

A Mathematical Framework for Analyzing Adaptive Incentive Protocols in P2P Networks

Bridge Qiao, Zhao[†] John C.S. Lui[†] Dah-Ming Chiu[‡]

Abstract—In P2P networks, incentive protocol is used to encourage cooperation among end nodes so as to deliver a scalable and robust service. However, the design and analysis of incentive protocols have been ad hoc and heuristic at best. The objective of this paper is to provide a simple, yet general framework to analyze and design incentive protocols. We consider a class of incentive protocols which can learn and adapt to other end nodes' strategies. Based on our analytical framework, one can evaluate the expected performance gain, and more importantly, the system robustness of a given incentive protocol. To illustrate the framework, we present two adaptive learning models and three incentive policies and show the conditions in which the P2P networks may collapse and the conditions in which the P2P networks can guarantee a high degree of cooperation. We also show the connection between evaluating incentive protocol and evolutionary game theory so one can easily identify robustness characteristics of a given policy. Using our framework, one can also gain the understanding on the price of altruism and system stability. This framework can help protocol designers to quickly evaluate the correctness of their incentive policies and to explore the proper incentive mechanism to achieve cooperation.

I. Introduction

Incentive protocols play a crucial role in many networking environments. For example, consider a wireless mesh network wherein a node needs other nodes to assist in its packet forwarding. Since packet forwarding increases the energy consumption, therefore unless there is some built-in incentive mechanism, rational nodes will choose not to perform any packet forwarding. If enough wireless nodes behave in this selfish manner, the underlying wireless network will be partitioned and nodes will be unreachable. Another example is in P2P file sharing protocols where nodes rely on other nodes to perform uploading service. This mutual uploading service offloads the server and allows the system to scale. Again, without the incentive mechanism to encourage nodes to perform uploading service, the server will be overwhelmed and nodes may never be able to finish the file downloading process. The above examples illustrate one important point: *embedding incentive protocols to encourage cooperation among nodes is crucial so that the overall system performance can be improved*. However, the design and analysis of incentive protocols have been ad-hoc and heuristic at best.

It is important to point out that there is a natural tendency that nodes will not used a fixed strategy but instead, adapt from the environment. Authors in [2], [6] point out that there are

benefits for nodes to *learn* and *adapt* from neighboring nodes in a P2P network, e.g., nodes will provide uploading service to other nodes, but when they discover that there are other nodes that can free ride on their altruism and get good downloading service, then these nodes may choose to change its strategy and adapt a more selfish strategy. Therefore, to fully understand a given incentive protocol, we need to have a systematic and formal methodology to model the dynamic learning and adaptive behavior of cooperating/competing nodes, and to evaluate the robustness and effectiveness of the underlying incentive protocol. The aim of this paper is to provide a general analytical framework to design and analyze a large class of adaptive incentive protocols for P2P networks. Our contributions are:

- We propose a general analytical framework to evaluate the performance of adaptive incentive protocols in P2P networks.
- To illustrate the utility of our mathematical framework, two different learning models and derive their performance measures and robustness conditions.
- We carry out performance evaluation of the above three incentive protocols and show their performance gain and formally state that under what conditions the P2P network will be robust and under what conditions the P2P network may collapse.
- We show the connection between evaluating the robustness of incentive protocols and *evolutionary game theory*. We illustrate how one can map linear incentive policies to two-player games, and to give an efficient technique to identify the robustness characteristics of linear incentive policies.
- We quantify the performance and robustness of the system when there is cost associated with realizing an incentive protocol.
- We show that there is a tradeoff between altruism and system robustness and justify why one may want to limit the degree of altruism so as to encourage cooperation.

The outline of this paper is as follows. We present a general performance model of incentive policies for P2P systems in Section II. In Section III, we present two learning models for strategy adaptation. In Section IV, we present two incentive policies and any incentive policy in a generalized incentive policy class and show how to use our the framework to analyze these protocols. In Section V, we derive the performance measures such as system gain and the expected gain for individual strategy, as well as the robustness conditions for the given incentive policies. Results of performance evaluation

[†]Department of Computer Science & Engineering, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong; {qzhao, cslui}@cse.cuhk.edu.hk.

[‡]Information Engineering Department, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong; dmchiu@ie.cuhk.edu.hk.

are presented in Section VI on the three incentive policies. In Section VII, we provide the connection between our framework and evolutionary game theory. We use a simple game-theoretical technique to identify robustness characteristics of linear policies. In Section VIII and IX, we present the price of altruism and how it relates to the network stability. Related work is given in Section X and Section XI concludes.

II. An Incentive Model for P2P Networks

In this section, we present a general mathematical model to study some incentive protocols in P2P networks. Given an incentive protocol, we show how to use this framework to evaluate (a) its evolution and robustness, and (b) its performance measures such as expected service received and service contributed of the given incentive protocol. For the incentive protocols we study, we have the following assumptions:

- **Finite strategies:** we consider incentive policies which have finite strategies. Given an incentive policy $\mathcal{P} = \{s_1, s_2, \dots, s_n\}$ where s_i is the i^{th} strategy in \mathcal{P} . Peers can choose to use any $s_i \in \mathcal{P}$. A peer using strategy s_i is called a type i peer. Potential strategies can range from altruism (e.g., willing to contribute) to egoism (e.g., refusing to contribute).
- **Service model:** we model a P2P network as a discrete time system. At the beginning of each time slot, each peer randomly selects another peer in the system and requests for service¹. A selected peer may choose to serve the request based on its current strategy. Let $g_i(j)$ denote the probability that a type i peer will provide service to a type j peer. Accordingly, one can define an $n \times n$ generosity matrix \mathbf{G} with $G_{ij} = g_i(j)$. At each time slot, a peer obtains $\alpha > 0$ points when it receives a service from another peer, while loses β points when it provides a service to another peer. Without loss of generality, we normalize $\beta = 1$.

Let $x_i(t)$ denote the fraction of type i peers in the system at time t . We define $E[R_i(t)]$ as the expected services that a type i peer can receive in one time slot, $E[R_i(t)]$ can be expressed as:

$$E[R_i(t)] = \sum_{j=1}^n x_j(t)g_j(i) \quad \text{for } i = 1, \dots, n. \quad (1)$$

Let $E[S_i(t)]$ denote the expected number of service units provided by a type i peer at time t , and this quantity can be derived as follows. Assume that at time t , there are $N(t)$ number of peers in the P2P network. Consider a tagged type i peer and denote \mathcal{N} as the set representing the other $N(t) - 1$ peers in the P2P network. Let $k \in \mathcal{N}$, then the probability that this tagged type i peer will provide service to this peer k is \mathcal{L} , which can be expressed as:

$$\begin{aligned} \mathcal{L} &= \text{Prob}[k \text{ selects this type } i \text{ peer}] \times \\ &\quad \text{Prob}[\text{type } i \text{ peer will serve } k] \\ &= \frac{1}{N(t)-1} \left(\sum_{j=1}^n \text{Prob}[k \text{ is of type } j]g_i(j) \right), \end{aligned}$$

¹this assumption is also made in several other P2P studies, e.g. [10], [18].

and

$$\text{Prob}[k \text{ is of type } j] = \begin{cases} \frac{x_j(t)N(t)}{N(t)-1} & \text{for } j \neq i, \\ \frac{x_i(t)N(t)-1}{N(t)-1} & \text{for } j = i. \end{cases}$$

Since $|\mathcal{N}| = N(t) - 1$, the expected number of service units provided by this tagged type i peer in one time slot is $E[S_i(t)] = [N(t) - 1]\mathcal{L}$. Combining the above expressions and by assuming that the number of peers $N(t)$ in a P2P system is relatively large, we have

$$E[S_i(t)] \approx \sum_{j=1}^n x_j(t)g_i(j) \quad \text{for } i = 1, 2, \dots, n. \quad (2)$$

Define $\tilde{p}_i(t)$ be the random variable denoting the performance gain of type i peer at time slot t and denote its expectation by $\bar{P}_i(t)$. Because a peer receives α points for each service it receives and loses $\beta = 1$ point for each service it provides, the expected performance gain per slot at time t is $\bar{P}_i(t)$, which can be expressed as:

$$\bar{P}_i(t) = \alpha E[R_i(t)] - E[S_i(t)] \quad i = 1, 2, \dots, n. \quad (3)$$

The above n equations can be expressed in matrix form and $\bar{\mathbf{P}}(t)$, the expected gain per time slot for the P2P network at time t is

$$\bar{\mathbf{P}}(t) = \sum_{i=1}^n x_i(t)\bar{P}_i(t) = (\alpha - 1)\mathbf{x}^T(t)\mathbf{G}\mathbf{x}(t), \quad (4)$$

where $\mathbf{x}(t)$ is a column vector of $[x_1(t), \dots, x_n(t)]$.

In summary, to evaluate the performance and robustness of a given incentive protocol, one has to first ‘‘determine’’ all values in matrix \mathbf{G} (i.e., all $g_i(j)$ for a given incentive policy). In Section IV, we will illustrate how to use this analytical framework to study several incentive protocols.

Note that for an incentive policy, it may include strategies such as serving other peers upon request, or refusing to serve upon request. A peer uses strategy s_i may choose to adapt to a new strategy s_j when this peer discovers that strategy s_j will provide a better performance gain, or $\bar{P}_j(t) > \bar{P}_i(t)$. How to discover and adapt to that strategy s_j with a higher gain than s_i can be modeled by the underlying learning process, which we will describe next.

