SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

November 2017, Vol. 60 110205:1–110205:17
doi: 10.1007/s11432-017-9165-7

Special Focus on Distributed Cooperative Analysis, Control and Optimization in Networks

Cooperation and distributed optimization for the unreliable wireless game with indirect reciprocity

Changbing TANG[1], Xiang LI[2,3*], Zhen WANG[4] & Jianmin HAN[1]

[1]*College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua* 321004*, China;*
[2]*Adaptive Networks and Control Lab, Department of Electronic Engineering, Fudan University, Shanghai* 200433*, China;*
[3]*Research Center of Smart Networks and Systems, Fudan University, Shanghai* 200433*, China;*
[4]*School of Cyberspace, Hangzhou Dianzi University, Hangzhou* 310018*, China*

Received April 21, 2017; accepted July 10, 2017; published online September 22, 2017

Abstract Cooperation in packet forwarding among users and operators of a distributed wireless network has been widely studied. However, because of the limited computational resources, users in wireless communication do not prefer to cooperate with others unless cooperation may improve their own performance. Therefore, the key problem in cooperation enforcement must be solved first to enable a wireless network to be efficient. Yet, most of the existing game-theoretic cooperation stimulation approaches assume that the interactions between any pair of players (users) are long-lasting. In this paper, we apply game theory to optimize the communication efficiency of a distributed wireless network with finite number of interactions between any pair of players. Based on the mechanism of indirect reciprocity, we theoretically analyze the optimal action rule with the method of dynamic programming, and derive the approximate threshold of benefit-to-cost ratio to achieve the optimal action rule. Furthermore, we adopt the replicator dynamics to assess the evolutionary stability of the optimal action rule against the perturbation effect. Numerical illustrations verify the performance of the proposed method on wireless cooperation.

Keywords wireless cooperation, distributed optimization, game theory, indirect reciprocity, optimal action, evolutionary stability

Citation Tang C B, Li X, Wang Z, et al. Cooperation and distributed optimization for the unreliable wireless game with indirect reciprocity. Sci China Inf Sci, 2017, 60(11): 110205, doi: 10.1007/s11432-017-9165-7

1 Introduction

Over the past decade, we have witnessed tremendous development in distributed wireless networks, where all the networking functions rely on the contribution of the participants [1]. Meanwhile, the increasing and diversified demand of wireless services nowadays requires high spectrum efficiency and data rate transmission. As an example, the users of a network must make mutual contribution to packet forwarding ensuring an operable communication of the network. However, since the users usually have limited computational resources (such as battery, memory and processing capacity), selfish users may refuse to be cooperative [2]. It is well-known in the literature that the performance of an entire system

* Corresponding author (email: lix@fudan.edu.cn)

can be degraded dramatically by these selfish behaviors [3, 4]. Additionally, punishing and identifying selfish users will decrease the throughput of cooperative users, which triggers the complete network disconnection [5]. Therefore, a key problem in distributed wireless networks is how to optimize the communication efficiency of the networks, i.e., developing incentive mechanisms to ensure cooperative packet forwarding among selfish users.

Many approaches have been proposed to stimulate cooperation in the services of distributed wireless networks. One is to adopt the payment-based scheme to stimulate cooperation [6–10], where a selfish user can be paid by another user given that the former helps forward the packets for the latter. The payments can be points, money, or objects of similar value. Even if the payment-based schemes can stimulate users to be cooperative, their potential applications have been greatly limited by the requirement of tamper-proof hardware or central billing services. Another approach is to adopt the reputation-based scheme to enforce cooperation [11–14]. The basic goal of reputation-based schemes is to evaluate each user's trustworthiness based on his behaviors, and distinguish misbehaving users from cooperative users according to their reputations. With the reputation-based schemes, each user can maintain a table recording of the reputation values, but it cannot reward high reputation users or punish low reputation users at different reputation levels.

Subsequent efforts have not deviated far from the above two approaches but have tried to use game theory [15–20], in which each vertex is regarded as an intelligent rational agent, and chooses his own action only with local information. In [21], Srinivasan et al. used generous "TIT-FOR-TAT" strategy to study cooperation among selfish users. While in [22], Félegyházi et al. developed a packet forwarding game model in autonomous mobile ad hoc networks, and derived the conditions for the Nash equilibrium of cooperation users. The authors of [23] proposed a two-player packet forwarding game model to analyze cooperation stimulation in autonomous mobile ad hoc networks, and derived the cheat-proof packet forwarding strategies. Ref. [24] proposed the solution to cooperatively deliver packets, where they presented a distributed algorithm to obtain the stable coalitions and evaluated the stable coalitional structures. In [25], the authors proposed a zero-determinant strategy for resource sharing in wireless cooperation, and found the maximum social welfare that the administrator of cooperation (AoC) can achieve with existence of participants of cooperation (PoCs). Besides, in [26–30], the authors employed game theory to analyze cooperation enforcement among selfish users, and poured attention to the updating interaction strategies of individuals with the behaviors of others. A thorough review of this topic can be referred to [31–33].

Nevertheless, most of the available game theoretical frameworks stand on the assumption that the channels of wireless communication are perfect. In reality, one of the prominent properties of channels in wireless communication is that the channels are unreliable. During the forwarding process, a link breakage or transmission errors may lead to dropped packets even if other players are willing to forward the packets [34, 35]. Accordingly, how to enforce cooperation among selfish users in the scenarios with unreliable channels is still open in wireless communication. Furthermore, another inadequacy in the aforementioned game theoretical frameworks lies that they assume fixed partners for the players in the game. With the famous Prisoner's Dilemma, the optimal strategy in such a situation is always to defect. Consequently, one mechanism for cooperation stimulation is the direct reciprocity, in which the action of each player towards his/her opponent is purely determined by the actions of the opponent. Yet, the optimal action is not effective against the "trembling hands" or "fuzzy minds". In fact, players will periodically update their partners for better performance due to mobility and variation of environment. Apparently, within such a scenario, there are no incentives for all players to take the cooperative action as their behaviors will not be evaluated by other players except their opponents.

Consequently, "indirect reciprocity" is proposed to stimulate the players playing cooperatively in the scenario where the interactions are limited with unreliable channels in wireless communication. Indirect reciprocity is captured in the principle: "I help you not because you have helped me but because you have helped others" [36]. In [37], the authors proposed a game modelling of indirect reciprocity for cooperation enforcement in cognitive networks. They formulated the problem to find the optimal action rule as a Markov Decision Process (MDP), showing the optimal action rule to be an evolutionarily stable strategy

(ESS) with an appropriate cost-to-gain ratio. However, they only gave the numerical solution of the cost-to-gain ratio with extensive simulations. Furthermore, they did not discuss the unreliable properties of channels during the forwarding process. Ref. [38] dealt with the enforcement cooperation of packet forwarding game for the unreliable wireless channels with the indirect reciprocity mechanism. However, the packet forwarding model they used is a peer-to-peer game between the two users, who forwards the packet for each other at each time slot, while the event of packet forwarding is not bilateral in reality. Besides, they only discussed the packet forwarding problem from the viewpoint of the evolutionary stability, and ignored the optimal strategy chosen by the self-interest players during the process of packet forwarding. The contributions of this paper are as follows.

- We model the packet forwarding process as a game for the scenario with a limited number of interactions between any pair of players, and stimulate cooperation among selfish players in unreliable wireless communications with the indirect reciprocity scheme.
- We theoretically analyze the optimal action rule based on the method of dynamic programming, and derive the approximate threshold of benefit-to-cost ratio to achieve the optimal action rule.
- We adopt the replicator dynamics to evaluate the evolutionary stability of the optimal action rule against the perturbation effect. We also show that the optimal action will quickly spread over the whole population when the benefit-to-cost ratio exceeds the certain threshold, and once the whole population uses it, there will be no deviation.
- We verify the performance of the optimal action with the indirect reciprocity scheme through a multitude of simulations. Also, the network throughput performance is offered with different channel loss probabilities and reputation updating errors.

The organization of the rest part is as follows. Section 2 models the packet forwarding problem as an indirect reciprocity game, and introduces some preliminaries of the paper. In Section 3, we analyze the stationary reputation distribution of the whole population, and figure out the threshold of benefit-to-cost ratio to reach the optimal action rule. Besides, we analyze the evolutionary stability of the optimal action rule using the replicator dynamics. Section 4 presents numerical simulations to illustrate the efficiency and effectiveness of the optimal action rule. Finally, Section 5 concludes the paper.

