Cooperation Stimulation in Cooperative Communications: An Indirect Reciprocity Game

Yang Gao, Yan Chen and K. J. Ray Liu
Department of Electrical and Computer Engineering,
University of Maryland, College Park, MD 20742, USA
E-mail: {yanggao, yan, kjrliu}@umd.edu

Abstract—**The viability of cooperative communications largely depends on the willingness of users to help. However, in future wireless networks where users are rational and pursue different objectives, they will not help relay information for others unless this can improve their own utilities. Therefore, it is very important to study the incentive issues when designing cooperative communication systems. In this paper, we propose a cooperation stimulation scheme for multiuser cooperative communications using indirect reciprocity game. By introducing the notion of reputation and social norm, rational users who care about their future utilities get the incentive to cooperate with others. Different from existing works on reputation based schemes that mainly rely on experimental verifications, we prove theoretically that cooperating with users having good reputation can be sustained as an equilibrium. Moreover, by modeling the action spreading as an evolutionary game, we show through simulations that the equilibria we found are evolutionarily stable and can be reached with proper initial conditions. Finally, simulation results are shown to verify the efficiency and effectiveness of the proposed scheme.**

I. Introduction

In recent years, cooperative communications [1] have been viewed as a promising transmit paradigm for future wireless networks. Through the cooperation of relays, cooperative communications can build a virtual antenna array for the source node and therefore enable single-antenna users to achieve spatial diversity.

However, most existing works assume by default that users are altruistic and willing to help unconditionally, regardless of their own utilities, which appears to be unrealistic in wireless networks where users are rational, intelligent and often do not serve a common objective. They will and have the capabilities to make intelligent decisions based on their own preferences. Moreover, since relaying others' information consumes valued resources such as power and frequency, users have no incentive to help and tend to act selfishly as "free-riders". In such a case, cooperative communication protocols will fail to achieve good social outcomes without considering incentive issues. It is therefore of great interest to design effective incentive schemes that can stimulate cooperation among selfish users.

Many efforts have been made in the literature to stimulate cooperative behaviors in communication networks with rational and selfish users, which can be broadly categorized into the following three classes [2]: payment based, direct reciprocity based and reputation based schemes. Payment based methods have been widely used to stimulate cooperation for wireless networks [3] [4]. However, the implementation of payment based schemes requires an infrastructure for billing services and some temper-proof hardware, which is impractical for many applications. Direct reciprocity based schemes that rely on repeated prisoner's dilemma model can also be employed to sustain cooperation [5]. Nevertheless, this model assumes that the interaction between a pair of users lasts for infinite many times, which is generally not true in multiuser cooperative communications. Instead of having a fixed relay, source nodes select different relay nodes at each time to achieve higher order of spatial diversity and thus better performance. Reputation is also an effective tool for cooperation stimulation [6] [7]. However, effectiveness of these reputation based schemes is demonstrated only through experimental results while there is no theoretical justification.

In this paper, we propose to employ indirect reciprocity game [8] to stimulate cooperation among selfish users in a multiuser cooperative communication network. Indirect reciprocity is a key concept in explaining the evolution of human cooperation and has drawn great attentions in both areas of economics [9] and evolutionary biology [10]. The basic idea behind indirect reciprocity is that through building up a reputation and social judgement system, cooperation can lead to a good reputation and expect to be rewarded by others in the future. Moreover, based on the indirect reciprocity game modeling, we can theoretically justify the use of reputation in stimulating cooperation, which is lacked in the current literature.

The rest of the paper is organized as follows. In Section II, we describe the problem formulation and introduce basic components in our system model. Then, the steady state analysis is presented in details in Section III. We model action spreading at the transient state as an evolutionary game in Section IV. Finally, we show the simulation results in Section V and draw conclusions in Section VI.

II. System Model

A. Physical Layer Model with Relay Selection

We consider a TDMA network consisting of N nodes numbered $1, 2, ..., N$ having their own information to be delivered to a base station (BS) d. Nodes are assumed to be rational in the sense that they will act to maximize their own utilities. Throughout this paper, we will use user, node and player interchangeably.

