

Cooperation Stimulation for Multiuser Cooperative Communications Using Indirect Reciprocity Game

Yang Gao, *Student Member, IEEE*, Yan Chen, *Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

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 theoretically demonstrate the effectiveness of the proposed scheme in two steps. First, we conduct steady state analysis of the game and show that cooperating with users having good reputation can be sustained as an equilibrium when the cost-to-gain ratio is below a certain threshold. Then, by modeling the action spreading at transient states as an evolutionary game, we show that the equilibria we found in the steady state analysis are stable and can be reached with proper initial conditions. Moreover, we introduce energy detection to handle possible cheating behaviors of users and study its impact to the proposed indirect reciprocity game. Finally, simulation results are shown to verify the effectiveness of the proposed scheme.

Index Terms—Cooperation stimulation, cooperative communications, evolutionarily stable strategy, game theory, indirect reciprocity, Markov decision process.

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 improve communication capacity, speed, and performance; reduce battery consumption and extend network lifetime; increase throughput and stability region for multiple access schemes; expand transmission coverage area; and provide cooperation tradeoff beyond source-channel coding for multimedia communications [1].

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 ad hoc networks [3] [4] [5] and peer-to-peer networks [6] [7]. The cooperation stimulation problem has been studied in multiuser cooperative communication networks [8], where a two-level Stackelberg game was used to jointly address the incentive issue, relay selection and resource allocation problems in a distributed manner. In [9], the pricing game was studied under scenarios where channel state information (CSI) was held privately. 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 [10] [11]. In [12], Yu and Liu derived a set of optimal cooperation strategies for users in ad hoc networks using optimality criteria, such as Pareto optimality, fairness, and cheat-proofing. In [13], the cooperation stimulation problem has been studied for mobile ad hoc networks under scenarios where noisy and imperfect observations exist. Nevertheless, direct reciprocity based schemes assume implicitly that the interaction between a pair of users lasts for infinite many times, which is generally not true for 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 [14] [15]. In [16], a local reputation system was first set up based on shared history among the neighborhood nodes and then used to identify and punish non-cooperative nodes. The work in [17] proposed to enforce cooperation through a global reputation mechanism. 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 [18] to stimulate cooperation among selfish users in a multiuser cooperative communication network. Indirect reci-

Paper approved by E. G. Larsson, the Editor for Game Theory and Communications Systems Optimization of the IEEE Communications Society. Manuscript received October 7, 2011; revised March 21 and May 9, 2012.

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA (e-mail: {yanggao, yan, kjrlu}@umd.edu).

Digital Object Identifier 10.1109/TCOMM.2012.091212.110678

procuity is a key concept in explaining the evolution of human cooperation and was first studied under the name of “third party altruism” in 1971 [19]. Later, such a concept drew great attentions in both areas of economics [20] and evolutionary biology [21] [22]. 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 main contributions of this paper are summarized as follows.

- We propose a game-theoretic scheme to jointly consider the cooperation stimulation and relay selection for multiuser cooperative communications based on indirect reciprocity game. With the proposed scheme, selfish users have the incentive to cooperate and the full spatial diversity can be achieved when global CSI is available.
- We conduct steady state analysis of the indirect reciprocity game by formulating the problem of finding the optimal action rule at the steady state as a Markov Decision Process (MDP). We analyze mathematically all equilibrium steady states of the game and show that cooperating with users having good reputation can be sustained as an equilibrium when the cost to gain ratio is less than a certain threshold.
- To study the transient state of the game, we further model the action spreading at transient states as an evolutionary game. Then, we show that the equilibria we found are stable and demonstrate with simulation results that they can be reached given proper initial conditions.
- To deal with possible cheating behaviors of users, we introduce energy detection at the base station (BS) and study its impact to the indirect reciprocity game.

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 using MDP is presented in details in Section III. We model action spreading at the transient state as an evolutionary game in Section IV. In Section V, energy detection at the BS is introduced to deal with cheating behaviors and its impact to the indirect reciprocity game is studied. Finally, we show the simulation results in Section VI and draw conclusions in Section VII.

II. SYSTEM MODEL

In this section, we first present our physical layer model which employs the amplify-and-forward (AF) cooperation protocol and relay selection. Then we show the proposed incentive scheme using indirect reciprocity game and analyze its overhead. Finally, the payoff function is discussed.

A. Physical Layer Model with Relay Selection

As shown in Fig. 1 (a), we consider a TDMA based multi-user cooperative communication network that consists of N nodes numbered $1, 2, \dots, N$. All nodes have their own information to be delivered to a base station (BS) d . Without loss of generality, the transmitted information can be represented by

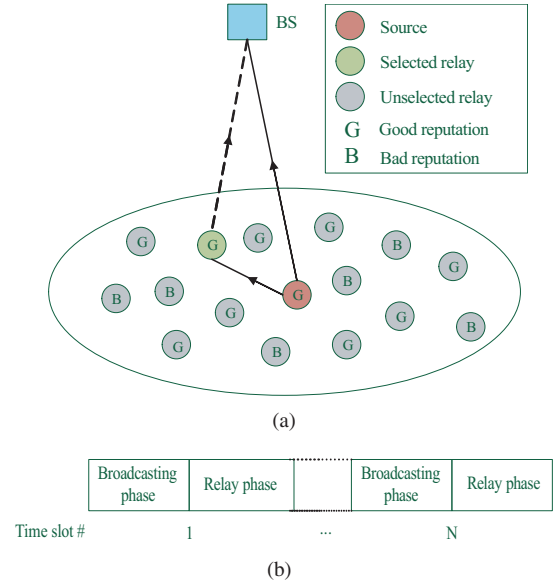


Fig. 1. Multi-user cooperative communication system: (a) system model, (b) time frame structure.

symbols, while nodes in practice will transmit the information in packets that contains a large number of symbols. 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, as shown in Fig. 1 (b). At each time slot, only one prescribed node is allowed to transmit and all the remaining $N - 1$ nodes can serve as potential relays. AF protocol is employed in the system model. As a result, every time slot will consists of two phases. In phase 1, the source node broadcasts its information to the BS and all other nodes. Assuming that node i acts as the source node, then the received signals $y_{i,d}^{(1)}$ and $y_{i,j}^{(1)}$ at the BS and node j respectively can be expressed as

$$y_{i,d}^{(1)} = \sqrt{P_s} h_{i,d} x_i + n_{i,d}, \quad (1)$$

$$y_{i,j}^{(1)} = \sqrt{P_s} h_{i,j} x_i + n_{i,j}, \quad (2)$$

where P_s is the transmitted power at the source node, x_i is the transmitted symbol with unit energy, $h_{i,d}$ and $h_{i,j}$ are channel coefficients from user i to the BS and user j respectively, and $n_{i,d}$ and $n_{i,j}$ are additive noise. 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 $i, j \in \{1, 2, \dots, N\}$. We also assume quasi-static channel in our system model, which means channel conditions remain the same within each time slot and vary independently from time slot to time slot.

In phase 2, a relay node is selected to amplify the received signal and forward it to the destination with transmitted power P_r . The received signal at the destination in phase 2 can be

written as

$$y_{j,d}^{(2)} = \frac{\sqrt{P_r P_s} h_{i,j} h_{j,d}}{\sqrt{P_s |h_{i,j}|^2 + N_0}} x_i + \frac{\sqrt{P_r} h_{j,d}}{\sqrt{P_s |h_{i,j}|^2 + N_0}} n_{i,j} + n_{j,d}. \quad (3)$$

Based on (3), we can calculate the relayed SNR by relay node j for source node i as

$$\Gamma_{i,j,d} = \frac{P_r P_s |h_{i,j}|^2 |h_{j,d}|^2}{P_r |h_{j,d}|^2 N_0 + P_s |h_{i,j}|^2 N_0 + N_0^2}. \quad (4)$$

We adopt two relay selection schemes based on the availability of CSI. If the BS is assumed to have the global CSI, e.g. BS can collect CSI from all potential relays through feedback channels, 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] [23] [24]. On the other hand, if the BS does not know the global CSI, a random relay selection (RRS) is employed, in which the BS will randomly choose one node as the relay from all potential relays with equal probability. 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

$$\Gamma_i^c = \begin{cases} \max_{j \neq i} \Gamma_{i,j,d} & \text{for ORS,} \\ \Gamma_{i,j,d} & \text{for RRS if node } j \text{ is selected.} \end{cases} \quad (5)$$

Note that for RRS, the required CSI of MRC can be obtained by the BS through channel estimations after the relay selection. In case of the selected relay node choosing not to help, we assume that the source node will not retransmit its packet and the system will remain idle during that phase.