2 Game modelling for packet forwarding with indirect reciprocity

2.1 Problem formulation

Consider a distributed wireless network with a sufficiently large number of players (nodes). Due to the limited communication range, the service of source providers cannot arrive directly at the destination, i.e., the communication relies on the packet forwarding by other intermediate players (see Figure 1). At each time slot, one player will act as a service provider, and the communication needs the relay player to forward the provider's packet to its destination. During the process of communication, the relay will choose his strategy, X, from the strategy set $\mathbb{S} = \{F, D\}$, where F and D are packet forwarding and dropping, respectively. We summarize the notations in Table 1.

As a result of channel noise, there are imperfect observations usually in such wireless communication networks, where each player launches some traffic monitoring mechanisms to track strategies of its neighbors [12]. Consider that the receiver of each player catches a private signal from the set $\Theta = \{f, d\}$, where f and d are the observations of packet forwarding and dropping, respectively. With imperfect observation, the forwarding strategy F of one player may be observed as signal d by the other player because of the link breakage or transmission errors. Such a channel loss probability is denoted as p_e. As shown in Figure 1, node S_j forwards the packet for S_i to destination node D_i, but the channel noise might fail the forwarding strategy. Accordingly, the probability for the receiver D_i to observe the signal of node S_j to be f is $1 - p_e$, or d with probability p_e. While the observed signal of node S_j from D_i is d with S_j dropping the packet.

Since most of the devices of wireless communication network are assumed to rely on limited computation resources, players are reluctant to packet forwarding for other players. Such behaviors can be a serious

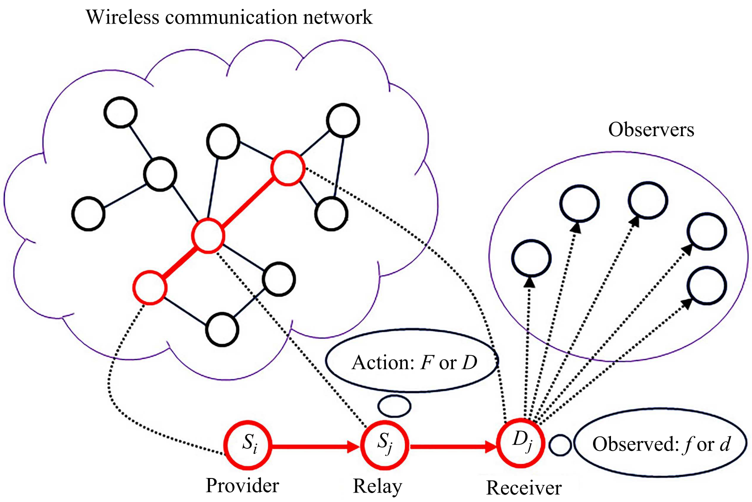

Figure 1 (Color online) The illustration of the packet forwarding on distributed wireless network. At each time slot, the packets of the provider will be forwarded or dropped by the relay to the receiver depending on the providers reputation under channel loss probability p_e. Afterwards, reputation of the relay will get updated with the observed receiver signal under reputation updating error μ. Then, the relays reputation is spread through a channel of noisy gossip from the receiver and the observers to the whole network. After the interaction, each participant returns to the network with probability ω, or leaves the network with probability $1-\omega$ without return.

threat to the efficient communication of a network [3, 4], which is captured by the story of the Prisoner's Dilemma (PD) [34, 35]. After one round of communication, the gain of a provider is b given that the packets are successfully delivered to its destination with probability $1-p_e$. Meanwhile, the forwarding effort of relay players will increase a certain cost, denoted as c. Thus, the payoff matrix M between F and D is described as

$$\begin{array}{cc} & \begin{array}{cc} F & \qquad D \end{array} \\ \begin{array}{c} F \\ D \end{array} & \begin{pmatrix} b(1-p_e)-c & -c \\ b(1-p_e) & 0 \end{pmatrix}. \end{array} \tag{1}$$

The payoff structure yields a unique Nash equilibrium (NE). When the external stimulation is absent, no matter what strategy of the opponent is, the best choice of each player is not to forward the packets, i.e., both players will not forward the packet for others.

2.2 Indirect reciprocity mechanism

Most of the studies concerning the wireless cooperation based on the mechanism of direct reciprocity, in which the action of a relay towards a provider is decided only by the recorded ways whose opponents treats to him/her. It is obvious that under such a scenario, all relays get no motive to transform packets since their behaviors will not be assessed by other players except their opponents. To stimulate the cooperation among the selfish players in the wireless communication network described in Figure 1, we use the mechanism of indirect reciprocity for cooperation enforcement [36]. The essential concept of indirect reciprocity is: "I help you and somebody else helps me". In the context of indirect reciprocity, an action between one player and its opponent is observed by a subset of the population. Therefore, a crucial problem in the indirect reciprocity mechanism is to establish the reputation system, which is based on the history of players' actions.

In this paper, we consider a binary reputation system, in which a binary reputation (good (G) and bad (B)) is endowed to each player. As shown in Figure 1, the packets of the provider will be forwarded or dropped by the relay to the receiver depending on the providers' reputation with probability p_e. After the interaction, reputation of the relay will be updated with the observed receiver signal, while the reputation of the provider remains the same. In some cases, the traffic monitoring mechanism of reputation collection

Table 1 Notations in the paper

Notation	Physical meaning
$\mathbf{S}$	The strategy set of packet forwarding
F	The strategy forwarding
D	The strategy dropping
Θ	The set of observed signal
f	The observations of packet forwarding
d	The observations of packet dropping
b	Gain of a player as a provider
c	Cost of a player as a relay
$i(j)$	The index of player
$I(J)$	The reputation of player
$G(B)$	The player with good (bad) reputation
p_e	The channel loss probability
μ	The reputation updating error
ω	The discounting factor of the furture
$\mathbb{A}$	The action rule of relay
$s_G(s_B)$	The strategy for a relay towards a good (bad) provider
Q	The social norm
q_{iJ}	The reputation assigned to a relay who has taken the strategy i towards a provider whose reputation is J
q	The social resolution of players
$R'(R,X)$	The new reputation of a relay who takes the strategy X towards a provider with reputation R
N	The set of all game players
x_g	The frequency of good players
$W_{I,J}$	Maximum payoff of a player can gain from this interaction to future, I is currently reputation and J is a reputation of opponent
λ	Advantage of being a good player
m	Index of action rule (m=1,2,3)
x_m	Frequency of strategy m
η	Scale factor of dynamical evolution
P_m	Expected payoff of action rule m
$\bar{P}$	Average payoff of three action rules

cannot be reliable, resulting in false reports occasionally [27]. Therefore, the reputation system must be fault tolerant. In our model, the parameter μ ($0 \leqslant \mu \leqslant 1/2$) is introduced to capture this uncertainty, i.e., a correct reputation is assigned with probability $1-\mu$; an incorrect reputation is assigned with probability μ. Finally, the relay's reputation spreads through a channel of noisy gossip from the receiver and the observers to the whole network.

Besides, due to mobility and changes of environment, the players will periodically update their partners to achieve a better performance, i.e., the number of interactions between any pair of players is finite. After the interactions, each participant returns to the network with probability ω, or leaves the network with probability $1-\omega$ without return. Here, the parameter ω reflects a discounting factor of the future. In exchange for each player leaving the network, a new player enters with either a good or bad reputation depending on the proportion of reputation in the current network. Thus, the total network size remains unchanged.

2.3 Social norms

A matrix Q used for updating the reputation of players is called as a social norm. In the social norm, each element q_{iJ} stands for the reputation assigned to a relay taking strategy i toward a provider with reputation J. Thus, the norm is second-order, which means that all matters are the pairs between the action of the relay and the reputation of the provider. That is to say, even when players are equipped with higher cognitive capacities, they rely on a simple norm as a key for evolutionary success. Without loss of generality, we assume all players in the population sharing the same norm. To simplify the analysis, we only consider the case of two strategies of relay ($i = \{F, D\}$) and binary reputation of provider ($J = \{G, B\}$). Certainly, the results can be extended to the case of multi-strategies and multi-reputations.

