Evolution of Cooperation: A Case with Interference-Aware Cooperative Spectrum Sensing in Cognitive Radio Networks

You-En Lin[1] and Hung-Yun Hsieh[1,2]

[1] Graduate Institute of Communication Engineering
[2] Department of Electrical Engineering,
National Taiwan University, Taipei, Taiwan 106
hyhsieh@cc.ee.ntu.edu.tw

Abstract. We consider in this work a group of secondary users with backlogged traffic to transmit in the primary network. To avoid interfering with the primary user, each secondary user must perform interference-aware spectrum sensing before transmission. Unlike conventional sensing techniques, interference-aware spectrum sensing allows a secondary user to adjust its sensing parameters for optimal performance depending on the probability of interfering with the primary user. While interference-aware sensing can achieve better performance for individual users, challenges arise when secondary users collaborate with each other for cooperative spectrum sensing due to their unequal interference probabilities that result in a conflict for setting the optimal sensing parameters. To model this problem, we consider an interference-aware cooperative sensing game and analyze player behaviors under such a game. We find that there is a unique pure Nash equilibrium of the game, but it tends to deviate from the desirable solution of social optimum. We then design a repeated game based on evolutionary game theory to address this problem. Players in the repeated game have the chance to revenge "uncooperative" players in ensuing repetitions for driving the equilibrium to the social optimum. We show through numerical results that the proposed game of evolution does achieve the desirable performance for interference-aware cooperative sensing in dynamic spectrum access.

1 Introduction

The concept of dynamic spectrum access allows unlicensed secondary users to opportunistically fill in the white space left by the licensed primary user for better utilization of the spectrum. To enable such opportunistic spectrum access, an important step for the secondary user is to sense the spectrum before transmission to detect the presence of the primary user [17,4]. Due to the stochastic nature of channel fading and/or background noise, however, it is possible that the signal detector at the secondary user makes an inaccurate detection decision in terms of false alarm and missed detection. While occurrences of false alarm and missed detection are equally undesirable that should ideally be minimized

R. Jain and R. Kannan (Eds.): GameNets 2011, LNICST 75, pp. 530–545, 2012.

whenever possible, the fundamental tradeoffs between the two types of errors prevent the detector from operating under arbitrarily small probabilities of false alarm and missed detection [10,14].

To operate the signal detector in the desired operation region, related work has proposed several sensing techniques and optimization frameworks for optimizing the detector performance. For example, the authors in [12] aim to design an optimal detector under non-Gaussian noise that can minimize the probability of false alarm while ensuring that the probability of missed detection is below the target threshold. While it is reasonable to limit the maximum probability of missed detection while minimizing the probability of false alarm for ensuring that the primary user is properly protected without suffering from undesirable interference [12,11,18,9], such a spectrum sensing strategy can become *conservative* and result in sub-optimal performance for opportunistic spectrum access. The reason is that different secondary users could potentially incur different degrees of interference on the primary user when missed detection occurs. If secondary users are required to provide the same level of protection to the primary user, a secondary user with a lower probability of interference should ideally be allowed to have a larger probability of missed detection during spectrum sensing. Probability of spectrum hole discovery and hence spectrum utilization can thus be improved by employing such an interference-aware spectrum sensing technique.

While interference-aware spectrum sensing can potentially improve the performance for individual users in opportunistic spectrum access, challenges arise when secondary users collaborate with each other for cooperative spectrum sensing. In cooperative spectrum sensing, a secondary transmitter cooperates with a group of secondary users to make the transmission decision through fusing the local decisions of individual signal detectors. It has been shown in related work that cooperative sensing can indeed improve the sensing performance for individual detectors under channel artifacts such as fading and noise [8,6,13]. In interference-aware cooperative spectrum sensing, however, since the optimal sensing parameters depend on the interference probability of the concerned transmitter, it is possible that the final decision is not optimal for other cooperating users. In such a case, a secondary user may not have sufficient incentive to perform local spectrum sensing and contribute to the fusion process only to obtain a final decision that does not improve its performance. If secondary users are allowed to make their own decisions without considering the resultant benefit of the whole group, interference-aware cooperative spectrum sensing cannot be guaranteed to achieve the desirable performance gain due to cooperation.

To model this problem and gain more insights on finding the solution, we resort to game theory and focus on the negotiation process for sensing parameters in interference-aware cooperative spectrum sensing. A secondary user acts as a player in the game and it participates in the game to determine its sensing parameters for maximizing its payoff. Under such a game model, as we show in Section 3, there exists a pure Nash equilibrium where no player can change its sensing parameters unilaterally without decreasing its payoff. The Nash equilibrium, however, fails to provide the socially optimal performance for all players due to the selfish actions of individual players. To address the problem of sub-optimality due to non-cooperation, we then design a repeated game based on evolutionary game theory. In the repeated game, a stage game is played in repetitions where each player can learn from history the actions of other players to adapt its action in ensuing repetitions. Four different roles including *Solitary, Leader, Follower, and Glutton* that players in the game can choose to act are proposed. To stimulate cooperation, a special role of *Avenger* is also included so disadvantaged player can "punish" non-cooperative players with the goal of driving the equilibrium of the game to social optimum. We show through numerical results that the proposed repeated game can indeed stimulate cooperation through evolution of player roles and achieve the desirable performance for interference-aware cooperative sensing in opportunistic spectrum access.

The paper is organized as follows. Section 2 presents the network scenario as well as the spectrum sensing and sharing models. Based on the system models, Section 3 formulates a game for interference-aware cooperative spectrum sensing and then presents the best responses of the players and the Nash equilibrium thus obtained. Due to the selfish behavior of individual players, it is also shown in Section 3 that the Nash equilibrium will deviate form the social optimum. To address the problem, a repeated game based on evolutionary game theory is proposed in Section 4 to stimulate cooperation among secondary users and drive the equilibrium to the social optimum. Evaluation results in Section 5 show that the proposed evolutionary game can indeed achieve the desired performance with significant performance gains.

