

Analysis of Adaptive Incentive Protocols for P2P Networks

Bridge Q. Zhao John C.S. Lui Dah-Ming Chiu
The Chinese University of Hong Kong

Abstract—Incentive protocols play a crucial role to encourage cooperation among nodes in networking applications. The aim of this paper is to provide a *general analytical framework* to analyze and design a large family of incentive protocols. We consider a class of incentive protocols wherein peers can distributively learn and adapt their actions. Using our analytical framework, one can evaluate the expected performance gain and system robustness of a given incentive protocol. To illustrate the framework, we present three incentive policies and two learning (or adaptive) models. We show under what conditions the network may collapse (e.g., no cooperation in the system) or the incentive protocol can guarantee a high degree of cooperation. In particular, we formally show the connection between evaluating incentive protocols and evolutionary game theory so to identify robustness characteristics of an incentive policy.

I. Introduction

Designing a correct and robust *incentive protocol* is crucial in many network applications. To illustrate, consider a P2P file sharing network in which peers rely on other peers to perform uploading service. This mutual uploading service offloads the server and allows the system to scale. Without an incentive mechanism to encourage peers to perform uploading service, the server will be overwhelmed and peers may take a long time to download a file. The above example demonstrates one important point: *embedding incentive protocols to encourage cooperation is crucial for the overall system performance*. A formal framework to analyze and evaluate various incentive protocols is important for proving the incentive protocol can indeed encourage collaboration and that the system is robust under perturbation.

As pointed out in [2], [6], there is a natural tendency or even beneficial for nodes to *learn* from the environment and *adapt* their actions. For example, a P2P node may choose to provide uploading service to other nodes, but when it discovers that there are many free riders (e.g., those nodes which enjoy the downloading service without providing any upload contribution) in the system, then this node may choose to change its behavior and adapt to 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 incentive protocols for networking applications. Our contributions are:

- We present a general analytical framework to evaluate incentive protocols in which peers have distributed learning and adaptive capabilities (Section II, III).
- We illustrate our mathematical framework by presenting three incentive policies (Section IV) under two different learning models and derive their performance measures and robustness conditions (Section V).
- We present the performance evaluation to illustrate the performance gain of using strategy adaptation and state formally under what conditions the system will be robust and under what conditions the system may collapse (Section VI).
- We show the connection between evaluating the robustness of incentive protocols and *evolutionary game theory*. We show how to map linear incentive policies to two-player games, and to give an efficient technique to identify the robustness characteristics of linear incentive policies. (Section VII).

II. A Generalized P2P Incentive Model

In this section, we present a general mathematical model to study some incentive protocols for P2P networks. Given an incentive protocol, we show how to use this framework to evaluate (1) its evolution and robustness, and (2) various performance measures such as expected service received and service contributed of the given incentive protocol. In our study, we make 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 willing to contribute (altruism), to refusing to contribute (egoism).

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 and requests for service (this assumption is made in several studies, e.g. [10], [16]). 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, we define $n \times n$ generosity matrix \mathbf{G} with $G_{ij} = g_i(j)$. At each time slot, a peer gains $\alpha > 0$ 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 set $\beta = 1$.

Let $x_i(t)$ denote the fraction of type i peers in the system at time t , and $E[R_i(t)]$ be the expected services that a type

i peer can receive in one time slot. We can express $E[R_i(t)]$ 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 system. Let $k \in \mathcal{N}$, then the probability that this tagged type i peer will provide service to this peer k is \mathcal{L} , where:

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

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 $N(t)$ is sufficiently large, we have

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

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

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

One can express the above n equations in matrix form and derive $\bar{P}(t)$, the expected gain per time slot for the whole P2P network at time t as

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

where $\mathbf{x}(t)$ is a column vector of $[x_1(t), \dots, x_n(t)]$. The key point is that to model and analyze an incentive protocol, one has to “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 framework by studying 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 who 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 in the following section.