It is intuitional that forwarding packets to the provider with good reputation or dropping packets to the provider with bad reputation set up a good reputation, and will be rewarded by others. In this paper, we adopt a binary reputation system which is called stern-judging [39], which is in accord with the above intuition. Under the social norm of stern-judging, helping a good player and refusing to help a bad one lead to a good reputation, whereas refusing to help a good individual and helping a bad one leads to a bad reputation. Therefore, the reputation updating of relay is given as the following rule:

$$\begin{array}{ccc} & G & B \\ F & G & B \\ D & B & G. \end{array} \tag{2}$$

Here, a new reputation $R'(R, X)$ (abbreviated representation R', $R' \in \{G, B\}$) of a relay with the strategy X ($X \in \{F, D\}$) towards a provider with reputation R ($R \in \{G, B\}$) will be assigned. The success and simplicity of stern-judging social norm relies on never being morally dubious: for each type of encounter, there are one G reputation and one B reputation. Moreover, it is always possible for anyone to be promoted to the best standard possible in a single reputation. Therefore, the stern-judging can promote cooperation effectively under indirect reciprocity [40].

To distinguish the ability of players between the good and bad players, a parameter $q = 1 - 2\mu$ of social resolution is introduced. Denoted y_G and y_B as the fraction of players with a good and bad reputation in the absence of errors, respectively. And, denoted y_g and y_b as the perceived fraction of good and bad players in the presence of errors, respectively. Clearly, $y_G + y_B = 1$ and $y_g + y_b = 1$. Thus, we have

$$\begin{cases} y_g = (1-\mu)y_G + \mu y_B, \\ y_b = \mu y_G + (1-\mu)y_B. \end{cases} \tag{3}$$

From (3), we obtain $y_g - y_b = q(y_G - y_B)$. And the perceived difference between good and bad players, $y_g - y_b$, equals q times of the actual difference $y_G - y_B$. Therefore, q can be interpreted as the social resolution between good and bad players. If $q = 0$, there is no distinction between good and bad players; while the social resolution is perfect when $q = 1$.

2.4 Action rules

With the social norm of stern-judging, each player will choose a new action $\mathbb{A}$ according to the provider's reputation. Formally, an action rule, $\mathbb{A}$, is an action table of the transmitter, i.e., $\mathbb{A}$ means that the relay takes strategy $s(G)$ for a good provider and takes strategy $s(B)$ for a bad one. Each of $s(G)$ and $s(B)$ can be either F or D. Thus, the action rule, $\mathbb{A}$, has $2^2 = 4$ possible elements: $\mathbb{A} = \{s_G s_B | FF, FD, DF, DD\}$. As an example, FF means that the transmitter adopts strategy F towards a good provider, and also adopts strategy F towards a bad provider. Since strategy DF is illogical in practice, we only consider three of these actions in this paper, i.e., FF, FD and DD.

3 Game analysis of optimal action rule

To explore the cooperative and effective character for proposed action rules, we want to find an action rule that is cooperative and optimal under the above second-order social norm. For a game of N players, the strategy profile $\boldsymbol{s} = (s_1, s_2, \ldots, s_N)$ describes how different players choose their strategies. s_i denotes the strategy of player i which is chosen from the action rule set $\mathbb{A}$. Denote the complemental-strategy profile of player i as $s_{-i} = (s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_N)$, so that $\boldsymbol{s} = (s_i, s_{-i})$. Let $U_i(s) = U_i(s_i, s_{-i})$ be the player's utility function of strategy s_i. The NE describes a "steady state" strategy profile of a game in which no player can be better off by deviating from it, given the strategies taken by all other players remain fixed. That is, the best strategy s_i^* is the best response of player i to s_{-i} iff for all other $s_i \in \mathbb{A}$, $U_i(s_i^*, s_{-i}) \geqslant U_i(s_i, s_{-i})$. More formally, we have the definition.

Definition 1. A NE is a strategy profile that corresponds to the mutual best response. A strategy profile $s^* = (s_1^*, s_2^*, \ldots, s_N^*)$ of a game is an NE if $U_i(s_i^*, s_{-i}^*) \geqslant U_i(s_i, s_{-i}^*)$, $\forall\ i \in N$ and $\forall\ s_i \neq s_i^*$.

Besides, we consider the stability of the optimal action. As an example, the action rule DD is evolutionarily stable, since the best choice of one player is to deny forwarding and choosing strategy D has no cost against the population with DD action rule. However, DD is not a cooperative action rule as all game interactions are defection. In this case, the network performance of packet forwarding is damaged terrifically.

Generally, the stability of the optimal action rule can be depicted by the concept of evolutionarily stable strategy [41], 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 this paper, we adopt another substitutional evolutionary stability definition.

Definition 2. An action is Cooperative Evolutionarily Stable (CES) iff it satisfies the following two criteria [42]:

(I) Cooperativity (CO). More than half of the interactions in game are cooperative.

(II) Evolutionary stability (ES). Under the social norm of stern-judging, action φ is evolutionarily stable against any other action ϕ ($\varphi \neq \phi$).

3.1 Stationary reputation distribution

Denote the fraction of good players as x_g. For convenience, we define δ_G and δ_B as

$$\delta_G = \begin{cases} 1, & \text{if} \quad R' = G, \\ 0, & \text{if} \quad R' = B, \end{cases}$$

and

$$\delta_B = \begin{cases} 1, & \text{if} \quad R' = G, \\ 0, & \text{if} \quad R' = B, \end{cases}$$

where δ_G (δ_B) equals one if a player adopting strategy s_G (s_B) obtains a good reputation after playing with a good (bad) provider. Otherwise, δ_G (δ_B) should be zero.

At each time interval, $2\Delta t$ proportion of the players is selected to participate the game, half of which are providers, and the others are relays. Since only the relay's reputation will be updated after the interaction, we only need to consider the reputation evolution of relays. Among the relays, there are $x_g\Delta t$ proportion with good reputation, $(1-x_g)\Delta t$ proportion with bad reputation. After one interaction, the proportion of relays with good reputation becomes $x_g[\mu(1-\delta_G)+(1-\mu)\delta_G]\Delta t + (1-x_g)[\mu(1-\delta_B)+(1-\mu)\delta_B]\Delta t$. Each participant stays in the population with probability ω, or leaves the population with probability $1-\omega$. In exchange for players who leave the population, $(1-\omega)\Delta t$ proportion of the players enters the population. Assume that the proportion of new players with good reputation is the same as the current reputation proportion, which equals x_g. Thus, after Δt time interaction, the proportion of relay with good reputation becomes $\omega[(1-2\mu)\{x_g\delta_G + (1-x_g)\delta_B\} + \mu]\Delta t + (1-\omega)x_g\Delta t$, which leads to the differential equation governing the evolution of players with good reputation as

$$\frac{\mathrm{d}x_g}{\mathrm{d}t} = \omega\left[(1-2\mu)\left\{x_g\delta_G + (1-x_g)\delta_B\right\} + \mu - x_g\right]. \tag{4}$$

Therefore, the stationary reputation distribution x_g^* of good reputation is the solution to $\frac{\mathrm{d}x_g}{\mathrm{d}t}=0$, which is calculated as

$$x_g^* = \begin{cases} 1-\mu, & \text{if} \quad (\delta_G, \delta_B) = (1,1), \\ 1/2, & \text{if} \quad (\delta_G, \delta_B) = (1,0) \text{ or } (0,1), \\ \mu, & \text{if} \quad (\delta_G, \delta_B) = (0,0). \end{cases} \tag{5}$$

If $x_g^* = 1/2$, then a player meets a good provider with probability 1/2 and meets a bad one with probability 1/2. For seeking the condition of cooperation in **CO** criterion, the action rule must take cooperation with both types of providers, i.e., $s_G = s_B = F$. This implies that the action rule is FF, which is obviously susceptible to the invasion by DD action rule. Thus, it does not satisfy the **ES** criterion.

To satisfy the **CO** criterion, two possibilities are allowable: (i) $(\delta_G, \delta_B) = (1,1)$ and $s_G = F$, or (ii) $(\delta_G, \delta_B) = (0,0)$ and $s_B = F$. Since there is a complete symmetry between two labels, "good" and "bad", we can swap them without changing anything. Note that the argument above is theoretically equivalent to restrict the action rules cooperating with good players, i.e., $s_G = F$. Thus, we only consider the situation of $(\delta_G, \delta_B) = (1,1)$, $s_G = F$.

