An Indirect Reciprocity Game Theoretic Framework for Dynamic Spectrum Access

Biling Zhang[1,2], Yan Chen[1], and K. J. Ray Liu[1]
[1]Department of Electrical and Computer Engineering,
University of Maryland, College Park, MD 20742, USA.
[2]School of Network Education,
Beijing University of Posts and Telecommunications, Beijing, 100876, P. R. China
E-mail:{bzhang15, yan, kjrliu}@umd.edu

***Abstract*—In this paper, we propose a spectrum access framework to address the efficient allocation of channels of a base station (BS) by stimulating cooperation between primary users (PUs) and secondary users (SUs). We model the cooperation stimulation problem as an indirect reciprocity game, where SUs help PUs relay information and gain reputations to access the vacant channels in the future. We design a reputation updating policy under time-varying channels and prove the existence of stationary reputation distribution. We further formulate the decision making of an SU as a Markov Decision Process (MDP) and use a modified value iteration algorithm to find the optimal action rule. Moreover, we theoretically derive the condition under which the optimal action rule is an evolutionarily stable strategy (ESS). Finally, simulation results are shown to verify the effectiveness of the proposed scheme.**

***Index Terms*—Indirect reciprocity, dynamic spectrum access, game theory, Markov decision process, evolutionarily stable strategy**

I. INTRODUCTION

Cooperation stimulation among selfish nodes in dynamic spectrum access (DSA) networks is an important problem. Since primary users (PUs) and secondary users (SUs) generally belong to different operators or service providers and have different objectives, they may not cooperate if cooperation cannot bring them benefits. Game theory, a well-developed mathematical tool that studies the interaction among rational users [1] [2], has been used to analyze the cooperation behaviors among nodes in DSA networks [3] [4] [5] [6]. However, most of the existing game theoretic frameworks for DSA networks are based on the direct reciprocity model and repeated games where the underlying assumption is that the game among a group of nodes is played for infinite times. Nevertheless, this assumption is not true in reality since players need to change their partners frequently due to mobility or changes of environment. In such a case, the only optimal strategy for them is to always play non-cooperatively. Moreover, without punishment for the cheating behavior, players may cheat even after they have agreed to cooperate.

Another two categories of cooperation stimulation schemes are payment-based and reputation-based schemes. Although the payment-based schemes can achieve promising results, the requirement of tamper-proof hardware or central billing services greatly limits their potential applications. Until now few reputation-based mechanisms have been proposed for DSA networks, and there is no theoretical analysis about using reputation to stimulate cooperation in DSA networks.

Indirect reciprocity has recently drawn a lot of attentions in the area of social science and evolutionary biology [7] [8]. Chen and Liu [9] have showed that indirect reciprocity game is a new paradigm to stimulate cooperation among cognitive nodes. However, the framework proposed in [9] is too general and cannot be directly applied to DSA networks. Therefore a careful design for the resource allocation by taking into account the overall system performance is needed.

In this paper, we propose a spectrum access framework to address the fair and efficient allocation of both used and unused channels by incorporating indirect reciprocity based incentive mechanism to stimulate cooperation between primary and secondary users and thus improve the overall system performance. We formulate the SUs' decision making as a Markov Decision Process (MDP) and use a modified value iteration algorithm to find the optimal action rule. We show that within an appropriate cost-to-gain ration, the optimal strategy for SUs is to use the maximum power to relay if the source-relay-destination channel does not encounter an outage. We also find that this optimal action rule will lead to a "good" society where almost all the SUs have the highest reputation, i.e., perfect cooperation between the primary system and the secondary system. Finally, we derive theoretically the stable condition of the optimal action rule.

The rest of the paper is organized as follows. Section II describes in details our system model. In Section III we introduce the reputation updating model, characterize the equilibrium of the indirect reciprocity game, and analyze the stability of optimal action rule. Finally, we show the simulation results in Section IV and draw conclusions in Section V.

II. SYSTEM MODEL

As shown in Fig. 1, we consider the system where PUs far away from the BS are experiencing low achievable rates while some SUs between the PUs and the BS have more favorable channel conditions and are searching opportunities to use the licensed spectrum.