III. Learning Models for P2P Networks

Learning and adapting to the environment are natural behavior of a rational individual. Peers may get information from external environment, and adjust their strategy so as to obtain better performance. This process can be spontaneous and gradual. The rate of adaptation depends on the *truthfulness* of information received by peers and the *sensitivity* of peers toward this information. Since peers learn and adapt naturally, one can consider adding a layer of software so as to guide peers to learn so the system will operate at a desirable point. In short, learning activities do exist and it is worthwhile to promote in P2P systems. As a result, such learning behavior has a significant impact on the evolution and dynamics of the system. In this paper, we will present two learning models and will study how these learning models can affect the dynamics of incentive policies in P2P networks.

A. Current-best Learning Model (CBLM)

One learning abstraction is that peers discover the best strategy at the current time and adapt to it. We call this the “*current-best learning model*” and it can be described as follows: at the end of a time slot, a peer can choose to switch (or adapt) to another strategy $s' \in \mathcal{P}$ with probability γ_a which we called the adapting rate. To decide which strategy to switch to, a peer needs to “*learn*” from other peers. Let $s_h(t)$ be the strategy that has the highest expected gain among all $s \in \mathcal{P}$ at the end of the time slot t (or $h \in \operatorname{argmax}_j \{\bar{P}_j(t)\}$). Then a peer using strategy s_i will switch to strategy s_h at time slot $t + 1$ with probability $\gamma_s(\bar{P}_h(t) - \bar{P}_i(t))$, where γ_s is the sensitivity to the performance gap. We call the product $\gamma = \gamma_a \gamma_s$ as the learning rate. Under this learning model, peers will adapt to the current best strategy, and the probability of adaptation to this current winning strategy is *proportional* to the performance gap of the expected gain.

Note that there are many ways to realize this learning abstraction, and one approach is the following: a P2P system can distributively elect a leader and all peers report their current performance gain to this leader. The leader is responsible for computing the average gain for all strategies. Peers can query this leader about the current best strategy $s_h(t)$. Note that when γ is sufficiently small, the leader will not be overwhelmed by the query workload.

Let us illustrate how the current-best learning model can affect the system dynamics of a given incentive policy. One can express $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]$, where $x_i(t)$ is the fraction of peers using strategy s_i at time t , using the following difference equations:

$$x_i(t+1) = x_i(t) - \gamma x_i(t) (\bar{P}_h(t) - \bar{P}_i(t)), \quad i \neq h,$$

and for $s_h(t)$, the strategy that has the highest expected gain, we have

$$x_h(t+1) = x_h(t) + \gamma \sum_{i=1, i \neq h}^n x_i(t) (\bar{P}_h(t) - \bar{P}_i(t)).$$

We can transform the above difference equations to a continuous model² as:

$$\begin{aligned} \frac{dx_h(t)}{dt} &= \gamma \sum_{i \neq h} x_i(t) (\bar{P}_h(t) - \bar{P}_i(t)) \\ &= \gamma \left(\bar{P}_h(t) - \sum_{i=1}^n x_i(t) \bar{P}_i(t) \right) = \gamma (\bar{P}_h(t) - \bar{P}(t)), \end{aligned} \quad (5)$$

$$\frac{dx_i(t)}{dt} = -\gamma x_i(t) (\bar{P}_h(t) - \bar{P}_i(t)), \quad i \neq h. \quad (6)$$

In summary, given an incentive policy, we first need to determine all entries in the corresponding generosity matrix $\mathbf{G} = \{g_i(j)\}$, then we can evaluate the dynamics of the system using the above differential equations.

²Informally, the transformation can be carried out by assuming that (1) the peer-request process is a Poisson process with rate equal to 1, (2) the number of adapting events is a Poisson process with rate γ_a . (3) In each event, the sensitivity to performance gap is γ_s , and we have $\gamma = \gamma_a \gamma_s$.

B. Opportunistic Learning Model (OLM)

The current-best learning abstraction requires each peer to update its type and its gain to a data collecting node (or leader), and this node needs to compute the average gain for all peers in a P2P network. Therefore, the computational requirement may be high and the data collecting node needs to be resourceful or else one will face the scalability problem. Here, we propose another learning abstraction which we called the “*opportunistic learning model*”. This learning abstraction can be described as follows: at the end of a time slot, each peer randomly chooses another peer in the network as its *teacher* with probability γ_a . If the teacher is of a different type and has a better performance gain, the peer adapts to the teacher’s strategy with sensitivity γ_s to their performance gap. One interesting note is that this learning abstraction does not require frequent access to shared global history and can be realized in a fully distributed fashion.

Let us illustrate how this learning abstraction can affect the system dynamics of a given incentive policy. Let $f_i(\tilde{p}_i(t))$ be the probability density function (pdf) of random variable \tilde{p}_i , $\rho_{ij}(t)$ be the rate that type i peers will switch to type j peers at time t , then:

$$\begin{aligned} \Pr[\text{type } i \text{ peer switches to type } j \text{ when } j \text{ is a teacher}] &= \\ &= \int_{\tilde{p}_i(t) < \tilde{p}_j(t)} \gamma_s (\tilde{p}_j(t) - \tilde{p}_i(t)) f_i(\tilde{p}_i(t)) f_j(\tilde{p}_j(t)) d\tilde{p}_i(t) d\tilde{p}_j(t). \end{aligned}$$

Since the fraction of type i peers is $x_i(t)$, the teacher will be of type j with probability $x_j(t)$ and adapting rate is γ_a , thus the rate that type i peers switches to type j peers is:

$$\begin{aligned} \rho_{ij}(t) &= \gamma x_i(t) x_j(t) \times \\ &= \int_{\tilde{p}_i(t) < \tilde{p}_j(t)} (\tilde{p}_j(t) - \tilde{p}_i(t)) f_i(\tilde{p}_i(t)) f_j(\tilde{p}_j(t)) d\tilde{p}_i(t) d\tilde{p}_j(t). \end{aligned}$$

Similarly, the rate that type j peers switches to type i :

$$\begin{aligned} \rho_{ji}(t) &= \gamma x_j(t) x_i(t) \times \\ &= \int_{\tilde{p}_j(t) < \tilde{p}_i(t)} (\tilde{p}_i(t) - \tilde{p}_j(t)) f_j(\tilde{p}_j(t)) f_i(\tilde{p}_i(t)) d\tilde{p}_j(t) d\tilde{p}_i(t). \end{aligned}$$

Therefore, the total rate of population flow from type i to type j is $\delta_{ij}(t)$, where:

$$\begin{aligned} \delta_{ij}(t) &= \rho_{ij}(t) - \rho_{ji}(t) = \gamma x_i(t) x_j(t) E[\tilde{p}_j(t) - \tilde{p}_i(t)] \\ &= \gamma x_i(t) x_j(t) [\bar{P}_j(t) - \bar{P}_i(t)]. \end{aligned}$$

The total in-flow to type j , which we denote as \hat{x}_j , is:

$$\frac{dx_{ij}(t)}{dt} = \sum_{i=1}^n \delta_{ij} = \gamma x_j(t) [\bar{P}_j(t) - \bar{P}(t)] \quad j = 1, \dots, n. \quad (7)$$

In summary, given an incentive policy \mathcal{P} , we first need to derive all entries in the corresponding generosity matrix \mathbf{G} , then we can use Equations (5)-(7) to study the system dynamics and determine the robustness condition. In the following section, we present several incentive policies to illustrate this analytical framework.

IV. Incentive Policies for P2P Networks

Let us now present several incentive protocols and illustrate how to use the mathematical framework to analyze their evolution and determine their robustness conditions and performance measures. For simplicity of illustration, we classify peers in a P2P network according to their *behavior* upon receiving a request [2]. These include:

- 1) **Cooperator:** a peer has an altruistic behavior and it always serve other peers independent of whether other peers provide service or not.
- 2) **Defector:** a peer has a selfish behavior and it always refuse to serve any request from other peers in the network.
- 3) **Reciprocator:** a peer has a reciprocative behavior when it serves the requester according to the requester's service strategy, e.g., if the requester is a cooperator (defector), this peer will serve (deny) the request. The rationale of this type of behavior is to make a fair exchange of service.

One interesting question is how to design a proper incentive policy so as to keep the P2P networks as scalable and robust as possible. Let us proceed to illustrate using our methodology.

A. Mirror Incentive Policy \mathcal{P}_{mirror}

The first policy we consider is called the *mirror incentive policy* \mathcal{P}_{mirror} . For this policy, when a reciprocative peer receives a request for service, this peer infers (e.g., similar to the tit-for-tac operation in BT) the requester's reputation, and it will only provide service with the *same probability* as this requester serves other peers in the system. For example, suppose peer k received 100 requests from other peers and served 60 of them, then when peer k requests a reciprocator for service, peer k will only get the service with probability of 0.6. Hence, if the requester is a cooperator (defector or reciprocator), the receiving peer will act exactly like a cooperator (defector or reciprocator) to the requester. This is the reason why we coin it the mirror incentive strategy.