We divide time into time frames and each time frame is further divided into N time slots. At each time slot, only one prescribed node is allowed to transmit and all the remaining $N-1$ nodes can serve as potential relays. Amplify-and-forward (AF) protocol is employed in the system model. As a result, we can calculate the relayed SNR by relay node j for source node i as

$$\Gamma_{i,j,d} = \frac{P_r P_s |h_{i,j}|^2 |h_{j,d}|^2}{P_r |h_{j,d}|^2 N_0 + P_s |h_{i,j}|^2 N_0 + N_0^2}. \tag{1}$$

where P_s is the transmitted power of the source node, P_r is the transmitted power of the relay node, $h_{i,d}$ and $h_{i,j}$ are channel coefficients from user i to the BS and user j respectively. Without loss of generality, we model the additive noise for all links as i.i.d. zero-mean, complex Gaussian random variables with variance N_0. Moreover, homogeneous channel condition is considered in this work, where we model channel coefficients $h_{i,d}$ and $h_{i,j}$ as zero-mean, complex Gaussian random variables with variance σ_1^2 and σ_2^2 respectively for all $i,j \in \{1,2,...N\}$. We also assume quasi-static channel in our system model, which means channel condition remains the same within each time slot and varies independently from time slot to time slot.

We adopt two relay selection schemes based on the availability of CSI. If the BS is assumed to have the global CSI, then we employ optimal relay selection (ORS), in which the relay node that can provide the best relayed SNR will be selected to assist the source node. Since the best relay is selected at each time slot, source nodes can achieve full spatial diversity if the relay nodes choose to cooperate [1]. On the other hand, if the BS does not know the global CSI, a random relay selection (RRS) is employed, i.e. the BS will randomly choose one relay node. Once a relay is selected, it will decide whether to help according to a certain action rule which maximizes its own payoff and send its decision back to the BS. If the selected relay node chooses to help, then the received SNR increment at the BS after the maximal-ratio combining (MRC) can be expressed as

$$\Gamma_i^c = \begin{cases} \max_{j \neq i} \Gamma_{i,j,d} & \text{for ORS,} \\ \Gamma_{i,j,d} & \text{for RRS if node j is selected.} \end{cases} \tag{2}$$

B. Incentive Schemes Based on Indirect Reciprocity Game

In order to stimulate the selected relay node to cooperate, we employ an incentive scheme based on indirect reciprocity game. Reputation and social norm are two key concepts in indirect reciprocity game modeling. In particular, a reputation score is assigned to each user at the end of every time slot that reflects the social assessment toward this user. In this paper, we adopt a binary reputation score, where users can have either good reputation or bad reputation which are denoted by G and B respectively. Social norm is a function used for updating reputation. In our system model, only the selected relay node's reputation will be updated while the reputation for source node and unselected relays remains unchanged. Unless otherwise specified, we will simply use relay or relay node to indicate the selected relay node in the rest of this paper. We design

k \ i,j	GG	GB	BG	BB
C	1	λ	$1-\lambda$	0
D	λ	1	0	$1-\lambda$

the social norm Q as a function of relay's current reputation, source node's reputation and the relay's action as

$$Q : \{G,B\} \times \{G,B\} \times \{C,D\} \mapsto [0,1], \tag{3}$$

where C and D stand for cooperation and defection of the relay respectively. The value of the social norm is designed to be the probability of assigning a good reputation to the relay. More specifically, for any $i,j \in \{G,B\}$ and $k \in \{C,D\}$, $Q(i,j,k)$ stands for the probability of having a good reputation at the end of this time slot for the relay that currently has reputation i and chooses action k towards the source node with reputation j. Values of the proposed social norm are shown in Table I, where $\lambda \in [0,1]$ is a parameter that controls the weight of current reputation in determining the new reputation.

An action rule, $\mathbf{a} = [\begin{array}{cccc} a_{G,G} & a_{G,B} & a_{B,G} & a_{B,B} \end{array}]^T$, is an action table of the relay, where element $a_{i,j}$ stands for the probability of cooperation given the relay's reputation i and the source's reputation j. In our system model, every user decides its action rule at the beginning of each time frame, based on the social norm and reputation distribution of the network.