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. Although more complicated reputation scores can be considered here, we will show in the rest of this paper that a binary reputation score is sufficient in sustaining cooperation among rational users. Social norm is a function used for updating reputation, which specifies what new reputation users will have according to their performed actions and current 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. Moreover, all reputation updates will be

TABLE I
SOCIAL NORM

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

Algorithm 1 : Proposed Indirect Reciprocity Game in One Time Frame

1. BS notifies users the reputation distribution of the population.
2. Users decide their action rules based on the social norm and reputation distribution.
3. for time slots $i=1,2,\dots,N$
 - User i broadcasts to the BS and other users.
 - BS selects one relay node using ORS or RRS and notifies the selected relay the source node's reputation.
 - The selected relay decides whether to cooperate according to his/her action rule and reports his/her decision to the BS.
 - The selected relay amplifies and forwards signals for the source if chooses to cooperate or remains silence if not.
 - BS updates the selected relay's reputation.

performed at the BS, who maintains the reputation information of all users.

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

$$Q : \{G, B\} \times \{G, B\} \times \{C, D\} \mapsto [0, 1], \quad (6)$$

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. When λ gets smaller, the relay's new reputation will become less relevant to its current reputation and therefore depend more on the immediate reputation that is determined by the relay's action and the source's reputation.

An action rule, $\mathbf{a} = [a_{G,G} \ a_{G,B} \ a_{B,G} \ a_{B,B}]^T$, 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 . For the special case of pure action rules, elements in the action table can only take values of 0 or 1. 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.

Finally, we summarize in Algorithm 1 the proposed indirect reciprocity game for one time frame.

C. Overhead of The Proposed Scheme

In the following, we would like to briefly analyze the overhead of the proposed scheme. The main overhead introduced by relay selection is the effort paid for channel estimations. If RRS is employed, two additional channel estimations need to be performed in each time slot to obtain CSI between the BS

and the selected relay as well as that between the source and the selected relay. This results in a complexity of $\mathcal{O}(1)$, which is with the same order as the traditional TDMA scheme. If ORS is employed, CSI between the BS and all potential relays as well as that between the source and all potential relays must be estimated, which leads to a complexity of $\mathcal{O}(N)$.

Moreover, at each time slot, the BS needs to first notify the reputation score of the source node to the selected relay node and then update the selected relay's reputation at the end. Since only binary reputation scores are considered in this paper, we can represent each reputation score efficiently using one bit. Therefore, the communication overhead of reputation update is just 2 bits per time slot, which is almost negligible compared with the size of users' packets.

D. 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, which is defined as

$$\mathcal{C} = P_r c, \quad (7)$$

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

$$\mathcal{G} = E_h[\Gamma_i^c] \cdot g, \quad (8)$$

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

$$\mathcal{G} \approx \begin{cases} \frac{P_r P_s \sigma_1^2 \sigma_2^2 g}{P_r \sigma_1^2 N_0 + P_s \sigma_2^2 N_0} \sum_{n=1}^{N-1} \frac{1}{n} & \text{for ORS,} \\ \frac{P_r P_s \sigma_1^2 \sigma_2^2 g}{P_r \sigma_1^2 N_0 + P_s \sigma_2^2 N_0} & \text{for RRS.} \end{cases} \quad (9)$$

Proof: See Appendix. ■

In practice, the gain can be estimated either using (9) 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 games. Therefore, one important aspect of the network state in indirect reciprocity game modeling is the reputation distribution among

the whole population. In this subsection, we first derive the reputation distribution updating rule. Then we determine the stationary reputation distribution and define the steady state of the game.

Let x_t represents the probability of a user to 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

$$\begin{aligned} x_{t+1} &= x_t [x_t d_{G,G} + (1-x_t) d_{G,B}] \\ &\quad + (1-x_t) [x_t d_{B,G} + (1-x_t) d_{B,B}], \\ &= (d_{G,G} - d_{G,B} - d_{B,G} + d_{B,B}) x_t^2 \\ &\quad + (d_{G,B} + d_{B,G} - 2d_{B,B}) x_t + d_{B,B}, \\ &\triangleq f_{\mathbf{a}}(x_t), \end{aligned} \quad (10)$$

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.

$$d_{i,j} = a_{i,j} Q(i, j, C) + (1 - a_{i,j}) Q(i, j, D). \quad (11)$$

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. According to Table I and (11), we have

$$\begin{cases} d_{G,G} = a_{G,G}(1-\lambda) + \lambda, \\ d_{G,B} = -a_{G,B}(1-\lambda) + 1, \\ d_{B,G} = a_{B,G}(1-\lambda), \\ d_{B,B} = -a_{B,B}(1-\lambda) + (1-\lambda). \end{cases} \quad (12)$$

Based on the reputation distribution updating rule in (10), 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

$$x_{\mathbf{a}} = f_{\mathbf{a}}(x_{\mathbf{a}}). \quad (13)$$

Proof: First, according to (10), the stationary reputation distribution $x_{\mathbf{a}}$ given action rule \mathbf{a} , if exists, must be the solution to (13). Next, in order to show the existence of the stationary reputation distribution, we need to verify that equation (13) has a solution in the interval $[0, 1]$. Let $\tilde{f}_{\mathbf{a}}(x) = f_{\mathbf{a}}(x) - x$. We have $\tilde{f}_{\mathbf{a}}(0) = d_{B,B} \geq 0$ and $\tilde{f}_{\mathbf{a}}(1) = d_{G,G} - 1 \leq 0$. Since (13) is a quadratic equation, there must exist a solution in the interval $[0, 1]$. ■

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

$$\mathbf{H}_{\mathbf{a}} = \begin{bmatrix} (1+d_{G,G})x_{\mathbf{a}} & (1+d_{G,G})(1-x_{\mathbf{a}}) & (1-d_{G,G})x_{\mathbf{a}} & (1-d_{G,G})(1-x_{\mathbf{a}}) \\ (1+d_{G,B})x_{\mathbf{a}} & (1+d_{G,B})(1-x_{\mathbf{a}}) & (1-d_{G,B})x_{\mathbf{a}} & (1-d_{G,B})(1-x_{\mathbf{a}}) \\ d_{B,G}x_{\mathbf{a}} & d_{B,G}(1-x_{\mathbf{a}}) & (2-d_{B,G})x_{\mathbf{a}} & (2-d_{B,G})(1-x_{\mathbf{a}}) \\ d_{B,B}x_{\mathbf{a}} & d_{B,B}(1-x_{\mathbf{a}}) & (2-d_{B,B})x_{\mathbf{a}} & (2-d_{B,B})(1-x_{\mathbf{a}}) \end{bmatrix}, \quad (19)$$

$$\mathbf{b}_{\mathbf{a}} = \frac{1}{2} [(\mathcal{G} - \mathcal{C})a_{G,G} \quad \mathcal{G}a_{B,G} - \mathcal{C}a_{G,B} \quad \mathcal{G}a_{G,B} - \mathcal{C}a_{B,G} \quad (\mathcal{G} - \mathcal{C})a_{B,B}]^T, \quad (20)$$

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

$$u_{i,j}^r(a_{i,j}) = -\mathcal{C}a_{i,j} + \delta [d_{i,j}x_{\mathbf{a}}v_{G,G} + d_{i,j}(1-x_{\mathbf{a}})v_{G,B} + (1-d_{i,j})x_{\mathbf{a}}v_{B,G} + (1-d_{i,j})(1-x_{\mathbf{a}})v_{B,B}], \quad (14)$$