2 System Model

We consider a simple network scenario with one primary transmitter serving as the base station of a cellular network or a TV broadcast station with service range R. To leverage the spectrum holes in the frequency bands used by primary users, secondary users can sense the activity of the primary cell and reuse the spectrum when primary users are not active. Even when primary users are active, secondary users can still reuse the spectrum as long as no undesirable interference (as explained in Section 2.2) is incurred on primary users.

In such a scenario, a challenging issue is to achieve reliable spectrum sensing when secondary users are outside the cell and the signal from the primary transmitter is relatively weak. To mitigate the problem of channel artifacts such as fading and noise, we assume that a group of secondary users $\mathcal{G}$ with backlogged traffic cooperate for spectrum sensing. Secondary users operate on a TDMA-basis by following the frame structure shown in Figure 1. In each frame, secondary users in $\mathcal{G}$ first negotiate the sensing parameters (e.g. detection threshold for an energy detector) to be used during the following sensing period. After negotiation, each secondary user performs local spectrum sensing and then broadcasts the local decision to be fused for the final decision in cooperative sensing. If transmission opportunity (spectrum hole) is identified based on the sensing result, secondary users send in turns during the transmission slots until the next frame starts. In the following, we explain in more details the spectrum sensing and sharing models used for opportunistic spectrum access.

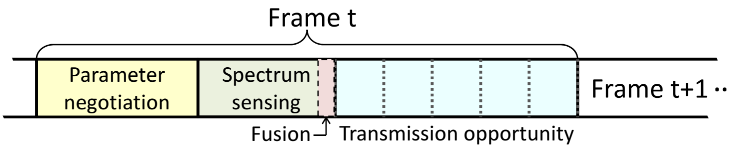

Fig. 1. Frame structure for secondary users

2.1 Spectrum Sensing Model

We consider the use of an energy detector for detecting the presence of the primary user due to its applicability to a wide range of signals and mathematical amenity compared to other detectors [5,7]. A secondary user first senses the local spectrum and takes N samples of the signal measurements. Denote H_0 as the case when the primary transmitter is inactive (no transmission) and H_1 as the case when the primary transmitter is active. Sample $Y_j[n]$ taken at time n by secondary user j can then be written as follows:

$$Y_j[n] = \begin{cases} W_j[n], & H_0, \\ X_j[n] + W_j[n], & H_1, \end{cases} \tag{1}$$

where $W_j[n]$ is the sample of the additive white Gaussian noise (AWGN) with variance σ^2, and $X_j[n]$ is the sample of the primary signal that we model as a Gaussian signal [11]. To decide whether the primary signal is present, secondary user j first computes a test statistic as $Z_j = \frac{1}{N}\sum_{n=1}^{N} |Y_j[n]|^2$. Based on the test statistic Z_j, secondary user j makes a positive decision ($\mathcal{D} = 1$) regarding the presence of the primary user if $Z_j \geq \theta_j$, where θ_j denotes its detection threshold. Otherwise, the spectrum is considered empty ($\mathcal{D} = 0$).

Due to the stochastic nature of the background noise, it is possible that the decision regarding the spectrum usage is not correct. False alarm occurs when the primary user is inactive but determined to be active by the secondary user, whereas missed detection occurs when the primary user is active but determined as being inactive. With sufficiently large number of samples N, the probability of false alarm $\mathbb{P}_{\mathrm{FA}}^{(j)}$ and missed detection $\mathbb{P}_{\mathrm{MD}}^{(j)}$ observed by secondary user j can be derived as follows [11]:

$$\mathbb{P}_{\mathrm{FA}}^{(j)} = Q\left(\frac{\theta_j - \sigma^2}{\sqrt{2/N}\sigma^2}\right), \tag{2}$$

$$\mathbb{P}_{\mathrm{MD}}^{(j)} = 1 - Q\left(\frac{\theta_j - (P_j + \sigma^2)}{\sqrt{2/N}(P_j + \sigma^2)}\right), \tag{3}$$

where P_j is the mean power of the received primary signal at secondary user j, and $Q(\cdot)$ is the complementary CDF of the standard Gaussian distribution.

After the decision of the local spectrum sensing is made by each secondary user, cooperative spectrum sensing is performed using decision fusion. We consider in this paper the use of the OR rule for fusing local decisions from individual detectors. The probabilities of missed detection and false alarm for the cooperative sensing set $\mathcal{G}$ can then be expressed as $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})} = \prod_{j\in\mathcal{G}} \mathbb{P}_{\mathrm{MD}}^{(j)}$ and $\mathbb{P}_{\mathrm{FA}}^{(\mathcal{G})} = 1 - \prod_{j\in\mathcal{G}}(1 - \mathbb{P}_{\mathrm{FA}}^{(j)})$. It is clear that $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ depends on detection thresholds θ_j of individual detectors and $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ is smaller than $\mathbb{P}_{\mathrm{MD}}^{(j)}$ of any user j in the cooperative set.

2.2 Spectrum Sharing Model

As we mentioned earlier, in opportunistic spectrum access, a secondary user is allowed to use the spectrum as long as it does not interfere with the primary user. We thus define for secondary user j the probability of interfering with the primary user as $\mathbb{P}_{\mathrm{I}}^{(j)} = \mathrm{Prob}(\gamma < \gamma_t)$, where γ is the mean SINR during the reception of a packet with M symbols for a primary user at the closest point to the secondary user. Assuming Gaussian primary and secondary signals similar to [11], we can derive the probability of interference for sufficiently large value of M as follows:

$$\mathbb{P}_{\mathrm{I}}^{(j)} = Q\left(\frac{P_p/\gamma_t - P_c - \sigma^2}{\sqrt{2/M\,(P_p/\gamma_t + P_c^2 + \sigma^4)}}\right), \tag{4}$$

where P_p and P_c are the mean powers of primary and secondary signals at the primary receiver respectively.

