{max} , the complex channel gain of secondary link g , miss-detection probability ϵ and false-alarm probability ω , and parameter c .

1. Initialize : $t = 0$

- each SU $j \in M$ randomly take action $x^j[0]$, then get $\mathbf{X}^j[0]$;
- estimate the expected utility $u^j(x^j[0])$ according to (11);

2. Choose a strategy at time $t - 1$:

- choose a channel $x^j[t - 1] \in N$ to access according to the mixed strategy $\mathbf{X}^j[t - 1]$;
- calculate the utility $u^j(x^j[t - 1])$ by (11);

3. Learn :

- all the active nodes update $\mathbf{X}^j[t]$ strategy according to (33);
 - choose a channel $x^j[t] \in N$ to access according to the mixed strategy $\mathbf{X}^j[t]$;
 - if $\mathbf{X}^j[t] = \mathbf{X}^j[t - 1]$ is a stationary point, stop;
 - otherwise, go to Step 2 until convergence occurs.
-

Without information exchange, we define the process of strategy updating as follows

$$\mathbf{X}^j[t + 1] = \mathbf{X}^j[t] + cw^j[t] \cdot (\mathbf{e}_{x^j}^\dagger - \mathbf{X}^j[t]), \quad (33)$$

where $0 < c < 1$ is a parameter for controlling the process of strategy updating, and \mathbf{e}_{x^j} is a unit vector of appropriate dimension with x^j th component unity. It is easy to see from Eq. (33) that DLES algorithm is completely distributed. Since the updating rule is only dependent on their individual action-utility experiences. It is also noted that it neither needs any information exchange, nor monitors the actions taken by other users.

Theorem 4. Consider the sequence of processes $\{\mathbf{X}^j[t]\}, \{\mathbf{X}^j[t]\}$ converges weakly, as $c \rightarrow 0$.

Proof. Let $\mathcal{X}[t] = (\mathbf{X}^1[t], \dots, \mathbf{X}^m[t])^\dagger$ denote the state of the team at t . Under Algorithm 3 and (33), $\{\mathcal{X}[t], t \geq 0\}$ is a Markov process. Besides, note that $\{\mathcal{X}[t], t \in [t, t+1)\}$ is a piecewise-constant, which is right continuous and have left hand limits. First, we define a function f satisfies

$$\mathcal{X}[t+1] = \mathcal{X}[t] + cf(\mathcal{X}[t], \mathbf{U}[t], \mathbf{x}[t]), \quad (34)$$

where $\mathbf{U} = (u^1, \dots, u^m)^\dagger$ is the resulting utility of SUs. This function represents the updating specified by (33), and is bounded and continuous and it does not depend on c . Then define another function ϕ by

$$\phi(\mathcal{X}) = E[f(\mathcal{X}[t], \mathbf{U}[t], \mathbf{x}[t]) | \mathcal{X}[t] = \mathcal{X}]. \quad (35)$$

Consider an ordinary differential equation

$$\begin{aligned} \frac{d\mathcal{X}}{dt} &= \phi(\mathcal{X}), \\ \text{s.t. } \mathbf{X}^j[0] &\in [0, 1]^n \text{ and } \sum_{i=1}^n X_i^j[0] = 1. \end{aligned} \quad (36)$$

Hence by [22] (Theorem 3.1), the ordinary differential Eq. (36) has a unique solution, and the sequence $\{\mathcal{X}[t], t \in [t, t+1)\}$ converges weakly as $c \rightarrow 0$ to the solution of Eq. (36). \square

The convergence of functions implied by weak convergence ensured by Theorem 4, along with the knowledge of the nature of the solutions of the ordinary differential Eq. (36), enables us to understand the long term behavior of \mathcal{X} . Then we focus on the solution to the ordinary differential Eq. (36), and a theorem is given as follows:

Theorem 5. If parameter c is sufficiently small, the solution of ordinary differential Eq. (36) in Theorem 4 is an ESS of the proposed game.

Proof. We first denote Υ as a bounded set $[0, 1]^{m \times n}$. Some properties can be got from [22] (Theorem 3.2): 1) all corners of Υ are stationary points, 2) all Nash equilibria are stationary points, 3) all stationary points that are not NE are unstable. So all the stationary points are NEs, and 4) all corners of Υ that are strict Nash equilibria (in pure strategies) are asymptotically stable.

