

Cooperation Stimulation in Cognitive Networks Using Indirect Reciprocity Game Modelling

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

Abstract—In cognitive networks, since nodes generally belong to different authorities and pursue different goals, they will not cooperate with others unless cooperation can improve their own performance. Thus, how to stimulate cooperation among nodes in cognitive networks is very important. However, most of existing game-theoretic cooperation stimulation approaches rely on the assumption that the interactions between any pair of players are long-lasting. When this assumption is not true, according to the well-known Prisoner's Dilemma and the backward induction principle, the unique Nash equilibrium (NE) is to always play non-cooperatively. In this paper, we propose a cooperation stimulation scheme for the scenario where the number of interactions between any pair of players are finite. The proposed algorithm is based on indirect reciprocity game modelling where the key concept is “*I help you not because you have helped me but because you have helped others*”. We formulate the problem of finding the optimal action rule as a Markov Decision Process (MDP). Using the packet forwarding game as an example, we show that with an appropriate cost-to-gain ratio, the strategy of forwarding the number of packets that is equal to the reputation level of the receiver is an evolutionarily stable strategy (ESS). Finally, simulations are shown to verify the efficiency and effectiveness of the proposed algorithm.

Index Terms—Indirect reciprocity, cognitive network, game theory, Markov decision process.

I. INTRODUCTION

In cognitive networks [1], since nodes generally belong to different authorities and pursue different goals, fully cooperative behaviors cannot be taken for granted. Instead, nodes will only cooperate with others when cooperation can improve their own performance. We regard the nodes with such behaviors as selfish nodes. Therefore, a key problem in cognitive networks is how to stimulate cooperation among selfish nodes.

One way to stimulate cooperation among selfish nodes is to use payment based methods [2] [3]. However, the requirement of tamper-proof hardware or central billing services greatly limits their potential applications. Another way to stimulate cooperation among selfish nodes is to use reputation-based methods. Marti et. al [4] propose to identify the misbehaving nodes and deflect the traffic around them. Buchegger and Boudec [5] propose to further punish the misbehaving nodes by denying forwarding packets. Nevertheless, there is no theoretical justification about the optimality of such approaches.

Recently, efforts have been made to mathematically analyzing cooperation in cognitive networks using game theory. Srinivasan et al. [6] propose to use generous TIT-FOR-TAT strategy while Urpi et al. [7] propose to use Bayesian games. In [8], Yu and Liu propose a game theoretic framework

to jointly analyze cooperation stimulation and security in autonomous mobile ad hoc networks. However, most of the existing game theoretical frameworks rely on the assumption that the game between a pair of players is directly played for infinite times. In reality, due to mobility or changes of environment, nodes will periodically update their partners to achieve better performance, which means that any pair of players are supposed to play for only finite times with the termination time are either known or can be estimated by both players. Note that every player can experience infinite times with many players but never always with the same partner. In such a case, the only optimal strategy is to always play non-cooperatively [9]. The major reason causing such a non-cooperative optimal strategy is the implicit assumption of direct reciprocity in most games, where the action of a player taking towards his/her opponent is purely determined by the history of how the opponent treats him/her. Obviously, under such a scenario, all players have no incentive to play cooperatively since their behaviors will not be evaluated by other players except their opponents.

To stimulate the plays' incentive to play cooperatively, not only the evaluations from the opponents but also the evaluations from other observers should be taken into account, which leads to the notion of “indirect reciprocity”. Indirect reciprocity is a key mechanism for the evolution of human cooperation and has recently drawn a lot of attentions in the area of social science and evolutionary biology [10] [11]. The key concept of indirect reciprocity is “*I help you not because you have helped me but because you have helped others*”. In this paper, we propose to use indirect reciprocity game modelling to stimulate cooperation among selfish nodes for the scenario where the number of interactions between any pair of players are finite. We formulate the problem of finding the optimal action rule as a Markov Decision Process (MDP). Using the packet forwarding game as an example, we show that with an appropriate cost-to-gain ratio, the strategy of forwarding the number of packets that is equal to the reputation level of the receiver is an evolutionarily stable strategy (ESS). Finally, simulations are shown to verify the efficiency and effectiveness of the proposed algorithm.