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

$$u_{i,j}^s(a_{j,i}) = \mathcal{G}a_{j,i} + \delta [x_{\mathbf{a}}v_{i,G} + (1-x_{\mathbf{a}})v_{i,B}]. \quad (15)$$

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

$$v_{i,j} = \frac{1}{2} u_{i,j}^r(a_{i,j}) + \frac{1}{2} u_{i,j}^s(a_{j,i}). \quad (16)$$

Substituting (14) and (15) into (16), we have

$$v_{i,j} = \frac{1}{2} \{ -\mathcal{C}a_{i,j} + \delta [d_{i,j}x_{\mathbf{a}}v_{G,G} + d_{i,j}(1-x_{\mathbf{a}})v_{G,B} + (1-d_{i,j})x_{\mathbf{a}}v_{B,G} + (1-d_{i,j})(1-x_{\mathbf{a}})v_{B,B}] \} + \frac{1}{2} \{ \mathcal{G}a_{j,i} + \delta [x_{\mathbf{a}}v_{i,G} + (1-x_{\mathbf{a}})v_{i,B}] \}. \quad (17)$$

Let $\mathbf{V} = [v_{G,G} \quad v_{G,B} \quad v_{B,G} \quad v_{B,B}]^T$ denote the long-term expected payoff vector. The following proposition can be derived.

Proposition 3: In the proposed indirect reciprocity game, the long-term expected payoff vector in a steady state $(\mathbf{a}, x_{\mathbf{a}})$ can be obtained as

$$\mathbf{V} = (\mathbf{I} - \frac{\delta}{2} \mathbf{H}_{\mathbf{a}})^{-1} \mathbf{b}_{\mathbf{a}}, \quad (18)$$

where $\mathbf{H}_{\mathbf{a}}$ is defined in (19), $\mathbf{b}_{\mathbf{a}}$ is defined in (20) and \mathbf{I} is a 4 by 4 identity matrix.

Proof: By rearranging (17) into the matrix form, we have

$$(\mathbf{I} - \frac{\delta}{2} \mathbf{H}_{\mathbf{a}}) \mathbf{V} = \mathbf{b}_{\mathbf{a}}. \quad (21)$$

To prove (18), it suffices to show that matrix $(\mathbf{I} - \frac{\delta}{2} \mathbf{H}_{\mathbf{a}})$ is invertible. Since the row sum of $\frac{1}{2} \mathbf{H}_{\mathbf{a}}$ is 1 for every row and

$0 < \delta < 1$, by the Gerschgorin theorem and the definition of spectral radius in [25], we have

$$\mu(\frac{\delta}{2} \mathbf{H}_{\mathbf{a}}) < 1, \quad (22)$$

where $\mu(\cdot)$ represents the spectral radius. Then, the Corollary C.4 in [26] establishes the invertibility of $(\mathbf{I} - \frac{\delta}{2} \mathbf{H}_{\mathbf{a}})$. ■

C. Equilibrium Steady State

From above analysis, we can see that each player's utility depends heavily on other players' actions. Therefore, as a rational decision-maker, every player will condition his/her action on others' actions. For example, from the social norm in Table 1, we can see that the relay node will have good reputation with a larger probability by choosing cooperation than by choosing defection when the source node has good reputation. In such a case, if other players' action rules favor players with good reputation, the relay node will choose to help in the current time slot since he/she will benefit from others' help in the future. On the other hand, if other players help good reputation players with a very low probability, then the relay node may choose not to help since cooperation is costly.

To study these interactions theoretically, 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 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:

- 1) $(\mathbf{a}, x_{\mathbf{a}})$ is a steady state;
- 2) \mathbf{a} is the best response of any user, given that the reputation distribution is $x_{\mathbf{a}}$ and all other users are adopting action rule \mathbf{a} , i.e. the system is in steady state $(\mathbf{a}, x_{\mathbf{a}})$.

From the definition above, we can see that no user can benefit from any unilateral deviations in an equilibrium steady state. Moreover, determining whether a steady state is an equilibrium is equivalent to the problem of finding the best response of users 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 $(\mathbf{a}, x_{\mathbf{a}})$. Furthermore, since the transition probability and the reward function remain unchanged for a given steady state, the proposed MDP is stationary [26].

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

$$v_{i,j} = \max_{\hat{a}_{i,j}} \left[\frac{1}{2} u_{i,j}^r(\hat{a}_{i,j}) + \frac{1}{2} u_{i,j}^s(a_{j,i}) \right], \quad (23)$$

which can be solved numerically using the well-known value iteration algorithm [26]. 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 4: 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: One necessary condition for a steady state to be an equilibrium is that any single user has no incentive to deviate from the specified action rule for one interaction, which can be mathematically expressed as

$$\frac{1}{2}u_{i,j}^r(a_{i,j}) + \frac{1}{2}u_{i,j}^s(a_{j,i}) \geq \frac{1}{2}u_{i,j}^r(\hat{a}_{i,j}) + \frac{1}{2}u_{i,j}^s(a_{j,i}) \quad (24)$$

for all $i, j \in \{G, B\}$ and $\hat{a}_{i,j} \in [0, 1]$. In (24), $\{a_{i,j}\}$ is the steady state action rule that is employed by all other players and $\{\hat{a}_{i,j}\}$ is an alternative action rule for the player. The second terms on both sides are identical, which is due to the fact that only relay's actions will affect the payoffs. Moreover, since only one-shot deviation is considered here, the long-term expected payoffs starting from next interaction remain unchanged. After substituting (14) into (24), we can rewrite (24) as

$$\begin{aligned} C(\hat{a}_{i,j} - a_{i,j}) \geq & \delta[\Delta d_{i,j}x_{\mathbf{a}}v_{G,G} + \Delta d_{i,j}(1 - x_{\mathbf{a}})v_{G,B} \\ & - \Delta d_{i,j}x_{\mathbf{a}}v_{B,G} - \Delta d_{i,j}(1 - x_{\mathbf{a}})v_{B,B}], \end{aligned} \quad (25)$$

where $\Delta d_{i,j} = \hat{d}_{i,j} - d_{i,j}$ and $\hat{d}_{i,j}$ is the reputation updating probability of user using action rule $\hat{a}_{i,j}$. By substituting (12) into (25) and rearranging the equations, we have

$$[C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] (\hat{a}_{G,G} - a_{G,G}) \geq 0, \forall \hat{a}_{G,G} \in [0, 1]. \quad (26)$$

$$[C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] (\hat{a}_{G,B} - a_{G,B}) \geq 0, \forall \hat{a}_{G,B} \in [0, 1]. \quad (27)$$

$$[C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] (\hat{a}_{B,G} - a_{B,G}) \geq 0, \forall \hat{a}_{B,G} \in [0, 1]. \quad (28)$$

$$[C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] (\hat{a}_{B,B} - a_{B,B}) \geq 0, \forall \hat{a}_{B,B} \in [0, 1]. \quad (29)$$

In (26)-(29), \mathbf{V} is the long-term expected payoff vector which can be computed by (18) and $\mathbf{r} = [x_{\mathbf{a}} \ 1 - x_{\mathbf{a}} \ -x_{\mathbf{a}} \ -1 + x_{\mathbf{a}}]^T$.

Two coefficient terms, $[C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}]$ and $[C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}]$, are critical here in evaluating the steady state. According to (18), we can see that $C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V} = 0$ and $C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V} = 0$ are two linear equations of ρ , each of which can have at most one solution. Therefore, if an steady state is an equilibrium for more than one possible ρ , it must satisfy (26)(28) when $C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V} \neq 0$ holds and (27)(29) when $C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V} \neq 0$ holds.