To improve the PUs' QoS, some SUs are selected as relays by the BS based on a certain relay selection strategy. The SUs who are selected as relays need to make decision whether to

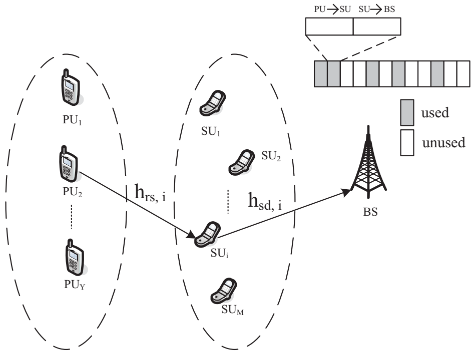

Fig. 1. The System Model

help PUs relay the information, and if help, how much power to use for relaying. Although our scheme is applicable to any cooperation protocol, here, we assume SUs use Decode-and-Forward (DF) protocol to relay PUs' signals. Since there may be more than one SU selected as relays for a PU, a time frame is divided into several time slots, in the first of which the PU transmits and the selected SUs relay PU's information in the following. After the transmission, the BS will update the selected SUs' reputations. Later, when SUs apply for the usage of vacant channels, they will be given the right for accessing certain amount of vacant channels by the BS according to their reputations. From the discussion above, we can see that the concept of indirect reciprocity adopted in our model is that the BS helps SUs because SUs have helped PUs.

A. Action

In our model, action $a \in \mathcal{A}$ stands for the power level used by an SU when relaying the data of a PU, where $\mathcal{A} = \{0, 1, ..., N\}$ is the action set that an SU can choose its action from when it is selected as a relay. In the action set $\mathcal{A}$, "0" represents that an SU uses zero power to relay, i.e., the SU denies cooperation, while "N" denotes that an SU relays a PU's data with the maximum power P_{max}. Then $a_{ij} \in \mathcal{A}$ is used to describe the action performed by an SU with reputation i under source-relay-destination channel condition j.

B. Social Norms

A social norm Q is a matrix used for assigning the immediate reputation of players. In social norm Q, the element $Q_{i,j}$ represents the reputation assigned to an SU immediately after it takes an action i based on the source-relay-destination channel quality j. Without loss of generality, we assume that all SUs in the system share the same norm. In our model, $Q_{i,j}$ is defined as

$$Q_{i,j} = \begin{cases} 0, & j = 0, \\ i, & j \neq 0. \end{cases} \tag{1}$$

The social norm (1) is designed to encourage an SU to relay the PU's signal using a higher power by assigning the SU a larger reputation as long as the channel is not in outage.

C. Power level and Relay Power

In this subsection, we will discuss how to quantize the relay power to $N+1$ levels based on the outage probability. According to [11], the maximum average mutual information for the repetition-coding DF protocol is

$$\begin{aligned} I_{DF} \;=\; & 0.5\min\{log(1+SNR_{sr}|h_{sr}|^2), \\ & log(1+SNR_{sd}|h_{sd}|^2+SNR_{rd}|h_{rd}|^2)\}, \end{aligned} \tag{2}$$

where $SNR_{sd} = SNR_{sr} = P_s/\sigma_n^2$, $SNR_{rd} = P_r/\sigma_n^2$, P_s is the transmission power of the PU, P_r is the relay power of the SU and σ_n^2 is the variance of additive Gaussian white noise. Here we assume σ_n^2 is the same for all S-R, R-D and S-D channels.

Assuming that all S-R, R-D and S-D channels are complex Gaussian, i.e. $h_{sr} \sim \mathcal{CN}\left(0, \sigma_{sr}^2\right)$, $h_{rd} \sim \mathcal{CN}\left(0, \sigma_{rd}^2\right)$ and $h_{sd} \sim \mathcal{CN}\left(0, \sigma_{sd}^2\right)$. Given the required transmission rate R, the outage probability at the BS with DF relay protocol can be derived according to [11]. Therefore given R, if we quantize the required outage probability interval $[p_{min}^{out}, p_{max}^{out}]$ into $N-1$ levels $\mathcal{P}^{out} = \{p_1^{out}, p_2^{out}, ..., p_{N-1}^{out}\}$, then $\forall p_i^{out} \in \mathcal{P}^{out}$, the actual power at each power level $P_{r,i}$ an SU should use can be obtained according to the outage probability. For $p^{out} > p_{max}^{out}$, we quantize it to the $0th$ level and set $P_{r,0} = 0$, and for $p^{out} < p_{min}^{out}$, we quantize it to the Nth level and set $P_{r,N} = P_{max}$.