C. Payoff Functions

In this subsection, we discuss payoff functions in the proposed game. In each time slot, if the relay chooses to decline the request, both source and relay will receive a payoff of 0. On the other hand, if the relay chooses to cooperate, then the source node will receive a gain $\mathcal{G}$ while the relay suffers a cost $\mathcal{C}$. Since the realization of channel is not available to users when they determine their action rules, payoff functions should be measured in an average sense. In this work, we choose the cost as a linear function of transmitted power as

$$\mathcal{C} = P_r c,$$

where c is the cost per unit power. For the gain function, we design it to be a linear function of the averaged SNR increment as

$$\mathcal{G} = E_h[\Gamma_i^c] \cdot g,$$

where g is the gain per unit SNR increment. Here, user i is assumed to be the source node and the expectation is taken over the joint distribution of all channel coefficients. Note that other forms of payoff functions can also be similarly considered and put into the framework of this paper.

Proposition 1: Based on the channel models in Section II.A and assuming $P_s/N_0 \gg 1$ and $P_r/N_0 \gg 1$, the gain function can be estimated by

$$\mathcal{G} \approx \begin{cases} \frac{P_r P_s \sigma_1^2 \sigma_2^2 g}{P_r \sigma_1^2 N_0 + P_s \sigma_2^2 N_0} \sum_{n=1}^{N-1} \frac{1}{n} & \text{for ORS,} \\ \frac{P_r P_s \sigma_1^2 \sigma_2^2 g}{P_r \sigma_1^2 N_0 + P_s \sigma_2^2 N_0} & \text{for RRS.} \end{cases} \tag{4}$$

$$\mathbf{H_a} = \begin{bmatrix} (1+d_{G,G})x_\mathbf{a} & (1+d_{G,G})(1-x_\mathbf{a}) & (1-d_{G,G})x_\mathbf{a} & (1-d_{G,G})(1-x_\mathbf{a}) \\ (1+d_{G,B})x_\mathbf{a} & (1+d_{G,B})(1-x_\mathbf{a}) & (1-d_{G,B})x_\mathbf{a} & (1-d_{G,B})(1-x_\mathbf{a}) \\ d_{B,G}x_\mathbf{a} & d_{B,G}(1-x_\mathbf{a}) & (2-d_{B,G})x_\mathbf{a} & (2-d_{B,G})(1-x_\mathbf{a}) \\ d_{B,B}x_\mathbf{a} & d_{B,B}(1-x_\mathbf{a}) & (2-d_{B,B})x_\mathbf{a} & (2-d_{B,B})(1-x_\mathbf{a}) \end{bmatrix}. \tag{12}$$

$$\mathbf{b_a} = \frac{1}{2}\begin{bmatrix} (\mathcal{G}-\mathcal{C})a_{G,G} & \mathcal{G}a_{B,G} - \mathcal{C}a_{G,B} & \mathcal{G}a_{G,B} - \mathcal{C}a_{B,G} & (\mathcal{G}-\mathcal{C})a_{B,B} \end{bmatrix}^T. \tag{13}$$

Proof: Due to page limitation, we show the proof in the supplementary information [11]. ■

In practice, the gain can be estimated either using (4) or through experiments conducted at the BS. Let $\rho = \frac{\mathcal{C}}{\mathcal{G}}$ represent the cost to gain ratio of the game, which can greatly influence user behaviors. Intuitively, it would be more likely for users to cooperate if ρ is smaller. In this work, we restrict that $0 < \rho < 1$.

III. STEADY STATE ANALYSIS USING MDP

A. Stationary Reputation Distribution

Reputation is a key concept in indirect reciprocity game. Therefore, one important aspect of the network state in indirect reciprocity game modeling is the reputation distribution among the whole population. Let x_t represent the ratio of users that have good reputation at time frame t. Then by assuming an action rule $\mathbf{a}$ is employed by all users in the network, we have

$$x_{t+1} = x_t[x_t d_{G,G} + (1-x_t)d_{G,B}] + (1-x_t)[x_t d_{B,G} + (1-x_t)d_{B,B}]$$
$$\overset{\Delta}{=} f_\mathbf{a}(x_t), \tag{5}$$

where $d_{i,j}$ with $i,j \in \{G,B\}$ is the reputation updating probability which stands for the probability that the relay will have a good reputation after one interaction, given that it currently has reputation i and the source's reputation is j. The $d_{i,j}$ can be calculated based on the social norm in Table I as follows,

$$d_{i,j} = a_{i,j}Q(i,j,C) + (1-a_{i,j})Q(i,j,D). \tag{6}$$

Clearly, $d_{i,j}$ is a function of the action $a_{i,j}$ and we use $d_{i,j}$ instead of $d_{i,j}(a_{i,j})$ just for notation simplicity.