If $[C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] > 0$, for (26) and (28) to be valid, we must have $a_{G,G} = 0$ and $a_{B,G} = 0$. On the

other hand, if $[C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] < 0$, (26) and (28) will lead to $a_{G,G} = 1$ and $a_{B,G} = 1$. Similarly, from (27) and (29), we will have $a_{G,B} = 0$ and $a_{B,B} = 0$ if $[C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] > 0$ as well as $a_{G,B} = 1$ and $a_{B,B} = 1$ if $[C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] < 0$. Moreover, since $[C - \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] < 0$ and $[C + \delta(1 - \lambda)\mathbf{r}^T\mathbf{V}] < 0$ can not be satisfied simultaneously, there are only three potential equilibrium action rules. The corresponding reputation distributions can then be calculated respectively according to Proposition 2. ■

Results in Proposition 4 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:

Since the formulated MDP for each steady state is stationary, then, according to Theorem 6.2.7 in [26], it suffices to consider only stationary action rules in order to find the optimal action rule. At a steady state $(\mathbf{a}, x_{\mathbf{a}})$, we can express the long-term expected payoff that a user choosing action rule $\hat{\mathbf{a}}$ can receive while others are adopting action rule \mathbf{a} as in (30).

The matrix form of (30) can be written as

$$\mathbf{V}(\hat{\mathbf{a}}, \mathbf{a}) = \frac{\delta}{2}\mathbf{H}_{\hat{\mathbf{a}}}\mathbf{V}(\hat{\mathbf{a}}, \mathbf{a}) + \mathbf{b}(\hat{\mathbf{a}}, \mathbf{a}), \quad (31)$$

where $\mathbf{H}_{\hat{\mathbf{a}}}$ is defined in (19) with the subscript emphasizing its dependence on action rule $\hat{\mathbf{a}}$, and $\mathbf{b}(\hat{\mathbf{a}}, \mathbf{a}) = \frac{1}{2}\mathcal{G}[a_{G,G} \ a_{B,G} \ a_{G,B} \ a_{B,B}]^T - \frac{1}{2}C\hat{\mathbf{a}}$. Applying results in Proposition 3, we have

$$\mathbf{V}(\hat{\mathbf{a}}, \mathbf{a}) = (\mathbf{I} - \frac{\delta}{2}\mathbf{H}_{\hat{\mathbf{a}}})^{-1}\mathbf{b}(\hat{\mathbf{a}}, \mathbf{a}). \quad (32)$$

Moreover, the sufficient and necessary condition for the steady state $(\mathbf{a}, x_{\mathbf{a}})$ to be an equilibrium can be written as

$$\mathbf{V}(\mathbf{a}, \mathbf{a}) \geq \mathbf{V}(\hat{\mathbf{a}}, \mathbf{a}) \quad (33)$$

for all $\hat{\mathbf{a}} = [\hat{a}_{G,G} \ \hat{a}_{G,B} \ \hat{a}_{B,G} \ \hat{a}_{B,B}] \in [0, 1]^4$.

In the following, we solve (33) based on (32) for each of the three steady states in Theorem 1 respectively.

$$v_{i,j}(\hat{\mathbf{a}}, \mathbf{a}) = -\frac{1}{2}\mathcal{C}a_{i,j} + \frac{1}{2}\delta \left[\hat{d}_{i,j}x_{\mathbf{a}}v_{G,G}(\hat{\mathbf{a}}, \mathbf{a}) + \hat{d}_{i,j}(1-x_{\mathbf{a}})v_{G,B}(\hat{\mathbf{a}}, \mathbf{a}) + (1-\hat{d}_{i,j})x_{\mathbf{a}}v_{B,G}(\hat{\mathbf{a}}, \mathbf{a}) \right. \\ \left. + (1-\hat{d}_{i,j})(1-x_{\mathbf{a}})v_{B,B}(\hat{\mathbf{a}}, \mathbf{a}) \right] + \frac{1}{2}\mathcal{G}a_{j,i} + \frac{1}{2}\delta [x_{\mathbf{a}}v_{i,G}(\hat{\mathbf{a}}, \mathbf{a}) + (1-x_{\mathbf{a}})v_{i,B}(\hat{\mathbf{a}}, \mathbf{a})] \quad \forall i, j \in \{G, B\}. \quad (30)$$

- 1) When $\mathbf{a} = [0 \ 0 \ 0 \ 0]^T$ and $x_{\mathbf{a}} = 1/2$, we have $\mathbf{V}(\mathbf{a}, \mathbf{a}) = \mathbf{0}$ and $\mathbf{b}(\hat{\mathbf{a}}, \mathbf{a}) = -\frac{1}{2}\mathcal{C}\hat{\mathbf{a}}$. Therefore, (33) is equivalent to

$$\mathcal{C}(\mathbf{I} - \frac{\delta}{2}\mathbf{H}_{\hat{\mathbf{a}}})^{-1}\hat{\mathbf{a}} \geq \mathbf{0}. \quad (34)$$

Since all elements in matrix $\mathbf{H}_{\hat{\mathbf{a}}}$ and vector $\hat{\mathbf{a}}$ are non-negative, we have $(\mathbf{H}_{\hat{\mathbf{a}}})^n\hat{\mathbf{a}} \geq \mathbf{0}$ for all integer n and all action rule $\hat{\mathbf{a}}$. Then, applying the identity $(\mathbf{I} - \frac{\delta}{2}\mathbf{H}_{\hat{\mathbf{a}}})^{-1} = \sum_{n=0}^{\infty} (\frac{\delta}{2}\mathbf{H}_{\hat{\mathbf{a}}})^n$, we can see that (34) holds for all $0 < \rho < 1$. Therefore, $(\mathbf{a}, x_{\mathbf{a}})$ is an equilibrium steady state for all $0 < \rho < 1$.

- 2) When $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ and $x_{\mathbf{a}} = 1$, based on (32), we can have

$$v_{G,G}(\hat{\mathbf{a}}, \mathbf{a}) = \frac{2(1-\delta)(\mathcal{G}-\mathcal{C}\hat{a}_{G,G})+\delta(1-\lambda)(\mathcal{G}-\mathcal{C})\hat{a}_{B,G}}{2(1-\delta)(2-\delta(1+\lambda+(1-\lambda)(\hat{a}_{G,G}-\hat{a}_{B,G})))}, \\ v_{G,B}(\hat{\mathbf{a}}, \mathbf{a}) = \frac{\psi_1 + \psi_2\hat{a}_{G,G} + \psi_3\hat{a}_{G,B} + \psi_4\hat{a}_{B,G}}{2(1-\delta)(2-\delta(1+\lambda+(1-\lambda)(\hat{a}_{G,G}-\hat{a}_{B,G})))}, \\ v_{B,G}(\hat{\mathbf{a}}, \mathbf{a}) = \frac{(\delta(1-\lambda)\mathcal{G} - (2-\delta-\delta\lambda)\mathcal{C})\hat{a}_{B,G}}{2(1-\delta)(2-\delta(1+\lambda+(1-\lambda)(\hat{a}_{G,G}-\hat{a}_{B,G})))}, \\ v_{B,B}(\hat{\mathbf{a}}, \mathbf{a}) = \frac{\delta(1-\delta)(1-\lambda)(\mathcal{G}-\mathcal{C}\hat{a}_{G,G})+\psi_5\hat{a}_{B,G}+\psi_3\hat{a}_{B,B}}{2(1-\delta)(2-\delta(1+\lambda+(1-\lambda)(\hat{a}_{G,G}-\hat{a}_{B,G})))},$$

where

$$\begin{cases} \psi_1 = [2-\delta(1+\lambda)-\delta^2(1-\lambda)]\mathcal{G}, \\ \psi_2 = -\delta(1-\delta)(2\mathcal{C}+\mathcal{G}(1-\lambda)), \\ \psi_3 = -(1-\delta)[\delta(1-\lambda)\mathcal{G} \\ \quad + (2-\delta(1+\lambda+2(1-\lambda)(\hat{a}_{G,G}-\hat{a}_{B,G})))\mathcal{C}], \\ \psi_4 = \delta(1-\lambda)(\mathcal{G}-\delta\mathcal{C}), \\ \psi_5 = \delta^2(1-\lambda)\mathcal{G}-\delta(1+\lambda-2\lambda\delta)\mathcal{C}. \end{cases}$$

Since $\psi_3 < 0$ and the denominator $2(1-\delta)(2-\delta(1+\lambda+(1-\lambda)(\hat{a}_{G,G}-\hat{a}_{B,G}))) > 0$, the long-term expected payoffs are maximized when $\hat{a}_{G,B} = 0$ and $\hat{a}_{B,B} = 0$. Then, fixing $\hat{a}_{G,B} = 0$ and $\hat{a}_{B,B} = 0$ and maximizing the long-term expected payoffs with respect to $\hat{a}_{G,G} \in [0, 1]$ and $\hat{a}_{B,G} \in [0, 1]$, we can show that the payoff functions are maximized at the boundary point where $\hat{a}_{G,G} = 1$ and $\hat{a}_{B,G} = 1$ when $\rho = \frac{\mathcal{C}}{\mathcal{G}} \leq \frac{\delta(1-\lambda)}{2-\delta-\lambda\delta}$.