D. Channel Quality Distribution

The quality of source-relay-destination channel is also quantized to $N+1$ levels, i.e., 0, 1, ..., N. Here the level "0" means the channel has encountered an outage while the level "1" to the level "N" stand for different channel qualities. Next we will discuss how to acquire the channel quality distribution H with the kth element $H(k)$ being the probability of channel quality at the level k.

In our work, an SU which has the "best" source-relay-destination channel, i.e., $h^* = \max_i \min(|h_{sr,i}|^2, |h_{rd,i}|^2)$ will be selected. Suppose there are M SUs in the secondary system. Since the S-R channel and the R-D channel for all SUs are independent identical distribution (i.i.d) complex Gaussian channels, the PDF of h^* can be found as follow

$$f_{h^*}(x) = M(\sigma_{sr}^{-2}+\sigma_{rd}^{-2})\left(1-e^{-(\sigma_{sr}^{-2}+\sigma_{rd}^{-2})x}\right)^{M-1} e^{-(\sigma_{sr}^{-2}+\sigma_{rd}^{-2})x}. \tag{3}$$

Finally, by dividing the domain of h^* into $N+1$ intervals and integrating $f_{h^*}(x)$ over each interval, we obtain the channel quality distribution H.

III. Proposed Algorithm and Analysis

In this section we first propose a reputation updating model by incorporating the impact of time-varying channel on SU's action. Then we analyze the existence of stationary reputation distribution of the whole population. Finally, we characterize the equilibrium of our indirect reciprocity game and the stable condition for the optimal action rule.

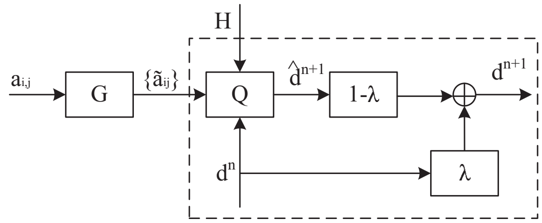

Fig. 2. Reputation Update Model

A. Reputation Updating Model

In order to establish the reputation of an SU with the social norm, we develop a reputation updating model as shown in Fig 2. Since an SU's action may be distorted by the time-varying R-D channel and hence falsely detected by the BS, here we use G, which will be discussed in the next subsection, as the power detection transition matrix to describe the impact of R-D channel on SU's action.

We assign a reputation distribution d for each SU to capture all the likelihoods of the SU's reputation. Let $d = [d(0),\ d(1), ..., d(N)]^T$ be the reputation distribution for a specific SU, then $d(k)$ denotes the probability of the SU's being assigned with reputation k. Suppose, at time index n, an SU with reputation i is selected under channel quality j and performs action $a_{i,j}$. Since the action may be distorted during the transmission, the BS may detect $a_{i,j}$ as $\tilde{a}_{i,j}$. In such a case, an immediate reputation distribution $\hat{d}^n$ will be assigned to the SU according to the social norm. Finally, the BS updates the SU's reputation distribution in time index $n+1$, d^{n+1}, using a linear combination of the SU's original reputation distribution d^n and the immediate reputation distribution $\hat{d}^n$ with a weight λ. Here the weight λ can be treated as a discount factor of the past reputation.

In a simple case, an SU's reputation distribution is e_i where e_i is the standard basis vector. According to Fig 2, after the SU takes action $a_{i,j}$, its reputation distribution $d_{i,j,a_{i,j}}$ can be updated as

$$d_{i,j,a_{i,j}} = (1-\lambda)\sum_{l=0}^{N}\left[e_{Q_{l,j}} \cdot G_{a_{i,j},l}\right] + \lambda e_i. \quad (4)$$

where $G_{a_{i,j},l}$ is the probability that the SU's action is $a_{i,j}$ but detected as l, and $Q_{l,j}$ is the immediate reputation assigned to the SU when it takes an action l with channel quality j.