III. Learning Models

Learning is a natural behavior for rational peers since they get information from the external environment, and adjust their strategy to get better performance. When learning mechanisms are implemented in the system, learning and adapting to a better strategy will have higher performance than non-adaptive system. Here, we present two learning models and study how these learning models can affect the dynamics of incentive policies in P2P networks.

Current-best Learning Model (CBLM): A simple way to learn is that all peers find the “current best” strategy in the system and switch to it. This 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 $h \in \arg\max_j \{\bar{P}_j(t)\}$). 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.

For 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 implement this learning procedure, 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 this learning model can affect the system dynamics for a given incentive policy. We 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:

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

For computational efficiency, one can transform the above difference equations to a continuous model¹ as:

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

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

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

Again, given an incentive policy, if we can determine its $g_i(j), \forall i, j$, then we can evaluate the dynamics of the system using the above differential equations.

Opportunistic Learning Model (OLM): this learning model 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. Note that this learning model does not require frequent access to shared global history and can be realized in a distributed fashion.

Let us show how this learning model 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:

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

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:

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

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

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

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 \dot{x}_j , is:

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

In summary, given an incentive policy \mathcal{P} and its generosity matrix \mathbf{G} , 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

Let us illustrate how to use the analytical framework to analyze different incentive protocols. In general, one can classify peers in a P2P network according to their *behavior* upon receiving a request [2], they are:

- *Cooperator*: a peer has a cooperative (or altruistic) behavior when it serves other peers unconditionally.

- *Defector*: a peer has a defective (or egoistic) behavior when it refuses to serve any request from other peers.
- *Reciprocator*: a peer has a reciprocative behavior when it serves the requester according to the requester's service strategy. In short, it tries to make a fair service exchange.

The above behaviors are often implemented into a P2P software. One interesting question is how to design a proper incentive policy so as to keep the system scalable and robust? In the following, we will address this important question.

Mirror Incentive Policy (\mathcal{P}_{mirror}): under this policy, when a reciprocative peer receives a request for service, this peer infers (e.g., like the tit-for-tac 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 another peer for service, peer k will only get the service with probability of 0.6. Hence, if the requester is a cooperator (defector or reciprocator), this 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.

The \mathcal{P}_{mirror} has three pure strategies: (1) pure cooperation, or s_1 , (2) mirror reciprocation, or s_2 , and (3) pure defection, or s_3 . To model \mathcal{P}_{mirror} , we need to derive entries of 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:

$$\begin{aligned} g_2(2) &= \text{Prob}[\text{a reciprocator will grant a request}] \\ &= \sum_{i=1}^3 \text{Prob}[\text{the requester is of type } i] \times \\ &\quad \text{Prob}[\text{granting the request} | \text{type } i \text{ requests}] \\ &= x_1(t)g_2(1) + x_2(t)g_2(2) + x_3(t)g_2(3) \\ &= x_1(t) + x_2(t)g_2(2) = \frac{x_1(t)}{1 - x_2(t)}. \end{aligned} \quad (8)$$

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 like this: peers using s_2 serve the requester (say type j) with the probability equal to the requester's contribution/consumption ratio, or $E[S_j]/E[R_j]$. In case 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 $g_2(3) = 0$. As for $g_2(2)$, we can derive the following expressions:

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

Since $E[R_2(t)] = E[S_2(t)]$, we have $g_2(2) = 1$. The other values of $g_i(j)$ are the same as that in \mathcal{P}_{mirror} .

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.

Linear Incentive Policy Class (\mathcal{C}_{LIP}): this represents a “family” of incentive policies. 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 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:

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

It is easy to see that the proportional incentive policy belongs to the linear incentive policy class ($\mathcal{P}_{prop} \in \mathcal{C}_{LIP}$) 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 $x_2(t)$.

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, robustness of an incentive protocol implies that the system will finally stay at a high contribution level (e.g., most peers are cooperators or reciprocators) and the network is immune to system perturbation such as peer arrivals or departures.