Then we let $\mathcal{X}[0]$ be a corner of Υ that is a NE, and \mathcal{X}^{opt} be the solution of (36). Based on the Theorem 4, the $\{\mathcal{X}[t], t \geq 0\}$ converges weakly to \mathcal{X}^{opt} . Then \mathcal{X}^{opt} is a stationary point and also strict NE in Υ . So \mathcal{X}^{opt} is an ESS of the proposed game in pure strategies. \square

5. Simulation

We conduct a set of simulations to evaluate the performance of the strategies in this section. We consider a CRN with $n = 5$ channels and its bandwidth of channels $W = 11Mbps$, and the interference threshold $I = 0.5$, $a = 1$ and $b = 10$. We assume that complex channel gain of all secondary links' are the same and equal to $g = 0.3$. Miss-detection probability and false-alarm probability of all SUs are assumed to be the same. For the purposes of comparison, we present a scenario as benchmark scheme: SUs randomly set their strategies, regardless of the existence of the others, which is denoted as random spectrum access (RAND) model.

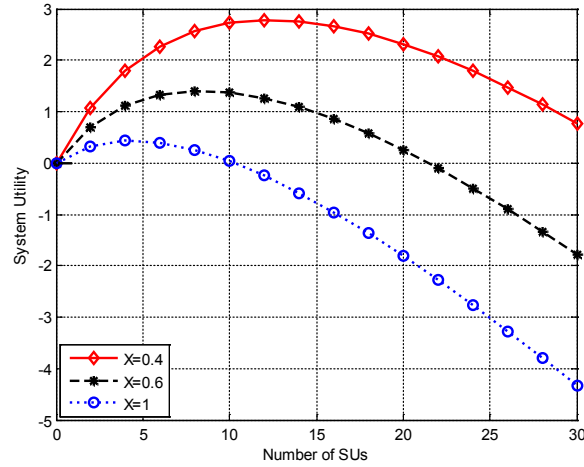


Fig. 3. System utility versus the number of SUs based on accurate spectrum state estimation.

Fig.3 shows the system utility versus the number of SUs based on accurate spectrum state estimation, where X represents the average access probability of each SU. It is obvious that the function of utility increases nonlinearly with the number of SUs m . The reason is that the spectrum opportunities are limited, and then it seems disadvantageous for SUs to access spectrum with a high probability in a time slot. In particular, in the case of $X = 1$, it is noted that the throughput of network is negative when $N^j > 10$. In short, each SU will get a higher utility, if it can take a rational strategy.

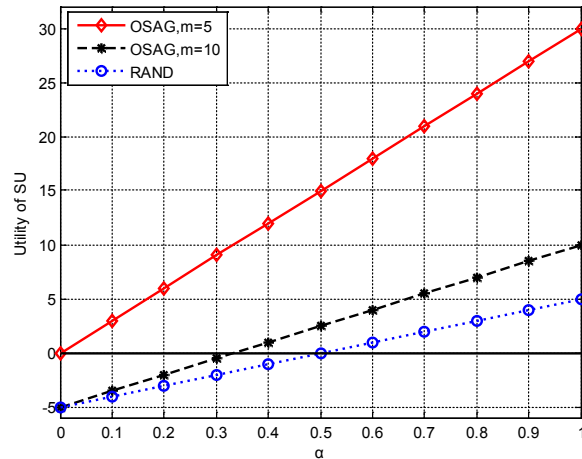


Fig. 4. Utility of SU versus α based on accurate spectrum state estimation.

The comparison results of the proposed model (OSAG) and RAND model based on accurate spectrum state estimation with different number of SUs is shown in Fig.4. Specially, Fig.4 also shows the impact of α on the utility values of node j , $j \in M$ with $\varphi_{max} = 20$. We can observe that the utility values of node increases as α increasing from $\alpha = 0$ and $\alpha = 1$. For SU j , the proposed model with $m = 5$ leads to the highest utility values compared to the proposed model

with $m = 10$ and RAND scheme. It is because j in the OSGA model prepares for the worst case where the other SUs are selfish and competing with j on the same channel. One can also observe from Fig.4 that for a specific α , the utility value of the OSAG corresponding to $m = 5$ is strictly higher than that of $m = 10$. The reason is that the more SUs the network has, the more likely the collisions they have, which decreases the utility.