Under the \mathcal{P}_{mirror} policy, there are three pure strategies: (1) pure cooperation, or s_1 , (2) mirror reciprocation, or s_2 , and (3) pure defection, or s_3 . To evaluate the performance and robustness of \mathcal{P}_{mirror} , we need to derive all entries in the generosity matrix \mathbf{G} , or $g_i(j)$, which is the probability that a peer of type i will serve a peer of type j . Based on the definition of the mirror policy, it is easy to see that $g_1(j) = 1$ and $g_3(j) = 0$ for $j \in \{1, 2, 3\}$, $g_2(1) = 1$ and $g_2(3) = 0$. The remaining issue is the expression of $g_2(2)$, and it can be derived as follows:

$$\begin{aligned} g_2(2) &= \text{Prob}[\text{a reciprocator will grant a request}] \\ &= \sum_{i=1}^3 \text{Prob}[\text{the requester is of type } i] \times \\ &\quad \text{Prob}[\text{granting the request} | \text{type } i \text{ requests}] \\ &= x_1(t)g_2(1) + x_2(t)g_2(2) + x_3(t)g_2(3) \\ &= x_1(t) + x_2(t)g_2(2). \end{aligned}$$

Solving the above equation, we have

$$g_2(2) = \frac{x_1(t)}{1 - x_2(t)} = \frac{x_1(t)}{x_1(t) + x_3(t)}. \quad (8)$$

In other words, the probability for a reciprocator to serve another reciprocator is close to 1 when the fraction of defector is close to zero, but as the probability will approach zero when the fraction of defector increases.

B. Proportional Incentive Policy \mathcal{P}_{prop}

We consider another incentive policy which we called the proportional incentive policy \mathcal{P}_{prop} . This incentive policy was proposed in [2] in which results were obtained via simulation only. Reciprocative strategy s_2 in \mathcal{P}_{prop} is defined as follows: peers using s_2 serve the requester (say type j) with the probability equal to the requester's contribution/consumption ratio, or $E[S_j]/E[R_j]$. When the ratio is larger than one, the probability to serve the requester is equal to one. By definition, if the requester is a cooperator, its ratio can be larger than one. Thus, we have $g_2(1) = 1$. If the requester is a defector, its ratio is zero, hence $g_2(3) = 0$. The remaining issue is the expression for $g_2(2)$, which can be derived as follows:

$$\begin{aligned} E[R_2(t)] &= x_1(t)g_1(2) + x_2(t)g_2(2) + x_3(t)g_3(2) \\ &= x_1(t) + x_2(t)g_2(2), \\ E[S_2(t)] &= x_1(t)g_2(1) + x_2(t)g_2(2) + x_3(t)g_2(3) \\ &= x_1(t) + x_2(t)g_2(2). \end{aligned}$$

Since $E[R_2(t)] = E[S_2(t)]$, we have $g_2(2) = 1$, or a reciprocator will always serve another reciprocator. The other values of $g_i(j)$ are $g_1(j) = 1$ and $g_3(j) = 0$ for $j \in \{1, 2, 3\}$.

Comparing to the mirror strategy, the proportional strategy takes into account the services consumed by requesters and reciprocators can enjoy more service from other reciprocators. Also, the reciprocator will serve another reciprocator *independent* on the state of the system.

C. Linear Incentive Policy Class \mathcal{C}_{LIP}

The proportional incentive policy belongs to a "class" of incentive policies which we called the *linear incentive policy class*, \mathcal{C}_{LIP} . Any policy in \mathcal{C}_{LIP} has a constant generosity matrix $\mathbf{G} = [G_{ij}]$ where $G_{ij} = g_i(j)$. In here, constant implies that $g_i(j)$ is independent of any $x_i(t)$. It is easy to see that the performance of each strategy is a linear function of $x_i(t)$ for a linear policy.

To implement a policy in \mathcal{C}_{LIP} , one can first design a classifier for reciprocative peers to infer the types of requesters. For example, suppose there are three strategies available: cooperation, reciprocation and defection. We can design the following classifier: it visits the shared history, and identifies those peers who never contribute as defectors. Those who serve the defectors are cooperators, and the rest are reciprocators. With such a classifier, a linear strategy can serve cooperators, reciprocators and defectors with different probabilities p_c , p_r , p_d as specified by the protocol designer. The generosity matrix \mathbf{G} is:

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 \\ p_c & p_r & p_d \\ 0 & 0 & 0 \end{bmatrix} \quad (9)$$

It is easy to see that the proportional incentive policy belongs to the linear incentive policy class ($\mathcal{P}_{prop} \in \mathcal{C}_{LIP}$) because $p_c = 1, p_r = 1$ and $p_d = 0$ while the mirror incentive policy is not ($\mathcal{P}_{mirror} \notin \mathcal{C}_{LIP}$) because its $g_2(2)$ depends on $x_1(t)$ and $x_3(t)$.

V. Performance and Robustness of Incentive Policies

In this section, we analyze and compare the performance and robustness of the three incentive policies described in the previous section. Informally, an incentive protocol of a P2P system is robust when the system will finally stay at a high contribution level (e.g., most peers are cooperators or reciprocators) and the P2P network is immune to system perturbation such as peer arrivals or departures.

A. Robustness Analysis of Mirror Incentive Policy using the current-best learning method

We first consider the mirror incentive policy \mathcal{P}_{mirror} using the current-best learning model (CBLM). Given the derivation of $g_i(j)$ of \mathcal{P}_{mirror} in Section IV, we substitute them into Eq. (3) and (4) to obtain $\bar{P}_i(t)$, the expected gain of using strategy s_i for $i = \{1, 2, 3\}$, as well as $\bar{P}(t)$, the expected gain of the system. These performance measures are:

$$\bar{P}_1(t) = \alpha(x_1(t) + x_2(t)) - 1, \quad (10)$$

$$\bar{P}_2(t) = (\alpha - 1) \frac{x_1(t)}{1 - x_2(t)}, \quad (11)$$

$$\bar{P}_3(t) = \alpha x_1(t), \quad (12)$$

$$\bar{P}(t) = (\alpha - 1) \frac{x_1(t)}{1 - x_2(t)}. \quad (13)$$

Let us consider their respective differences:

$$\bar{P}_3(t) - \bar{P}_1(t) = 1 - \alpha x_2(t),$$

$$\bar{P}_3(t) - \bar{P}_2(t) = \frac{x_1(t)(1 - \alpha x_2(t))}{1 - x_2(t)},$$

$$\bar{P}_2(t) - \bar{P}_1(t) = \frac{(1 - \alpha x_2(t))(1 - x_1(t) - x_2(t))}{1 - x_2(t)}.$$

Based on the above expressions, we have the following important observations:

- **Case A:** when $x_2(t) > 1/\alpha$, we have $\bar{P}_1(t) > \bar{P}_2(t) > \bar{P}_3(t)$, or cooperators always enjoy the best performance. Therefore defectors and reciprocative peers will continue to adapt their strategies to the cooperative strategy. Therefore, $x_2(t)$ and $x_3(t)$ will decrease until $x_2(t) = 1/\alpha$.
- **Case B:** when $x_2(t) = 1/\alpha$, the performance of these three strategies are the same and hence, there will not be any strategy adaptation in the system.
- **Case C:** when $x_2(t) < 1/\alpha$, we have $\bar{P}_3(t) > \bar{P}_2(t) > \bar{P}_1(t)$. In other words, defectors have the best performance and so cooperators and reciprocative peers will adapt their strategies to the defective strategy. Since $x_2(t) < 1/\alpha$ will continue to hold, the population of cooperators and reciprocative peers will keep decreasing until defectors dominate the system (e.g., most peers adapt the defective strategy). At this time, the P2P network *collapses* since no one wants to contribute service to others.

When a P2P system uses the \mathcal{P}_{mirror} incentive protocol under CBLM, the system has two equilibria: B and C respectively. At B , the fraction of reciprocative peers $x_2(t)$ will stay at the level $1/\alpha$. At C , the P2P network will be dominated by defectors. However, point B is *not* a stable equilibrium. Suppose the system is at B with $x_2(t) = 1/\alpha$, and $x_2(t)$ changes a little bit (e.g., due to arrival or departure of peers and these peers are of defective behavior). If the change is positive, the system will go to case A and then drop back to B . But if the change is negative, the P2P network will go to C and never return to B . Since we cannot control the arrival or departure of peers, the system will eventually go to case C and contribution will cease to exist. In summary, the \mathcal{P}_{mirror} incentive policy is *not robust* and eventually all peers will choose the defective strategy.

B. Robustness Analysis of Mirror Incentive Policy using the opportunistic learning method

Now, let us consider using \mathcal{P}_{mirror} under the opportunistic learning model (OLM). We have:

$$\bar{P}_3(t) - \bar{P}(t) = \frac{x_1(t)(1 - \alpha x_2(t))}{1 - x_2(t)},$$

$$\bar{P}_2(t) - \bar{P}(t) = 0,$$

$$\bar{P}_1(t) - \bar{P}(t) = -\frac{(1 - \alpha x_2(t))(1 - x_1(t) - x_2(t))}{1 - x_2(t)}.$$

Based on Eq. (7), the population of the reciprocative peers will not change. The final state of the system depends on the initial reciprocative population, which we classify into the following cases:

- **Case A:** when $x_2 > 1/\alpha$, we have $\bar{P}_3(t) < \bar{P}(t) < \bar{P}_1(t)$. From Eq. (7), we see that defectors will keep decreasing until they become extinct, and the P2P system will only have cooperative and reciprocative peers.
- **Case B:** when $x_2 = 1/\alpha$, we have $\bar{P}_3(t) = \bar{P}(t) = \bar{P}_1(t)$. The system is in an unstable equilibrium and will go to either case A or case C if there is any increase or decrease in $x_2(t)$ due to arrival or departure of reciprocative peers.
- **Case C:** when $x_2 < 1/\alpha$, we have $\bar{P}_3(t) > \bar{P}(t) > \bar{P}_1(t)$. Cooperators will become extinct and the system will eventually *collapse*.