The rest of this paper is organized as follows. In Section II and III, we describe the system model and show how to find the optimal action rule. In Section IV, we describe the action spreading algorithm. Finally, we show the simulation results in Section V and draw conclusions in Section VI.

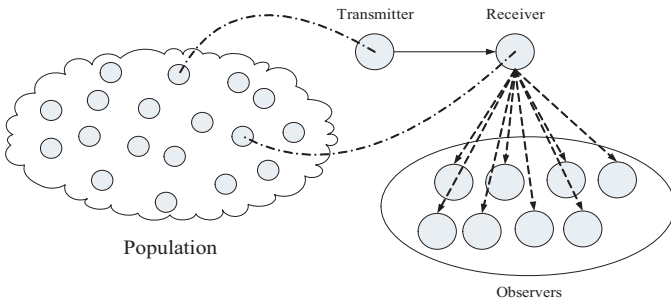


Fig. 1. System model.

II. THE SYSTEM MODEL

As shown in Figure 1, let us consider a cognitive network with sufficiently large population of nodes. Due to mobility and/or changes of environment, short interactions rather than long-lasting associations between anonymous partners are dominant. At each time slot, a fraction of players is chosen from the population to form pairs to forward packets. Within each pair, one player acts as a transmitter and the other player as a receiver. Let $\mathbf{A} = \{0, 1, \dots, L\}$ stand for the action set that the transmitter may choose, where the action $i \in \mathbf{A}$ stands for the transmitter forwards i packets to the receiver.

In the simplest model with $L = 1$, the receiver can obtain a gain g at a cost c to the transmitter. The payoff structure yields an instance of the well-known Prisoner's Dilemma game and the unique NE is defecting. Moreover, with backward deduction, the NE remains the same even the game is played finite times. Such a non-cooperative optimal strategy is mainly because of the use of direct reciprocity, where the action of a transmitter taking towards a receiver is purely determined by the history of how the receiver treats him/her.

To stimulate the cooperation within such a scenario, we use the indirect reciprocity game modelling, where the essential concept is: *"I help you not because you have helped me but because you have helped others"*. Therefore, a key concept in indirect reciprocity game is the establishment of the notion of reputation, which is the evaluation of the history of the players' action. We assume that the reputation is quantized to $L + 1$ levels with "0" being the worst reputation and "L" being the best reputation, i.e., the reputation set can be represented as $\mathbf{T} = \{0, 1, \dots, L\}$. Here, we also assume that everyone agrees on the reputation of an individual and no private opinions are allowed. However, errors in assigning reputation are possible.

A. Social Norms

A social norm, \mathbf{Q} , is a matrix used for updating the immediate reputation of players, where the immediate reputation is the reputation that a transmitter can immediately obtain by taking an action. Each element $Q_{i,j}$ in the social norm stands for the immediate reputation assigned to a transmitter who has taken the action i toward a receiver whose reputation is j . Without loss of generality, we assume that all players in the population share the same norm. Although the immediate reputation is only determined by the action of the transmitter and the reputation of the receiver, we can see from the later discussion, the final reputation updating rule also involves the reputation of the transmitter.

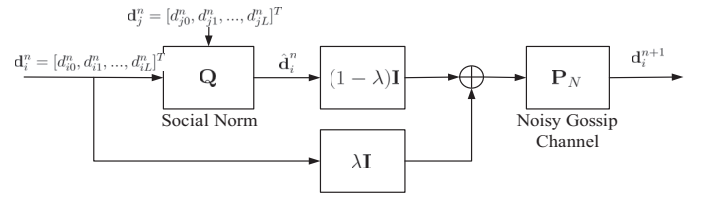


Fig. 2. Reputation updating policy.