Note that $\mathbb{P}_{\mathrm{I}}^{(j)}$ is the probability of causing interference when *both the primary and secondary transmitters are active.* That is, it is a conditional probability that applies when missed detection occurs. To allow for better spectrum reuse based on interference-aware sensing, we can formulate the following constraint

$$\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}\mathbb{P}_{\mathrm{I}}^{(j)} \leq \varepsilon \tag{5}$$

to ensure that the primary receiver at the worst point is not interfered (unable to correctly decode the primary signal) by the secondary user with a probability of at least $1-\varepsilon$. For simplicity, we say that *the secondary user does not interfere with the primary user when Inequality* (5) *holds.* If Inequality (5) does not hold, on the other hand, the secondary user is not allowed to send for spectrum sharing. Since different secondary users may incur unequal degrees of interference on the primary user, there exists a conflict of interests in setting individual detection thresholds for obtaining $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ in cooperative spectrum sensing. We elaborate the problem in the following section.

3 Interference-Aware Sensing Game

In this section, we first formulate a game to model the negotiation process for spectrum sensing parameters in interference-aware cooperative spectrum sensing. We then find the best response of each player and derive the unique Nash equilibrium and social optimum for the game.

3.1 Game Formulation

During the negotiation process of sensing parameters, each player (secondary user) negotiates the detection threshold (that affects the final probability of missed detection) to maximize its payoff. To determine the payoff, we note that a spectrum hole exists for a secondary user under the following two cases:

- Prob $(\mathcal{D} = 0 \mid H_0)$: The primary user is not active, and the secondary user does not make any false alarm about the spectrum usage. The expected throughput for this case can be expressed as $(1-\mathbb{P}_{\mathrm{FA}}^{(\mathcal{G})})(1-\mathbb{P}_{\mathrm{E}})C_0^{(j)}$, where $\mathbb{P}_{\mathrm{E}}$ is the probability that the primary transmitter uses the spectrum, i.e. $\mathbb{P}_{\mathrm{E}} = \mathrm{Prob}(H_1)$, and $C_0^{(j)}$ is the channel capacity at the corresponding receiver of secondary user j when the primary user is inactive.
- Prob $(\{\mathcal{D} = 0 \mid H_1\} \cap \{\gamma \geq \gamma_t\})$: The secondary user makes missed detection when the primary user is active, but transmission of the secondary user does not interfere with the primary receiver at the worst point $(\gamma \geq \gamma_t)$. The expected throughput for this case can be expressed as $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}\mathbb{P}_{\mathrm{E}}(1-\mathbb{P}_{\mathrm{I}}^{(j)})C_1^{(j)}$, where $C_1^{(j)}$ is the channel capacity at the corresponding receiver of secondary user j when the primary user is active.

We can therefore define the payoff for secondary user j as the overall expected throughput by summing the achievable throughput in the two cases:

$$C^{(j)} = \mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}\mathbb{P}_{\mathrm{E}}(1-\mathbb{P}_{\mathrm{I}}^{(j)})C_1^{(j)} + (1-\mathbb{P}_{\mathrm{FA}}^{(\mathcal{G})})(1-\mathbb{P}_{\mathrm{E}})C_0^{(j)}. \tag{6}$$

A rational secondary user j thus may adjust its detection threshold θ_j for maximizing the expected throughput. To ensure that primary users are properly protected, however, we define the utility function for each player as follows:

$$U^{(j)} = \begin{cases} C^{(j)}, & \mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}\mathbb{P}_{\mathrm{I}}^{(j)} \leq \varepsilon \\ 0, & \text{Otherwise} \end{cases}, \; j \in \mathcal{G}. \tag{7}$$

That is, a secondary user gets the payoff (expected throughput) only when Inequality (5) holds. Otherwise, the secondary user is not allowed to use the spectrum, and its payoff is zero.

3.2 Best Response

The utility function of secondary user j in Equation (7) is a function of $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$, and hence utility $U^{(j)}$ depends not only on the detection threshold θ_j of player j but also on those of other players. Note that maximization of the utility function with respect to θ_j as shown in Equation (7) belongs to monotonic optimization [15]. Hence, the optimal value of θ_j that maximizes the utility occurs on the boundaries of the constraints as follows:

– If $\left(\mathbb{P}_{\mathrm{I}}^{(j)} \prod_{k\in\mathcal{G},k\neq j} \mathbb{P}_{\mathrm{MD}}^{(k)}\right) > \varepsilon$, optimal θ_j occurs on the boundary of the interference constraint such that $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}\mathbb{P}_{\mathrm{I}}^{(j)} = \varepsilon$. Using Equation (3), we have

$$\theta_j^* = (\sigma^2 + P_j)\{1 - \sqrt{2/N}Q^{-1}[\varepsilon/(\mathbb{P}_{\mathrm{I}}^{(j)} \prod_{k\in\mathcal{G},k\neq j} \mathbb{P}_{\mathrm{MD}}^{(k)})]\}. \tag{8}$$

– If $\left(\mathbb{P}_{\mathrm{I}}^{(j)} \prod_{k\in\mathcal{G},k\neq j} \mathbb{P}_{\mathrm{MD}}^{(k)}\right) \leq \varepsilon$, optimal θ_j occurs when $\mathbb{P}_{\mathrm{MD}}^{(j)} = 1$ or $\mathbb{P}_{\mathrm{FA}}^{(j)} = 0$. It can be observed from Equations (2) and (3) that $\theta_j^* \to \infty$. This is a special case since the local sensing results contributed by other players are already sufficient for protecting the primary receiver. It is thus not necessary for player j to perform local sensing, and its best strategy to maximize the expected throughput thus is not to perform any local sensing but use the result from cooperative sensing.

A player in the game can therefore determine its detection threshold (best response) for maximizing its payoff given the detection thresholds of other players.

3.3 Nash Equilibrium

As discussed in Section 3.2, the maximum payoff for player j occurs on the constraint boundary, i.e. $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})} = \varepsilon/\mathbb{P}_{\mathrm{I}}^{(j)}$. While $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ is shared by all members in cooperative sensing, $\mathbb{P}_{\mathrm{I}}^{(j)}$ is different for different players in the group. For a player with a larger $\mathbb{P}_{\mathrm{I}}^{(j)}$, it would desire a smaller $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ to maximize its payoff without violating the interference constraint. It therefore needs to reduce the detection threshold θ_j such that $\mathbb{P}_{\mathrm{MD}}^{(j)}$ and $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ can be decreased. On the other hand, a player with a smaller $\mathbb{P}_{\mathrm{I}}^{(j)}$ would desire a larger $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ by increasing detection threshold θ_j and $\mathbb{P}_{\mathrm{MD}}^{(j)}$. Negotiation of detection thresholds comes to an end when no player wants to change its threshold unilaterally. As we show in the following, there is exactly one pure Nash equilibrium in the game.