- 3) The steady state with $\mathbf{a} = [0 \ 1 \ 0 \ 1]^T$ and $x_{\mathbf{a}} = 0$ is symmetric with the previous steady state. Therefore, the same result can be proved in a similar manner as in 2).

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 enforced by using the proposed indirect reciprocity game.

IV. EVOLUTIONARY MODELING OF THE INDIRECT RECIPROCITY GAME

A. Evolution Dynamics 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

$$\frac{dp_{\mathbf{a}}}{dt} = \eta \left(U_{\mathbf{a}} - \sum_{\mathbf{a} \in \mathbf{A}^D} p_{\mathbf{a}} U_{\mathbf{a}} \right) p_{\mathbf{a}}, \quad (35)$$

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 (35), we have

$$p_{\mathbf{a}}^{t+1} = \left[1 + \eta \left(U_{\mathbf{a}} - \sum_{\mathbf{a} \in \mathbf{A}^D} p_{\mathbf{a}} U_{\mathbf{a}} \right) \right] p_{\mathbf{a}}^t. \quad (36)$$

B. Evolutionarily Stable Strategy

An action rule is asymptotically stable to the replicator dynamics if and only if it is the Evolutionarily Stable Strategy [27], an equilibrium concept widely adopted in evolutionary game theory. Let $\pi(\mathbf{a}, \hat{\mathbf{a}})$ denote the payoff of a player using action rule \mathbf{a} against other players using action rule $\hat{\mathbf{a}}$. Then, we have the formal definition of an ESS as follows.

Definition 3: An action rule \mathbf{a}^* is an ESS if and only if, for all $\mathbf{a} \neq \mathbf{a}^*$,

- equilibrium condition: $\pi(\mathbf{a}, \mathbf{a}^*) \leq \pi(\mathbf{a}^*, \mathbf{a}^*)$, and
- stability condition: if $\pi(\mathbf{a}, \mathbf{a}^*) = \pi(\mathbf{a}^*, \mathbf{a}^*)$, $\pi(\mathbf{a}, \mathbf{a}) < \pi(\mathbf{a}^*, \mathbf{a})$.

According to the above definition of ESS, we have the following theorem.

Theorem 2: In the indirect reciprocity game, we have

- 1) For all $0 < \rho < 1$, action rule \mathbf{a}_1 is an ESS at the steady state $\{\mathbf{a}_1, x_{\mathbf{a}_1}\}$,
- 2) When $\rho < \frac{\delta(1-\lambda)}{2-\delta-\lambda\delta}$, action rule \mathbf{a}_2 and \mathbf{a}_3 are ESSs at steady states $\{\mathbf{a}_2, x_{\mathbf{a}_2}\}$ and $\{\mathbf{a}_3, x_{\mathbf{a}_3}\}$ respectively.

Proof: From the definition of ESS, in order to show an action rule is an ESS, it suffices to prove that the corresponding equilibrium is strict. When $\mathbf{a} = \mathbf{a}_1$, we know from the proof of Theorem 1 that (34) holds for all $0 < \rho < 1$ and all action rules $\hat{\mathbf{a}}$. Moreover, since the row sum of matrix $\frac{\delta}{2}\mathbf{H}_{\hat{\mathbf{a}}}$ is $\delta \in (0, 1)$ for every row, the equality in (34) holds if and only if $\hat{\mathbf{a}} = \mathbf{a}$. Therefore, $(\mathbf{a}_1, x_{\mathbf{a}_1})$ is a strict equilibrium steady state for all $0 < \rho < 1$. Similarly as in the proof of Theorem 1, we can also show that $(\mathbf{a}_2, x_{\mathbf{a}_2})$ and $(\mathbf{a}_3, x_{\mathbf{a}_3})$ are strict equilibrium steady states when $0 < \rho < \frac{\delta(1-\lambda)}{2-\delta-\lambda\delta}$. ■

From Theorem 2, we can see that, when ρ takes value at certain intervals, equilibrium steady states found in Theorem 1 are also stable in the sense that if such an action rule is adopted by the majority of the population, then no other action rule can spread among the population under the influence of replicator dynamics.

V. ENERGY DETECTION

The indirect reciprocity game discussed so far requires that the relay reports its action to the BS. However, due to the selfish nature of users, the selected relay will cheat if cheating can lead to a higher payoff. For example, when the source node has a good reputation, the relay may notify the BS that it will help but keeping silence at the relay phase. The system performance will degrade as a result. To overcome such limitations, we introduce energy detection at the BS to detect whether or not the source's signal is forwarded by the relay.

The hypothesis model of received signals at the relay phase is

$$H_0 : y(t) = n(t), \quad (37)$$

$$H_1 : y(t) = P_r hx(t) + n(t), \quad (38)$$

where $n(t)$ is the additive white Gaussian noise, $x(t)$ is the normalized signal forwarded by the relay, P_r represents the transmitted power of the relay and h is the channel gain from the relay to the BS. The detection statistics of the energy detector is the average energy of M observed samples

$$S = \frac{1}{M} \sum_{t=1}^M |y(t)|^2. \quad (39)$$

Then, the BS can decide whether the relay helped forward signals for the source by comparing the detection statistics S with a predetermined threshold S_0 . The probability of false alarm P_F and the probability of detection P_D for a given threshold are expressed as

$$P_F = \Pr\{S > S_0 | H_0\}, \quad (40)$$

$$P_D = \Pr\{S > S_0 | H_1\}, \quad (41)$$

which can be computed based on the receiver operating characteristic (ROC) curves in [28]. In this work, we regard

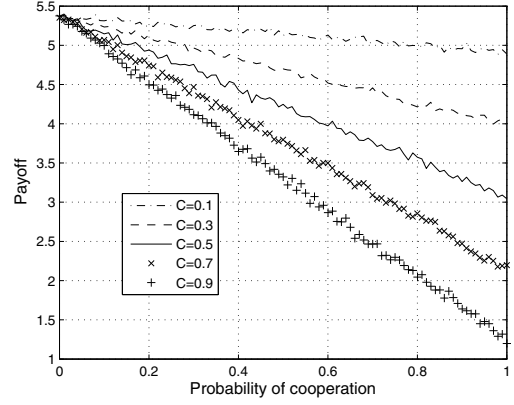


Fig. 2. The payoff versus the probability of cooperation in systems without incentive schemes.

P_F and P_D as system parameters and analyze their impact on user behaviors as follows.

With energy detection, the BS will no longer rely on reports from the relay and thus can prevent the performance degradation caused by cheating. On the other hand, however, reputation may be assigned incorrectly due to false alarm and missing detection. Therefore, after taking the effect of energy detection into account, the new reputation updating probability $d_{i,j}$ can be written as

$$d_{i,j} = [a_{i,j}P_D + (1 - a_{i,j})P_F] Q(i, j, C) + [a_{i,j}(1 - P_D) + (1 - a_{i,j})(1 - P_F)] Q(i, j, D). \quad (42)$$

Then, following the same analysis in Section III and Section IV, we study the indirect reciprocity game with energy detection and obtain the following results.