Since both the cardinalities of the action set and the reputation set are $L + 1$, there are $(L + 1)^{(L+1) \times (L+1)}$ possible social norms. Based on the intuition that forwarding packets to the receiver with good reputation or denying forwarding packets to the receiver with bad reputation should receive good reputation, in this paper, we assume that the immediate reputation $Q_{i,j}$ is defined as follows

$$Q_{i,j} = L - |i - j|. \quad (1)$$

For the case $L = 1$, the 2×2 social norm is

$$\mathbf{Q}^{2 \times 2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (2)$$

where "1" and "0" stand for good and bad reputation.

With (2), we can see that the transmitter can obtain a good immediate reputation by either forwarding packets to the receiver with good reputation or denying forwarding packets to the receiver with bad reputation. On the other hand, the transmitter will obtain a bad immediate reputation if he/she either denies forwarding packets to the receiver with good reputation or forwards packets to the receiver with bad reputation.

B. Action Rules

An action rule, \mathbf{a} , is an action table of the transmitter, where $a_{i,j}$ stands for the action of the transmitter based on his/her own reputation i and the corresponding receiver's reputation j . Since both the cardinalities of the action set and the reputation set are $L + 1$, there are $(L + 1)^{(L+1) \times (L+1)}$ possible action rules. The optimal action rule, \mathbf{a}^* , should be the one that maximizes the payoff function as discussed later.

III. OPTIMAL ACTION RULE

A. Reputation Updating Policy

A key concept in indirect reciprocity game is reputation [11]. There is a similar notion of trust [12], however, which is mostly based on direct reciprocity. Players monitor the social interactions and help others establish the reputation of being a helpful player. Therefore, one important step in indirect reciprocity game modelling is how to update reputation based on players' actions. In this subsection, we develop a reputation updating policy based on the action of the transmitter, the reputation of the transmitter and the reputation of the receiver.

To capture not only the mean behavior of the transmitter's reputation but also all likelihoods of the transmitter's reputation that may be, we assign a reputation distribution for each player. Let $\mathbf{d} = [d_0, d_1, \dots, d_L]^T$ be a reputation distribution for a specific player. Then d_i stands for the likelihood of the player being assigned with reputation i .

The proposed reputation updating policy is shown in Fig. 2. Suppose, at time index n , a transmitter with reputation distribution \mathbf{d}_i^n is matched with a receiver with reputation

distribution \mathbf{d}_j^n . By taking a certain action, the transmitter is assigned with an immediate reputation $\hat{\mathbf{d}}_i^n$

$$\hat{\mathbf{d}}_i^n = \sum_p \sum_{q: Q_{ap,q}=k} \mathbf{d}_i^n(p) \mathbf{d}_j^n(q). \quad (3)$$

Then, the receiver and the observers will update the transmitter's reputation distribution using a linear combination of the transmitter's original and immediate reputations, where the weight λ can be treated as a discounting factor.

$$\tilde{\mathbf{d}}_i^n = \lambda \mathbf{d}_i^n + (1 - \lambda) \hat{\mathbf{d}}_i^n. \quad (4)$$

Since the reputation is propagated among the population through gossip, which can be modelled as a noisy channel, the transmitter's new reputation distribution is

$$\mathbf{d}_i^{n+1} = \mathbf{P}_N \left(\lambda \mathbf{d}_i^n + (1 - \lambda) \hat{\mathbf{d}}_i^n \right), \quad (5)$$

where \mathbf{P}_N is the transition matrix of the noisy channel. Without loss of generality, we define \mathbf{P}_N as follows

$$\mathbf{P}_N(i, j) = \begin{cases} 1 - \mu, & i = j; \\ \mu/L, & \text{else.} \end{cases} \quad (6)$$

with $\mu \in [0, 0.5]$ being a constant. Note that the analysis in this paper are also applicable to the \mathbf{P}_N with other forms.