A. Robustness Analysis of Mirror Incentive Policy: let us first study the mirror incentive policy \mathcal{P}_{mirror} using the current-best learning model (CBLM). Since we have derived $g_i(j)$ of \mathcal{P}_{mirror} in Section IV, substituting them into Eq. (3) and (4), we obtain $\bar{P}_i(t)$, the expected gain of using strategy s_i for $i = \{1, 2, 3\}$, and $\bar{P}(t)$, the expected gain of the system. Now, let us consider their respective differences, which are:

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

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 achieve the best performance. Therefore defectors and reciprocative peers will continue to adapt their

strategies to the cooperative strategy 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 $x_2(t) < 1/\alpha$, $\bar{P}_3(t) > \bar{P}_2(t) > \bar{P}_1(t)$, and defectors have the best performance, therefore, 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. At this time, the P2P network *collapses* (or is not robust) since no one wants to provide service in the P2P network.

For \mathcal{P}_{mirror} 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 collapse. In summary, the mirror incentive policy is *not robust* and eventually all peers will have the defective behavior.

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

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

By the dynamics model 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. The system will achieve the optimal overall performance.

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 $x_2(t)$.

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: Using our methodology, we show that \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* an appropriate incentive protocol for P2P networks.

B. Robustness Analysis of Proportional Incentive Policy: for the policy \mathcal{P}_{prop} , we have derived $g_i(j)$ in Section IV. Let

us consider their respective differences:

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

Note that under the proportional incentive policy, reciprocal behavior is *always better* than cooperative behavior, and we have the following cases:

Case A: when $x_2(t) > \frac{1}{\alpha-1}x_1(t)$, $\bar{P}_2(t) > \bar{P}_3(t)$, so the fraction of reciprocal 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 $x_2(t) = \frac{1}{\alpha-1}x_1(t)$, $\bar{P}_3(t) = \bar{P}_2(t) > \bar{P}_1(t)$, so only cooperative peers will continue to adapt to either 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 system will eventually operate at case A, where the fraction of reciprocal 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.

Let us consider the policy \mathcal{P}_{prop} under the opportunistic learning model (OLM). 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 reciprocal 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.

Remark: Using our methodology, we show that the proportional strategy \mathcal{P}_{prop} is robust under both the current-best learning and the opportunistic learning processes.

C. Robustness Analysis for policy in the Linear Incentive

Class: the idea of reciprocal strategy is to infer the type of requester from the shared history and provide differentiated service. Proportional strategy requires 100% accuracy of such inference. 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} .

Consider a policy in \mathcal{C}_{LIP} with $g_2(1) = p_c$, $g_2(2) = p_r$, $g_2(3) = p_d$. 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:

$$p_d = 0 \quad (9)$$

$$p_r \geq p_c. \quad (10)$$

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. (9) and Eq. (10) are satisfied.

It is worth noting that by Eq. (10), when p_c is small, the system is more likely to be robust. This seems counter-intuitive since reciprocal 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 reciprocal peers, we have to limit the degree of altruism in the network.

Now we restrict our attention to linear strategies with $p_r, p_c > p_d > 0$. The robustness of these policies depends on the initial population, and this is especially true for the reciprocators. Let $c_{upper} = \frac{p_c}{(\alpha-1)(p_r-p_d)+p_c-p_d}$ and $c_{lower} = \frac{p_d}{(\alpha-1)(p_r-p_d)}$. It can be shown that for both learning models, when $x_2(0) > c_{upper}$, the system will be robust, and when $x_2(0) < c_{lower}$, the system will collapse in the current-best learning. As for other initial conditions, the robustness depends on the learning mechanism and the fraction of other strategies.

Figure 1 shows the robust region of the linear policy with $p_c = 0.9, p_r = 1, p_d = 0.3$. 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 opportunistic learning is *strictly larger* than the current-best learning, and the boundary curves intersect the vertical axis at $x_2 = c_{lower} = 0.0714$.