B. Power Detection and G-matrix

In this subsection we will discuss how the BS estimates the power level an SU uses through energy detection. Let x_i, $i = 1, 2, \dots N_T$ be the signals modulated with ASK, PSK or FSK and broadcasted by a PU in T frames. After relayed by an SU with the DF protocol, the signals received by the BS are

$$y_i = \sqrt{P_r} h_{rd} \hat{x}_i + n_i,\ i = 1, 2, \dots, N_T, \quad (5)$$

where $\hat{x}_i$ is the signal decoded by SU, P_r is the SU's relay power, $h_{rd} \sim \mathcal{CN}\left(0, \sigma_{rd}^2\right)$, $n_i \sim \mathcal{N}\left(0, \sigma_n^2\right)$, and $E\left(\hat{x}_i^2\right) = \sigma_x^2$. After coherent demodulation and being matched with the constellation map, $\hat{x}_i$ can be obtained, based on which the BS can further estimate the relay power the SU used.

Suppose that both the channel state information of the R-D channel and the relay power the SU used remain unchanged in T frames. To improve the accuracy of estimation, the BS combines N_T received signals $y_i = \sqrt{P_r} h_{rd} \hat{x}_i + n_i$, $i = 1, 2, \dots, N_T$ using Maximal Ratio Combining (MRC) and gets

$$z \ = \ h_{rd}^* \vec{x}^H \vec{y} = \sqrt{P_r} |h_{rd}|^2 ||\vec{x}||^2 + h_{rd}^* \vec{x}^H \vec{n}. \quad (6)$$

From (6) we can see that z follows a Gaussian distribution, i.e.,

$$z_i \sim \mathcal{N}\left(N_T \sigma_{rd}^2 \sigma_x^2 B_i,\ N_T \sigma_{rd}^2 \sigma_x^2 \sigma_n^2\right) = \mathcal{N}\left(\mu_i, \sigma^2\right), \quad (7)$$

where $\mu_i = N_T \sigma_{rd}^2 \sigma_x^2 B_i$, $\sigma = N_T \sigma_{rd}^2 \sigma_x^2 \sigma_n^2$ and $B_i = \sqrt{P_{r,i}}$.

Since all z_is are Gaussian, the optimal detection threshold would be $v_i = \frac{\mu_i + \mu_{i+1}}{2}$, $i = 0, 1, ...N-1$. With the optimal thresholds, we have the power detection probabilities $\Pr(B_j|B_i)\ \forall i, j$ which are used to construct the probability transition matrix G as follows.

$$G_{i,j} = \Pr(B_j|B_i) = \begin{cases} 1 - Q(\frac{\mu_0 + \mu_1 - 2\mu_i}{2\sigma}), & j = 0 \\ Q(\frac{\mu_{j-1} + \mu_j - 2\mu_i}{2\sigma}) - Q(\frac{\mu_j + \mu_{j+1} - 2\mu_i}{2\sigma}), & j \in (0, N) \\ Q(\frac{\mu_{N-1} + \mu_N - 2\mu_i}{2\sigma}), & j = N \end{cases} \quad (8)$$

where $G_{i,j} \in G$ denotes the probability of an SU using power level i to relay while being detected as power level j.

The transmission matrix G in (8) has two nice properties as described below in Lemma 1 and Lemma 2, and these two properties will be used in later analysis. Due to page limitation, the proof of these two Lemmas are shown in the supplementary information [12].

Lemma 1: $G_{N,j} < G_{i,j}$, $\forall\, 0 \le i < N$ and $0 \le j < i$.

Lemma 2: $G_{i,i} > G_{i,j}$, $\forall\, 0 \le i, j \le N$ and $i \neq j$.