B. Stationary Reputation Distribution

Let $\mathbf{x} = [x_0, x_1, \dots, x_L]^T$ stand for the reputation distribution of the entire population, where x_i is the portion of the population that have the reputation i . Since every pair of transmitter and receiver is randomly chosen from the population, given the transmitter with reputation i , the probability of matching with the receiver with reputation k is x_k . After the transmission, the reputation of the transmitter is updated using the policy shown in Fig.2. Therefore, the evolution of \mathbf{x} can be described by

$$\frac{d\mathbf{x}}{dt} = \mathbf{x}^{new} - \mathbf{x}, \quad (7)$$

where \mathbf{x}^{new} is the new reputation distribution

$$\mathbf{x}^{new} = \mathbf{P}_N (\lambda \mathbf{I} + (1 - \lambda) \mathbf{P}_T) \mathbf{x}, \quad (8)$$

with the matrix \mathbf{P}_T being defined as

$$\mathbf{P}_T(j, i) = \sum_{k: Q_{a_i^*, k}=j} x_k. \quad (9)$$

According to (7-9), the stationary reputation distribution \mathbf{x}^* is the solution to the following equation

$$\mathbf{P}_N (\lambda \mathbf{I} + (1 - \lambda) \mathbf{P}_T) \mathbf{x}^* = \mathbf{x}^*. \quad (10)$$

From (9) and (10), we can see that, given the optimal action \mathbf{a}^* , the stationary reputation distribution can be found by solving the nonlinear equations in (10).

C. Payoff Function

Suppose that the cost of forwarding a packet is a constant, c , the total cost of the transmitter with reputation i taking action $a_{i,j}$ towards a receiver with reputation j is given by

$$C(a_{i,j}) = a_{i,j} c. \quad (11)$$

Similarly, if the gain of receiving a packet is a constant, g , the total gain of the receiver with reputation i is

$$G(a_{j,i}) = a_{j,i} g, \quad (12)$$

where $a_{j,i}$ is the action of the transmitter with reputation j .

Let $W_{i,j}$ denote the maximum payoff that a player, currently having reputation i and being matched with a player with reputation j , can gain from this interaction to future. Obviously, if the player with reputation i serves as a transmitter, then the long-term expected payoff that he/she can obtain by taking action $a_{i,j}$ would be

$$f_1(a_{i,j}) = -a_{i,j} c + \delta \sum_k \sum_l \tilde{\mathbf{d}}_{i \rightarrow j}(k) \mathbf{x}^*(l) W_{k,l}, \quad (13)$$

$$\text{with } \tilde{\mathbf{d}}_{i \rightarrow j} = \mathbf{P}_N \left(\lambda \mathbf{e}_i + (1 - \lambda) \mathbf{e}_{Q_{a_{i,j}, j}} \right).$$

The first term $a_{i,j} c$ is the immediate cost the transmitter incurred by taking action $a_{i,j}$, and the second term $\sum_k \sum_l \tilde{\mathbf{d}}_{i \rightarrow j}(k) \mathbf{x}^*(l) W_{k,l}$ stands for the benefit he gains in the future with a discounting factor δ .

On the other hand, if the player with reputation i serves as a receiver, the expected payoff that he/she can obtain is

$$f_2 = a_{j,i}^* g + \delta \sum_l x_l^* W_{i,l}, \quad (14)$$

where the first term $a_{j,i}^* g$ is the immediate gain he/she can obtain when the transmitter taking the optimal action $a_{j,i}^*$, and the second term $\sum_l x_l^* W_{i,l}$ stands for the benefit he gains in the future with a discounting factor δ .

With each interaction, the play acts either as a transmitter or as a receiver with equal probability $\frac{1}{2}$. Therefore, the Bellman equation of $W_{i,j}$ can be written as

$$W_{i,j} = \max_{a_{i,j}} \frac{1}{2} \left(-a_{i,j} c + \delta \sum_k \sum_l \tilde{\mathbf{d}}_{i \rightarrow j}(k) \mathbf{x}^*(l) W_{k,l} \right) + \frac{1}{2} \left(a_{j,i}^* g + \delta \sum_l \mathbf{x}^*(l) W_{i,l} \right), \quad (15)$$

and the optimal action $a_{i,j}^*$ can be computed by

$$a_{i,j}^* = \arg \max_{a_{i,j}} \frac{1}{2} \left(-a_{i,j} c + \delta \sum_k \sum_l \tilde{\mathbf{d}}_{i \rightarrow j}(k) \mathbf{x}^*(l) W_{k,l} \right). \quad (16)$$