Based on the reputation distribution updating rule in (5), we study the stationary reputation distribution and have the following proposition.

Proposition 2: For any action rule $\mathbf{a}$, there exists a stationary reputation distribution, which is the solution to the following equation

$$x_\mathbf{a} = f_\mathbf{a}(x_\mathbf{a}). \tag{7}$$

Proof: Due to page limitation, we show the proof in the supplementary information [11]. ■

From Proposition 2, we can see that if an action rule $\mathbf{a}$ is employed by all users, then the stationary reputation distribution will be reached. As a consequence, the game will become stable, which leads to the steady state of the proposed indirect reciprocity game defined as follows.

Definition 1 (Steady State): $(\mathbf{a}, x_\mathbf{a})$ is a steady state of the indirect reciprocity game if $\mathbf{a}$ is an action rule that employed by all users and $x_\mathbf{a}$ is the corresponding stationary reputation distribution.

B. Long-Term Expected Payoffs at Steady States

In this subsection, we study the long-term expected payoff functions at the steady state. Assume that the indirect reciprocity game is in a steady state $(\mathbf{a}, x_\mathbf{a})$, i.e. all players choose action rule $\mathbf{a}$ and the reputation distribution remains stable at $x_\mathbf{a}$. Let $v_{i,j}$ with $i,j \in \{G,B\}$ denote the expected payoff that a player, currently having reputation i and being matched with a player with reputation j can have from this interaction to future. If the player acts as the relay, then its long-term expected payoff can be expressed as

$$u_{i,j}^r(a_{i,j}) = -\mathcal{C}a_{i,j} + \delta[d_{i,j}x_\mathbf{a}v_{G,G} + d_{i,j}(1-x_\mathbf{a})v_{G,B}$$
$$+ (1-d_{i,j})x_\mathbf{a}v_{B,G} + (1-d_{i,j})(1-x_\mathbf{a})v_{B,B}], \tag{8}$$

where the first term represents the cost incurred in the current interaction and the second term represents the future payoff, which is discounted by a discounting factor $\delta \in (0,1)$. On the other hand, if the player acts as the source, then the long-term expected payoff can be written as

$$u_{i,j}^s(a_{j,i}) = \mathcal{G}a_{j,i} + \delta[x_\mathbf{a}v_{i,G} + (1-x_\mathbf{a})v_{i,B}]. \tag{9}$$

Note that only relay's reputation will be updated. Moreover, by the homogeneous assumption, the probabilities of being the source and the relay for an arbitrary user are $\frac{1}{N}$ and $\frac{N-1}{N}\frac{1}{N-1} = \frac{1}{N}$ respectively. Therefore, given that the user is participating in the interaction, it will act as either the source or the relay with equal probability $1/2$. Therefore, the long-term expected payoff at the steady state can be written as

$$v_{i,j} = \frac{1}{2}u_{i,j}^r(a_{i,j}) + \frac{1}{2}u_{i,j}^s(a_{j,i}). \tag{10}$$

Let $\mathbf{V} = \begin{bmatrix} v_{G,G} & v_{G,B} & v_{B,G} & v_{B,B} \end{bmatrix}^T$ denote the long-term expected payoff vector. Then, by rearranging (10) into the matrix form, we have

$$(\mathbf{I} - \frac{\delta}{2}\mathbf{H_a})\mathbf{V} = \mathbf{b_a}, \tag{11}$$

where $\mathbf{H_a}$ and $\mathbf{b_a}$ are defined in (12) and (13) respectively. Moreover, it can be proved according to the Corollary C.4 in [12] that matrix $(\mathbf{I} - \frac{\delta}{2}\mathbf{H_a})$ is invertible. Therefore, in the proposed indirect reciprocity game, we can obtain the long-term expected payoff vector in a steady state $(\mathbf{a}, x_\mathbf{a})$ as

$$\mathbf{V} = (\mathbf{I} - \frac{\delta}{2}\mathbf{H_a})^{-1}\mathbf{b_a}. \tag{14}$$

C. Equilibrium Steady State

In this subsection, user behaviors at the steady state are studied. In particular, we first define a new concept of equilibrium steady state. Then, by modeling the problem of finding optimal action rule at the steady state as a Markov Decision Process (MDP), we characterize all equilibrium steady states of the proposed indirect reciprocity game mathematically.