In summary, we have obtained the stationary reputation distribution $x_g^* = 1-\mu$, which satisfies the **CO** criterion. At this point, $s_G = F$, $R'(G,F) = G$, and $R'(B, s_B) = G$, which coincides with the social norm of stern-judging.

3.2 Optimal action rule with dynamic programming

Let us explore the optimal action under the stationary reputation distribution $x_g^* = 1-\mu$. Consider a monomorphic population which satisfies $s(G) = F$ and the social norm of stern-judging. According to the payoff matrix M, the cost of forwarding packet is constant c. Consequently, the cost of the relay with reputation I ($I \in \{G,B\}$) adopting strategy X towards a provider with reputation J ($J \in \{G,B\}$) is given by

$$\alpha(X) = d_i \cdot c, \tag{6}$$

where parameter $d_i = 1$ when $X = F$, and $d_i = 0$ when $X = D$. Similarly, when the packets are successfully delivered to the destination, a provider gets a benefit $b(1-p_e)$, and the gain of the provider with reputation J is

$$\beta(X) = d_i \cdot b(1-p_e). \tag{7}$$

Denote $W_{I,J}$ as the maximum payoff of a player can get from this interaction to future, where I is the current reputation and J is a reputation of being matched player.

If the player acts as a relay with reputation I and is matched to a player with reputation J, then the long-term expected payoff that he/she can obtain by taking strategy X is

$$f_1(X) = -\alpha(X) + \omega \cdot W_{(1-\mu)R' + \mu\overline{R'}, (1-\mu)G+\mu B}. \tag{8}$$

Conveniently, we introduce the notation, $\overline{G} = B$, $\overline{B} = G$, and $W_{y_1G+y_2B, z_1G+z_2B} = y_1z_1W_{G,G} + y_1z_2W_{G,B} + y_2z_1W_{B,G} + y_2z_2W_{B,B}$. In (8), the first term $-\alpha(X)$ stands for the immediate cost of relay incurred by taking strategy X, and the second term $W_{(1-\mu)R'+\mu\overline{R'},(1-\mu)G+\mu B}$ represents the gains of relay that he/she can obtain in the future with a discounting factor ω. After the interactions, the relay stays in the population with probability ω. Besides, after taking strategy X, the reputation of relay changes to R' with probability $1-\mu$, and to $\overline{R'}$ with probability μ, which gives rise to $(1-\mu)R' + \mu\overline{R'}$ in (8). His/her opponent is randomly sampled from the population in the next round. From (5), we know that the fraction of good players equals to $x_g^* = 1-\mu$ in the population. Thus, the reputation of his/her next opponent is good with probability $1-\mu$ and bad with probability μ, resulting in $(1-\mu)G + \mu B$ in (8). The dynamic programming of this process is shown in Figure 2.

On the other hand, if the player with reputation I serves as a provider, then he/she can get the long-term expected payoff as

$$f_2 = \beta(X) + \omega \cdot W_{I,(1-\mu)G+\mu B}. \tag{9}$$

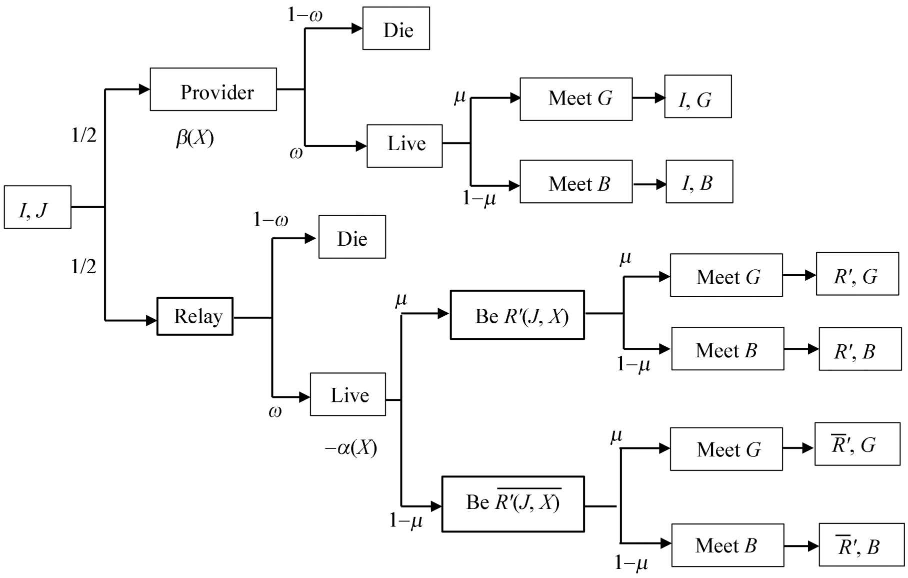

Figure 2 The illustration of dynamic programming.

Here the first term $\beta(X)$ reflects the immediate gain of provider towards the relay taking strategy X, and the second term $W_{I,(1-\mu)G+\mu B}$ is the benefit of provider in the future with a discounting factor ω. As a provider, the reputation will be unchanged after the interaction. Besides, the opponent of a provider is randomly sampled from the population in the next round, leading to the reputation of opponent be good with probability $1-\mu$, and bad with probability μ.

With each interaction, the player acts as a provider or as a relay with an equal probability, which triggers the Bellman equation of $W_{I,J}$ as follows:

$$W_{I,J} = \max_{X=F,D} \left(\frac{1}{2} f_1(X) + \frac{1}{2} f_2 \right). \tag{10}$$

Note that the solution to maximize (10) is independent of reputation I. Therefore, the best response action to $\tilde{X}$, denoted by $\tilde{X}^*$, can be evaluated by

$$\begin{aligned} \tilde{X}^*(J) &= \arg\max_{X=F,D} W_{I,J} \\ &= \arg\max_{X=F,D} \left[-\frac{1}{2}\alpha(X) + \frac{1}{2}\omega \cdot W_{(1-\mu)R'+\mu\overline{R'},(1-\mu)G+\mu B} \right]. \end{aligned} \tag{11}$$

To solve (11), define $\lambda \equiv W_{G,(1-\mu)G+\mu B} - W_{B,(1-\mu)G+\mu B}$, which reflects the "advantage of being a good player". According to (10), it is easily calculated that

$$\begin{aligned} \lambda &= W_{G,(1-\mu)G+\mu B} - W_{B,(1-\mu)G+\mu B} \\ &= \frac{1}{2}\Big\{ G[S(G)] - G[S(B)] + \omega (W_{G,(1-\mu)G+\mu B} - W_{B,(1-\mu)G+\mu B}) \Big\} \\ &= \frac{1}{2}\Big\{ G[S(G)] - G[S(B)] + \omega\lambda \Big\}. \end{aligned}$$

Thus, we obtain

$$\lambda = \frac{G[S(G)] - G[S(B)]}{2-\omega}. \tag{12}$$

Rewrite (11) as (13), where $\varphi_G[G]=1$, $\varphi_G[B]=0$.

$$\tilde{X}^*(J) = \begin{cases} \arg\max_{X=F,D} \left[\frac{1}{2}\{-\alpha(X) + \omega \cdot W_{(1-\mu)G+\mu B,(1-\mu)G+\mu B}\} \right], & \text{if } R'(J,X)=G, \\ \arg\max_{X=F,D} \left[\frac{1}{2}\{-\alpha(X) + \omega \cdot W_{(1-\mu)B+\mu B,(1-\mu)G+\mu G}\} \right], & \text{if } R'(J,X)=B, \end{cases}$$

$$
\begin{aligned}
&= \begin{cases} \arg\max\limits_{X=F,D} \left[\dfrac{1}{2}\{-\alpha(X) + \omega \cdot W_{(1-2\mu)G-(1-2\mu)B+\mu G+(1-\mu)B,(1-\mu)G+\mu B}\}\right], & \text{if } R'(J,X)=G, \\ \arg\max\limits_{X=F,D} \left[\dfrac{1}{2}\{-\alpha(X) + \omega \cdot W_{(1-\mu)B+\mu B,(1-\mu)G+\mu G}\}\right], & \text{if } R'(J,X)=B, \end{cases} \\
&= \arg\max_{X=F,D} \left[\frac{1}{2}\{-\alpha(X) + \omega\{(1-2\mu)(W_{G,(1-\mu)G+\mu B} - W_{B,(1-\mu)G+\mu B})\varphi_G[R'(J,X)]\right. \\
&\quad \left. + W_{(1-\mu)B+\mu B,(1-\mu)G+\mu G}\}\}\right] \\
&= \arg\max_{X=F,D} \left[\frac{1}{2}\{-\alpha(X) + \omega q\lambda\varphi_G[R'(J,X)]\} + \frac{\omega}{2} W_{(1-\mu)B+\mu B,(1-\mu)G+\mu G}\right] \\
&= \arg\max_{X=F,D} \left[\frac{1}{2}\{-\alpha(X) + \omega q\lambda\varphi_G[R'(J,X)]\}\right]. \qquad (13)
\end{aligned}
$$