C. Stationary Reputation Distribution

Let $\eta = \{\eta_0, \eta_1, ..., \eta_N\}$ be the reputation distribution of the entire population, where η_i is the portion of the population that has reputation i. After the transmission, the reputation of the SU is updated according to the reputation updating policy and the new reputation distribution of the entire population η^{new} can be computed by

$$\eta^{new} = (1-\lambda)\eta p_T + \lambda\eta \quad (9)$$

where p_T is the reputation transition matrix. Each element $p_{T_{i,j}}$ in p_T stands for the probability of reputation i turning into j after SU has taken an action. According to Fig. 2, $p_{T_{i,j}}$ can be calculated as

$$p_{T_{i,j}} = \sum_{k,l:Q_{l,k}=j} G_{a_{i,j}^*,l} H(k), \quad (10)$$

where $a_{i,j}^*$ is the optimal action an SU with reputation i takes under channel condition j.

In a steady state, we have $\eta^{new} = \eta$. Therefore, according to (9), if there exists a stationary reputation distribution η^*, it must be the solution to the following equation

$$\eta^* p_T = \eta^* \quad (11)$$

In the following Theorem 1, we will show that there is a unique stationary reputation distribution.

Theorem 1: There exists a unique stationary reputation distribution of the whole population for any given optimal action rule and the stationary reputation distribution is p_T's eigenvector corresponding to the eigenvalue 1.

Proof: Due to page limitation, the proof is shown in the supplementary information [12]. ■

D. Payoff Function and Equilibrium Analysis

In this subsection, we will discuss how an SU chooses the optimal action. By helping relay the PU's data, the SU will incur a cost that is determined by power consumption. Here, we assume a linear cost function $C(a_{i,j}) = ca_{i,j}$ with c being the price per unit power level.

On the other hand, the SU will gain benefit from the allocation of certain vacant channels in the future based on its newest reputation after relaying. Suppose there are Y channels and let p be the probability of a channel being unoccupied, then the expected amount of vacant channels is $N_v = pY$. In our model, the BS allocates the vacant channels to all M SUs in a proportionally fair manner. In such a case, the number of vacant channels $t_{i,j,a_{i,j}}$ that an SU will be allocated after it takes action $a_{i,j}$ is

$$t_{i,j,a_{i,j}} = n_t \frac{\sum_{n=0}^{N} n \cdot d_{i,j,a_{i,j}}(n)}{\sum_{n=0}^{N} n \cdot \eta^*(n)}, \tag{12}$$

where $n_t = N_c N_v / M$ and N_c is the maximal allowed number of SUs in one channel. Suppose the gain from accessing one channel is a constant g, then the total gain of an SU taking action $a_{i,j}$ is $T(a_{i,j}) = P_v g t_{i,j,a_{i,j}}$ with $P_v = 1 - (1-p)^Y$ being the probability of BS having vacant channels.

Let $W_{i,j}$ denote the maximum payoff an SU, currently with reputation i and taking action $a_{i,j}$ under channel condition j, can obtain from this interaction to future. Then, we have

$$W_{i,j} = \max_{a_{i,j} \in \mathcal{A}} -ca_{i,j} + P_v g t_{i,j,a_{i,j}} + \delta \sum_k \sum_l d_{i,j,a_{ij}}(k) H(l) W_{k,l} \tag{13}$$

The first term in (13) is the immediate cost an SU incurs by taking action $a_{i,j}$, the second term is the immediate gain it may obtain, and the last term is the long-term benefit it receives in the future with a discount factor δ.

Note that (13) is a Bellman equation of $W_{i,j}$ and thus finding the optimal action rule is an MDP, which can be solved using a modified value iteration algorithm [9]. By adjusting the initial conditions in the modified value iteration algorithm, we find two possible action rules as follows

$$a_1^*(i,j) = 0, \forall i, j. \tag{14}$$

and

$$a_2^*(i,j) = \begin{cases} 0, & j = 0; \\ N, & j \neq 0. \end{cases} \tag{15}$$

The first optimal action a_1^* is a non-cooperative equilibrium, where SUs always deny cooperation. The second optimal action a_2^* is the desired equilibrium where SUs always use their maximum power to relay PUs' data unless the source-relay-destination channels encounters an outage.