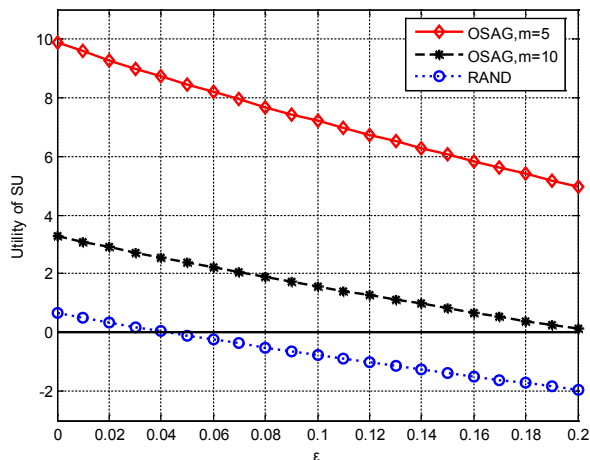


Fig. 5. Utility of SU versus miss-detection probability ϵ based on inaccurate spectrum state estimation.

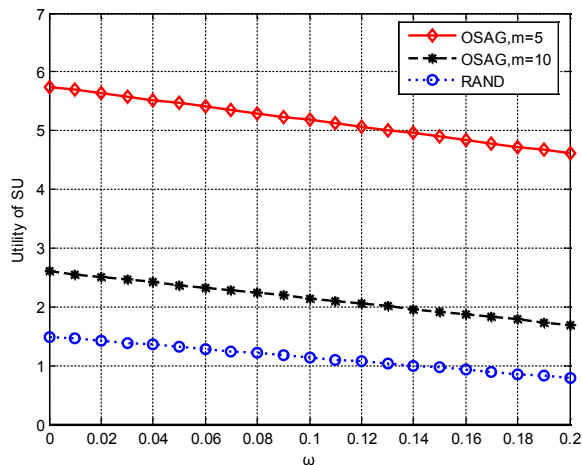


Fig. 6. Utility of SU versus false-alarm probability ω based on inaccurate spectrum state estimation.

Fig.5 shows the impact of ϵ on the utility values of SU j , $j \in M$ with $\alpha = 0.5$, $\varphi_{max} = 20$ and $\omega = 0.05$. The SU's utility value decreases as the miss-detection probability ϵ increases from $\epsilon = 0$ to $\epsilon = 0.2$. The reason is that the higher miss-detection probability of SU will cause more collisions with PUs. It is obviously that the proposed scheme corresponding to $m = 5$ performances best for the SUs comparing to that of $m = 10$ and RAND schemes. Besides, in Fig.5, the utility values of SU are all greater than 0 in the proposed scheme with $m = 5$ and $m = 10$. It shows that, in OSAG,

the SUs will always get profit in the spectrum access with ϵ increasing from $\epsilon = 0$ to $\epsilon = 0.2$. We can also observe from Fig.5 that $u^j < 0$ when $\epsilon > 0.04$ in the RAND scheme. It means that without a rational spectrum access strategy, the SU will obtain negative profit in the spectrum access, and OSAG performances better in the impact of miss-detection probability ϵ .

Similarly, Fig.6 shows the impact of ω on the utility values of SU j , $j \in M$ with $\alpha = 0.5$, $\rho_{max} = 20$ and $\epsilon = 0.05$. It is obviously, the proposed model obtains higher utility than the RAND scheme. Besides, the utility values of SU are all greater than 0 in the proposed model and RAND scheme. It means that the false-alarm probability ω causes lower impact than ϵ in the spectrum access, since the false-alarm of SUs don't cause any collision in the primary network.

In Fig.7, we show the dynamics for the proposed game using the JSFP based channel selection (JCS) algorithm shown in Algorithm 1. As expected, starting from a low initial probability of cooperation, the SUs tend to increase the degree of cooperation, as a result, the system tends to increase the utility value. Finally, the proposed game achieves a pure-strategy Nash equilibrium, which is shown in Fig.7.

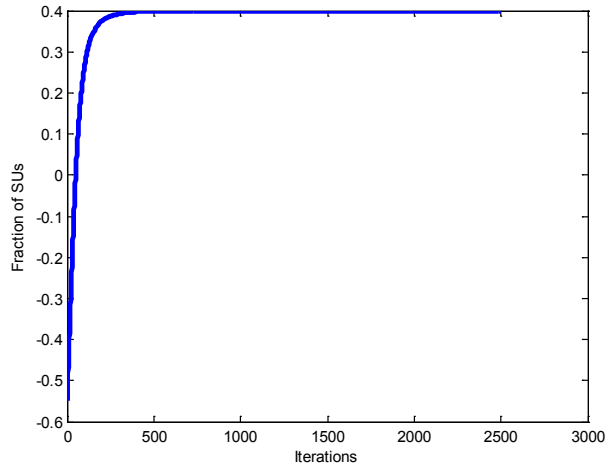


Fig. 7. The converging process of the JCS algorithm.