Interestingly, 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 system robust under the opportunistic learning but harm the system under the current-best learning model. Intuitively, we can explain this result as follows. In current-best learning, when reciprocators are overwhelmed by defectors, cooperators merely help defectors, the winner to increase, which makes the system even less robust. However, in opportunistic learning, cooperators compensate the loss for reciprocators and help them to increase, which makes the system more robust.

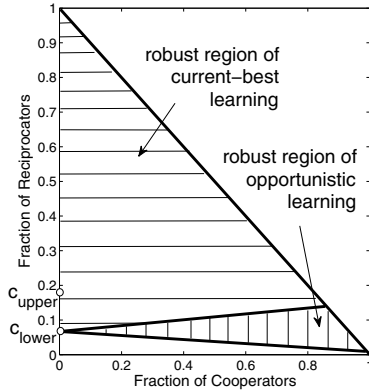


Fig. 1. Robust regions for the two learning models.

VI. Simulations and Performance Studies

In this section, we present the simulation results to validate our model, and to show the dynamics and performance of various incentive protocols. Unless we state otherwise, the parameters we use are: number of peers ($N = 500$), gain per service provided ($\alpha = 7$), cost per service consumed ($\beta = 1$), learning rate ($\gamma = 0.004$). The simulations are carried out via discrete time slots. In each time slot, every 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 switch to a new strategy according to a given probability and learning model, i.e., current-best or opportunistic learning.

Evolution of the Mirror Incentive Policy (\mathcal{P}_{mirror}): Figure 2 to 3 illustrate the population dynamics of \mathcal{P}_{mirror} with different initially conditions $\mathbf{x}(0)$ and different learning models. The solid lines are results from simulation and the dotted lines are results of the mathematical model. 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.

Figure 2 shows the dynamics under the current-best learning. 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.

Figure 3 shows the dynamics under opportunistic learning with the same initial conditions as 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 summary, we validate our mathematical model and confirm that the mirror incentive policy is not robust under the current-best learning but may survive when $x_2(0)$ is above a threshold.

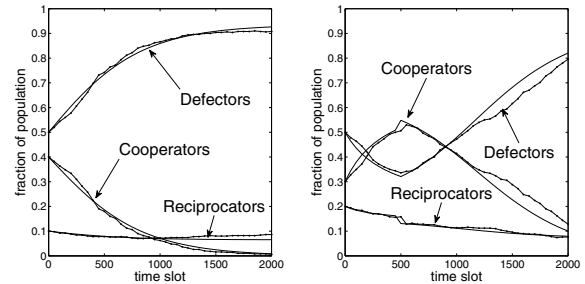


Fig. 2. Evolution of \mathcal{P}_{mirror} with current-best learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

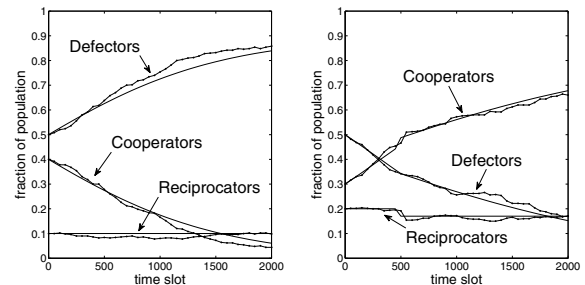


Fig. 3. Evolution of \mathcal{P}_{mirror} with opportunistic learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

Evolution of the Proportional Incentive Policy (\mathcal{P}_{prop}): Figure 4 to 5 illustrate the population dynamics of \mathcal{P}_{prop} under different initially conditions $\mathbf{x}(0)$ and different learning models. The simulation settings are the same as that of \mathcal{P}_{mirror} . Figure 4 shows the dynamics under the current-best learning. 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, the defectors die out and the system reaches a robust state. In the right graph, initially $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$, so the system starts 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 system is robust after this disturbance. In Figure 5, we repeat the same experiments under the opportunistic learning and the

results are similar. The performance gains at the end of each experiment are :

$\mathbf{x}(0)$	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, the proportional incentive policy \mathcal{P}_{prop} is more robust than the mirror incentive policy \mathcal{P}_{mirror} , and the simulation results show that our mathematical models are very accurate.