Remark: Based on the above analysis, one can conclude that the incentive policy \mathcal{P}_{mirror} is not robust under the current-best learning model or the opportunistic learning model. This result implies that \mathcal{P}_{mirror} is *not* a proper incentive protocol for P2P networks.

C. Robustness Analysis of Proportional Incentive Policy Using the current-best learning method

For the policy \mathcal{P}_{prop} , we have derived $g_i(j)$ in Section IV. The expected gain of the three strategies and the expected gain of the system are:

$$\bar{P}_1(t) = \alpha(x_1(t) + x_2(t)) - 1, \quad (14)$$

$$\bar{P}_2(t) = (\alpha - 1)(x_1(t) + x_2(t)), \quad (15)$$

$$\bar{P}_3(t) = \alpha x_1(t), \quad (16)$$

$$\bar{P}(t) = (\alpha - 1)(x_1(t) + x_1(t)x_2(t) + x_2^2(t)). \quad (17)$$

The performance difference of these strategies are:

$$\begin{aligned}\bar{P}_3(t) - \bar{P}_2(t) &= x_1(t) - (\alpha - 1)x_2(t), \\ \bar{P}_2(t) - \bar{P}_1(t) &= 1 - x_1(t) - x_2(t) \geq 0, \\ \bar{P}_3(t) - \bar{P}_1(t) &= 1 - \alpha x_2(t).\end{aligned}$$

One important note is that under the proportional incentive policy, reciprocative behavior is *always better* than cooperative behavior, and we have the following cases:

- **Case A:** when $x_2(t) > \frac{1}{\alpha-1}x_1(t)$, we have $\bar{P}_2(t) > \bar{P}_3(t)$. Therefore, the fraction of reciprocative peers $x_2(t)$ will keep increasing until they dominate in the P2P system. In this situation, the expected system gain $\bar{P}(t)$ reaches the maximum at $\alpha - 1$ and the system stabilizes at this point.
- **Case B:** when $x_2(t) = \frac{1}{\alpha-1}x_1(t)$, we have $\bar{P}_2(t) = \bar{P}_3(t) > \bar{P}_1(t)$. Therefore, only cooperative peers will continue to adapt to either strategy s_2 or s_3 . In this case, $x_1(t)$ will decrease but $x_2(t)$ will not. Therefore, the system will eventually go back to case A.
- **Case C:** when $x_2(t) < \frac{1}{\alpha-1}x_1(t)$, defective behavior has the highest performance so peers will adapt to this strategy. However, since s_2 has a higher performance than s_1 , $x_1(t)$ will decrease at a *faster* rate than $x_2(t)$ until the system reaches $x_2(t) = \frac{1}{\alpha-1}x_1(t)$ and the system will go to case B.

In summary, the P2P system is robust and the system will eventually operate at case A, where the fraction of reciprocative peers dominates the system. Moreover, the system achieves the optimal overall performance at this point. It is important for us to point out that this mathematical result agrees with the observation made in [2], which was obtained only via simulation.

D. Robustness Analysis of Proportional Incentive Policy Using the opportunistic learning method

Consider the policy \mathcal{P}_{prop} under the opportunistic learning model (OLM). From Eq.(14)-(17), we have:

$$\begin{aligned}\bar{P}_3(t) - \bar{P}(t) &= x_1(t) - (\alpha - 1)x_2(t)(1 - x_2(t)), \\ \bar{P}_2(t) - \bar{P}(t) &= (\alpha - 1)x_2(t)x_3(t) \geq 0, \\ \bar{P}_1(t) - \bar{P}(t) &= ((\alpha - 1)x_2(t) - 1)x_3(t).\end{aligned}$$

Since $\bar{P}_2(t) - \bar{P}(t) \geq 0$, then based on Eq. (7), the number of reciprocative peers will keep increasing until defectors become extinct, while cooperators will start increasing after $x_2(t)$ becomes larger than $1/(\alpha - 1)$. Finally, the system will achieve the optimal overall performance since the P2P system only has cooperators and reciprocators.

Remark: Based on the above analysis, we show that the proportional strategy \mathcal{P}_{prop} is robust under both the current-best learning and the opportunistic learning methods and this incentive protocol can encourage peers to contribute.

E. Robustness Analysis for Incentive Protocol in the Linear Incentive Class

The idea of reciprocative strategy is to infer the type of requester from the shared history and provide differentiated

service. Although it is feasible in theory, it may be difficult to implement in real systems due to incomplete history and delay in the reputation update. How will this inaccuracy affect the P2P network and what is the design margin? To answer these questions, let us now analyze the robustness of any incentive policy in the generalized linear incentive class \mathcal{C}_{LIP} .

For any incentive policy in \mathcal{C}_{LIP} , we have $g_2(1) = p_c$, $g_2(2) = p_r$, $g_2(3) = p_d$. The performance of these three strategies and the overall system are:

$$\begin{aligned}\bar{P}_1(t) &= \alpha(x_1(t) + p_c x_2(t)) - 1, \\ \bar{P}_2(t) &= \alpha(x_1(t) + p_r x_2(t)) - (p_c x_1(t) + p_r x_2(t) + p_d x_3(t)), \\ \bar{P}_3(t) &= \alpha(x_1(t) + p_d x_2(t)), \\ \bar{P}(t) &= (\alpha - 1)(x_1(t) + p_c x_1(t)x_2(t) + p_r x_2^2(t)).\end{aligned}$$

To analyze the robustness under the current-best learning model, the performance gaps between any two strategies are:

$$\begin{aligned}\bar{P}_3(t) - \bar{P}_2(t) &= p_c x_1(t) + p_d x_3(t) - (\alpha p_r - p_r - \alpha p_d)x_2(t), \\ \bar{P}_2(t) - \bar{P}_1(t) &= 1 - p_c x_1(t) - (p_r - \alpha p_r + \alpha p_c)x_2(t) - p_d x_3(t), \\ \bar{P}_3(t) - \bar{P}_1(t) &= 1 - \alpha(p_c - p_d)x_2(t).\end{aligned}$$

Comparing with the proportional strategy \mathcal{P}_{prop} , to make this policy robust, one sufficient condition is:

$$p_d = 0 \quad (18)$$

$$p_r \geq p_c. \quad (19)$$

Following the performance gap analysis for the proportional incentive policy, it is easy to check that the system is robust for both the current-best and opportunistic learning models when Eq. (18) and Eq. (19) are satisfied.

It is important for us to point out that by Eq. (19), when p_c is small, the system is more likely to be robust. It may seem counter-intuitive since reciprocative peers are punishing the altruistic cooperators. The explanation is that the blind altruism of cooperator helps defectors to survive thus damages the system. To protect reciprocative peers, we need to control the degree of altruism in the network. In later section, we also quantify the impact of altruism in a P2P system.

Let us now restrict our attention to linear strategies with $p_r, p_c > p_d > 0$. The robustness of these policies depends on the initial population $\mathbf{x}(0)$, and this is especially true for the reciprocators. Let us define

$$c_{upper} = \frac{p_c}{(\alpha - 1)(p_r - p_d) + p_c - p_d}, \quad (20)$$

$$c_{lower} = \frac{p_d}{(\alpha - 1)(p_r - p_d)}. \quad (21)$$

It is easy to show that for both learning methods, when $x_2(0) > c_{upper}$, the P2P system will be robust, and when $x_2(0) < c_{lower}$, the P2P system will collapse under the current-best learning method. As for other initial conditions, the robustness depends on the learning mechanism and the fraction of other strategies.

To illustrate the robustness region of the linear incentive policy, let us consider the case that $p_c = 0.9$, $p_r = 1$, $p_d = 0.3$ and the robustness region is depicted in Figure 1. The horizontal axis is the fraction of cooperators while the vertical

axis is the fraction of reciprocators. Since their sum is less than or equal to one, the whole state space is below the diagonal line. There is a boundary curve shown for each learning model and if the initial state is above the curve, the system will be robust for the corresponding learning model. We can observe that the robust region of the opportunistic learning method is *strictly larger* than that of the current-best learning method, and the boundary curves intersect the vertical axis at $x_2 = c_{lower} = 0.0714$.

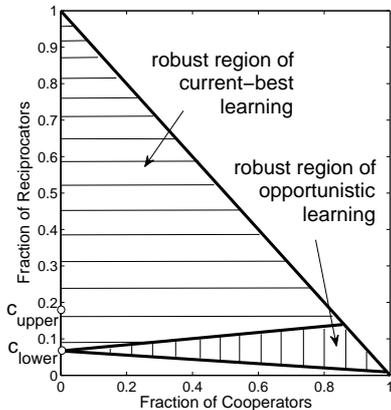


Fig. 1. Robust regions for the two learning methods.

It is interesting to note that the current-best boundary curve rises while opportunistic boundary curve drops as the initial fraction of cooperators increases. This implies that cooperators help in keeping the P2P system robust under the opportunistic learning but harm the P2P system under the current-best learning model. This can be explained as follows. Under the current-best learning, when reciprocators are overwhelmed by defectors, cooperators merely help defectors to increase, which makes the P2P system less robust. However, in the opportunistic learning, cooperators compensate the loss for reciprocators and help them to increase, which makes the system more robust.