The converging process of the evolutionary learning for achieving ESS in channel selection (ELES) algorithm is shown in Fig.8. During the iterations, the strategies cause a lower utility value of each SU in (11) than the average utility value of the system converge to 0 quickly. While the strategies with better influence in the utility value of SUs converge to 1. The converging process shows the strategies work better than the average utility will be promoted, and the strategies work worse than the average utility of the system will be died out during the iterations.

In the Fig.9, we show the dynamics for the proposed game using the distributed learning for achieving ESS in channel selection (DLES) algorithm shown in Algorithm 3. As expected, starting from different probabilities of cooperation, the SUs tend to increase the degree of cooperation, as a result, the system tends to increase the utility value. Finally, the proposed game achieves a mixed-strategy Nash equilibrium, which is shown in Fig.9.

The comparison result of the proposed scheme based on the accurate spectrum state estimation and inaccurate spectrum state estimation is shown in Fig.10. As shown in Fig.10, when the false-alarm probability increases from $\omega = 0$ to $\omega = 0.4$, the utility values of both cases decrease. We can also observe that the proposed scheme with considering the inaccurate spectrum state es-

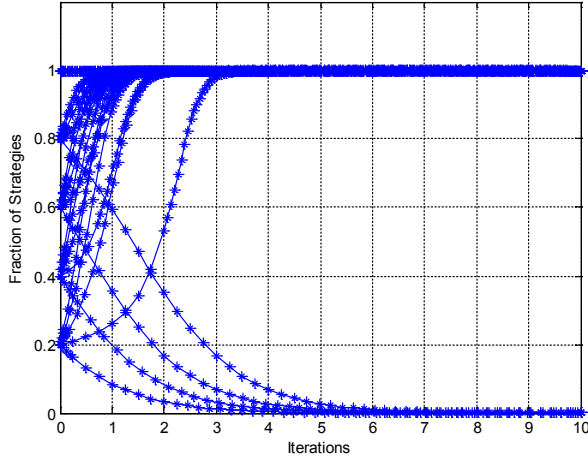


Fig. 8. The converging process of the ELES algorithm.

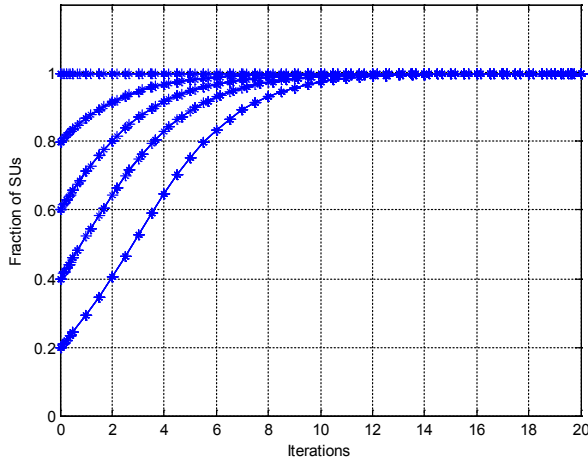


Fig. 9. The converging process of the DLES algorithm.

timization performs much better than that without considering the inaccurate spectrum state. It is shown in Fig.10 that the proposed scheme with considering the inaccurate spectrum state estimation performs better when false-alarm happens, and the proposed scheme without considering the inaccurate spectrum state estimation causes the utility's reduction quicker than the previous scheme.

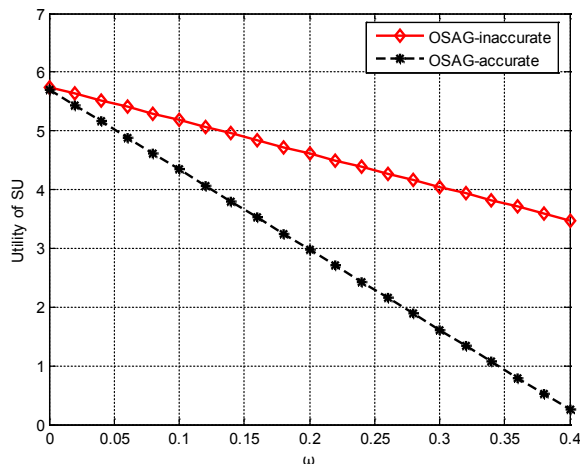


Fig. 10. Comparison of the proposed scheme based on the accurate spectrum state estimation (OSAG-accurate) and inaccurate spectrum state estimation (OSAG-inaccurate).