Definition 2 (Equilibrium Steady State): $(\mathbf{a}, x_{\mathbf{a}})$ is an equilibrium steady state of the indirect reciprocity game if $(\mathbf{a}, x_{\mathbf{a}})$ is a steady state and $\mathbf{a}$ is optimal in this steady state.

From the definition above, we can see that determining whether a steady state is an equilibrium is equivalent to the problem of finding the optimal action rule in this steady state, which can be modeled as a MDP. In this MDP formulation, the state is the reputation pair (i, j), the action is action rule $\mathbf{a}$, the transition probability is determined by $\{d_{i,j}\}$ and the reward function is determined by $\mathcal{C}$, $\mathcal{G}$ and the steady state $(\mathbf{a}, x_{\mathbf{a}})$.

Based on the MDP formulation, we can write the optimality equation as

$$v_{i,j} = \max_{\hat{a}_{i,j}} \left[\frac{1}{2} u_{i,j}^r(\hat{a}_{i,j}) + \frac{1}{2} u_{i,j}^s(a_{j,i}) \right], \quad (15)$$

which can be solved numerically using the well-known value iteration algorithm [12]. In this work, instead of solving the problem numerically, we will characterize the equilibrium steady states theoretically by exploring the structure of this problem. Note that the formulated MDP varies from steady state to steady state and there are infinitely many steady states, which makes the problem of finding all equilibria even harder. To make this problem tractable, we derive the following proposition, which successfully reduces the potential equilibria that are of the practical interests into the set of three steady states.

Proposition 3: In the proposed indirect reciprocity game, if $(\mathbf{a}, x_{\mathbf{a}})$ is an equilibrium steady state for more than one possible ρ, it must be one of the following steady states.

1) $(\mathbf{a_1}, x_{\mathbf{a_1}})$ with $\mathbf{a_1} = [\ 0\ \ 0\ \ 0\ \ 0\ ]^T$ and $x_{\mathbf{a_1}} = 1/2$
2) $(\mathbf{a_2}, x_{\mathbf{a_2}})$ with $\mathbf{a_2} = [\ 1\ \ 0\ \ 1\ \ 0\ ]^T$ and $x_{\mathbf{a_2}} = 1$
3) $(\mathbf{a_3}, x_{\mathbf{a_3}})$ with $\mathbf{a_3} = [\ 0\ \ 1\ \ 0\ \ 1\ ]^T$ and $x_{\mathbf{a_3}} = 0$.

Proof: Due to page limitation, we show the proof in the supplementary information [11]. ∎

Results in Proposition 3 show that steady states in the proposed indirect reciprocity game can be broadly categorized into two classes. In the first class, there are three steady states which are resistant to one-shot deviations and have the potential to be equilibria for a set of ρ. The second class consists of all remaining steady states, which either cannot be an equilibrium or can only be an equilibrium for a specific cost to gain ratio. However, such an equilibrium is not robust to estimation errors of system parameters, which is highly likely in a multiuser wireless network scenario, and thus is of no practical interests. Therefore, we only need to analyze three, instead of infinitely many, steady states to study practical equilibria of the indirect reciprocity game.

Next, we solve the optimality equations for the three steady states to show which of them are equilibria and under what conditions they will be. Our main results are summarized in the following theorem.

Theorem 1: In the proposed indirect reciprocity game, there are three equilibrium steady states, which can be given as follows.

1) $(\mathbf{a_1}, x_{\mathbf{a_1}})$ is an equilibrium for all $0 < \rho < 1$
2) $(\mathbf{a_2}, x_{\mathbf{a_2}})$ is an equilibrium if $0 < \rho \leq \frac{\delta(1-\lambda)}{2-\delta-\lambda\delta}$
3) $(\mathbf{a_3}, x_{\mathbf{a_3}})$ is an equilibrium if $0 < \rho \leq \frac{\delta(1-\lambda)}{2-\delta-\lambda\delta}$

Proof: Due to page limitation, we show the proof in the supplementary information [11]. ∎

From Theorem 1, we know that the proposed indirect reciprocity game can have three equilibria in practice. In the first equilibrium, users do not cooperate at all, which results in a reputation distribution of half and half. In the second equilibrium, users only cooperate with those having good reputation and all population have good reputation, while in the last equilibrium, users only collaborate with those having bad reputation and all population have bad reputation. Actually, it can be seen that the last two steady states are mutually symmetric states of the game, both of which lead to full cooperation but with different interpretations of reputation scores. Moreover, results in Theorem 1 show that, if the cost to gain ratio is below a certain threshold, cooperation can be sustained by using the proposed indirect reciprocity game.