VI. Performance Evaluation

We develop a simulator and carry out simulations to compare and validate our mathematical framework. We present the performance evaluation results in this section to show the dynamics and performance of various incentive protocols. Unless we state otherwise, the parameters we use are: shown in Table I. The simulations are carried out via discrete time

N	# of peers	500
α	gain per service provides	7
β	cost per service takes	1
γ	learning rate	0.004

TABLE I
PARAMETERS

slots. In each time slot, peer randomly selects another peer for service. The selected peer decides whether to grant a service or not according to a given incentive strategy (i.e.,

\mathcal{P}_{mirror} , \mathcal{P}_{prop} , or an incentive policy in \mathcal{C}_{LIP}). When all peers finish making their decisions, information is updated, then each peer learns to adapt to a new strategy according to a given probability and learning methods.

Exp. 1: Performance and Robustness of the Mirror Incentive Policy (\mathcal{P}_{mirror}): The population dynamics of \mathcal{P}_{mirror} under different initially conditions $\mathbf{x}(0)$ and different learning models are depicted in Figure 2 and 3. The solid lines are results from the simulation and the dotted lines are results of our mathematical framework. Again, $x_i(t)$ represents the fraction of peers using strategy s_i at time t , with s_1, s_2 and s_3 being cooperative, reciprocative and defective strategy respectively.

Let us consider the dynamics under the current-best learning method. Figure 2 depicts the result. In the left graph, the fraction of reciprocators at $t = 0$ is $x_2(0) < 1/\alpha$. One can observe that the fraction of cooperators, $x_1(t)$, gradually drops to zero and the system collapses. In the right graph, when $x_2(0) > 1/\alpha$, the system seems to be robust at first since there is a significant increase of the fraction of cooperator. At $t = 500^-$, we have $\mathbf{x}(t) = (0.51, 0.16, 0.33)^T$. However, at $t = 500$, we introduce a small disturbance where new cooperators arrive and some reciprocators leave the system, so $\mathbf{x}(500) = (0.54, 0.13, 0.33)^T$. Although there are more cooperative peers, this small disturbance causes the fraction of reciprocative peers to drop below the threshold $1/\alpha$, causing the system to collapse. As we observe at $t = 2000$, there is a significant fraction of defector in the system.

The dynamics of under the opportunistic learning method is depicted in Figure 3. The initial conditions is similar to those in Fig. 2. The left graph shows the system collapses due to small $x_2(0)$, in the right graph, the system survives the disturbance and is robust because the non-decreasing $x_2(t)$ leaves a generous margin for disturbance. The performance gains at $t = 2000$ for Figure 2 and 3 are:

$\mathbf{x}(0)$	P_0	P_1	P_2	P
$(0.4, 0.1, 0.5)^T$	-0.37	0.04	0.04	0.04
$(0.3, 0.2, 0.5)^T$	0.42	0.81	0.90	0.81
$(0.4, 0.1, 0.5)^T$	0.08	0.24	0.28	0.24
$(0.3, 0.2, 0.5)^T$	4.81	4.78	4.59	4.78

In conclusion, we validate our mathematical model and confirm that the \mathcal{P}_{mirror} policy is *not* robust under the current-best learning but may survive when $x_2(0)$ is above a threshold.

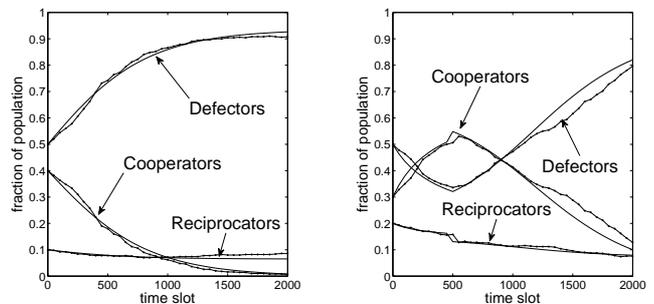


Fig. 2. Evolution of \mathcal{P}_{mirror} with current-best learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

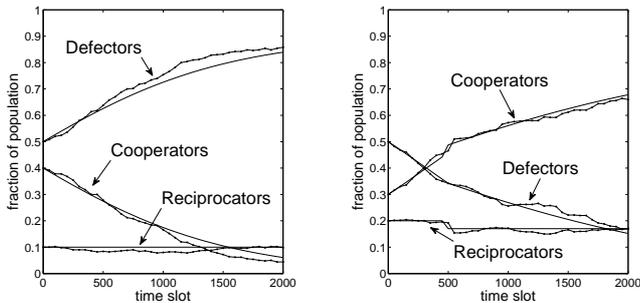


Fig. 3. Evolution of \mathcal{P}_{mirror} with opportunistic learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

Exp. 2: Performance and Robustness of the Proportional Incentive Policy (\mathcal{P}_{prop}):

The population dynamics of \mathcal{P}_{prop} under different initially conditions $\mathbf{x}(0)$ and different learning models are shown in Figure 4 and 5. The simulation settings are the same as that Exp. 1. The dynamic of the current-best learning is depicted in Figure 4. In the left graph, the initial condition is $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. We can see that after about 500 time slots, peers abandon the defective strategy and the P2P network reaches a robust state. In the right graph, initially we have $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$, so the P2P network begins with a large fraction of reciprocators. At $t = 500^-$, we have $\mathbf{x}(t) = (0.24, 0.75, 0.01)^T$. At $t = 500$, we introduce the same disturbance as before and we have $\mathbf{x}(500) = (0.27, 0.72, 0.01)^T$. We can see that the P2P network is robust after this disturbance. In Figure 5, we repeat the same experiments under the opportunistic learning and similar conclusion can be made. The performance gains at the end of each

$\mathbf{x}(0)$	P_0	P_1	P_2	P
$(0.4, 0.1, 0.5)^T$	5.97	5.97	-	5.97
$(0.3, 0.2, 0.5)^T$	5.96	5.97	-	5.97
$(0.4, 0.1, 0.5)^T$	5.97	5.99	-	5.98
$(0.3, 0.2, 0.5)^T$	5.98	5.98	-	5.98

In summary, our mathematical framework is very accurate and more importantly, we show that the proportional incentive policy \mathcal{P}_{prop} is *more robust* than the mirror incentive policy \mathcal{P}_{mirror} .

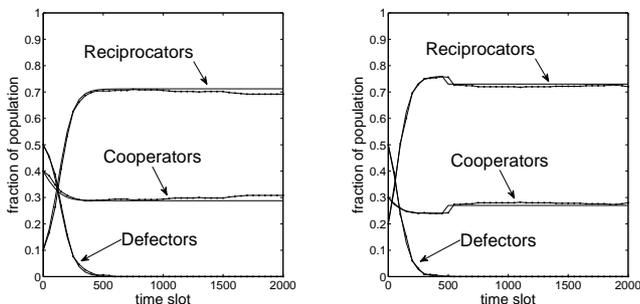


Fig. 4. Evolution of \mathcal{P}_{prop} with current-best learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

Exp. 3: Performance and Robustness of incentive policy in the Linear Incentive Class (\mathcal{C}_{LIP}):

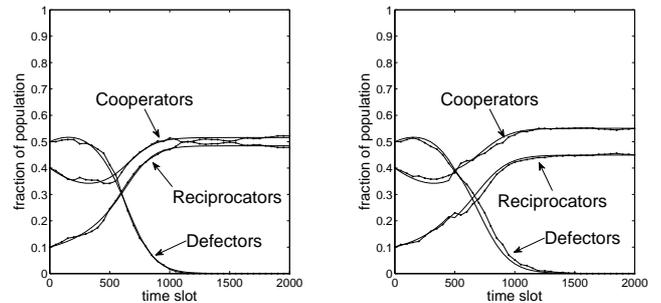


Fig. 5. Evolution of \mathcal{P}_{prop} with opportunistic learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

policy, we set $p_c = 0.9$, $p_r = 1.0$ and $p_d = 0.3$. Figure 6-7 show the population dynamics of this linear policy under different initial conditions and learning models.

Figure 6 shows the dynamics under the current-best learning. In the left graph, the initial population profile is $\mathbf{x}(0) = (0.7, 0.13, 0.17)^T$, which is in the robust region of current-best learning. We can see that reciprocators finally dominates the system. However, in the right graph, when the initial population profile $\mathbf{x}(0) = (0.7, 0.07, 0.23)^T$ is not in the robust region of current-best learning, the system collapses. Figure 7 shows the result under the opportunistic learning. The initial condition in the left graph is the same as that of the right graph of 6. However, the system becomes robust under the opportunistic learning. Note that $\mathbf{x} = (0.7, 0.07, 0.23)^T$ is between the two boundary curves. In the right graph of Figure 7, the system collapse when the initial state $\mathbf{x}(0) = (0.1, 0.05, 0.85)^T$ is not in the robust region of opportunistic learning. The performance gains for this policy are:

$\mathbf{x}(0)$	P_0	P_1	P_2	P
$(0.7, 0.13, 0.17)^T$	5.56	6.00	-	6.00
$(0.7, 0.07, 0.23)^T$	-	-0.24	0.04	0.04
$(0.7, 0.07, 0.23)^T$	5.37	6.00	-	6.00
$(0.1, 0.05, 0.85)^T$	-	-0.22	0.04	0.04

In conclusion, we validate our mathematical framework. When $p_d \neq 0$, the robustness of the linear policy depends on the *initial condition* $\mathbf{x}(0)$. When there are many reciprocators, the system tends to be robust, otherwise, it is more likely to collapse. Opportunistic learning is more robust than the current-best learning in the sense that it has a larger robust region.