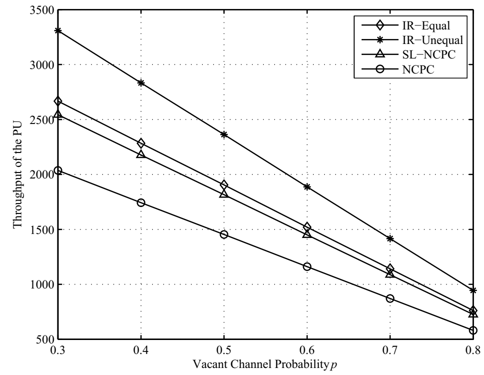

Fig. 3. Throughput of PU

E. Stability of Optimal Action Rule

In this subsection, we will discuss the stable condition for a_2^*. Let $F(a_{i,j}, a^*)$ be the payoff an SU can obtain when it takes action $a_{i,j}$ while others take the optimal action a_{ij}^*, then,

$$F(a_{i,j}, a^*) = -ca_{i,j} + P_v g t_{i,j,a_{i,j}} + \delta \sum_k \sum_l d_{i,j,a_{ij}}(k) H(l) W_{k,l}. \tag{16}$$

Then, according to the definition of ESS [10], a^* is an ESS if the following inequalities hold

$$F(a_{i,j}^*, a^*) > F(a_{i,j}, a^*) \qquad \forall a_{ij}. \tag{17}$$

According to (17), we can derive the stable condition for the optimal action a_2^* as described in Theorem 2.

Theorem 2: The optimal action a_2^* is an ESS if the cost-gain-ratio c/g satisfies $0 < c/g < \frac{P_v n_t (1-\lambda)}{N(1-\delta\lambda)} \min(G_{a_{i,j},a_{i,j}} + G_{N,N} - 1)$.

Proof: Due to page limitation, the proof is shown in the supplementary information [12]. ■

IV. SIMULATION

In this section, we perform simulations to evaluate the effectiveness of the proposed method. We consider a DSA network where the PU and BS locate at coordinate (0,0) and (1,0) respectively. There are five SUs randomly located in a circle centered at coordinate (0.5,0) with a radius of 0.25. The channels are complex Gaussian channels and the path loss factor γ is set to 4. The maximum transmission power of PU and SUs are $P_s = P_{r,i} = 10$, and the signal-to-noise ratio (SNR) is $SNR = P_s/\sigma_n = 10$dB.

We evaluate the system performance in terms of total throughput and outage probability by comparing the proposed scheme with two existing schemes. One is the "NCPC", where all SUs can only access unoccupied channels randomly and transmit their own data via a non-cooperative power control game. The second approach, which we denote as "SL-NCPC", is spectrum leasing proposed in [3].

In the first simulation, we evaluate the achievable end-to-end throughput of PU versus vacant channel probability p for different schemes. In our scheme, only one SU is selected as relay and we propose two strategies to allocate the transmission time between PU and the selected SU: "IR-Equal" where

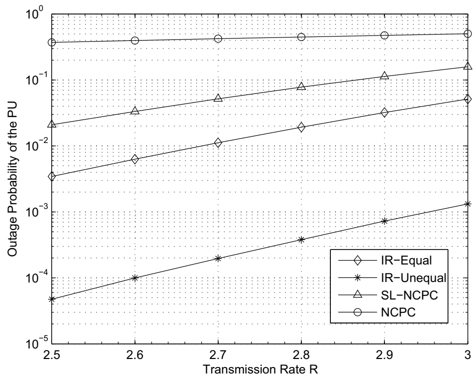

Fig. 4. Outage Probability of PU

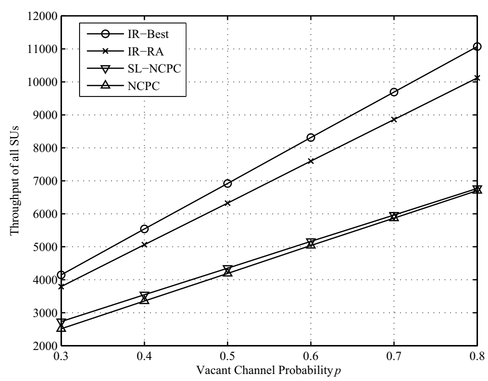

Fig. 5. Throughput of all SUs

one time slot is equally divided into two sub time slot, and "IR-Unequal" where the length of sub time slot, α, allocated to a PU's transmission is given by $\alpha = R_{rd}/(R_{sr} + R_{sd} - R_{rd})$.