Firstly, according to Equation (3), the probability of missed detection $\mathbb{P}_{\mathrm{MD}}^{(j)}$ and the detection threshold θ_j has 1-to-1 mapping. We can then use $\mathbb{P}_{\mathrm{MD}}^{(j)}$ in place of θ_j as the response of player j for notation clarity. Now let player k be the one with the maximum interference probability, i.e. $k = \arg\max_{j\in\mathcal{G}} \mathbb{P}_{\mathrm{I}}^{(j)}$. The unique pure Nash equilibrium of the interference-aware spectrum sensing game can then be written as $(1, \cdots, \varepsilon/\mathbb{P}_{\mathrm{I}}^{(k)}, \cdots, 1)$. Equivalently, in Nash equilibrium the probability of missed detection of all players is 1 except for player k such that $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})} = \mathbb{P}_{\mathrm{MD}}^{(k)} = \varepsilon/\mathbb{P}_{\mathrm{I}}^{(k)}$. In such a case, player k takes all responsibility of spectrum sensing because it has the most stringent interference constraint for protecting the primary user. Other players are simply free-riders who can obtain the final sensing decision without contributing any local sensing results. Therefore, under the Nash equilibrium, the cooperative spectrum sensing game turns out to non-cooperative, single-node spectrum sensing. The Nash equilibrium clearly is far from being social optimum in terms of cooperative spectrum sensing. This motivates us to develop rules for the game so that a socially optimal equilibrium can be achieved.

3.4 Social Optimum

To compare the Nash equilibrium against the social optimum, we first define the social utility as the sum of the expected throughput of all secondary users in $\mathcal{G}$. Finding the social optimum is then equal to finding the optimal solution of the following optimization problem:

$$\underset{\theta_j,\, j\in\mathcal{G}}{\text{Maximize}} \sum_{j\in\mathcal{G}} C^{(j)}, \text{ subject to } \mathbb{P}_{\text{MD}}^{(\mathcal{G})}\mathbb{P}_{\text{I}}^{(j)} \leq \varepsilon,\, \forall j \in \mathcal{G}. \tag{9}$$

Although Problem (9) is not a convex problem, it belongs to monotonic programming and hence the global optimal solution can be found.

To start, note that due to the monotonic increasing property of $\mathbb{P}_{\text{MD}}^{(\mathcal{G})}$ with respect to θ_j, the interference constraint for the user with the largest interference probability, say user $k = \arg\max_{j\in\mathcal{G}} \mathbb{P}_{\text{I}}^{(j)}$, will hit the boundary first. Hence, the $|\mathcal{G}|$ interference constraints can be replaced by $\mathbb{P}_{\text{MD}}^{(\mathcal{G})} = \varepsilon/\mathbb{P}_{\text{I}}^{(k)}$. In addition, since θ_j and $\mathbb{P}_{\text{MD}}^{(j)}$ has 1-to-1 mapping, the optimization problem can then be rewritten as:

$$\underset{\mathbb{P}_{\text{MD}}^{(j)},\, j\in\mathcal{G}}{\text{Maximize}} \prod_{j\in\mathcal{G}} \left(1 - \mathbb{P}_{\text{FA}}^{(j)}\right), \text{ subject to } \prod_{j\in\mathcal{G}} \mathbb{P}_{\text{MD}}^{(j)} = \frac{\varepsilon}{\mathbb{P}_{\text{I}}^{(k)}}. \tag{10}$$

As can be observed from Problem (10), the objective and constraint functions are products of $|\mathcal{G}|$ independent functions *each with only one variable as the parameter*. Such a property makes it amenable to a solution based on dynamic programming [1] that solves a complex problem by recursively breaking into simpler sub-problems. To detail, assume cooperative set $\mathcal{G} = \{1, 2, ..., |\mathcal{G}|\}$ without loss of generality. The original problem can be solved in $|\mathcal{G}|$ stages (sub-problems), and each stage can be solved using the following recursive equation:

$$\begin{aligned} f^{(j)}(S^{(j)}) = \underset{\mathbb{P}_{\text{MD}}^{(i)}}{\text{Maximize}} &\left(1 - \mathbb{P}_{\text{FA}}^{(j)}\right) \cdot f^{(j+1)}\left(S^{(j+1)}\right), \\ &\text{subject to } \mathbb{P}_{\text{MD}}^{(j)} \cdot S^{(j+1)} = S^{(j)}, \end{aligned} \tag{11}$$

where state $S^{(j)}$ indicates the value of the constraint function counting only detectors $\{j, ..., |\mathcal{G}|\}$ and $f^{(j)}\left(S^{(j)}\right)$ is the maximum cumulative payoff from stage j to $|\mathcal{G}|$ under state $S^{(j)}$.

Algorithm 1 thus shows the pseudo-code for finding the optimal thresholds in interference-aware cooperative sensing. While Problem (10) has continuous state in [0, 1], it can be solved through proper discretization of the state space [2]. Input to the algorithm includes δ_1 and δ_2 that are used to control the search granularity for state $S^{(j)}$ and variable $\mathbb{P}_{\text{MD}}^{(j)}$ respectively. The algorithm consists

Algorithm 1. Finding optimal detection thresholds

Input: Protection constraint $\varepsilon/\mathbb{P}_{\mathrm{I}}^{(k)}$, granularity δ_1 and δ_2
Output: Detection thresholds θ_j^*, $\forall j \in \mathcal{G} = \{1, 2, ..., |\mathcal{G}|\}$