According to (13), we know that the maximization problem is deduced to find the action X to maximize $-\alpha(X) + \omega q\lambda\varphi_G[R'(J,X)]$. The first term, $-\alpha(X)$, represents the immediate cost of strategy X. The second term, $\omega q\lambda\varphi_G[R'(J,X)]$, stands for the future benefit through becoming a good player towards strategy X, which is $\lambda\varphi_G[R'(J,X)]$, multiplied by discounting factor ω and social resolution q. Hence, we are able to derive $\tilde{X}^*(G)$ and $\tilde{X}^*(B)$. Based on the criterions of **CO**, it is easy to distinguish that action DD does not satisfy the **CO**. Therefore, we only discuss the actions of FD and FF.

Theorem 1. Under the social norm of stern-judging, when the benefit-to-cost ratio $\frac{b}{c} > \frac{2-\omega}{\omega q(1-p_e)}$, action FD is an optimal action.

Proof. For each player with action FD, there are two stability conditions to satisfy: (I) the optimal strategy should be F when he/she meets the opponents with good reputation; (II) the optimal strategy should be D when he/she meets the opponents with bad reputation. Specifically, condition (I) requires

$$
\begin{aligned}
\tilde{X}^*(G) &= \arg\max_{X=F,D} \left[\frac{1}{2}\{-\alpha(X) + \omega q\lambda\varphi_G[R'(G,X)]\}\right] \\
&= F,
\end{aligned} \qquad (14)
$$

i.e.,

$$
\frac{1}{2}\{-\alpha(F) + \omega q\lambda\varphi_G[R'(G,F)]\} > \frac{1}{2}\{-\alpha(D) + \omega q\lambda\varphi_G[R'(G,D)]\}. \qquad (15)
$$

Since $R'(G,F) = G$, and $R'(G,D) = B$, we simplify (16) to

$$
\begin{aligned}
& -\alpha(F) + \omega q\lambda > -\alpha(D) \\
\Rightarrow \quad & -c + \frac{\omega q b(1-p_e)}{2-\omega} > 0 \Rightarrow \frac{b}{c} > \frac{2-\omega}{\omega q(1-p_e)}.
\end{aligned} \qquad (16)
$$

On the other hand, condition (II) requires

$$
\begin{aligned}
\tilde{X}^*(B) &= \arg\max_{X=F,D} \left[\frac{1}{2}\{-\alpha(X) + \omega q\lambda\varphi_G[R'(B,X)]\}\right] \\
&= D,
\end{aligned} \qquad (17)
$$

i.e.,

$$
\frac{1}{2}\{-\alpha(D) + \omega q\lambda\varphi_G[R'(B,D)]\} > \frac{1}{2}\{-\alpha(F) + \omega q\lambda\varphi_G[R'(B,F)]\}. \qquad (18)
$$

For $R'(B,D) = G$, which leads to $\varphi_G[R'(B,D)] = 1$, and $-\alpha(D) = 0$, (18) is always true.

In summary, if the benefit-to-cost ratio, b/c, satisfies that $\frac{b}{c} > \frac{2-\omega}{\omega q(1-p_e)}$, the action FD is the best response to itself, i.e., FD is an optimal action.

Remark 1. Note that FF never becomes an optimal action under the social norm of stern-judging, for strategy F will never be the best response to itself when he/she meets the opponents with bad reputation. That is, the condition $\frac{1}{2}\{-\alpha(F) + \omega q\lambda\varphi_G[R'(B,F)]\} > \frac{1}{2}\{-\alpha(D) + \omega q\lambda\varphi_G[R'(B,F)]\}$ will never be satisfied.

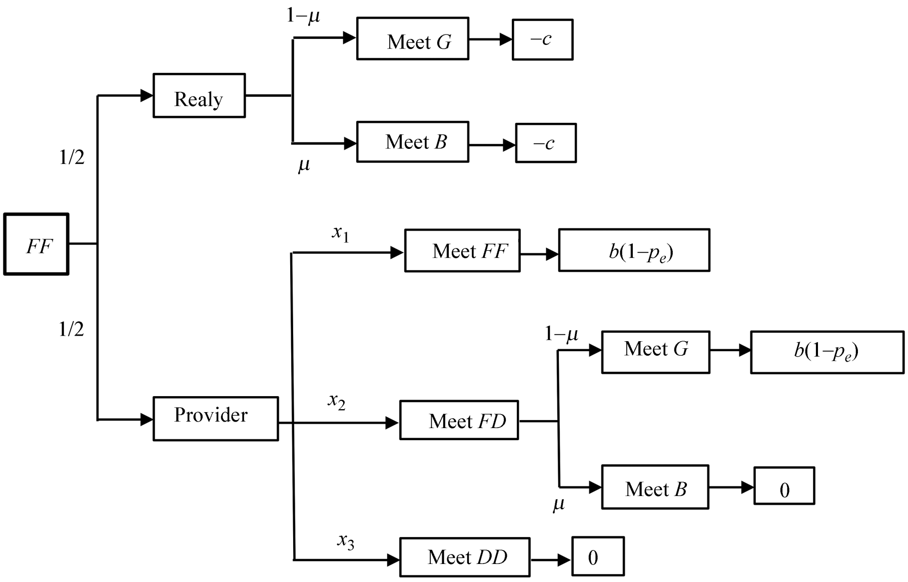

Figure 3 The illustration of expected payoffs' calculation to the actions FF.

3.3 Evolutionary stability of optimal action rule

Since the uncertainty of the system and/or the noisy parameters, players may take a non-optimal action rule during the forwarding process. Therefore, it is indispensable to take the perturbation effect into account, which motivates us to evaluate the evolutionary stability of the optimal action rule.

Denote x_1, x_2, and x_3 as the frequencies of strategy FF, FD, and DD, respectively. Then, we have $x_1 + x_2 + x_3 = 1$. Given the stationary reputation distribution $x_g^* = 1 - \mu$, we calculate the expected payoff of a strategy. For a FF player, he/she acts as a relay with probability $\frac{1}{2}$, and cooperate with cost c. With probability $\frac{1}{2}$ the player acts as a provider, who meets a FF, FD and DD player with probability x_1, x_2 and x_3, respectively. Thus, the expected gains of the provider are $b(1-p_e)$, $b(1-p_e)(1-\mu)$ and 0, respectively (see Figure 3).

Similarly, we can obtain the gain and cost of FD and DD players, which results in the expected payoffs of actions FF, FD and DD as

$$\begin{cases} P_1 = \dfrac{1}{2}(-c) + \dfrac{1}{2}[b(1-p_e)x_1 + b(1-p_e)(1-\mu)x_2], \\ P_2 = \dfrac{1}{2}(1-\mu)(-c) + \dfrac{1}{2}[b(1-p_e)x_1 + b(1-p_e)(1-\mu)x_2], \\ P_3 = \dfrac{1}{2}(0) + \dfrac{1}{2}[b(1-p_e)x_1 + b(1-p_e)(1-\mu)x_2], \end{cases} \tag{19}$$

where P_1, P_2, and P_3 are the expected payoffs of strategy FF, FD, and DD, respectively.