From Fig. 3 we can see that NCPC has the worst performance because there is no cooperation between the PU and SUs. Through cooperation with SUs using spectrum leasing, SL-NCPC can improve the throughput of PU. However, due to the selfish nature, SUs will use only part of their transmission power for relaying data, which leads to an inefficient cooperation. With the proposed indirect reciprocity game theoretic method, the selected SU is motivated to use its full power for relaying the PU's data. In such a case, the PU's throughput can be greatly improved. Moreover, from Fig. 3 we can see that the performance of the proposed scheme can be further improved with the "IR-unequal" time slot allocation.

We also study the performance of PU system in terms of outage probability. From Fig. 4 we can see that without cooperation, the outage probability of PU with NCPC is as high as 0.5 when R=2.5. With cooperation, the outage probability in SL-NCPC can drop to around 10^{-2}. With the proposed cooperation stimulation framework, the PU system can achieve the lowest outage probability.

In the third simulation, we evaluate the throughput of all SUs versus vacant channel probability and the results are shown in Fig. 5. In our proposed scheme there are two ways for BS to select the SU: "IR-RA" where the SU is randomly chosen and "IR-Best" where the SU with the best channel gain is chosen. From Fig. 5, we can see that the throughput of all SUs with the proposed scheme is up to 45% larger than that of SL-NCPC and NCPC. This is mainly because in our scheme, only one SU is assigned to access one channel so that the interference introduced by multiple SUs' transmission over one channel is avoided.

V. CONCLUSION

In this paper, we propose a novel spectrum accessing scheme to fairly and efficiently utilize both used and unused channels based on the indirect reciprocity game modeling. With the proposed indirect reciprocity model, the SUs are stimulated through reputation to use maximal power to help PUs relay data, which leads to a full cooperation between SUs and PUs. By theoretical analysis, we derive the stable condition for such a perfect cooperation strategy. We find that with a proper cost-to-gain ration, a full cooperation can be guaranteed. Simulation results show that the proposed scheme achieves much better performance compared with the state-of-the-art direct reciprocity spectrum access schemes.

REFERENCES

[1] M. J. Osborne and A. Rubenstein, A Course in Game Theory, MIT Press, 1994

[2] B. Wang, Y. Wu, Z. Ji, K.J.R. Liu, and T.C. Clancy, "Game Theoretical Mechanism Design Methods: Suppressing Cheating in Cognitive Radio Networks", *IEEE Signal Processing Magazine*, vol 25, no 6, pp.74-84, Nov 2008.

[3] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary Ad Hoc networks", *IEEE J. of Sel. Areas Commun.*, vol. 26, no. 1, pp.1–11, Jan. 2008.

[4] B. Wang and K.J.R. Liu, "Advances in Cognitive Radios: A Survey", *IEEE J. of Sel. Topics on Signal Processing*, vol 5, no 1, pp.5-23, Feb 2011.

[5] K.J.R. Liu and B. Wang, Cognitive Radio Networking and Security: A Game Theoretical View, Cambridge University Press, 2010.

[6] Y. Wu, B. Wang, K.J.R. Liu, and T.C. Clancy, "Repeated Open Spectrum Sharing Game with Cheat-Proof Strategies", *IEEE Trans. Wireless. Commun*, vol 8, no 4, pp.1922-1933, April 2009.

[7] H. Ohtsuki, Y. Iwasa, and M. A. Nowak, "Indirect reciprocity provides only a narrow margin for efficiency for costly punishment", *Nature*, vol. 457, pp. 79–82, 2009.

[8] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity", *Nature*, vol. 437, pp. 1291–1298, 2005.

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

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

[11] J. Laneman, D. Tse and G. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior", *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.

[12] Supplementary information: http://www.ece.umd.edu/∼yan/IRCR.pdf.