01: **for** $(j = |\mathcal{G}|;\ j \geq 1;\ j--)$ // iterate for stages
02: **for** $(\widehat{S} = \varepsilon/\mathbb{P}_{\mathrm{I}};\ \widehat{S} \leq 1;\ \widehat{S} = \widehat{S} + \delta_1)$ // iterate for states
03: **if** $(j == |\mathcal{G}|)$ // starting stage
04: $f^{(j)}\left(\widehat{S}\right) = 1 - \mathbb{P}_{\mathrm{FA}}^{(j)}\left(\widehat{S}\right),\ \mathbb{P}_{\mathrm{MD}}^{(j)}\left(\widehat{S}\right) = \widehat{S}$
05: **else**
06: **for** $(\widehat{\mathbb{P}}_{\mathrm{MD}} = \widehat{S};\ \widehat{\mathbb{P}}_{\mathrm{MD}} \leq 1;\ \widehat{\mathbb{P}}_{\mathrm{MD}} = \widehat{\mathbb{P}}_{\mathrm{MD}} + \delta_2)$
07: $\widehat{f} = \left[1 - \mathbb{P}_{\mathrm{FA}}^{(j)}\left(\widehat{\mathbb{P}}_{\mathrm{MD}}\right)\right] \cdot f^{(j+1)}\left(\frac{\widehat{S}}{\widehat{\mathbb{P}}_{\mathrm{MD}}}\right)$
08: **if** $\left(\widehat{f} > f^{(j)}\left(\widehat{S}\right)\right)$
09: $f^{(j)}\left(\widehat{S}\right) = \widehat{f},\ \mathbb{P}_{\mathrm{MD}}^{(j)}\left(\widehat{S}\right) = \widehat{\mathbb{P}}_{\mathrm{MD}}$
10: **end if**
11: **end for** // found optimal $\widehat{\mathbb{P}}_{\mathrm{MD}}$
12: **end if**
13: **if** $(j == 1)$ **break** // only one state for stage 1
14: **end for**
15: **end for**
16: **for** $(j = 1,\ \widehat{S} = \varepsilon/\mathbb{P}_{\mathrm{I}}^{(k)};\ j \leq |\mathcal{G}|;\ j++)$ // trace back
17: $\theta_j^* = (\sigma^2 + P_j)\left\{1 - \sqrt{\frac{2}{N}}Q^{-1}\left[\mathbb{P}_{\mathrm{MD}}^{(j)}\left(\widehat{S}\right)\right]\right\}$
18: $\widehat{S} = \widehat{S}/\mathbb{P}_{\mathrm{MD}}^{(j)}\left(\widehat{S}\right)$
19: **end for**

of three layers of loops: the first layer from Line 1 to Line 15 is used to iterate from stage $|\mathcal{G}|$ to stage 1. For each stage, each possible state is examined for calculating the cumulative payoff as shown between Line 2 and Line 14. The best payoff for each state is obtained by solving Problem (11) as shown between Line 6 and Line 11. Once the optimal $\mathbb{P}_{\mathrm{MD}}^{(j)}$ leading to the best payoff is obtained for each state, it is recorded to facilitate back-tracing when stage 1 reaches. Line 16 to Line 19 show how back-trace from stage 1 to stage $|\mathcal{G}|$ is done for retrieving the optimal sequence of states and detection thresholds. We note that while the feasible region of $\mathbb{P}_{\mathrm{MD}}^{(j)}$ of each detector j is between 0 and 1, in traversing states from stage $|\mathcal{G}|$ backwards, we can utilize the concept of the cutting plane method [3] to cut the feasible region and reduce the computation complexity. Briefly, since $S^{(1)} = \varepsilon/\mathbb{P}_{\mathrm{I}}^{(k)}$ and $S^{(j+1)} = S^{(j)}/\mathbb{P}_{\mathrm{MD}}^{(j)}$, we have $S^{(j)} \in [\varepsilon/\mathbb{P}_{\mathrm{I}}^{(k)}, 1]$ and $\mathbb{P}_{\mathrm{MD}}^{(j)} \in [S^{(j)}, 1]$. Limiting the search range for $S^{(j)}$ in each stage (Line 2) and $\mathbb{P}_{\mathrm{MD}}^{(j)}$ in each state (Line 6) can significantly reduce the computation complexity.

4 Evolution of Cooperation

The game presented in Section 3 can be regarded as a one-shot game that only models spectrum sensing within a frame in Figure 1, where secondary users pay the cost of spectrum sensing and earn the payoff from spectrum usage. For a group of backlogged secondary users, however, the demand for spectrum access can extend to several frames. Therefore, the one-shot game can be played repeatedly for continuous sensing and access of the spectrum. As we have shown in Section 3, the Nash equilibrium of the one-shot game is not the social optimum due to individual players selfishly maximizing their own utilities. To stimulate cooperation, we design a repeated game in this section with the goal of achieving the socially optimal equilibrium. Each secondary user in the repeated game maintains a list to remember payoff and corresponding actions in the past such that a player can adapt its best response based on the concept of evolution. In addition, we design the mechanism for cooperative players to punish selfish ones through repetitions of the game. In the following, we first analyze and classify possible strategies for a player to adopt in the repeated game, and then we present the evolutionary learning model.

4.1 Strategies in the Repeated Game

In a repeated game, a player may take into account the impact of its current action on future actions of other players. In addition, a player may have its own "personality" in choosing the best strategy of action during repetitions of the game. We classify four different strategies (roles) to model the behavior of players as follows:

Solitary. A solitary, say user j, is a conservative player who sets the detection threshold as in the case of single-node sensing. As a result, even if all other players are free-riders such that $\mathbb{P}_{\mathrm{MD}}^{(i)} = 1$, $i \neq j$, a solitary can still ensure that the final decision $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})} = \mathbb{P}_{\mathrm{MD}}^{(j)}$ will meet its interference constraint. Specifically, a solitary sets $\mathbb{P}_{\mathrm{MD}}^{(j)}$ as $\varepsilon/\mathbb{P}_{\mathrm{I}}^{(j)}$ to satisfy the interference constraint $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}\mathbb{P}_{\mathrm{I}}^{(j)} \leq \varepsilon$ while maximizing its utility. Clearly, without relying on the cooperation of others, a solitary can always get non-zero payoff as the case for single-node sensing. However, it cannot benefit from the increased payoff due to cooperative sensing. Hence, if not all player are free-riders, $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ is smaller than the expected value, and the utility of a solitary is not maximized.