From (15) and (16), we can see that the problem of finding the optimal action rule is a Markov Decision Process (MDP), where the state is the reputation pair (i, j) , the action is $a_{i,j}$, the transition probability is determined by $\tilde{\mathbf{d}}_{i \rightarrow j}$ and \mathbf{x}^* , and the reward is determined by c and g . Therefore, given the stationary reputation distribution, the optimal action can be found by solving (16) using value iteration algorithm.

IV. ACTION SPREADING DUE TO NATURAL SELECTION

Based on the previous analysis, we can find the optimal action rule and the stationary reputation distribution. However, we do not include the perturbation effect, where players may take non-optimal action rule due to uncertainty of the system and/or the incorrect (noisy) parameters. Taking the perturbation effect into account, we need to evaluate the stability of the optimal action rule. Here, we adopt the concept of evolutionarily stable strategy (ESS) [13], which is "a strategy such that, if all members of the population adopt it, then no mutant strategy could invade the population under the influence of natural selection". In the following, we discuss, by natural selection, how the action rules spread over the population using Wright-Fisher model [14].

Let M be the number of action rules, $\mathbf{a}_1, \dots, \mathbf{a}_M$, used in the population. Let p_i^t be the percentage of the population that

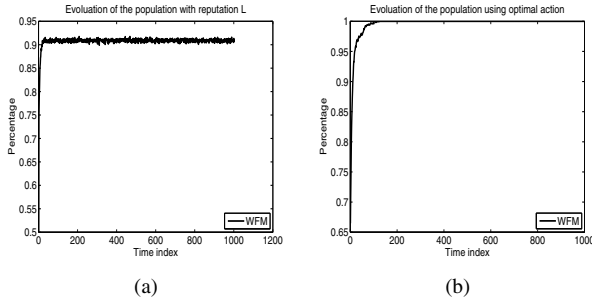


Fig. 3. The evolution using WFM when $g = 1$ and $c = 0.1$: (a) the percentage of the population with reputation $L = 1$; (b) the percentage of the population using optimal action shown in (19).

uses action rule \mathbf{a}_i at time t . Then, we have $\sum_{i=1}^M p_i^t = 1$. Let U_i^t be the average payoff by using action rule \mathbf{a}_i at time t .

With the Wright-Fisher Model, the probability of using action \mathbf{a}_i is proportional to the total payoff of the users using \mathbf{a}_i . Therefore, the action spreading equation can be written as

$$p_i^{t+1} = \frac{p_i^t U_i^t}{\sum_{j=1}^M p_j^t U_j^t}. \quad (17)$$

V. SIMULATION RESULTS

To verify the proposed algorithm, we simulate the packet forwarding game. We study a population of fixed size, $N = 1000$. Each new player receives an initial reputation, which is randomly chosen from $\{0, 1, \dots, L\}$ with equal probability $\frac{1}{L+1}$. Each player uses one of $(L+1)^{(L+1) \times (L+1)}$ possible action rules. Before any one elementary step of action updating, each individual has exactly 20 interactions with other randomly chosen individuals. Individuals act as transmitter and receiver on average 10 times each. After each interaction, the reputation of the transmitter is updated according to the reputation updating policy shown in Fig. 2. Then, the players in the population choose their new action rules according to previous payoff history of the whole population using Wright Fisher Model, which is denoted as “WFM”. In all the following simulations, the parameters λ , δ , and μ are set to be 0.5, 0.9, and 0.95 respectively.