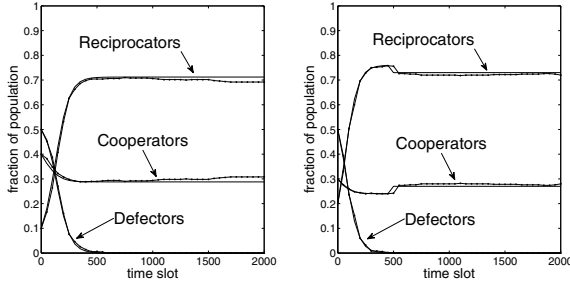


Fig. 4. Evolution of \mathcal{P}_{prop} with current-best learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

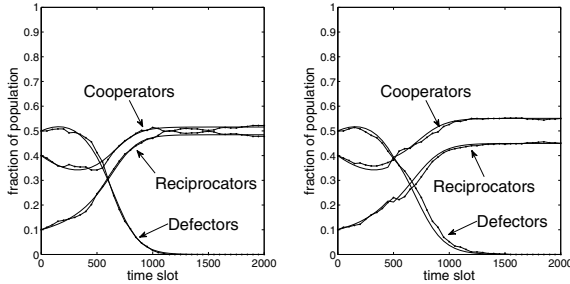


Fig. 5. Evolution of \mathcal{P}_{prop} with opportunistic learning. Left: $\mathbf{x}(0) = (0.4, 0.1, 0.5)^T$. Right: $\mathbf{x}(0) = (0.3, 0.2, 0.5)^T$.

Evolution of incentive policy in the Linear Incentive Class (\mathcal{C}_{LIP}): for the linear incentive 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 steady state performance gains for this policy is:

$\mathbf{x}(0)$	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 summary, when $p_d \neq 0$, the robustness of the linear policy depends on the *initial condition*. 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. The simulation results fit very well with our model.

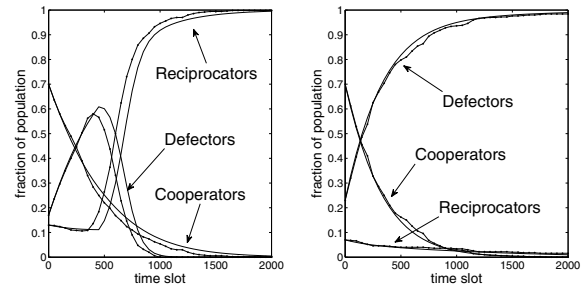


Fig. 6. Evolution of \mathcal{P}_{lin} with current-best learning. Left: $\mathbf{x}(0) = (0.7, 0.13, 0.17)^T$. Right: $\mathbf{x}(0) = (0.7, 0.07, 0.23)^T$.

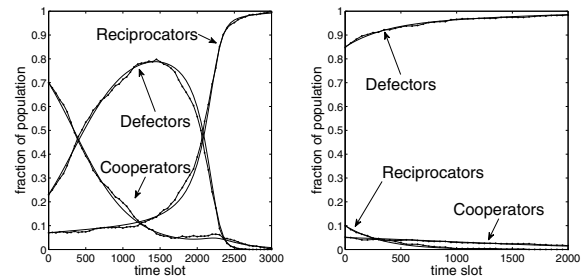


Fig. 7. Evolution of \mathcal{P}_{lin} with opportunistic learning. Left: $\mathbf{x}(0) = (0.7, 0.07, 0.23)^T$. Right: $\mathbf{x}(0) = (0.1, 0.05, 0.85)^T$.