In the following, we adopt the action spreading algorithm of replicator dynamics to demonstrate the evolution of frequency at Δt time interval [41], which means that the evolution of x_m ($m = 1, 2, 3$) is given by the following equation:

$$\begin{aligned} \Delta x_m =& [\omega a(x) + (1-\omega)b(x)]\Delta t - x_m \Delta t \\ =& \omega[\eta x_m(P_m - \bar{P}) - x_m]\Delta t, \end{aligned} \tag{20}$$

where $x = (x_1, x_2, x_3)^{\mathrm{T}}$, η is a scale factor controlling the speed of the evolution, P_m is the expected payoff of action rule m, and $\bar{P} = \sum_{m=1}^{3} x_m P_m$ is the average payoff of the three actions. Here, the first term $a(x) = \eta x_m(P_m - \bar{P})$ in (20) denotes the frequency variation caused by internal competition, which occurs with probability ω. And the second term $b(x) = x_m$ in (20) denotes the frequency variation caused by the external mobility, which happens with probability $1 - \omega$.

Define $\hat{P}_m = P_m - P_3$, and $\tilde{P} = \sum_{m=1}^{3} x_m \hat{P}_m$. We get the transformed deterministic dynamical evolution of frequency as

$$\begin{cases} \dot{x}_1 = \omega[\eta x_1(\hat{P}_1 - \tilde{P}) - x_1] \\ \quad = \omega[(-c\eta - 1)x_1 + c\eta x_1^2] \\ \qquad + \omega\eta \dfrac{[(1-2\mu)b(1-p_e)+c]x_1x_2 - (1-2\mu)b(1-p_e)x_1x_2(x_1+x_2)}{2 - \frac{1-2\mu}{1-\mu}(x_1+x_2)}, \\ \dot{x}_2 = \omega[\eta x_2(\hat{P}_2 - \tilde{P}) - x_2] \\ \quad = \omega(c\eta x_1x_2 - x_2) \\ \qquad + \omega\eta \dfrac{-cx_2 + [(1-2\mu)b(1-p_e)+c]x_2^2 - (1-2\mu)b(1-p_e)x_2^2(x_1+x_2)}{2 - \frac{1-2\mu}{1-\mu}(x_1+x_2)}. \end{cases} \tag{21}$$

We can investigate the stability of (21) to characterize the evolutionary stability of actions. Intuitively, if $b \gg c$, the potential cooperation gain will be larger than the immediate cooperation cost, i.e., results that each player is inclined to cooperate with other players. While $b \ll c$, each 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 value of benefit-to-cost ratio $\frac{b}{c}^*$ such that the optimal action rule is evolutionarily stable if $\frac{b}{c} > \frac{b}{c}^*$, which coincides with the result in Theorem 1.

Assume that parameter $\omega = 0.7$, the channel loss probability $p_e = 0.01$ and the reputation updating error $\mu = 0.01$. Consider a fixed-size population with $N = 300$, in which the players share the fixed social norm defined in (2). Initially, each player uses one of the possible action rules FF, FD and DD with randomly choosing reputation G or B. Before any one elementary step of action updating, each player has exactly 30 interactions with their opponents, among which every player acts as a provider and a relay 15 times on average. After all 30 interactions have taken place, all participants return to the population with probability ω or leave the population with probability $1 - \omega$. For each player who leaves, a new player enters the population to keep the total population size constant.

According to Theorem 1, we obtain the critical value of the benefit-to-cost ratio $\frac{b}{c}^* \approx 1.914$. The corresponding evolutionary stability of action FD is shown in Figure 4 by setting $b = 2.4$ and $c = 1$, in which the parameters of packet process satisfy the condition of Theorem 1, and action FD is the optimal action. In this case, the benefit-to-cost ratio $\frac{b}{c} = 2.4 > 1.914$, action FD will spread over the whole population when the initial frequency of action FD is chosen a suitable value (see Figure 4(a)). However, when decreasing the initial frequency of action FD, the percentage of the population with the optimal action FD no longer converges to 1, as shown in Figure 4(b). In fact, there is a feasible domain of the initial action distributions. On the other hand, when $b = 1.6$ and $c = 1$, the corresponding evolutionary stability of action FD is shown in Figure 5. In this case, the benefit-to-cost ratio $\frac{b}{c} = 1.6 < 1.914$, action FD will never converge to 1 regardless of the initial frequency of action FD (see Figure 5(a) and (b)).

4 Numerical results

4.1 Simulation setup

Consider an unreliable wireless network with 300 nodes randomly scattered in an area of 3600 m$\times$3600 m. Each node has a physical communication range of 100 m. The underlying MAC protocol is IEEE 802.11g CSMA/CA with a bandwidth of 4 Mbps. The unit slot time of the communication is 20 μs. SIFS and DIFS are 10 μs and 28 μs, respectively. The size of each data packet is 64 bytes and the ACK packet size is 32 bytes. At a constant bit rate of 2 packets per second, one datum is generated. Assumed that there is only one data session at each slot time.

In each time slot, one of the nodes in the communication network is picked as the relay to transmit the data packets for the provider. Comparatively, we define the “full cooperation” action rule, in which every node will unconditionally forward packets independent from other nodes’ reputation. Though the

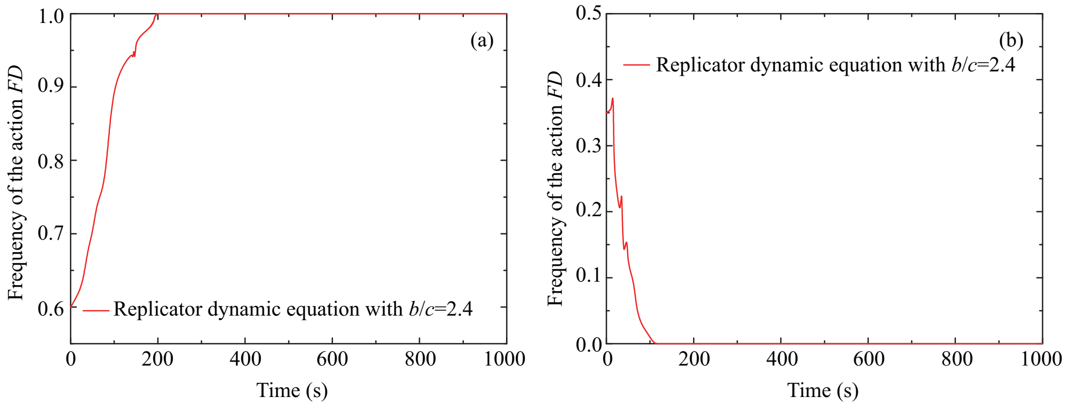

Figure 4 (Color online) The evolutionary stability of the action FD with $b = 2.4$ and $c = 1$. (a) The initial frequency of the action FD is setting to 0.6; (b) the initial frequency of the action FD is setting to 0.35.

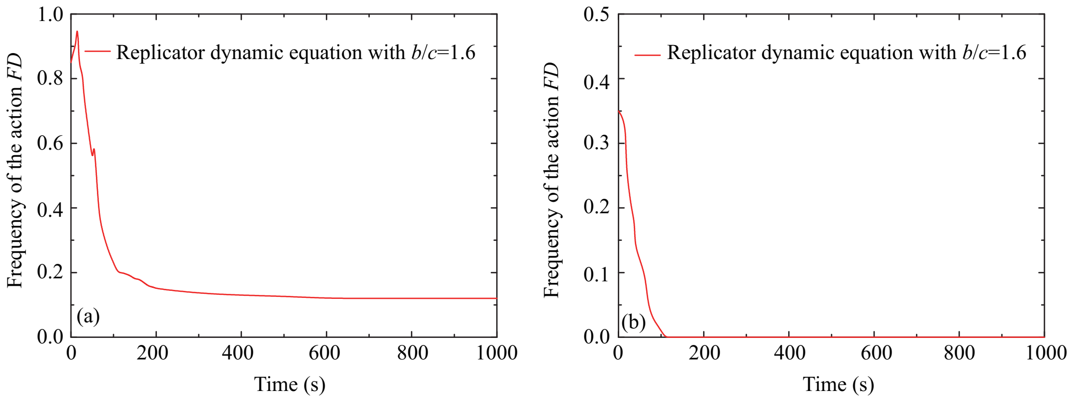

Figure 5 (Color online) The evolutionary stability of the action FD with $b = 1.6$ and $c = 1$. (a) The initial frequency of the action FD is setting to 0.85; (b) the initial frequency of the action FD is setting to 0.35.

"full cooperation" action rule is irrealizable in such wireless communication networks, it can act as a loose performance upper bound of other strategies.