Exp. 5: The Effect of Non-adaptive Peers: In this experiment, we consider the impact to the P2P networks when a “fixed portion” of peers do not adapt. The reason for carrying out this experiment is to understand the impact of maintaining some percentage of cooperators or reciprocators and to see whether the system can still be robust. As before, we denote the fraction of type i peers at time slot t by $x_i(t)$. Since the service model is independent with learning process, given the fraction of each type, the performance $\bar{P}_i(t)$ of type i peer is the same as before. Let the fraction of type i peers that do not learn be f_i , then for current-best learning, the system

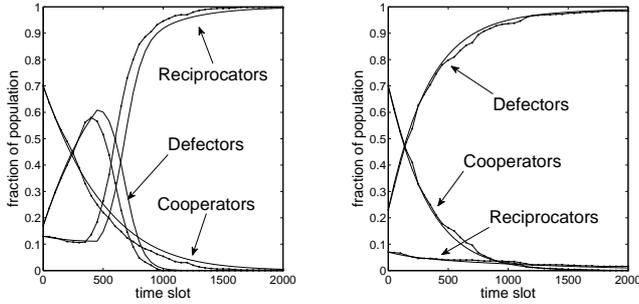


Fig. 6. Evolution of \mathcal{P}_{lin} with current-best learning. Left: $\mathbf{x}(0) = (0.7, 0.13, 0.17)^T$. Right: $\mathbf{x}(0) = (0.7, 0.07, 0.23)^T$.

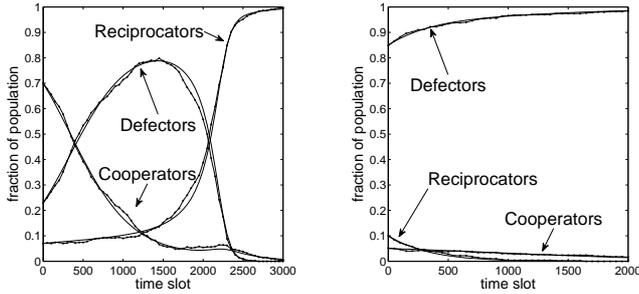


Fig. 7. Evolution of \mathcal{P}_{lin} with opportunistic learning. Left: $\mathbf{x}(0) = (0.7, 0.07, 0.23)^T$. Right: $\mathbf{x}(0) = (0.1, 0.05, 0.85)^T$.

dynamics is described by:

$$x_i(t+1) = x_i(t) - \gamma(x_i(t) - f_i)(\bar{P}_h(t) - \bar{P}_i(t)), i \neq h. \quad (22)$$

$$x_h(t+1) = x_h(t) + \gamma \sum_{i=1, i \neq h}^n (x_i(t) - f_i)(\bar{P}_h(t) - \bar{P}_i(t)). \quad (23)$$

Since the fraction of any type of non-adaptive peers does not change, to compare system performance and robustness under different policy and learning mechanism, we need to focus on the dynamics of adaptive peers. We say a system is robust if adaptive peers have high contribution level on average, otherwise, we say the system collapses. In here, we consider two policies under the current-best learning method and show how non-adaptive peers affect the system dynamics.

Exp. 5a: Mirror policy with CBL: Using the performance gap analysis, we have the following results:

- When $f_2 > 1/\alpha$, all adaptive defectors will switch to cooperative strategy, and the system is robust.
- When $f_2 < 1/\alpha$, all adaptive cooperators will switch to defective strategy, and the system collapses.

Remark: We see that the system robustness under the mirror policy and the current-best learning method depend only on the fraction of non-adaptive reciprocators.

Exp. 5b: Proportional policy with CBL: Using the performance gap analysis, we have the following results:

- if $x_2 < \frac{1}{1-\alpha}f_1$, all adaptive peers will become defectors and the system collapses.
- if $x_2 > \frac{1}{1-\alpha}x_1$, all adaptive defectors will adapt to reciprocators and the system is robust.

- if $\frac{1}{1-\alpha}f_1 < x_2 < \frac{1}{1-\alpha}x_1$, whether the system is stable or not depends on initial condition.

Remark: Here we can see that the non-adaptive cooperators, i.e., the seeders in a P2P system, has a *negative effect* on the system robustness under the proportional incentive policy with the current-best learning. The reason is that adaptive peers will have no motivation to contribute. However, if the seeders reach a high fraction, the P2P system can still have high average performance.

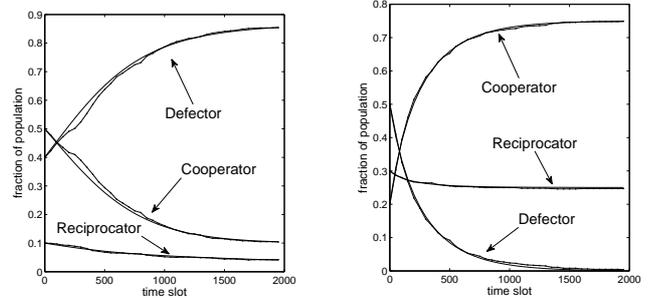


Fig. 8. Evolution of \mathcal{P}_{mirror} with current-best learning. Left: $\mathbf{x}(0) = (0.5, 0.1, 0.4)^T$. 0.1 fraction of fixed cooperators. Right: $\mathbf{x}(0) = (0.2, 0.3, 0.5)^T$. 0.1 fraction of fixed cooperators, 0.2 fraction of fixed reciprocators.

Figure 8 shows two scenarios for the the mirror incentive policy with the current-best learning. In the left graph, initially there are 0.5 fraction of cooperators and 0.1 fraction of cooperators are non-adaptive. We see that the defectors finally dominates the system and only the fixed cooperators are left. In the right graph, There are 0.2 fraction of non-adaptive reciprocators and we see that the defectors are driven out of the P2P system.

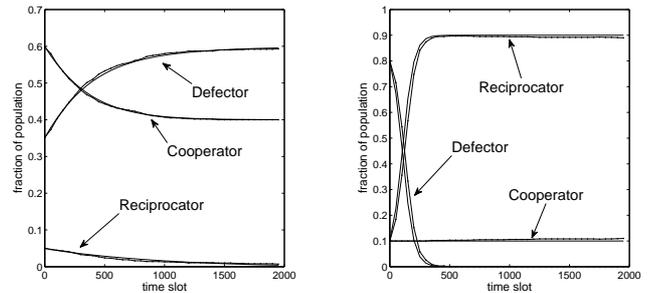


Fig. 9. Evolution of \mathcal{P}_{prop} with current-best learning. Left: $\mathbf{x}(0) = (0.6, 0.05, 0.35)^T$. 0.4 fraction of fixed cooperators. Right: $\mathbf{x}(0) = (0.1, 0.1, 0.8)^T$. 0.1 fraction of fixed cooperators.

Figure 9 shows the simulation results for the proportional incentive policy with the current-best learning. In the left graph, there are 0.4 fraction of non-adaptive cooperators and 0.05 fraction of reciprocators initially. The system finally collapses with only defectors and non-adaptive cooperators left. In the right graph, there are 0.1 fraction of non-adaptive cooperators and 0.1 fraction of reciprocators initially and the system is robust.

VII. Connection with Evolutionary Game Theory

It is important for us to show that there is a *connection* between our model and the evolutionary game theory. As a result, one can use a game-theoretic technique to identify important characteristics for system robustness.

Evolutionary game theory considers a population of individuals with pure strategy set \mathcal{S} . A *population profile* is a vector \mathbf{x} , where x_i is the probability that a peer in the population is using strategy $s_i \in \mathcal{S}$. The payoff of strategy σ in a population with profile \mathbf{x} is denoted as $\pi(\sigma, \mathbf{x})$. That is, the whole population acts as the second player. The evolution of the population profile is the key concern of evolutionary game theory. One central concept in evolutionary game theory is the *Evolutionary Stable Strategy* (ESS), which is the strategy that produces an equilibrium point in the evolution. We have the following definition.

Definition 1: A mixed strategy σ^* is an ESS if there exist an $\bar{\epsilon} > 0$ such that for every $0 < \epsilon < \bar{\epsilon}$ and every mixed strategy $\sigma \neq \sigma^*$, inequality $\pi(\sigma^*, \mathbf{x}_\epsilon) > \pi(\sigma, \mathbf{x}_\epsilon)$ holds. Here, \mathbf{x}_ϵ is the post-entry population profile with $\mathbf{x}_\epsilon = (1 - \epsilon)\sigma^* + \epsilon\sigma$, which depicts population profile after a small mutation from the strategy σ^* to σ .

One simple and important class of evolutionary game is the *pairwise contest population game*. In such a game, the payoff to a focal individual using σ in a population with profile \mathbf{x} is

$$\pi(\sigma, \mathbf{x}) = \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} p(s)x(s')\pi'(s, s'). \quad (24)$$

where $p(s)$ and $x(s')$ are the probability that the focal player and the selected player from the population using strategy s and s' respectively. The *associated two-player game* for this population game is given by the following payoff function:

$$\pi_1(s, s') = \pi_2(s', s) = \pi'(s, s'). \quad (25)$$

Remark: In a pairwise contest population game, for one player, the population with profile \mathbf{x} is indistinguishable from a single player that uses a mixed strategy σ with $\sigma(i) = x_i$. With this in mind, we use a strategy to denote a population profile and vice versa. To check and find ESSs in a pairwise contest population game, we have:

Theorem 1: σ^* is a ESS in a pairwise contest population game if and only if for any $\sigma \neq \sigma^*$, either one of the following two conditions holds:

- 1 $\pi'(\sigma^*, \sigma^*) > \pi'(\sigma, \sigma^*)$,
- 2 $\pi'(\sigma^*, \sigma^*) = \pi'(\sigma, \sigma^*)$ and $\pi'(\sigma^*, \sigma) > \pi'(\sigma, \sigma)$.