Corollary 1: In the indirect reciprocity game with energy detection, we have

- 1) The steady state with $\mathbf{a} = [0 \ 0 \ 0 \ 0]^T$ and $x_{\mathbf{a}} = 1/2$ is an equilibrium for all $0 < \rho < 1$
- 2) When $0 < \rho \leq \frac{\delta(1-\lambda)(P_D-P_F)}{2-\delta-\lambda\delta}$, the steady state with $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ and $x_{\mathbf{a}} = \frac{1-P_F}{2-P_D-P_F}$ and the steady state with $\mathbf{a} = [0 \ 1 \ 0 \ 1]^T$ and $x_{\mathbf{a}} = \frac{1-P_D}{2-P_D-P_F}$ are equilibria
- 3) Action rule $\mathbf{a} = [0 \ 0 \ 0 \ 0]^T$ is an ESS for all $0 < \rho < 1$
- 4) When $0 < \rho < \frac{\delta(1-\lambda)(P_D-P_F)}{2-\delta-\lambda\delta}$, action rules $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ and $\mathbf{a} = [0 \ 1 \ 0 \ 1]^T$ are ESSs

Proof: Following the same procedure as in the proof of Proposition 4 and Theorem 1, we can prove 1) and 2). Then, 3) and 4) can be proved in a similar manner as in the proof of Theorem 2. ■

VI. 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.

We first look into the case without reputation and social norm to show the necessity of cooperation stimulation

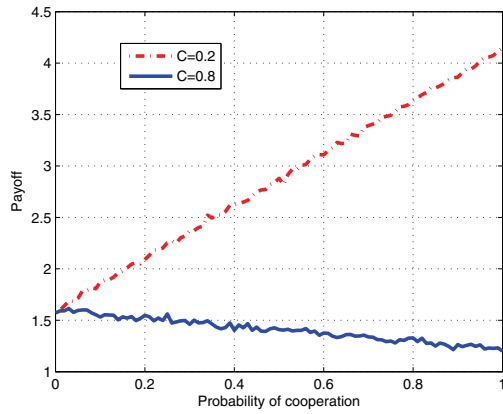


Fig. 3. Equilibrium evaluation of the game without energy detection.

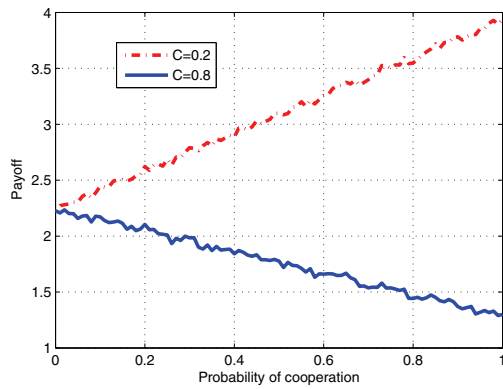


Fig. 4. Equilibrium evaluation of the game with energy detection.

schemes. We assume one particular user chooses to cooperate with probability p while all the other users choose to always cooperate. Fig. 2 shows payoffs of this particular user versus p under different cost values. From the figure, we can see that under all cost values, the user can always gain a higher payoff by cooperating with a lower probability. This is because cooperation is costly and no incentive scheme is employed. As a result, to maximize their payoffs, users will choose not to cooperate and act selfishly as free-riders, which leads to the failure of cooperative communication systems.

In the second simulation, we evaluate the performance of the proposed incentive scheme where λ is set to be 0.5. In Fig. 3, we assume that the game starts at steady state $(\mathbf{a}, x_{\mathbf{a}})$ with $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ and $x_{\mathbf{a}} = 1$. Then we show payoffs of a specified user that deviates to action rule $[p \ 0 \ p \ 0]^T$ under different cost values. From Fig. 3, we can see that as the probability of cooperation p increases, the user's payoff increases when $C = 0.2$ and decreases when $C = 0.8$. 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 Fig. 4, we evaluate the equilibrium steady state of the indirect reciprocity game with energy detection. In the simulation, the probability of false alarm P_F and the probability of detection P_D are set to be 0.1 and 0.9 respectively. Then, according to Corollary 1, the cost-to-gain ratio threshold

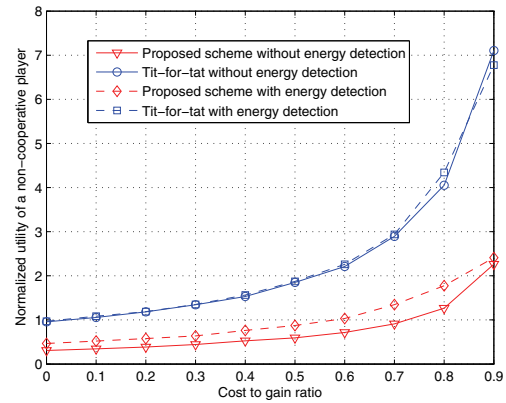


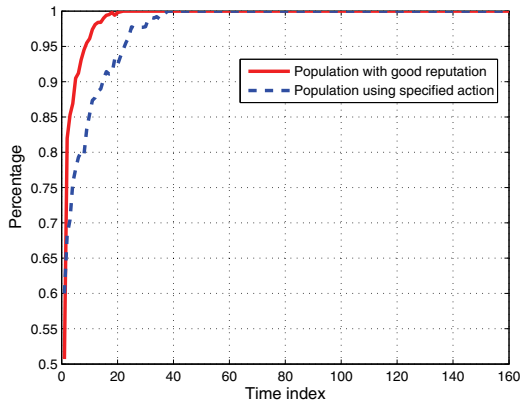
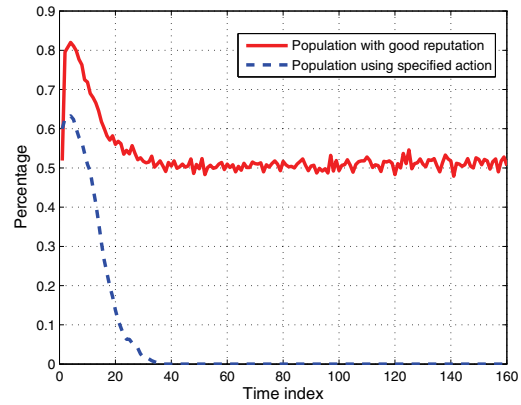
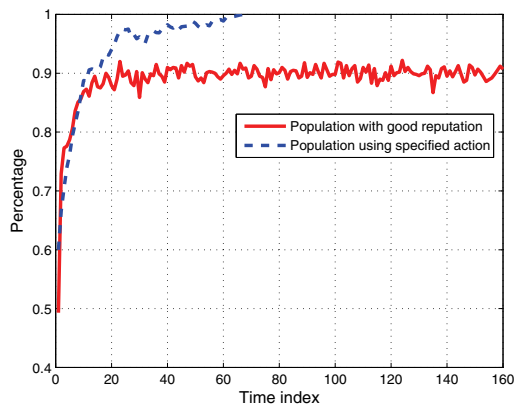
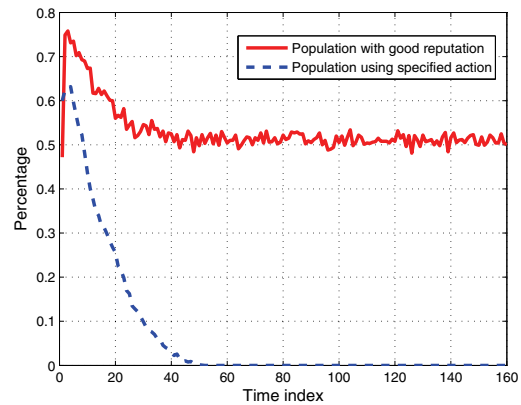
Fig. 5. Comparison of normalized utilities between the indirect reciprocity game and the tit-for-tat mechanism [10] in the case of unilateral deviations.

that enables cooperation becomes $\frac{\delta(1-\lambda)(P_D-P_F)}{2-\delta-\lambda\delta} = 36/65$ and the stationary reputation distribution that corresponds to $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ is $x_{\mathbf{a}} = \frac{1-P_F}{2-P_D-P_F} = 0.9$. Starting from such steady state, we show the payoff of a particular user that deviates to action rule $[p \ 0 \ p \ 0]^T$ under different cost values. From the figure, we can see that the user will have no incentive to deviate from the steady state action rule when $C = 0.2$ while it will not cooperate when $C = 0.8$ just as expected.