The comparison result of the proposed scheme (OSAG) and the random access game (RAG) [14] is shown in Fig.11. We denote the false-alarm probability as $\omega = 0.05$ and the miss-detection probability as $\epsilon = 0.05$. As shown in Fig.11, when the number of SUs increases from $m = 2$ to $m = 20$, the corresponding aggregate throughput value of both schemes decreases. It is also shown in Fig.11 that the proposed scheme performs better than that without considering the inaccurate spectrum state in the throughput of system.

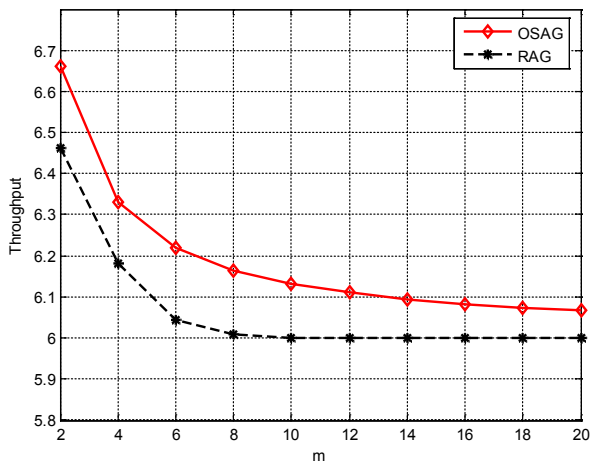


Fig. 11. Comparison of the proposed scheme (OSAG) and the random access game (RAG).

6. Conclusion

We proposed a potential game theoretic framework for opportunistic spectrum access, and three algorithms to learn the equilibria for the proposed game in cognitive radio networks. More specifically, JCS algorithm was proposed to achieve the pure Nash equilibrium based on accurate spectrum state estimation. It was shown that with the JCS algorithm, global optimization in terms of network throughput maximization and network collision minimization can be achieved with local information. In order to make the outcomes of game robust, we investigated evolutionary spectrum access mechanism using the evolutionary game theory with complete information based on inaccurate spectrum state estimation, which is denoted as ELES algorithm. Finally, without the need of information exchange, we proposed a distributed DLES algorithm based on stochastic learning automata to achieve the evolutionary stable strategy. Simulation results indicate that our scheme improves the network performance (high throughput and low collision) and achieves better fairness.

7. Acknowledgment

The authors would like to thank the autonomous reviewers for their helpful comments to improve this paper. This work is partially sponsored by National Natural Science Foundation of China (NOs. 61072080, U1405255), Fujian Normal University Innovative Research Team (NO. IRTL1207), and Major Science and Technology Project in Fujian Province (NO. 2014H61010105).

8. References

- [1] M. Maskery, V. Krishnamurthy, and Q. Zhao: ‘Decentralized Dynamic Spectrum Access for Cognitive Radios: Cooperative Design of A Non-Cooperative Game’, *IEEE transactions on communications*, 2009, **57**, (2), pp: 459-469.
- [2] M. J. Piran, M. E. Ahmed, A. Ali, J. B. Song, and D. Y. Suh: ‘Channel Allocation Based on Content Characteristics for Video Transmission in Time-Domain-Based Multichannel Cognitive Radio Networks’, *Mobile information systems*, 2015, DOI:org/10.1155/2015/934715.
- [3] Y. Xu, A. Anpalagan, Q. Wu, L. Shen, Z. Gao, and J. Wang: ‘Decision-Theoretic Distributed Channel Selection for Opportunistic Spectrum Access: Strategies, Challenges and Solutions’, *IEEE communications surveys and tutorials*, 2013, **15** (4).
- [4] E. Z. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. A. Siris: ‘Spectrum Assignment in Cognitive Radio Networks: A Comprehensive Survey’, *IEEE communications surveys and tutorials*, 2013, **15** (3).
- [5] D. I. Kim, L. B. Le and E. Hossain: ‘Joint Rate and Power Allocation for Cognitive Radios in Dynamic Spectrum Access Environment’, *IEEE transactions on wireless communications*, 2008, **7**, (12).
- [6] M. G. Khoshkholgh, K. Navaie and H. Yanikomeroğlu, ‘Access Strategies for Spectrum Sharing in Fading Environment: Overlay, Underlay, and Mixed’, *IEEE transactions on mobile computing*, 2010, **9**, (12).