4.2 Performance evaluation

We first aim at the one-hop packet forwarding scenario in unreliable communication networks, in which the two-player packet forwarding game is based on the mechanism of indirect reciprocity. Assume that the gain per unit is 2 and the cost per unit is 1, i.e., $b = 2$ and $c = 1$. The parameter $\omega = 0.7$.

Figures 6 and 7 present the average node payoff of the optimal action with the indirect reciprocity mechanism for different p_e and μ compared with the "full cooperation" action. In Figure 6, we set the reputation updating error as $\mu = 0.01$, while in Figure 7, we set the channel loss probability as $p_e = 0.01$. Figures 6 and 7 provide three insights: (i) When the channel becomes more unreliable (large p_e) and the reputation becomes more undistinguishable (large μ), the average node payoff drops; (ii) Comparing to the unconditionally cooperative payoff, the optimal action based on indirect reciprocity can stimulate cooperation with only a small performance loss, in which the parameters of packet process satisfy Theorem 1; (iii) As the increase of p_e and μ, the performance loss of the optimal action compared to the unconditionally cooperative action becomes large.

To further evaluate the performance of the optimal action, we consider the network performance for multiple hops packet forwarding. Assume every hop on a data route is independent. When all the generated packets are successfully delivered from a source to the destination, the state is denoted as "1". As shown in Figures 8 and 9, with a small channel loss probability ($p_e = 0.01$) and a small reputation updating error ($\mu = 0.01$), the optimal action based on the indirect reciprocity mechanism achieves almost

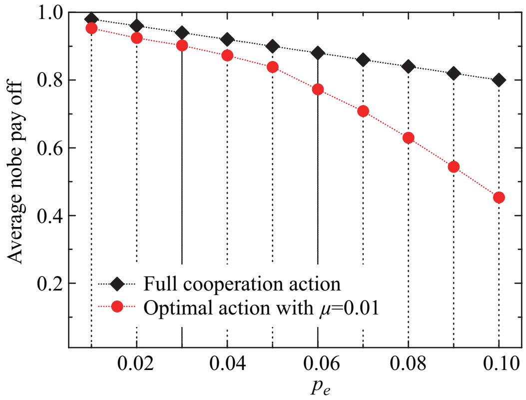

Figure 6 (Color online) The average node payoffs of the full cooperation action and the optimal action with different channel loss probability when $\mu = 0.01$.

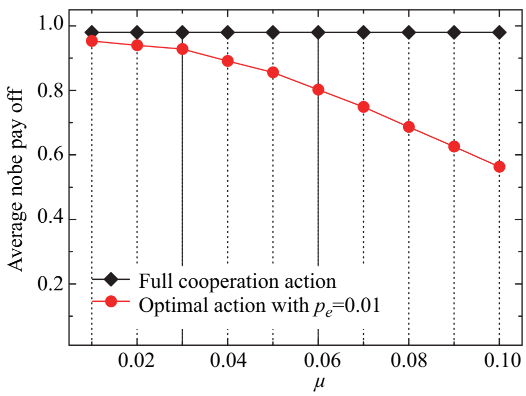

Figure 7 (Color online) The average node payoffs of the full cooperation action and the optimal action with different reputation updating error when $p_e = 0.01$.

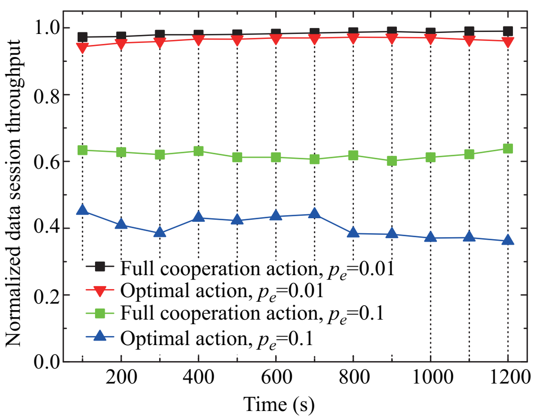

Figure 8 (Color online) The normalized data session throughput of the full cooperation action and optimal action with different channel loss probability when $\mu = 0.01$.

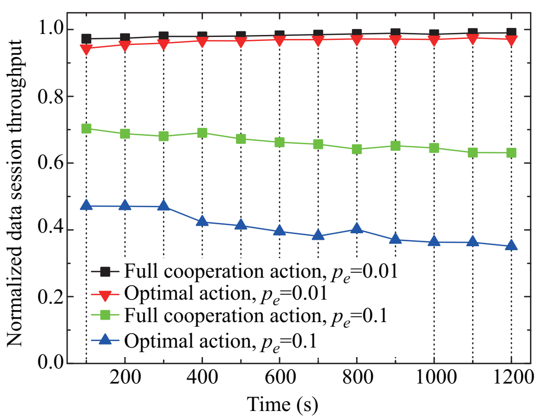

Figure 9 (Color online) The normalized data session throughput of the full cooperation action and optimal action with different reputation updating error when $p_e = 0.01$.

the same throughput as that of the fully cooperative action. On the other hand, when the parameters of packet process do not satisfy in Theorem 1 ($p_e = 0.1$ and $\mu = 0.1$, respectively), the throughput difference between the unconditional cooperation state and the optimal action becomes larger.

Besides, we illustrate the normalized data session throughput with different effects of hop count, channel unreliability and reputation updating error. From Figures 10 and 11, we know that the throughput drops when the channel becomes more unreliable, the reputation becomes more undistinguishable and the hop count increases. These observations also suggest that the optimal action yields the throughput performance is very close to that of the situation when all the nodes are unconditionally cooperative, in which the parameters of packet process satisfy Theorem 1.

5 Conclusion

Wireless communications require that all nodes in a network should complete a task cooperatively, in which the limited battery resources and the lack of a single authority are likely to trigger a noncooperative behavior for the packet forwarding. In this paper, we have modeled the unreliable packet forwarding among participants as an indirect reciprocity game with finite number of interactions between any pairs of players. We have theoretically analyzed the optimal action rule with the method of dynamic programming,

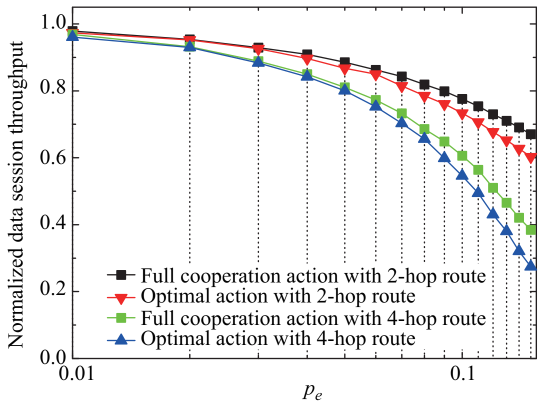

Figure 10 (Color online) The normalized data session throughput of hop count with different channel loss probability when $\mu = 0.01$.

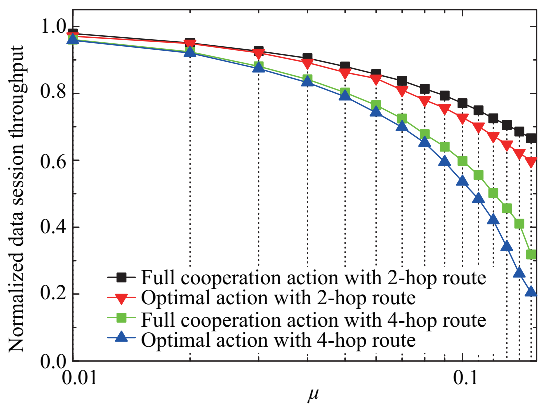

Figure 11 (Color online) The normalized data session throughput of hop count with different reputation updating error when $p_e = 0.01$.

and derived the approximate threshold of benefit-to-cost ratio to achieve the optimal action rule. Besides, we have investigated the evolutionary stability of the optimal action with the replicator dynamics. When the benefit-to-cost ratio exceeds the critical value $\frac{b}{c}^*$, the optimal action rule FD is evolutionarily stable with the suitable initial frequency of the action FD, i.e., the optimal action rule will quickly spread over the whole population. However, when the benefit-to-cost ratio is smaller than the critical value $\frac{b}{c}^*$, the optimal action rule FD will never be evolutionarily stable. Furthermore, we have illustrated that the performance of the wireless communications is observably improved by the optimal action in the packet forwarding process with small channel unreliability and reputation updating error.