A. Binary Reputation Scenario

We first evaluate the binary reputation scenario where $L = 1$. We assume that $g = 1$ and $c = 0.1$. According to section III, with different initial conditions, we can find three pairs of stationary reputation distribution \mathbf{x}^* and the optimal action rule \mathbf{a}^* , which are

$$\mathbf{x}_1^* = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \quad \mathbf{a}_1^* = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (18)$$

$$\mathbf{x}_2^* = \begin{pmatrix} 0.0909 \\ 0.9091 \end{pmatrix}, \quad \mathbf{a}_2^* = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}. \quad (19)$$

$$\mathbf{x}_3^* = \begin{pmatrix} 0.9091 \\ 0.0909 \end{pmatrix}, \quad \mathbf{a}_3^* = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}. \quad (20)$$

With $(\mathbf{x}_1^*, \mathbf{a}_1^*)$, no one will transmit any packet to the receiver regardless his/her own reputation and the corresponding receiver’s reputation. Obviously, it is a bad strategy since, with such a strategy, there is no cooperation and the payoff

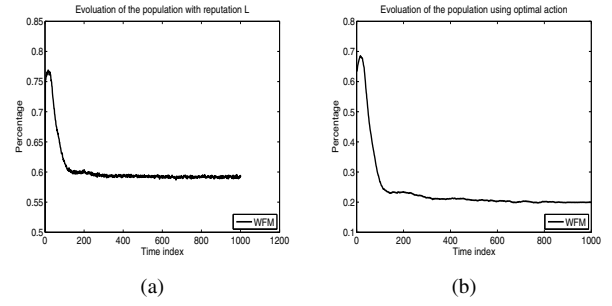


Fig. 4. The evolution using WFM when $g = 1$ and $c = 0.6$: (a) the percentage of the population with reputation $L = 1$; (b) the percentage of the population using optimal action shown in (19).

of every player is zero. The pairs $(\mathbf{x}_2^*, \mathbf{a}_2^*)$ and $(\mathbf{x}_3^*, \mathbf{a}_3^*)$ are symmetry where with the former pair, the transmitter will always cooperate with the receiver with good reputation, and with the latter pair, the transmitter will always cooperate with the receiver with bad reputation. We can also find that the pair $(\mathbf{x}_2^*, \mathbf{a}_2^*)$ leads to a population with more than 90 percentage of the population are good reputation while $(\mathbf{x}_3^*, \mathbf{a}_3^*)$ leads to a population with more than 90 percentage of the population are bad reputation. Here, we prefer $(\mathbf{x}_2^*, \mathbf{a}_2^*)$ since it leads to a “good” society with more than 90 percentage of the population are good reputation.

Then, we evaluate the evolutionary stability of $(\mathbf{x}_2^*, \mathbf{a}_2^*)$. In the simulation, the initial frequency of the optimal action rule \mathbf{a}^* shown in (19) is set to be 0.6. The initial frequencies of the other action rules are randomly chosen. The initial reputation of new players is randomly chosen from $\{0, 1\}$ with equal probability $\frac{1}{2}$. In Fig. 3, we show the evolutionary result using WFM. From Fig. 3, we can see that, the reputation distribution converges to the stationary reputation distribution \mathbf{x}_2^* and the action rule \mathbf{a}_2^* spreads over the whole population. And once the whole population adopt \mathbf{a}_2^* , no one will deviate. Therefore, the action rule \mathbf{a}_2^* is an ESS in this case.

From (15), we can see that the optimal action rule is determined by the values of g and c . Intuitively, if $g \gg c$, every player is willing to cooperate with other players since in such a scenario, the potential cooperation gain will be greater than the immediate cooperation cost. On the other hand, if $c \gg g$, every player tends not to cooperate with other players since the potential cooperation gain will be smaller than the immediate cooperation cost in such a scenario. Therefore, there should exist a critical cost-to-gain ratio γ such that the optimal action rule \mathbf{a}_2^* is stable if $c < \gamma g$ and is not stable otherwise.

By setting \mathbf{a}_2^* as the initial action rule \mathbf{a}^0 in the value iteration algorithm and varying g and c , we find that if $\frac{c}{g} \leq 0.582$, the optimal action rule found by value iteration is \mathbf{a}_2^* . On the other hand, if $\frac{c}{g} > 0.582$, the optimal action rule changes to be \mathbf{a}_1^* . Therefore, the critical cost-to-gain ratio γ is equal to 0.582 in this case.