Next, we compare the performance of the proposed indirect reciprocity game with that of the tit-for-tat incentive mechanism, which is employed by the BitTorrent file-distribution system to stimulate cooperative behaviors among users [10]. The specified user strategy in the tit-for-tat incentive mechanism is to choose cooperation unless the opponent choose to defect in the previous round. We compare, in Fig. 5, utilities of a user using the action of pure defection normalized by those of using the desired actions between the two schemes. All other users are assumed to adopt the desired actions respectively in both schemes. If the normalized utility is greater than 1, then deviating from the desired action to the action of pure defection is profitable. From the results, we can see that the proposed scheme can enforce cooperation over a much larger range of cost to gain ratios than the tit-for-tat mechanism. This is because the direct reciprocity model that underlies the tit-for-tat incentive mechanism assumes implicitly that the interaction between a pair of users lasts for a long time, which no longer holds in the multi-user cooperative communications scenario.

In the forth simulation, we study the evolutionary properties of the indirect reciprocity game. The initial probability of choosing a specified action rule $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ is set to be 0.6¹ 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

¹This simulation is intended to show that the specified action rule $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ is an ESS under low cost case and thus is resistant to the invasion of any other action rules. In practice, the BS can guide users to adopt the action rule \mathbf{a} at the very beginning by assigning the initial reputation according to the stationary reputation distribution $x_{\mathbf{a}}$.

Fig. 6. Population evolution of the game without energy detection ($C=0.2$).Fig. 8. Population evolution of the game without energy detection ($C=0.8$).Fig. 7. Population evolution of the game with energy detection ($C=0.2$).Fig. 9. Population evolution of the game with energy detection ($C=0.8$).

first study the low cost case where we set $C = 0.2$. From Theorem 2, we know that the specified action rule is an ESS for both games with and without energy detection. In Fig. 6, we show the evolutionary results for the indirect reciprocity game without energy detection. From Fig. 6, we can see that the game converges to the equilibrium steady state that corresponds to the action rule $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$ and remains stable once converges. Therefore, the specified action rule is verified to be an ESS for low cost case.

Evolutionary results for the game with energy detection are shown in Fig. 7. From the figure, we can see that the specified action rule is quickly spread over the whole population. Moreover, the reputation distribution converges to 0.9 and then remains stable just as indicated by Corollary 1.

We then study the high cost case where we set $C = 0.8$. In such case, the cost-to-gain ratio ρ is larger than the thresholds for both games with and without energy detection. Therefore, the specified action rule is no longer an ESS and the cooperation can not be sustained. To verify our theoretical results, we show in Fig. 8 and Fig. 9 the evolutionary results under $C = 0.8$ for games with and without energy detection respectively. From Fig. 8 and Fig. 9, we can see that for both games the reputation distributions converge to 1/2 and the probabilities of choosing the specified action eventually become zero. Therefore, the specified action rule is not an

ESS at the high cost case, which coincides with our results in Theorem 2 and Corollary 1.

Finally, we compare the proposed scheme with the tit-for-tat mechanism from the perspective of population evolution. We set $C = 0.5$ and assume the initial probability of choosing the desired action is 0.6 for both schemes. Moreover, for the proposed scheme, the initial reputation distribution of the population is set as 1/2. For the tit-for-tat mechanism, we consider two actions other than the tit-for-tat strategy: pure cooperation and pure defection, each of which has an initial probability of 0.2. In Fig. 10, we show the population evolution for the case without energy detection. From the results, we can see that the desired action, $\mathbf{a} = [1 \ 0 \ 1 \ 0]^T$, in our scheme is evolutionarily stable while the tit-for-tat strategy is vulnerable to invasions of other actions, which again shows that the proposed indirect reciprocity game is more effective than direct reciprocity based methods. Results with energy detection are of the similar form as in Fig. 10 and therefore are skipped due to page limitations.

VII. CONCLUSIONS AND DISCUSSIONS

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

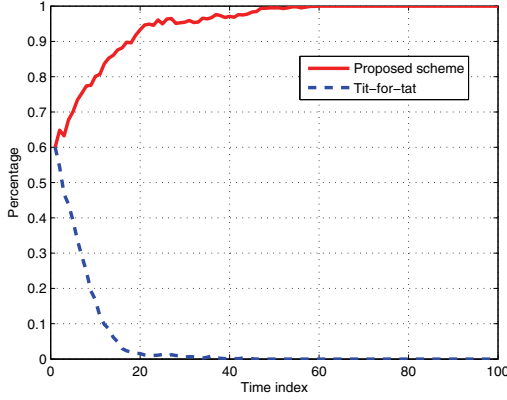


Fig. 10. Comparison of population evolution between the indirect reciprocity game and the tit-for-tat mechanism [10].

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 justify the use of reputation in stimulating cooperation. In particular, we 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 that at low cost case the action rule of relaying information for users with good reputation is an ESS and therefore resistant to the mutation of any other action rules. To take possible cheating behaviors of users into consideration, we also introduce energy detection at the BS and analyze its impact to the indirect reciprocity game. Simulation results show the effectiveness of the proposed scheme in stimulating cooperation among rational and selfish users.

Since there are multiple equilibria in the game, different initial conditions may lead to different outcomes of the game. For example, if the network starts with really bad conditions, e.g., the majority of users choose not to cooperate at the very beginning, then the game will converge to the non-cooperative equilibrium steady state. However, we would like to emphasize that, to achieve full cooperation, our requirements for the initial condition are not stringent. For example, we find from simulation results that even starting with only 60 percentage of population adopting the action rule of cooperating with good reputation users, such an action rule will quickly spread over the whole population and leads to a cooperative society. Moreover, in the proposed scheme, to ensure that the network starts with good initial conditions, one possible way is to let the BS assign the initial reputation to users according to the stationary reputation distribution x_a of the desired action rule a . In such a case, it becomes an equilibrium for all rational users to adopt the action rule a , since (a, x_a) is an equilibrium steady state. Then, even if a small portion of users choose other actions by mistake, the game can still converge to the desired steady state since it is evolutionarily stable.

APPENDIX

PROOF OF PROPOSITION 1

For ORS, let $Y = \max_{j \neq i} \frac{P_r P_s |h_{i,j}|^2 |h_{j,d}|^2}{P_r |h_{j,d}|^2 N_0 + P_s |h_{i,j}|^2 N_0}$. According to [[24], (16)], the cumulative distribution function (CDF) of Y can be written as

$$P_Y(y) = \left[1 - 2y \sqrt{\beta_1 \beta_2} e^{-y(\beta_1 + \beta_2)} K_1(2y \sqrt{\beta_1 \beta_2}) \right]^{N-1}, \quad (43)$$

where $\beta_1 = N_0/P_r \sigma_1^2$, $\beta_2 = N_0/P_s \sigma_2^2$ and $K_1(x)$ is the first-order modified Bessel functions of the second kind, defined in [[29], (9.6.22)]. Moreover, since $Y \geq 0$, we can calculate the expectation of Y as

$$\begin{aligned} E[Y] &= \int_0^\infty [1 - P_Y(y)] dy, \\ &= \sum_{n=1}^{N-1} \binom{N-1}{n} (-1)^{n-1} \int_0^\infty (2y \sqrt{\beta_1 \beta_2})^n e^{-y(\beta_1 + \beta_2)n} \\ &\quad \cdot (K_1(2y \sqrt{\beta_1 \beta_2}))^n dy, \quad (44) \end{aligned}$$

$$\approx \frac{1}{\beta_1 + \beta_2} \sum_{n=1}^{N-1} \binom{N-1}{n} \frac{(-1)^{n-1}}{n}, \quad (45)$$

$$= \frac{1}{\beta_1 + \beta_2} \sum_{n=1}^{N-1} \frac{1}{n}. \quad (46)$$

Note from (44) to (45), we approximated $K_1(x)$ as given in [[29], (9.6.9)] by $K_1(x) \approx 1/x$ and (46) is obtained using the identity in [[30], (0.155, 4)]. Finally, when $P_s/N_0 \gg 1$ and $P_r/N_0 \gg 1$, we have for ORS

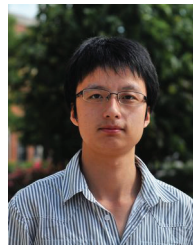
$$\mathcal{G} \approx E[Y] \cdot g \approx \frac{P_r P_s \sigma_1^2 \sigma_2^2 g}{P_r \sigma_1^2 N_0 + P_s \sigma_2^2 N_0} \sum_{n=1}^{N-1} \frac{1}{n}. \quad (47)$$

Since the estimate of \mathcal{G} under RRS can be regarded as a special case of that under ORS with $N-1=1$, results for RRS follow directly from (47).