$\pi'(\cdot, \cdot)$ is the payoff function of the associated two-player game.

Proof: please refer to [13]. ■

The implication of this theorem is that all ESSs of a pairwise contest population game are Nash equilibria and can be found from the strategic form of the associated two-player game.

Note that payoff function alone does not tell us how the population evolves. Payoff has to be interpreted to define a dynamic model of evolution. In replica dynamics [13], the payoff is interpreted as the number of offsprings as the result of certain strategy. Let $\bar{\pi}(\mathbf{x})$ be the average payoff, then the

system can be describe by a system of differential equations:

$$\dot{x}_i = (\pi(s_i, \mathbf{x}) - \bar{\pi}(\mathbf{x}))x_i \quad s_i \in \mathcal{S}. \quad (26)$$

Now we give the definition of an important stability concept in dynamic systems and its connection with ESS.

Definition 2: An asymptotically stable fixed point (ASF) of a dynamic system is a fixed point that any small deviation from it is eliminated by the dynamics as $t \rightarrow \infty$.

Theorem 2: For pairwise contest population games, the ESS of the associated two-player game is an asymptotically stable fixed point (ASF) of Eq. (26).

Proof: please refer to [13]. ■

In short, Theorem 2 justifies ESS as the evolutionary end point.

Let us apply the game-theoretic results to our incentive model. To find the underlying game, we derive the payoff function first. By Eq (1)-(3), the payoff of strategy s_i against the population profile $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]$ at time t is:

$$\begin{aligned} \pi(s_i, \mathbf{x}) &= \bar{P}_i(t) = \alpha \sum_{j=1}^n x_j(t)g_j(i) - \sum_{j=1}^n x_j(t)g_i(j) \\ &= \sum_{j=1}^n (\alpha g_j(i) - g_i(j))x_j(t), \quad i = 1, \dots, n. \end{aligned} \quad (27)$$

Note that the payoff $\pi(s_i, \mathbf{x})$ and generosity $g_j(i)$ are actually functions of t . Here we ignore argument t for readability. Comparing Eq. (27) with Eq. (24), we can construct the associated two player game with payoff

$$\pi_1(s_i, s_j) = \pi_2(s_j, s_i) = \alpha g_j(i) - g_i(j). \quad (28)$$

The physical meaning of the above payoff expression is as follows: it is the points gained by using strategy s_i after exchanging a pair of services with a peer using strategy s_j .

In summary, one can map our mathematical framework into a pairwise contest game if the generosity matrix \mathbf{G} is constant, and the associated two-player game is just an exchange of service. Both our model and the pairwise contest game have the field payoff described by Eq. (27), so their dynamical properties are the same. Formally, we have:

Theorem 3: A linear incentive policy can be mapped to a two-player symmetric game, and the ESS of this game is an ASF of its opportunistic learning dynamics described by Eq. (7). If the ESS is a pure strategy $s_e \in \mathcal{S}$, it is also an ASF of its current-best learning dynamics described by Eq. (5)-(6).

Proof: The mapping is given by Eq. (28). More accurately, the payoff matrix of first player is $\alpha\mathbf{G} - \mathbf{G}^T$, and the payoff matrix of the second player is $\alpha\mathbf{G}^T - \mathbf{G}$.

To prove the first part of the theorem, by Theorem 2, the ESS of the pairwise contest is an ASF of Eq. (26). Note that $\pi(s_i, \mathbf{x}) = \bar{P}_i(t)$ and $\bar{\pi}(\mathbf{x}) = \bar{P}(t)$, therefore Eq. (26) is the same as Eq. (7) up to a constant factor λ and hence they have the same set of ASF. So the ESS is also an ASF of Eq. (7).

For the second part, consider a small deviation from the pure s_e population. By Theorem 1, one can show that s_e 's payoff $\pi(s_e, \mathbf{x}_\epsilon) > \pi(\sigma, \mathbf{x}_\epsilon)$ for any other strategies σ . Since $\pi(s_i, \mathbf{x}) = \bar{P}_i(t)$, we have $\bar{P}_e(t) > \bar{P}_i(t)$ for any $i \neq e$. So s_e is the winner, and by Eq. (5)-(6), x_e will increase, which eliminates the deviation. Therefore s_e is an ASF. ■

The above theorem provides a simple game-theoretical method to analyze the robustness of our incentive model. For example, consider the linear incentive policy in section IV with $p_c = 0.9, p_r = 1, p_d = 0.3$. When $\alpha = 7$, the associated two-player game has payoff table as follows:

	s_1	s_2	s_3
s_1	6, 6	5.3, 6.1	-1, 7
s_2	6.1, 5.3	6, 6	-0.3, 2.1
s_3	7, -1	2.1, -0.3	0, 0

To find ESSs, we find Nash equilibria first. Noticing that s_1 is dominated by s_2 , we can reduce the payoff table to

	s_2	s_3
s_2	6, 6	-0.3, 2.1
s_3	2.1, -0.3	0, 0

This game has three Nash equilibria: s_2 and s_3 and a mixed strategy: $(c_{lower}, 1 - c_{lower})$. Here $c_{lower} = 0.0714$ (see Section IV). Using Theorem 1, we can check that the first two pure Nash equilibria are ESSs and the mixed Nash equilibria will collapse under small disturbance. By Theorem 3, for both learning models, the system has two stable states: one that all peers use the linear strategy and another that all peers decide to have the defective behavior. Which ESS the system will converge to depends on the initial state $x(0)$. When there is no cooperators, the players in the associated game will converge to s_2 if $x_2 > c_{lower}$ and to s_3 if $x_2 < c_{lower}$. This result agrees with our analysis in Section IV and is verified by our experiments in Section VI.

VIII. The Effect of Protocol Cost

In previous sections, we assume that there is a reputation service maintained by the system. In a distributed system, such service is specified by the incentive protocol and maintained by those who follow this protocol, i.e., the reciprocators. Reciprocators will provide their private history of transactions to this service and in return, they can access the type information of any requester, and avoid the cost of serving defectors. If the reputation service is implemented via the distributed hash table (DHT) method, each reciprocator also has to contribute local storage, bandwidth and computing power to manage the information. In short, there is additional protocol cost for reciprocators. We denote this cost as θ . In previous analysis of the proportional incentive strategy, we ignored θ and reach the conclusion that the fraction of reciprocator will finally dominate the P2P system and defectors will extinct. However, the result may be different if we take θ into account.

Here we analyze the effect of cost θ on the proportional policy with the current best learning method. Using similar approach, we can derive the performance gaps as:

$$\begin{aligned}\bar{P}_3(t) - \bar{P}_1(t) &= 1 - \alpha x_2(t), \\ \bar{P}_3(t) - \bar{P}_2(t) &= 1 - \alpha x_2(t) + \theta - x_3(t), \\ \bar{P}_2(t) - \bar{P}_1(t) &= x_3(t) - \theta.\end{aligned}$$

First, let us see what happens when there are only two types of users (e.g., (a) reciprocators and cooperators; (b) reciprocators and defectors). We have four cases:

- $x_3(t) = 0$. The cooperators will dominate the system and the overall average performance \bar{P} reaches the optimal value of $\alpha - 1$.
- $x_2(t) = 0$. The system will be overwhelmed by defectors, $\bar{P} = 0$.
- $x_1(t) = 0, x_2(0) > \frac{\theta}{\alpha-1}$. The system will be dominated by reciprocators and $\bar{P} = \alpha - 1 - \theta$.
- $x_1(t) = 0, x_2(0) < \frac{\theta}{\alpha-1}$. The system will be dominated by defectors and $\bar{P} = 0$.

If we have all three types of users initially, it turns out that the P2P system will either *collapse* or *oscillates*. We can divide the system dynamics into four phases:

- P1:** Reciprocation is the best strategy. In this phase, $x_1(t)$ and $x_3(t)$ decrease while $\bar{P}_1(t)$ increases. When $x_3(t)$ is less than θ , cooperation becomes better than reciprocation and the system enters phase P2.
- P2:** Cooperation is the best strategy. $x_2(t)$ and $x_3(t)$ decrease while \bar{P}_3 increases. Once $x_2(t)$ is less than $\frac{1}{\alpha}$, defectors get the best performance and the system enters phase P3.
- P3:** Defection is the best strategy, $x_2(t) > \frac{\theta}{\alpha-1}$. In this phase, $x_1(t)$ and $x_2(t)$ decrease. As $x_3(t)$ rises, reciprocation regains the advantage. If $x_2(t)$ does not drop below the threshold $\frac{\theta}{\alpha-1}$, reciprocation finally becomes the best strategy and the system enters phase P1 again, starting another cycle. If $x_2(t)$ drops below the threshold, the system enters phase P4, in which the system will collapse.
- P4:** Defection is the best strategy, $x_2(t) < \frac{\theta}{\alpha-1}$. In this phase, reciprocators keep decreasing and defectors finally dominates the system.

Remark: Whether the system collapses or oscillates depends on θ and the initial condition. In general, systems with larger θ are more likely to collapse because of the cost of realizing an incentive protocol. On the other hand, if we can prevent cooperators from getting better performance than reciprocators, the system will stay in P1 and be robust.