Acknowledgements This work was supported by National Science Fund for Distinguished Young Scholar of China (Grant No. 61425019), Key Projects of National Natural Science Foundation of China (Grant No. 71731004), National Natural Science Foundation of China (Grant Nos. 61403059, 61503342, 11572288, 61672468), and Zhejiang Provincial Natural Science Foundation of China (Grant Nos. LY15F020013, LY16F030002).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Song L, Niyato D, Han Z, et al. Wireless Device-to-Device Communications and Networks. Cambridge: Cambridge University Press, 2015. 1–32

2 Liu K J R, Sadek A K, Su W F, et al. Cooperative Communications and Networking. Cambridge: Cambridge University Press, 2009. 1–45

3 Xu Q, Zheng R, Saad W, et al. Device fingerprinting in wireless networks: challenges and opportunities. IEEE Commun Surv Tut, 2016, 18: 94–104

4 Ku M L, Li W, Chen Y, et al. On energy harvesting gain and diversity analysis in cooperative communications. IEEE J Sel Areas Commun, 2015, 33: 2641–2657

5 Jaramillo J J, Srikant R. DARWIN: distributed and adaptive reputation mechanism for wireless ad-hoc networks. In: Proceedings of 13th Annual ACM International Conference on Mobile Computing and Networking, Montreal, 2007. 87–97

6 Buttyán L, Hubaux J P. Stimulating cooperation in self-organizing ad hoc networks. ACM/Kluwer Mobile Netw Appl, 2003, 8: 579–592

7 Crowcroft J, Gibbens R, Kelly F, et al. Modelling incentives for collaboration in mobile ad hoc networks. Perform Eval, 2004, 57: 427–439

8 Zhong S, Chen J, Yang Y R. Sprite: a simple, cheat-proof, credit-based system for mobile ad-hoc networks. In: Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications, San Francisco, 2003. 1987–1997

9 Janzadeh H, Fayazbakhsh K, Dehghan M, et al. A secure credit-based cooperation stimulating mechanism for MANETs using hash chains. Future Generat Comput Syst, 2009, 25: 926–934
10 Wang L, Liao X K, Xue J L, et al. Enhancement of cooperation between file systems and applications-VFS extensions for optimized performance. Sci China Inf Sci, 2015, 58: 092104
11 Michiardi P, Molva R. CORE: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks. In: Advanced Communications and Multimedia Security. Boston: Springer, 2002. 107–121
12 He Q, Wu D, Khosla P. SORI: a secure and objective reputation-based incentive scheme for ad-hoc networks. In: Proceedings of IEEE Wireless Communications and Networking Conference, Atlanta, 2004. 825–830
13 Balakrishnan K, Deng J, Varshney V K. TWOACK: preventing selfishness in mobile ad hoc netwotks. In: Proceedings of IEEE Wireless Communications and Networking Conference, New orleans, 2005. 2137–2142
14 Refaei M T, DaSilva L A, Eltoweissy M, et al. Adaptation of reputation management systems to dynamic network conditions in ad hoc networks. IEEE Trans Comput, 2010, 59: 707–719
15 Mejia M, Peña N, Muñoz J L, et al. DECADE: distributed emergent cooperation through adaptive evolution in mobile ad hoc networks. Ad Hoc Netw, 2012, 10: 1379–1398
16 Akkarajitsakul K, Hossain E, Niyato D. Coalition-based cooperative packet delivery under uncertainty: a dynamic Bayesian coalitional game. IEEE Trans Mob Comput, 2013, 12: 371–385
17 Duarte P B F, Fadlullah Md Z, Vasilakos A V, et al. On the partially overlapped channel assignment on wireless mesh network backbone: a game theoretic approach. IEEE J Sel Areas Comm, 2012, 30: 119–127
18 Yang Y H, Chen Y, Jiang C X, et al. Wireless network association game with data-driven statistical modeling. IEEE Trans Wirel Commun, 2016, 15: 512–524
19 Xiao Y, Niyato D, Chen K C, et al. Enhance device-to-device communication with social-awareness: a belief-based stable marriage game framework. IEEE Wirel Commun, 2016, 23: 36–44
20 Xu C B, Zhao Y L, Zhang J F. Decision-implementation complexity of cooperative game systems. Sci China Inf Sci, 2017, 60: 112201
21 Srinivasan V, Nuggehalli P, Chiasserini C F, et al. Cooperation in wireless ad hoc networks. In: Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications, San Francisco, 2003. 808–817
22 Félegyházi M, Hubaux J P, Buttyan L. Nash equilibria of packet forwarding strategies in wireless ad hoc networks. IEEE Trans Mobile Comput, 2006, 5: 463–476
23 Yu W, Liu K J R. Game theoretic analysis of cooperation stimulation and security in antonomous mobile ad hoc networks. IEEE Trans Mobile Comput, 2007, 6: 507–521
24 Akkarajitsakul K, Hossain E, Niyato D. Cooperative packet delivery in hybrid wireless mobile networks: a coalitional game approach. IEEE Trans Mobile Comput, 2013, 12: 840–854
25 Zhang H Q, Niyato D, Song L Y, et al. Zero-determinant strategy for resource sharing in wireless cooperations. IEEE Trans Wirel Commun, 2016, 15: 2179–2192
26 Li Z, Shen H. Game-theoretic analysis of cooperation incentive strategies in mobile ad hoc networks. IEEE Trans Mobile Comput, 2012, 11: 1287–1303
27 Seredynski M, Bouvry P. Evolutionary game theoretical analysis of reputation-based packet forwarding in civilian mobile ad hoc networks. In: Proceedings of IEEE International Symposium on Parallel & Distributed Processing, Rome, 2009. 1–8
28 Song L, Niyato D, Han Z, et al. Game-theoretic resource allocation methods for device-to-device (D2D) communication. IEEE Wirel Commun, 2014, 21: 136–144
29 Jiang C X, Chen Y, Liu K J R. Multi-channel sensing and access game: bayesian social learning with negative network externality. IEEE Trans Wirel Commun, 2014, 13: 2176–2188
30 Xiao Y, Niyato D, Han Z. Dynamic energy trading for energy harvesting communication networks: a stochastic energy trading game. IEEE J Sel Areas Commun, 2015, 33: 2718–2734
31 Akkarajitsakul K, Hossain E, Niyato D, et al. Game theoretic approaches for multiple access in wireless networks: a survey. IEEE Commun Surv Tut, 2011, 13: 372–395
32 Khan M A, Tembine H, Vasilakos A V. Evolutionary coalitional games: design and challenges in wireless networks. IEEE Wirel Commun, 2012, 19: 50–56
33 Hoang D T, Lu X, Niyato D, et al. Applications of repeated games in wireless networks: a survey. IEEE Commun Surv Tut, 2015, 17: 2102–2135
34 Ji Z, Yu W, Liu K J R. A belief evaluation framework in autonomous MANETs under noisy and imperfect observation: vulnerability analysis and cooperation enforcement. IEEE Trans Mobile Comput, 2010, 9: 1242–1254

35 Wang W, Chatterjee M, Kwiat K. Cooperation in wireless networks with unreliable channels. IEEE Trans Commun, 2011, 59: 2808–2817

36 Nowak M A, Sigmund K. Evolution of indirect reciprocity. Nature, 2005, 437: 1291–1298

37 Chen Y, Liu K J R. Indirect reciprocity game modelling for cooperation stimulation in cognitive networks. IEEE Trans Commun, 2011, 59: 159–168

38 Tang C B, Li A, Li X. When reputation enforces evolutionary cooperation in unreliable MANETs. IEEE Trans Cybern, 2015, 45: 2190–2201

39 Tanabe S, Suzuki H, Masuda N. Indirect reciprocity with trinary reputations. J Theor Biol, 2013, 317: 338–347

40 Pacheco J M, Santos F C, Chalub F A C C. Stern-Judging: a simple, successful norm which promotes cooperation under indirect reciprocity. PLoS Comput Biol, 2006, 2: 1634–1638

41 Hofbauer J, Sigmund K. Evolutionary Games and Population Dynamics. Cambridge: Cambridge University Press, 1998. 92–95

42 Ohtsuki H, Iwasa Y, Nowak M A. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. Nature, 2009, 457: 79–82