Leader. In contrast to a solitary, a leader aims to fully utilize the gain of cooperative sensing through active coordination of the negotiation process among players. Effectively, a leader sets its constraint as the most stringent one such that $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})} = \varepsilon/\mathbb{P}_{\mathrm{I}}^{(j)}$. It then applies Algorithm 1 to compute the optimal detection thresholds and then broadcasts the results to all players. In this way, its utility is maximized if all other players follow the coordination; if some players do not follow the coordination, however, its utility is not maximized.

Follower. Instead of coming up with the "optimal" threshold for itself, a follower simply uses the suggestion from other players. If there is only one leader in the game, a follower simply follows the command of the former. If there are more than one leaders in the game, a follower adopts the command from the leader with the minimum detection threshold value. Finally, if no leader exists in the game, a follower copies the minimum threshold among other players.

Glutton. A glutton, say player j, at each repetition maximizes its utility by pushing $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})}$ to its constraint boundary such that $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})} = \varepsilon/\mathbb{P}_{\mathrm{I}}^{(j)}$. To do so, a glutton collects the responses from other players and then makes its decision as $\mathbb{P}_{\mathrm{MD}}^{(j)} = \min\{1, \varepsilon/(\mathbb{P}_{\mathrm{I}}^{(j)} \prod_{i\in\mathcal{G}, i\neq j} \mathbb{P}_{\mathrm{MD}}^{(i)})\}$, where $\mathbb{P}_{\mathrm{MD}}^{(i)}$, $i \neq j$ are responses of other players. It is possible, however, that a glutton fails to collect the correct responses from all other players such that estimation of responses is needed. While over-estimation of responses results in less payoff, under-estimation results in zero payoff due to violation of the interference constraint.

In addition to the aforementioned four strategies in the repeated game, a special strategy called **Avenger** is proposed to stimulate cooperation. An avenger is a special role used for punishing non-cooperation (selfishness) of other players. If the final result of the negotiation fails to satisfy the interference constraint of a player, the "unsatisfied" player deviates from the original role by always claiming a positive decision ($\mathcal{D} = 1$) regardless of its sensing result. Consequently, $\mathbb{P}_{\mathrm{MD}}^{(j)} = 0$ and hence $\mathbb{P}_{\mathrm{MD}}^{(\mathcal{G})} = 0$ for the whole group. Equivalently, the expected payoff is zero for all players in the group, including the avenger itself. Since no player including selfish ones can get non-zero payoff, it is possible that in the next repetition non-cooperative players will be less "selfish" for improving its utility.

We have identified four possible strategies (roles) in the repeated game, but a secondary user may still need some rule for determining a suitable role for maximizing its utility. We present in the next section how the theory of evolution can be used for role learning.

4.2 Evolution of Strategies

As in the theory of evolution, a player selects the fittest strategy to be applied (for survival) in the game. If a strategy can bring more payoff than others to a player, the player has tendency to use it in ensuing game repetitions and hence the probability (time ratio) of choosing that strategy will increase. The replicator equation used in evolutionary game theory [16] for modeling the population of a species under *Nature* selection can then be applied to govern the dynamic increase and decrease of the probability distribution.

To start, Algorithm 2 shows the flow of how a secondary user plays the repeated game and learns its role based on the replicator equation. Let $\mathbb{P}_{\mathrm{S}}^{(j)}$, $\mathbb{P}_{\mathrm{L}}^{(j)}$, $\mathbb{P}_{\mathrm{F}}^{(j)}$ and $\mathbb{P}_{\mathrm{G}}^{(j)}$ denote the probabilities (time ratios) of player j choosing Solitary, Leader, Follower and Glutton respectively as the strategy. Initially, a player assigns equal probability to each strategy (Line 1). After the game starts, in each repetition, a player randomly chooses a strategy as the action based on the current distribution of strategies (Line 3). When a strategy is selected, a player follows the rule of the strategy to determine its $\mathbb{P}_{\mathrm{MD}}$ and detection threshold as presented in Section 4.1 (Line 4). After the negotiation process is complete,

Algorithm 2. Role evolution of player $j \in \mathcal{G}$

Input: Protection constraint $\varepsilon/\mathbb{P}_{\mathrm{I}}^{(j)}$ and learning step size α
Output: $\mathbb{P}_{\mathrm{S}}^{(j)}$, $\mathbb{P}_{\mathrm{L}}^{(j)}$, $\mathbb{P}_{\mathrm{F}}^{(j)}$ and $\mathbb{P}_{\mathrm{G}}^{(j)}$

01: $\mathbb{P}_{\mathrm{S}}^{(j)} = \mathbb{P}_{\mathrm{L}}^{(j)} = \mathbb{P}_{\mathrm{F}}^{(j)} = \mathbb{P}_{\mathrm{G}}^{(j)} = 0.25$ and $t = 1$
02: **while** $(j \in \mathcal{G})$
03: $\quad A[t] = \mathrm{rand}(\mathbb{P}_{\mathrm{S}}^{(j)}, \mathbb{P}_{\mathrm{L}}^{(j)}, \mathbb{P}_{\mathrm{F}}^{(j)}, \mathbb{P}_{\mathrm{G}}^{(j)})$
04: $\quad$ Claim $\mathbb{P}_{\mathrm{MD}}^{(j)}$ according to role $A[t]$
05: $\quad$ **if** $\prod_{i \in \mathcal{G}} \mathbb{P}_{\mathbf{MD}}^{(i)} > \varepsilon/\mathbb{P}_{\mathbf{I}}^{(j)}$
06: $\qquad$ Become an Avenger
07: $\quad$ **end if**
08: $\quad$ Apply corresponding θ_j to sense the spectrum
09: $\quad$ Earn the corresponding utility $U^{(j)}[t]$ as shown in Equation (7)
10: $\quad$ Record the best response up to t: $A^* = A[\arg \max_{1 \le \tau \le t} U^{(j)}[\tau]]$
11: $\quad$ Update fitness: $\phi_j(A^*) = \phi_j(A^*) + 1$
12: $\quad$ Update strategy distribution:
$\quad \mathbb{P}_a^{(j)} = \left\{(1+\alpha)\left(\phi_j(a) - \bar{\phi}_j\right)\right\} \mathbb{P}_a^{(j)}, a \in \{\mathrm{S, L, F, G}\}$
13: $\quad t = t + 1$
14: **end while**

a player checks $\mathbb{P}_{\mathrm{MD}}$ claimed by all other players. If the player fails to access the spectrum due to violation of its interference constraint, it switches to an Avenger; otherwise, it uses the sensing parameter as claimed to sense the spectrum (Line 5 to Line 8), and then evaluates its payoff using the utility function shown in Equation (7). A player maintains a data structure to remember its responses and the correspond payoff in history (Line 10). The player can then update the distribution of strategies using the replicator equation as follows.