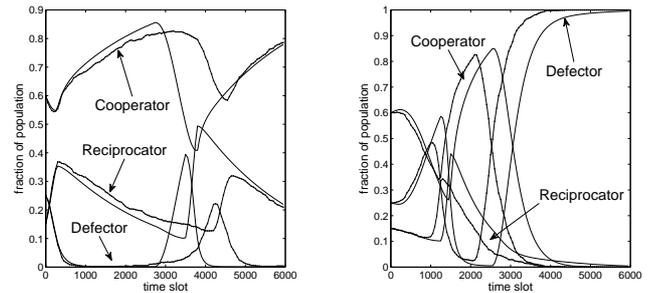


Fig. 10. Evolution of \mathcal{P}_{prop} with current-best learning. Left: $x(0) = (0.6, 0.15, 0.25)^T, \theta = 0.1$. Right: $\theta = 0.3$.

Figure 10 shows the simulation and the mathematical results. In the left graph, the initial condition is $x(0) = (0.6, 0.15, 0.25)^T$ and $\theta = 0.1$. We can see that the system oscillates with a period of about 4500 time slots. Due to the error accumulated for all simulation slots, there is a little offset between the curves of simulation results and the curves of our model prediction. However, our model still predicts the right shapes and trends of these performance curves. In the right

graph, the initial condition is the same but $\theta = 0.3$. We can see that the system collapses after one cycle.

IX. The Tradeoff between Altruism and System Robustness

In the previous section, we see that when there is a protocol cost, the proportional incentive policy may not be able to eliminate defectors under the current-best learning method. Defectors can revive and degrades the system performance periodically or permanently. This is mainly caused by the blind altruism of cooperators in providing services to others so that other peers will not have incentive to contribute. To prevent the periodic or permanent performance recession, reciprocators should limit or constrain cooperators so that cooperators will not overrun reciprocators when there are very few defectors. One interesting question is how much should reciprocators constrain the cooperators? Applying heavy limitation may not be fair to cooperators and may hinder cooperation between reciprocators and cooperators, while little limitation may not be enough to ensure the robustness of P2P systems. To address this problem, we consider and analyze a linear policy with $p_c = \rho, p_r = 1, p_d = 0$. Using the similar approach, we can derive the performance gaps as:

$$\begin{aligned}\bar{P}_3(t) - \bar{P}_1(t) &= 1 - \alpha\rho x_2(t), \\ \bar{P}_2(t) - \bar{P}_3(t) &= (\alpha - 1)x_2(t) - \rho x_1(t) - \theta, \\ \bar{P}_2(t) - \bar{P}_1(t) &= 1 + (\alpha - 1 - \alpha\rho)x_2(t) - \rho x_1(t) - \theta.\end{aligned}$$

Based on the performance gap analysis, we have:

- When $x_2 < \frac{\theta}{\alpha-1}$, reciprocators keep decreasing and the system will eventually collapse.
- When $x_2 > \max\{\frac{\rho+\theta}{\alpha-1}, \frac{\rho+\theta-1}{(\alpha-1)(1-\rho)}\}$, reciprocators finally dominates the system.

Remark: The max expression in the second scenario decreases as ρ increases. So the protocol cost θ and ethics level ρ are all reversely related to system robustness.

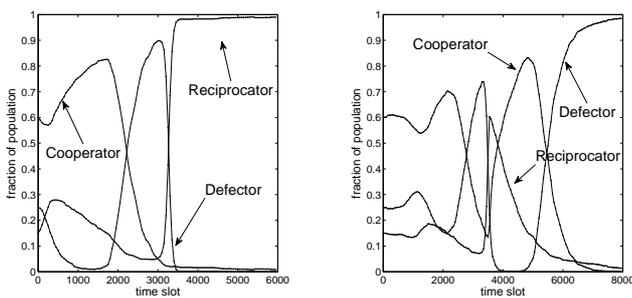


Fig. 11. Evolution of \mathcal{P}_{prop} with current-best learning. Left: $\mathbf{x}(0) = (0.6, 0.15, 0.25)^T$, $\rho = 0.96$. Right: $\rho = 0.99$.

Figure 11 shows the effect of of altruism ρ by the cooperators. In the left graph, we set $\rho = 0.96$ and we see that the system is robust. In the right graph, the initial condition is the same but we set $\rho = 0.99$, which is only a little bit higher. We see that the little difference of ρ leads to completely different result and the system eventually collapses.

X. Related Work

The earliest work on how to encourage cooperation in P2P networks is via micro-payment [4]. In essence, it uses a centralized approach to issue virtual currency. When a node provides service to another node, virtual currency is exchanged. Authors in [7]–[9] present the incentive issues and service differentiation in P2P networks. In [1], [14], [16], [17], authors also present their study of incentive issues in wireless networks. In [12], authors show that shared history based incentives can overcome the scalability problem of private history based mechanisms. Furthermore, one can use DHT to implement the shared history incentive mechanism. One example of shared history based incentive mechanism is the reciprocative strategy [2], [6]. Each node makes decisions according to the reputation of requesters and is studied via simulation only. As for learning mechanisms, Q-learning [4] and Slacer [5] are two learning methods and their performance study was carried out via simulation or via small scale prototyping only. This paper focus on the general mathematical framework to analyze the robustness and properties on adaptive incentive protocols with different learning mechanisms.

There are also some models developed to help in designing incentives mechanism. Authors in [3] assume that each peer has a fixed strategy set with a certain distribution while we assume peers can adapt their strategies. In [15], authors show that a proportional strategy can lead to market equilibria but the result does not generalize to multiple strategies. Authors in [11] analyze a reputation based reciprocative strategy and its evolution dynamics in a biological context. Our paper focuses on the robustness of distributed learning mechanisms.

XI. Conclusion

The main contribution of this paper is on introducing a general mathematical framework to model and evaluate the performance and robustness of incentive policies in P2P networks. We assume that peers are rational and they adapt their strategy based on the behavior of other peers. To illustrate our mathematical framework, we present two incentive policies and show that the mirror incentive policy \mathcal{P}_{mirror} may lead to a complete system collapse, while the proportional incentive policy \mathcal{P}_{prop} , which takes into account of service consumption and contribution, can lead to a robust system. We also analyze a general class of incentive policies (the linear incentive policy class) and show that, for a system to be robust, we have to assure certain fraction of reciprocators in the P2P system. Peers can learn about the payoff of other strategies via distributed learning mechanism. We also present two learning mechanisms and how they can be evaluated in our mathematical framework. We show that the current-best learning is less robust than the opportunistic learning, altruism may have detrimental impact on the system, and when the cost of realizing an incentive mechanism is high, the overall system may not be robust. In general, learning mechanism is worthwhile and one may consider incorporating this feature into the incentive protocol design so as to encourage peers to adapt and cooperate. This way, the P2P system can quickly converge to the desirable operating point.

Acknowledgement: This report is supported in part by the RGC Grant 415309.

REFERENCES

- [1] S. Buchegger and J.-Y. L. Boudec. Performance analysis of the confidant protocol. In *MobiHoc '02*. ACM, 2002.
- [2] M. Feldman, K. Lai, I. Stoica, and J. Chuang. Robust incentive techniques for peer-to-peer networks. In *ACM EC'04*, 2004.
- [3] M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica. Free-riding and whitewashing in peer-to-peer systems. In *Workshop on Practice & theory of incentives in networked systems*, 2004.
- [4] P. Golle, K. Leyton-Brown, and I. Mironov. Incentives for sharing in P2P networks. In *3rd ACM Conf. on Electronic Commerce*, 2001.
- [5] D. Hales and S. Arteconi. Slacer: a self-organizing protocol for coordination in peer-to-peer networks. *Intelligent Systems, IEEE*, 2006.
- [6] K. Lai, M. Feldman, I. Stoica, and J. Chuang. Incentives for cooperation in P2P networks. In *Workshop on Economics of P2P Systems*, 2003.
- [7] T. B. Ma, C. M. Lee, J. C. S. Lui, and K. Y. Yau. A Game Theoretic Approach to Provide Incentive and Service Differentiation in P2P Networks. In *ACM Sigmetrics*, 2004.
- [8] T. B. Ma, C. M. Lee, J. C. S. Lui, and K. Y. Yau. An Incentive Mechanism for P2P Networks. In *IEEE ICDCS*, 2004.
- [9] T. B. Ma, C. M. Lee, J. C. S. Lui, and K. Y. Yau. Incentive and Service Differentiation in P2P Networks: A Game Theoretic Approach. *IEEE/ACM Trans. on Networking*, 14(5), 2006.
- [10] L. Massoulié and M. Vojnović. Coupon replication systems. In *SIGMETRICS*. ACM, 2005.
- [11] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 1998.
- [12] V. Vishnumurthy, S. Chandrakumar, and E. Sirer. Karma: A secure economic framework for peer-to-peer resource sharing. In *Workshop on Economics of Peer-to-Peer Networks*, 2003.
- [13] J. N. Webb. Game theory: Decisions, interaction and evolution. *Springer*, pages 139–185, 2006.
- [14] F. Wu, T. Chen, S. Zhong, L. E. Li, and Y. R. Yang. Incentive-compatible opportunistic routing for wireless networks. In *ACM Mobicom*, 2008.
- [15] F. Wu and L. Zhang. Proportional response dynamics leads to market equilibrium. In *ACM STOC*, 2007.
- [16] S. Zhong, J. Chen, and Y. Yang. Sprite: a simple, cheat-proof, credit-based system for mobile ad-hoc networks. In *INFOCOM, IEEE*, 2003.
- [17] S. Zhong and F. Wu. On designing collusion-resistant routing schemes for non-cooperative wireless ad hoc networks. In *ACM Mobicom*, 2007.
- [18] Y. Zhou, D. M. Chiu, and J. Lui. A simple model for analyzing p2p streaming protocols. *ICNP*, 2007.