IV. EVOLUTIONARY MODELING OF THE INDIRECT RECIPROCITY GAME

The indirect reciprocity game is highly dynamic before it reaches the steady state. Since the reputation distribution of the whole population and actions adopted by different users are changing constantly, all users are uncertain about the network state and each other's actions. In such transient states, to improve their utilities, users will try different strategies in every play and learn from the strategy interactions using the methodology of understand-by-building. Moreover, since a mixed action rule is a probability distribution over pure action rules, users will adjust the probability of using a certain pure action rule as the network state evolves. Such an evolution process can be modeled by replicator dynamics in evolutionary game theory. Specifically, let $p_{\mathbf{a}}$ stand for the probability of users using pure action rule $\mathbf{a} \in \mathbf{A}^D$, where $\mathbf{A}^D$ represents the set of all pure action rules. Then, by replicator dynamics, the evolution of $p_{\mathbf{a}}$ is given by the following equation

$$\frac{dp_{\mathbf{a}}}{dt} = \eta \left(U_{\mathbf{a}} - \sum_{\mathbf{a} \in \mathbf{A}^D} p_{\mathbf{a}} U_{\mathbf{a}} \right) p_{\mathbf{a}}, \quad (16)$$

where $U_{\mathbf{a}}$ is the average payoff of users using action rule $\mathbf{a}$ and η is a scale factor controlling the speed of the evolution. After discretizing the replicator dynamic equation in (16), we have

$$p_{\mathbf{a}}^{t+1} = \left[1 + \eta \left(U_{\mathbf{a}} - \sum_{\mathbf{a} \in \mathbf{A}^D} p_{\mathbf{a}} U_{\mathbf{a}} \right) \right] p_{\mathbf{a}}^t. \quad (17)$$
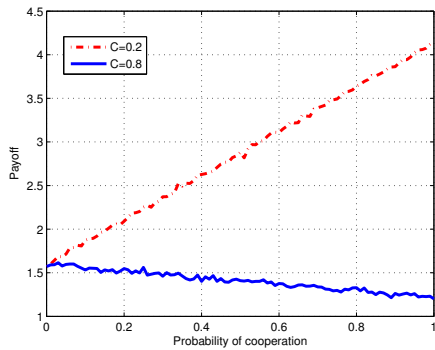
Fig. 1. Equilibrium evaluation of the indirect reciprocity game.

Fig. 2. Population evolution of the indirect reciprocity game: (a) C=0.2; (b) C=0.8.

An action rule is asymptotically stable to the replicator dynamics if and only if it is the Evolutionarily Stable Strategy [13], an equilibrium concept widely adopted in evolutionary game theory. In section V, we will examine by simulations the stability of the equilibria that we derived in Theorem 1.

V. SIMULATION RESULTS

In this section, we conducted numerical simulations to evaluate the proposed indirect reciprocity game. A fixed-size population with $N = 100$ is considered and the discounting factor δ of each user is set as 0.9. We assume $\mathcal{G} = 1$ in our simulations.

In the first simulation, we evaluate the performance of the proposed incentive scheme where λ is set to be 0.5. In Fig. 1, we assume that the game starts at the steady state $(\mathbf{a_2}, x_{\mathbf{a_2}})$. Then we show payoffs of a specified user that deviates to action rule $\begin{bmatrix} p & 0 & p & 0 \end{bmatrix}^T$ under different cost values, where $p \in [0, 1]$ is the probability of cooperation with users having good reputation. From Fig. 1, we can see that as p increases, the user's payoff increases when $\mathcal{C} = 0.2$ and decreases when $\mathcal{C} = 0.8$. This agrees with the our theoretic derivations in Theorem 1 since the threshold $\frac{\delta(1-\lambda)}{2-\delta-\lambda\delta} = 9/13$ according to the simulation settings.