In the replicator equation, the increasing rate of strategy a of player j can be modeled as:

$$\dot{\mathbb{P}}_a^{(j)} = \left[\phi_j(a) - \bar{\phi}_j\right] \mathbb{P}_a^{(j)},\ a \in \{\mathrm{S, L, F, G}\}, \tag{12}$$

where $\phi_j(a)$ is the fitness of strategy $a \in \{\mathrm{S, L, F, G}\}$ and $\bar{\phi}_j$ is the mean fitness of strategies of player j that can be written as $\bar{\phi}_j = \sum_{a \in \{\mathrm{S,L,F,G}\}} \mathbb{P}_a^{(j)} \phi_j(a)$. If a strategy has larger fitness than the mean fitness, its increasing rate $\dot{\mathbb{P}}_a^{(j)}$ is positive. Otherwise, the rate decreases. In addition, since the population of the next generation also depends on the current population, the increasing rate $\dot{\mathbb{P}}_a^{(j)}$ is proportional to the current value of $\mathbb{P}_a^{(j)}$. To reflect the objective of a player, the fitness of a strategy is designed as the times of bringing maximum payoff through history. Player j can then update the probability of each strategy in each repetition based on the following equation:

$$\mathbb{P}_a^{(j)} = \mathbb{P}_a^{(j)} + \alpha \dot{\mathbb{P}}_a^{(j)},\ a \in \{\mathrm{S, L, F, G}\}, \tag{13}$$

where α is a positive constant indicating the learning step size. It can be easily shown that the sum of $\mathbb{P}_a^{(j)}$ is equal to 1 after the update because the sum of

$\mathbb{P}_a^{(j)}$ is one initially and the sum of $\dot{\mathbb{P}}_a^{(j)}$ is zero. As repetition proceeds, a role with the largest probability for a player becomes the best role for itself.

5 Evaluation Results

In this section, we first show the convergence of the repeated game based on Algorithm 2 and then show its performance gain compared to the Nash equilibrium of the stage game under different scenarios.

5.1 Evolution of Strategies

We consider a network of three secondary users with different levels of interference to the primary user. $\mathbb{P}_{\mathrm{I}}$ of three users are 0.99, 0.9 and 0.6 respectively, and the requirement of the protection threshold ε is 0.05. Figure 2 thus shows the evolution of roles for players 1 and 3. Initially, each player assigns equal weight to each role. After 100 repetitions of the sensing game, the weight of each role becomes different, where player 1 has more probability to be a Glutton or a Solitary and player 3 enjoys the benefits of being a Leader. After another 100 repetitions of the game, the distributions exhibit quite different behaviors. Since player 1 has the most stringent constraint, it is okay for it to be a Leader while the other two act as Followers. On the other hand, it would be problematic if player 3 becomes a Leader or player 1 becomes a Glutton. Thus, the probability of (Leader, Follower, Follower) for players 1, 2, and 3 respectively shows clear increase. After 400 repetitions, the optimal roles for individual players become clear where player 1 acts as the Leader to solve the detection thresholds for all players, and other players follow the coordination of the Leader. Since detection thresholds are solved with the tightest constraint by Algorithm 1, it is also a social optimal solution. On the other hand, if the game is played without the Avenger, players quickly realize that greedy and risky strategies bring zero payoff. Consequently, all players become solitary and act conservatively. The results without the Avenger are not presented due to lack of space.

5.2 Gain of Evolution

Figure 3 shows the results when the cooperative set varies from 2 to 6 nodes. The right Y-axis shows the total expected throughput of the secondary users, and the left Y-axis shows the price of anarchy (PoA) defined as the achieved value over the social optimal value. It can be observed from the figure that in all cases the total expected throughput in the repeated game approaches the optimal performance and outperforms the Nash equilibrium. In turns of the PoA, it can be observed that the PoA of the Nash equilibrium decreases from 73% to 65% as the number of players increases. The reason is that under the Nash equilibrium, all players rely on the one with the most stringent constraint to sense the spectrum, and the gain of cooperation due to node diversity is not fully utilized.

