2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

User Adaptation and Long-run Evolution in Online Communities

Yu Zhang, *Member, IEEE*, Mihaela van der Schaar, *Fellow, IEEE*

Abstract—**This paper focuses on analyzing the interactions emerging between users in online communities. Network utility maximization and other methods are not effective when the communities are composed of intelligent and self-interested users (multimedia social communities, social networks etc.), because the interests of the individual users may be in conflict. In our prior work, we propose to design protocols in a stationary community to provide users incentives to voluntarily operate according to pre-determined social norms and provide services. In this paper, we extend the study to analyze the interactions of self-interested users under a social norm in an online community of finite population, where the stationary property of the community does not hold. To optimize their long-term performance based on their knowledge, users adapt strategies to play their best response by solving individual stochastic control problems. Understanding the evolution of a community provides protocol designers guidelines for designing social norms in which no user will have the incentive to adapt and deviate from the prescribed protocol, which in turn encourages cooperative behavior among users and achieves the optimal social welfare of the community.**

I. Introduction

The proliferation of social networking services has permeated our social and economic lives and created online social communities where individuals interact with each other. However, online communities in general rely on voluntary contribution of services by the individual users, and are therefore vulnerable to intrinsic incentive problems which lead to prevalent free-riding behaviors among users, at the expense of the collective social welfare of the community [1][2][4].

The incentive mechanisms that have been proposed to encourage cooperation in online communities mainly rely on the idea of reciprocity, which use differential services to provide incentives and can be further classified into direct reciprocity and indirect reciprocity [3][10][14]. In our prior work [9], we have developed an indirect reciprocity framework based on social norms. As pointed out in [5], in such highly interconnected online communities there is a natural tendency for users to adapt their strategies based on their interactions with the environment to maximize their long-term utility. For example, in P2P networks, a peer might provide upload services voluntarily if there are many peers who reciprocate by providing services in return; whereas, in a network where free-riders are the majority, the peer will choose to change its sharing behavior and adopt a more selfish strategy. Therefore, it is of critical importance to understand the users' adaptation behavior and how it influences the long-term evolution of the community, which can provide essential insights to facilitate the design of incentive mechanisms that can improve the efficiency and survivability of online communities.

The study of adaptation processes of rational users has attracted much attention in evolutionary game theory [6][7][10]. In these studies, players are bounded-rational and not necessarily able to guess the other players' choices correctly. Moreover, they adjust their choices of strategies over time as they observe on the other players' choices.

In this paper, we extend the study above to analyze the interactions of self-interested users under a social norm in an online community. Given the social norm implemented by the protocol designer of the community, we study how self-interested users dynamically adapt their service strategies to play the best response depending on their observations. We formalize the adaptation process of a self-interested user as a Markov Decision Process (MDP), and first prove that users' best response adaptation is always harmful to the community because it decreases the social welfare. In addition, this impact becomes more severe in communities with small populations. Then, we present guidelines for the design of social norms in which no user will have the incentive to adapt and deviate from the social strategy. Next, we study the stochastic stability of the community by analyzing its evolution when the operation error goes to zero. The long-run behavior of the community is characterized and we prove that the community will in most of cases converge to a unique stationary distribution. Understanding how a community evolves over time and understanding its long-run behavior can help protocol designers of online communities select appropriate social norms based on the community characteristics, in order to encourage cooperative behavior among users and maximize the social welfare of the community.

The remainder of this paper is organized as follows. In Section II, we introduce our proposed social norm based framework for indirect reciprocity in online communities. In Section III, we study users' adaptation behavior and the long-run evolution of the community. Section IV presents our experimental results and the conclusions are drawn in Section V where directions for future research are also outlined.

II. System Model

A. Repeated Game Formulation

We consider a social community consisting of N users.

The authors are with Department of Electrical Engineering, University of California, Los Angeles. (e-mail: yuzhang@ucla.edu, mihaela@ee.ucla.edu).

The community is modeled as a discrete-time system with time divided into periods. At the beginning of a period, each user selects another user to request a service [4] and receives a service request from one user [1]. The matching is uniformly random such that all users in the community have an equal probability to be chosen by a particular user.