In the second simulation, we study the evolutionary properties of the indirect reciprocity game. The initial probability of choosing the specified action rule $\mathbf{a_2}$ is set to be 0.6 while the initial probabilities of choosing other pure action rules are set equally as $0.4/15$. The initial reputation distribution of the population is assumed to be $1/2$. Moreover, we use $\eta = 0.1$ in the replicator dynamics equation. We first study the low cost case where we set $\mathcal{C} = 0.2$ and show the evolutionary results in Fig. 2 (a). From the figure, we can see that the game converges to the equilibrium steady state $(\mathbf{a_2}, x_{\mathbf{a_2}})$ and remains stable once converges. Therefore, the action rule $\mathbf{a_2}$ is verified to be an ESS for the low cost case.

We then study the high cost case where we set $\mathcal{C} = 0.8$. In such case, the cost-to-gain ratio ρ is larger than the threshold. Therefore, the steady state $(\mathbf{a_2}, x_{\mathbf{a_2}})$ is no longer an equilibrium and the cooperation can not be sustained. We show in Fig. 2 (b) the evolutionary results under $\mathcal{C} = 0.8$.
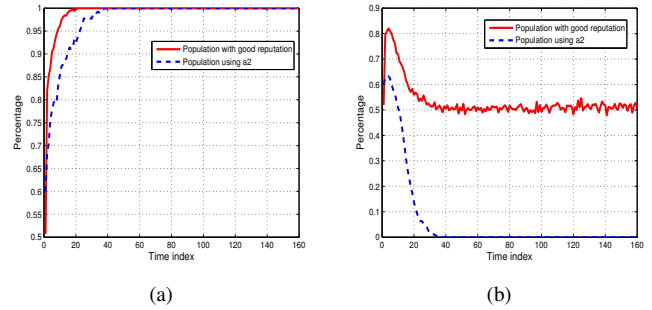
From the figure, we can see that the reputation distribution converges to $1/2$ and the probability of choosing the action rule $\mathbf{a_2}$ eventually becomes zero. Therefore, the action rule $\mathbf{a_2}$ is not an ESS at the high cost case.

VI. CONCLUSIONS

In this paper, we propose a cooperation stimulation scheme for multiuser cooperative communications using indirect reciprocity game. With the concept of reputation and social norm, our proposed scheme does not rely on the assumption that the number of interactions between a pair of users are infinite and therefore can be incorporated with any optimal relay selection algorithms to achieve full spatial diversity. Moreover, different from experimental verifications in existing works, we theoretically prove that cooperation with users having good reputation can be sustained as an equilibrium given that the cost to gain ratio is under a certain threshold. By modeling the action spreading as an evolutionary game, we further show by simulations that such an equilibrium action rule is also an ESS and therefore resistant to the mutation of any other action rules.

REFERENCES

[1] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking, Cambridge University Press*, 2008.
[2] M. Feldman, J. Chuang, "Overcoming free-riding behavior in peer-to-peer systems," *SIGecom Exch.*, 5(4):41-50, 2005.
[3] M. Neely, "Optimal pricing in a free market wireless network," in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pp. 213-221, May 2007.
[4] B. Wang, Z. Han, and K. J. R. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using stackelberg game," *IEEE Transactions on Mobile Computing*, vol. 8, no. 7, pp. 975-990, 2009.
[5] B. Cohen, "Incentives build robustness in BitTorrent," in *Proc. of the 1st Workshop on Economics of Peer-to-Peer Systems*, June 2003.
[6] M. Gupta, P. Judge, and M. Ammar, "A reputation system for peer-to-peer networks," in *Proc. of ACM NOSSDAV*, California, June 2003.
[7] S. Buchegger and J. L. Boudec, "Performance analysis of the confidant protocol," in *Proc. ACM MobiHoc*, 2002, pp. 226-236.
[8] Y. Chen and K. J. R. Liu, "Indirect Reciprocity Game Modelling for Cooperation Stimulation in Cognitive Networks," *IEEE Transactions on Communications*, Jan. 2011.
[9] M. Kandori, "Social Norms and Community Enforcement," *Rev. Economic Studies*, vol. 59, No. 1, pp. 63-80, Jan. 1992.
[10] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, pp. 1291-1298, 2005.
[11] Y. Gao, Y. Chen and K. J. R. Liu. Supplementary information. [Online]. Available: http://www.sig.umd.edu/yanggao/ICC2012SupplementaryInfo.pdf
[12] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, 1994.
[13] J. W. Weibull, *Evolutionary Game Theory*, MIT Press, 1995.