We verify the above statement by evaluating the stability of \mathbf{a}_2^* when $g = 1$ and $c = 0.6$. The corresponding evolutionary results are shown in Fig. 4. From Fig. 4, we can see that when $\frac{c}{g} = 0.6 > 0.582$, the percentage of the population using action rule \mathbf{a}_2^* does not converge to 1. Therefore, \mathbf{a}_2^* is not stable in this case. Correspondingly, we can also see that

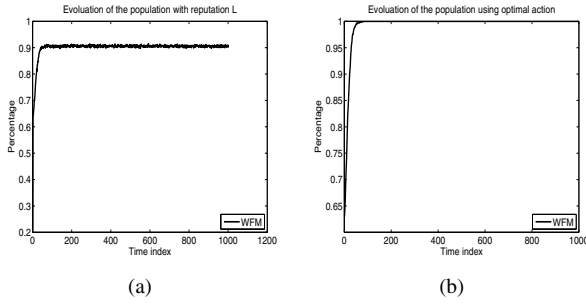


Fig. 5. The evolution using WFM when $g = 1$ and $c = 0.5$: (a) the percentage of the population with reputation $L = 4$; (b) the percentage of the population using optimal action shown in (19) when $L = 4$.

the reputation distribution does not converge to \mathbf{x}_0^* in this case.
B. Multi-Level Reputation Scenario

For the multi-level reputation scenario where $L \geq 2$, due to the large dimension of the action space $((L+1)^{(L+1) \times (L+1)})$, it is difficult to find all the possible pairs of \mathbf{x}^* and \mathbf{a}^* . However, based on the results in the binary reputation scenario, we can infer that one possible optimal action rule \mathbf{a}_0^* is to forward i packets to the receiver with reputation i , i.e.,

$$\mathbf{a}_0^*(i, j) = j. \quad (21)$$

Similar to the binary reputation scenario, we can find the corresponding stationary reputation distribution \mathbf{x}_0^* . For the special case with $L = 4$, \mathbf{x}_0^* is

$$\mathbf{x}_0^* = (0.0235 \ 0.0235 \ 0.0235 \ 0.0235 \ 0.906)^T. \quad (22)$$

Then, we obtain the stable region for the optimal action rule \mathbf{a}_0^* . By setting \mathbf{a}_0^* as the initial value in the value iteration algorithm and varying g and c , we find that if $\frac{c}{g} \leq 0.622$, the optimal action rule found by value iteration is still \mathbf{a}_0^* . On the other hand, if $\frac{c}{g} > 0.622$, the optimal action rule changes. Therefore, the critical cost-to-gain ratio in this case is equal to 0.622, which means that the stable region for \mathbf{a}_0^* is $\frac{c}{g} \leq 0.622$.

We then verify the above statement by simulating the packet forwarding game with two different cost-to-gain ratio setting. One is $g = 1$ and $c = 0.5$, i.e. $\frac{c}{g} = 0.5 < 0.622$, and the other is $g = 1$ and $c = 0.7$, i.e. $\frac{c}{g} = 0.7 > 0.622$. The evolutionary results for the former setting are shown in Fig. 5. From Fig. 5, we can see that when the cost-to-gain ratio is set to be $\frac{c}{g} = 0.5 < 0.622$, the reputation distribution converges to \mathbf{x}_0^* and the optimal action rule \mathbf{a}_0^* spread over the whole population, which verifies that \mathbf{a}_0^* is an ESS in this case. The evolutionary results for the latter cost-to-gain ratio setting are different and shown in Fig. 6. From Fig. 6, we can see that when the cost-to-gain ratio is set to be $\frac{c}{g} = 0.7 > 0.622$, the action rule \mathbf{a}_0^* does not spread over the whole population and the reputation distribution does not converge to \mathbf{x}_0^* . Therefore, \mathbf{a}_0^* is not stable in this case.