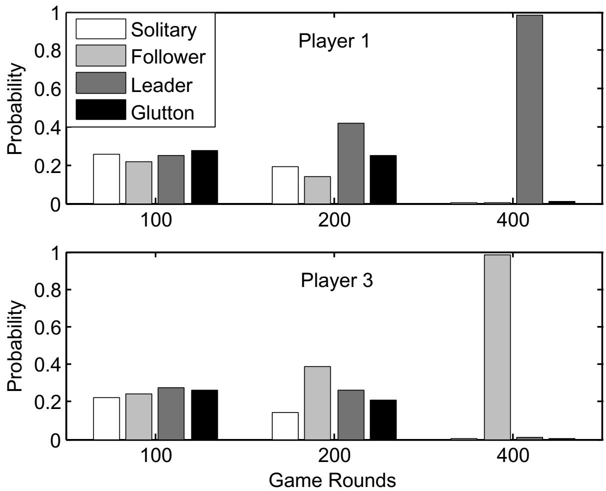

Fig. 2. Evolution of strategies in the repeated game

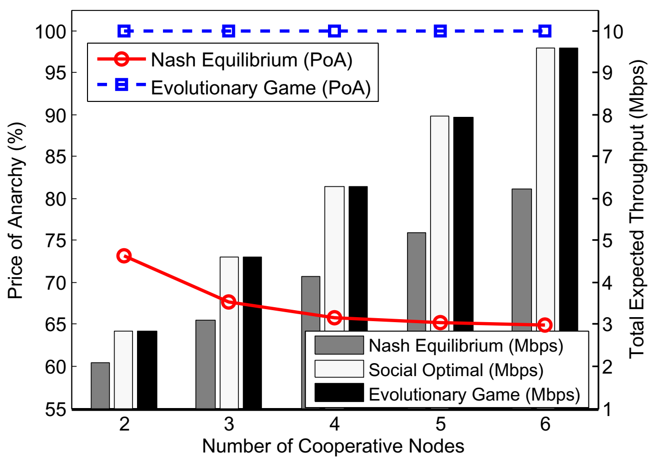

Fig. 3. Performance for different numbers of players

Figure 4 shows the results as the external interference varies from 0 to 1 (1 means that additional interference equal to the background noise is added during spectrum sensing). Note that as the external interference increases, it is more difficult to correctly detect the activity of the primary user and cooperation among nodes becomes more important. As the figure shows, the performance gain (compared to the result of the Nash equilibrium) of the repeated game increases as external interference increases and as the number of nodes increases. This substantiates that the proposed repeated game can indeed address the non-cooperative problem in the original one-shot game.

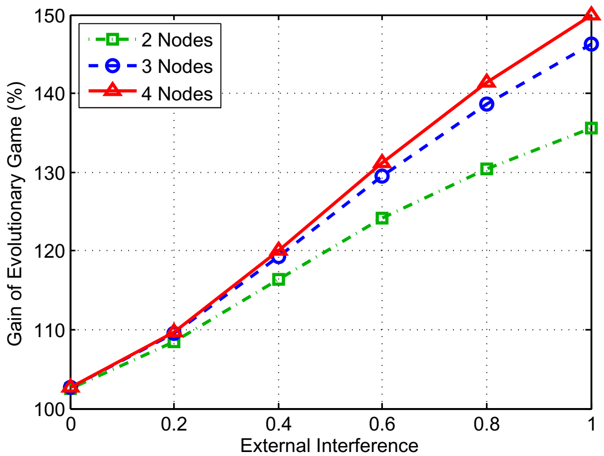

Fig. 4. Performance gains for varying external interference

6 Conclusions

In this work, we investigated the problem of interference-aware spectrum sensing for opportunistic spectrum access in cognitive radio networks. We showed that because different secondary users may have different levels of interference to the primary user, there is a conflict in setting the optimal sensing parameters for cooperative spectrum sensing. An interference-aware sensing game does not solve the problem since its Nash equilibrium will deviate from the social optimum. To address this problem, we designed a repeated game based on evolutionary game theory so players have the chance to revenge "non-cooperative" players in ensuing repetitions for driving the equilibrium to the social optimum. We showed through numerical results that the proposed repeated game does achieve the desirable performance for interference-aware cooperative sensing in opportunistic spectrum access.

Acknowledgements. This work was supported in part by funds from the National Science Council under Grants NSC-99-2219-E-002-020 and NSC-99-2911-I-002-001.

References

1. Bellman, R.: The theory of dynamic programming. Bulletin of the American Mathematical Society 60, 503–516 (1954)
2. Chow, C.S., Tsitsiklis, J.: An optimal multigrid algorithm for continuous state discrete time stochastic control. In: Proceedings of the 27th IEEE Conference on Decision and Control, vol. 3, pp. 1908–1912 (December 1988)
3. Cornuejols, G.: Valid inequalities for mixed integer linear programs. Mathematical Programming 112(1), 3–44 (2008)

4. Datla, D., Wyglinski, A., Minden, G.: A spectrum surveying framework for dynamic spectrum access networks. IEEE Transactions on Vehicular Technology 58(8), 4158–4168 (2009)
5. Digham, F.F., Alouini, M.S., Simon, M.K.: On the energy detection of unknown signals over fading channels. IEEE Transactions on Communications 55(1), 21–24 (2007)
6. Ganesan, G., Li, Y.: Cooperative spectrum sensing in cognitive radio networks. In: Proceeding of IEEE DySPAN, Baltimore, MD, USA, pp. 137–143 (November 2005)
7. Gardner, W.: Signal interception: A unifying theoretical framework for feature detection. IEEE Transactions on Communications 36(8), 897–906 (1988)
8. Ghasemi, A., Sousa, E.: Collaborative spectrum sensing for opportunistic access in fading environments. In: IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, pp. 131–136 (November 2005)
9. Huang, S., Liu, X., Ding, Z.: Optimal transmission strategies for dynamic spectrum access in cognitive radio networks. IEEE Transactions on Mobile Computing 8(12), 1636–1648 (2009)
10. Kay, S.M.: Fundamentals of Statistical Signal Processing: Detection Theory. Prentice Hall (February 1998)
11. Liang, Y.C., Zeng, Y., Peh, E., Hoang, A.T.: Sensing-throughput tradeoff for cognitive radio networks. IEEE Transactions on Wireless Communications 7(4), 1326–1337 (2008)
12. Moghimi, F., Nasri, A., Schober, R.: LP-norm spectrum sensing for cognitive radio networks impaired by non-Gaussian noise. In: Proceedings of IEEE Global Telecommunications Conference (GLOBECOM), pp. 1–6 (November 2009)
13. Peh, E., Liang, Y.C., Guan, Y.L., Zeng, Y.: Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view. IEEE Transactions on Vehicular Technology 58(9), 5294–5299 (2009)
14. Sahai, A., Tandra, R., Mishra, S.M., Hoven, N.: Fundamental design tradeoffs in cognitive radio systems. In: Proceedings of the First International Workshop on Technology and Policy for Accessing Spectrum (TAPAS), New York, NY, USA (2006)
15. Tuy, H.: Monotonic optimization: Problems and solution approaches. SIAM Journal on Optimization 11(2), 464–494 (2000)
16. Weibull, J.W.: Evolutionary Game Theory. MIT Press (1995)
17. Yucek, T., Arslan, H.: A survey of spectrum sensing algorithms for cognitive radio applications. IEEE Communications Surveys and Tutorials 11(1), 116–130 (2009)
18. Zhang, T., Wu, Y., Lang, K., Tsang, D.: Optimal scheduling of cooperative spectrum sensing in cognitive radio networks. IEEE Systems Journal 4(4), 535–549 (2010)