VII. Connection with Evolutionary Game Theory

In this section, we 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

$$\pi(\sigma, \mathbf{x}) = \sum_{s \in \mathcal{S}} \sum_{s' \in \mathcal{S}} p(s)x(s')\pi'(s, s'). \quad (11)$$

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:

$$\pi_1(s, s') = \pi_2(s', s) = \pi'(s, s'). \quad (12)$$

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 use the following result [13]:

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 $\pi'(\sigma^*, \sigma) > \pi'(\sigma, \sigma)$.

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

By this theorem, all ESSs of a pairwise contest population game are Nash equilibria and can be found directly 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:

$$\dot{x}_i = (\pi(s_i, \mathbf{x}) - \bar{\pi}(\mathbf{x}))x_i \quad s_i \in \mathcal{S}. \quad (13)$$

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. (13).

Theorem 2 justifies ESS as the evolutionary end point. For the complete proof, we refer the readers to [13].

One can 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:

$$\begin{aligned} \pi(s_i, \mathbf{x}) &= \bar{P}_i(t) = \alpha \sum_{j=1}^n x_j(t)g_j(i) - \sum_{j=1}^n x_j(t)g_i(j) \\ &= \sum_{j=1}^n (\alpha g_j(i) - g_i(j))x_j(t), \quad i = 1, \dots, n. \end{aligned} \quad (14)$$

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. (14) with Eq. (11), we can construct the associated two player game with payoff

$$\pi_1(s_i, s_j) = \pi_2(s_j, s_i) = \alpha g_j(i) - g_i(j). \quad (15)$$

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. (14), so their dynamical properties are the same.

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. (15). 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. (13). Note that $\pi(s_i, \mathbf{x}) = \bar{P}_i(t)$ and $\bar{\pi}(\mathbf{x}) = \bar{P}(t)$, therefore Eq. (13) 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 importance of this theorem is that it 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 shown in Table I.

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_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

TABLE I
ASSOCIATED TWO-PLAYER GAME

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

TABLE II
REDUCED GAME

Table II. 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 cooperator, 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. Related Work

Micro-payment [4] is probably the earliest work on designing incentive protocol for P2P networks. It relies on a centralized server and uses virtual currency to provide incentive for resource sharing. Since then, much efforts are focused on incentive mechanisms for P2P systems [7]–[9] and wireless networks [1], [15]. Authors in [12] show that shared history based incentives can overcome the scalability problem of private history based mechanisms and one can use DHT to implement the shared history incentive mechanism. One shared history based incentive is the reciprocative strategy [2], [6]. It 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 examples and are studied via simulation or via small scale prototyping. This paper focus on the analytical framework to analyze the robustness and properties on adaptive incentive protocols with learning capability.

There are some existing work on designing a particular incentive mechanism. Authors in [3] assume that each peer has a fixed strategy with a certain distribution while we assume peers can adapt their strategies. In [14], 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 dynamics via the learning behavior in P2P systems.

IX. Conclusion

In this paper, we present a general mathematical framework to model and evaluate the evolution and performance of incentive policies. Peers are assumed to be rational and are able to learn about the behavior of other peers. We compare 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. We introduced distributed learning mechanism into incentive protocol analysis and study two learning models. We show that the current-best learning is less robust than the opportunistic learning, and altruism may have detrimental impact on the system.

We show that learning mechanism is worthwhile and one may consider introducing a layer of software in the P2P system so as to provide such learning mechanism to users. This way, the software can guide the users in the decision process and helps the P2P system to quickly converge to the desirable operating point. Nevertheless, what is the optimal learning rate and how to promote trustworthy information exchange are interesting and we will consider them in our future work.

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 '04*.
- [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 '05*. ACM.
- [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. pages 139–185, 2006.
- [14] F. Wu and L. Zhang. Proportional response dynamics leads to market equilibrium. In *ACM STOC*, 2007.
- [15] S. Zhong, J. Chen, and Y. Yang. Sprite: a simple, cheat-proof, credit-based system for mobile ad-hoc networks. In *INFOCOM, IEEE*, 2003.
- [16] Y. Zhou, D. M. Chiu, and J. Lui. A simple model for analyzing p2p streaming protocols. *ICNP 2007*.