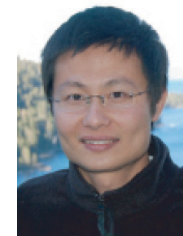
REFERENCES

- [1] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, 2008.
- [2] M. Feldman and J. Chuang, "Overcoming free-riding behavior in peer-to-peer systems," *SIGecom Exch.*, vol. 5, no. 4, pp. 41–50, 2005.
- [3] S. Zhong, J. Chen, and Y. R. Yang, "Sprite: a simple, cheat-proof, credit-based system for mobile ad-hoc networks," in *2003 Annual Joint Conf. IEEE Comput. Commun. Societies*.
- [4] P. Marbach and Y. Qiu, "Cooperation in wireless ad hoc networks: a market-based approach," *IEEE/ACM Trans. Netw.*, vol. 13, pp. 1325–1338, Dec. 2005.
- [5] M. Neely, "Optimal pricing in a free market wireless network," in *Proc. 2007 IEEE International Conf. Comput. Commun.*, pp. 213–221.
- [6] P. Golle, K. Leyton-Brown, and I. Mironov, "Incentives for sharing in peer-to-peer networks," in *Proc. 2001 ACM Symp. Electron. Commerce*.
- [7] S. C. V. Vishnumurthy and E. G. Sirer, "Karma: a secure economic framework for p2p resource sharing," in *Proc. 2003 Workshop Economics Peer-to-Peer Syst.*
- [8] B. Wang, Z. Han, and K. J. R. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using Stackelberg game," *IEEE Trans. Mobile Comput.*, vol. 8, no. 7, pp. 975–990, 2009.
- [9] H. Xiao and E. M. Yeh, "Pricing games with incomplete information in parallel relay networks," in *Proc. 2011 IEEE International Conf. Commun.*
- [10] B. Cohen, "Incentives build robustness in BitTorrent," in *Proc. 2003 Workshop Economics Peer-to-Peer Syst.*

- [11] S. Jun and M. Ahamad, "Incentives in BitTorrent induce free riding," *2005 ACM SIGCOMM Workshop Economics Peer-to-Peer Syst.*
- [12] W. Yu and K. J. R. Liu, "Game theoretic analysis of cooperation and security in autonomous mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 5, pp. 459–473, May 2007.
- [13] Z. Ji, W. Yu, and K. J. R. Liu, "A belief evaluation framework in autonomous MANETs under noisy and imperfect observation: vulnerability analysis and cooperation enforcement," *IEEE Trans. Mobile Comput.*, vol. 9, pp. 1242–1254, Sep. 2010.
- [14] P. Michiardi and R. Molva, "Core: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks," *Commun. Multimedia Security*, pp. 107–121, Sept. 2002.
- [15] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The eigentrust algorithm for reputation management in p2p networks," in *Proc. 2003 Int. World Wide Web Conf.*, pp. 640–651.
- [16] S. Buchegger and J. L. Boudec, "Performance analysis of the confidant protocol," in *Proc. 2002 ACM MobiHoc*, pp. 226–236.
- [17] M. Gupta, P. Judge, and M. Ammar, "A reputation system for peer-to-peer networks," in *Proc. 2003 ACM NOSSDAV*.
- [18] Y. Chen and K. J. R. Liu, "Indirect reciprocity game modelling for cooperation stimulation in cognitive networks," *IEEE Trans. Commun.*, Jan. 2011.
- [19] R. L. Trivers, "The evolution of reciprocal altruism," *Quarterly Rev. Biology*, vol. 46, pp. 35–57, 1971.
- [20] M. Kandori, "Social norms and community enforcement," *Rev. Economic Studies*, vol. 59, no. 1, pp. 63–80, Jan. 1992.
- [21] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, pp. 1291–1298, 2005.
- [22] 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.
- [23] Y. Zhao, R. S. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," in *Proc. 2006 IEEE Int'l Symp. Inf. Theory*, pp. 1234–1238.
- [24] A. S. Ibrahim, A. K. Sadek, W. Su, and K. J. R. Liu, "Cooperative communications with relay selection: when to cooperate and whom to cooperate with?" *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2814–2827, July 2008.
- [25] P. Lancaster and M. Tismenetsky, *The Theory of Matrices: With Applications*, 2nd edition. Academic Press, 1985.
- [26] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.
- [27] J. W. Weibull, *Evolutionary Game Theory*. MIT Press, 1995.
- [28] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, pp. 523–531, Apr. 1967.
- [29] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th edition. Dover Publications, 1970.
- [30] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, 5th edition. Academic Press, 1994.



Yang Gao (S'12) received the B.S. in Electronic Engineering from Tsinghua University, Beijing, China in 2009. Now he is a Ph.D. student in the Department of Electrical and Computer Engineering at University of Maryland, College Park. His current research interests are in the area of cooperation stimulation in communication and networking systems, crowdsourcing, smart grid and network economics. He received the silver medal of the 21st National Chinese Physics Olympiad, the national scholarship from the Ministry of Education of People's Republic of China in 2008, the honor of Outstanding Graduate of Tsinghua University in 2009 and the University of Maryland Future Faculty Fellowship in 2012.



Yan Chen (S'06) received the Bachelor's degree from University of Science and Technology of China in 2004, the M. Phil degree from Hong Kong University of Science and Technology (HKUST) in 2007, and the Ph.D. degree from University of Maryland College Park in 2011. He is currently a research associate in the Department of Electrical and Computer Engineering at University of Maryland College Park. His current research interests are in social learning and networking, smart grid, cloud computing, crowdsourcing, network economics, and multimedia signal processing. Dr. Chen received the University of Maryland Future Faculty Fellowship in 2010, Chinese Government Award for outstanding students abroad in 2011, University of Maryland ECE Distinguished Dissertation Fellowship Honorable Mention in 2011, and was the Finalist of A. James Clark School of Engineering Dean's Doctoral Research Award in 2011.



K. J. Ray Liu (F'03) was named a Distinguished Scholar-Teacher of University of Maryland, College Park, in 2007, where he is Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of signal processing and communications with recent focus on cooperative communications, cognitive networking, social learning and networks, and information forensics and security.

Dr. Liu is the recipient of numerous honors and awards including IEEE Signal Processing Society Technical Achievement Award and Distinguished Lecturer. He also received various teaching and research recognitions from University of Maryland including university-level Invention of the Year Award; and Poole and Kent Senior Faculty Teaching Award and Outstanding Faculty Research Award, both from A. James Clark School of Engineering. An ISI Highly Cited Author in Computer Science, Dr. Liu is a Fellow of IEEE and AAAS.

Dr. Liu is President of IEEE Signal Processing Society where he has served as Vice President - Publications and Board of Governor. He was the Editor-in-Chief of *IEEE Signal Processing Magazine* and the founding Editor-in-Chief of *EURASIP Journal on Advances in Signal Processing*.