VI. CONCLUSIONS

In this paper, we propose a cooperation stimulation scheme for cognitive networks using indirect reciprocity game modelling. Different from the existing game theoretic approaches, our proposed scheme does not rely on the assumption that the number of interactions between a pair of players is infinite.

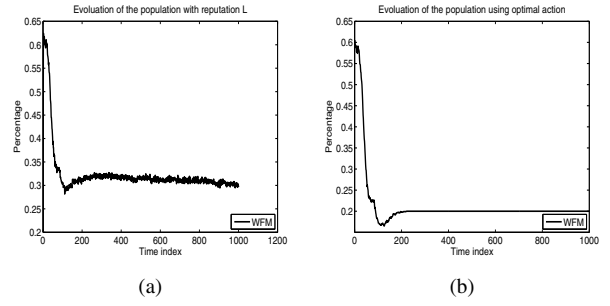


Fig. 6. The evolution using WFM when $g = 1$ and $c = 0.7$: (a) the percentage of the population with reputation $L = 4$; (b) the percentage of the population using optimal action shown in (19) when $L = 4$.

From simulation results, we can see that with a proper cost-to-gain ratio, the action rule of forwarding i packets to the receiver with reputation level i is an ESS. Even starting with only 60 percentage of population adopting the optimal strategy, the optimal strategy will quickly spread over the whole population by natural selection. And once the whole population use the optimal strategy, no one will deviate. Moreover, such an ESS will lead to a “good” society where more than 90 percentage of the population have good reputation.

REFERENCES

- [1] R. W. Thomas, D. H. Friend, L. A. DaSilva, and A. B. MacKenzie, “Cognitive networks: Adaptation and learning to achieve end-to-end performance objectives,” *IEEE Commun. Mag.*, vol. 44, pp. 51–57, December 2006.
- [2] S. Zhong, L. Li, Y. G. Liu, and Y. R. Yang, “On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks - an integrated approach using game theoretical and cryptographic techniques,” in *Proc. ACM MobiCom*, 2005.
- [3] L. Anderegg and S. Eidenbenz, “Ad hoc-veg: A truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents,” in *Proc. ACM MobiCom*, 2003.
- [4] S. Marti, T. J. Giuli, K. Lai, and M. Baker, “Mitigating routing misbehavior in mobile ad hoc networks,” in *Proc. ACM MobiCom*, 2000.
- [5] S. Buchegger and J. Y. L. Boudec, “Performance analysis of the confidant protocol,” in *Proc. ACM MobiHoc*, 2002.
- [6] V. Srinivasan, P. Nuggehalli, C. F. Chiasserini, and R. R. Rao, “Cooperation in wireless ad hoc networks,” in *Proc. IEEE INFOCOM*, 2003.
- [7] A. Urpi, M. Bonuccelli, and S. Giordano, “Modeling cooperation in mobile ad hoc networks: A formal description of selfishness,” in *Proc. Workshop Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOPT’03)*, 2003.
- [8] W. Yu and K. J. R. Liu, “Game theoretic analysis of cooperation stimulation and security in autonomous mobile ad hoc networks,” *IEEE Trans. Mobile Comput.*, vol. 6, pp. 459–473, May 2007.
- [9] M. J. Osborne and A. Rubinsten, *A Course in Game Theory*. The MIT Press, 1994.
- [10] M. A. Nowak and K. Sigmund, “Evolution of indirect reciprocity,” *Nature*, vol. 437, pp. 1291–1298, 2005.
- [11] 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.
- [12] Y. Sun, W. Yu, Z. Han, and K. J. R. Liu, “Information theoretic framework of trust modelling and evaluation for ad hoc networks,” *IEEE Journal of Selected Areas in Communications, special issue on Security in Wireless Ad Hoc Networks*, vol. 24, pp. 305–317, 2006.
- [13] J. M. Smith, *Evolutionary and the theory of games*. Cambridge University Press, 1982.
- [14] R. Fisher, *The genetical theory of natural selection*. Clarendon Press, Oxford, 1930.