- [7] Q. Zhao, B. Krishnamachari and K. Liu: ‘On Myopic Sensing for Multi-Channel Opportunistic Access: Structure, Optimality, and Performance’, *IEEE transactions on wireless communications*, 2008, **7**, (12).
- [8] N. Hasan and W. Tu: ‘On energy-efficient performance-guaranteed channel assignment in cognitive radio-based wireless mesh networks’, *International journal of communication systems*, 2015, **28**, pp: 2197-2213.
- [9] N. Hasan, W. Ejaz, N. Ejaz, H. S. Kim, A. Anpalagan, and M. Jo, ‘Network Selection and Channel Allocation for Spectrum Sharing in 5G Heterogeneous Networks’, *IEEE access*, 2016.
- [10] Kehao Wang, Quan Liu, Fangmin Li, Lin Chen, Xiaolin Ma: ‘Myopic policy for opportunistic access in cognitive radio networks by exploiting primary user feedbacks’, *IET Communications*, 2015, **9**, (7), pp: 1017–1025.
- [11] Kehao Wang, Lin Chen, Quan Liu: ‘On Optimality of Myopic Policy for Opportunistic Access With Nonidentical Channels and Imperfect Sensing’, *IEEE Transactions on Vehicular Technology*, 2014, **63**, (5), pp: 2478–2483.
- [12] Kehao Wang, lin chen, Quan Liu, Khaldoun Al Agha: ‘On Optimality of Myopic Sensing Policy with Imperfect Sensing in Multi-Channel Opportunistic Access’, *IEEE Transactions on Communications*, 2013, **61**, (9), pp: 3854–3862.
- [13] Z. Ji, K. J. R. Liu: ‘Cognitive Radios for Dynamic Spectrum Access–Dynamic Spectrum Sharing: A Game Theoretical Overview’, *IEEE communications magazine*, 2007, **45**, (5), pp: 88–94.
- [14] L. Chen, S. H. Low, J. C. Doyle: ‘Random access game and medium access control design’, *IEEE/ACM transactions on networking*, 2010, **18**, (4), pp: 1303–1316.
- [15] H. Fang, L. Xu, C. Huang: ‘Dynamic Opportunistic Spectrum Access of Multi-channel Multi-radio Based on Game Theory in Wireless Cognitive Network’. *Proc. IEEE ninth international conference on mobile ad-hoc and sensor networks (MSN)*, 2013, pp: 127–132.
- [16] M. Cagalj, S. Ganeriwal, I. Aad, and J. P. Hubaux: ‘On Selfish Behavior in CSMA/CA Networks’. *Proc. IEEE INFOCOM*, 2005, pp. 2513–2524.
- [17] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, Y. D. Yao: ‘Opportunistic Spectrum Access in Cognitive Radio Networks: Global Optimization Using Local Interaction Games’, *IEEE journal of selected topics in signal processing*, 2012, **6**, (2), pp: 180–194.
- [18] X. Chen, J. Huang: ‘Spatial Spectrum Access Game’, *IEEE transaction on mobile computing*, 2015, **14**, (3), pp: 646–659.
- [19] X. Feng, G. Sun, X. Gan, F. Wang: ‘Cooperative Spectrum Sharing in Cognitive Radio Networks: A Distributed Matching Approach’, *IEEE transaction on communications*, 2014, **62**, (8), pp: 2651–2664.
- [20] ‘Spectrum Policy Taskforce Report,’ technical report, Fed. Comm. Commission, http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-228542A1.pdf, Nov. 2002.
- [21] J. R. Marden, G. Arslan, J. S. Shamma: ‘Joint Strategy Fictitious Play with Inertia for Potential Games’, *IEEE transactions on automatic control*, 2009, **54**, (2), pp: 208–220.

- [22] P. S. Sastry, V. V. Phansalkar, M. A. L. Thathachar: ‘Decentralized Learning of Nash Equilibria in Multi-Person Stochastic Games With Incomplete Information’, IEEE transaction on systems, man, and cybernetics, May 1994, **24**, (5).
- [23] W. R. Wade: ‘An Introduction to Analysis’, 4th ed. Pearson, 2010.
- [24] B. Wang, K. J. R. Liu, T. C. Clancy: ‘Evolutionary Cooperative Spectrum Sensing Game: How to Collaborate?’, IEEE transactions on communications, March 2010, **58**, (3).
- [25] J. M. Smith: ‘Evolution and the Theory of Games’, Cambridge university, 1982.
- [26] C. E. Shannon: ‘A Mathematical Theory of Communication’. Bell syst’, Tech. J., 1948, **27**, pp: 379-423.