Similar to [9], we model the interaction between a pair of matched users as a one-period asymmetric gift-giving game [5]. The user who requests services is called a *client* and the user whose services are requested is called a *server*. Upon receiving the request, the server selects its action, i.e. the level of contribution to the client. We assume that the server's choice of actions is binary [2], as $z \in \mathcal{Z} = \{0,1\}$, where $z = 1$ indicates that the server provides the requested service to the client and $z = 0$ indicates that the server refuses to provide the service. Users' utilities in a one-period game are determined by the server's action. When the server provides a service by choosing $z = 1$, it consumes a cost of c , and the client gains a benefit of b by receiving the requested service; whereas both users receive a utility of 0 when the server chooses $z = 0$.

In the repeated game, each user is tagged with a reputation $\theta \in \Theta = \{0,1,2,\cdots,L\}$ representing its social status. The service strategy that a user adopts in the repeated game is reputation-based and represented as a mapping $\sigma : \Theta \to \mathcal{Z}$, where each term $\sigma(\tilde{\theta}) \in \mathcal{Z}$ is the contribution level of this user when it is matched with a client of reputation $\tilde{\theta} \in \Theta$. We consider a finite set of threshold-based strategies in this paper. That is, a behavioral strategy $\sigma \in \Gamma$ can be characterized by a service threshold $h_\sigma \in \{0,1,\cdots,L+1\}$. A user adopting ϕ only provides service to the users whose reputations are larger than or equal to h_σ .

We consider a social norm κ that consists of a social strategy and a reputation scheme. A social strategy $\phi : \Theta \times \Theta \to \mathcal{Z}$ represents the approved behavior of the server within the community, while a reputation scheme updates the reputations of users depending on their past actions as servers. As an example, we consider social rules which satisfy the following property in the design of protocols:

$$\phi\left(\theta,\tilde{\theta}\right) = \begin{cases} 1 & if\ \theta \geq h\ and\ \tilde{\theta} \geq h \\ 1 & if\ \theta < h \\ 0 & otherwise \end{cases} . \qquad (1)$$

As a result, we have the service threshold for servers of a reputation θ to be $h_\theta = 0$ when $\theta < h$ and $h_\theta = h$ when $\theta \geq h$. Here, we restrict our attention on $h \in \{1,\cdots,L\}$ without loss of generality. Hence, differential services are provided to users of different reputations according to the value of h , which is called the *social threshold* for convenience. Particularly, users of reputation less than h are called as bad users and users of reputation no less than h are called as good users.

After a server takes an action, its client reports the action to some trustworthy third-party managing device in the community (e.g. the tracker in P2P networks). In practice, a system is continually being subjected to small perturbations that arise due to various types of operation errors in the community. To formalize the effect of such perturbations, we assume that the client's report is subject to a small error probability ε of being reverted. That is, $z = 0$ is reported to the managing device with probability ε while the server actually plays $z = 1$, and vice versa. Formally, a reputation scheme is a mapping $\tau : \Theta \times \Theta \times \mathcal{Z} \to \Theta$, in which $\tau(\theta,\tilde{\theta},z_o)$ is the reputation of the server in the next period given its current reputation θ , the client's reputation $\tilde{\theta}$, and the server's reported action $z_o \in \mathcal{Z}$. As an example, we consider the following simple reputation scheme in this paper:

$$\tau(\theta,\tilde{\theta},z_o) = \begin{cases} \min\{L,\theta+1\} & if\ z_o = \phi(\theta,\tilde{\theta}) \\ 0 & if\ z_o \neq \phi(\theta,\tilde{\theta}) \end{cases} . \qquad (2)$$

B. Utility Function

Due to their limited processing capabilities, users can only form simple beliefs about the strategies deployed by other users in the community [12]. We assume that each user maintains a belief that all users of reputation higher than 0 will follow the social strategy ϕ and all users of reputation 0 will provide no service. Let $\mu = \{n(\theta)\}_{\theta=0}^{L}$ denote the community configuration, in which $n(\theta)$ represents the number of users of reputation θ in the community. Note that since a user can never be matched with itself, it is sometimes more convenient to compute the expected utility of a particular user of reputation $\hat{\theta}$ by employing the configuration of all users other than itself, which is called the opponent configuration and is denoted as $\eta = \{m(\theta)\}_{\theta=0}^{L}$ with $m(\theta) = n(\theta)$ for all $\theta \neq \hat{\theta}$ and $m(\theta) = n(\theta) - 1$ if $\theta = \hat{\theta}$. We can thus compute the expected one-period utility $v_\kappa(\sigma,\theta,\mu)$ of a user of reputation θ and following the strategy ϕ as:

$$v_{\kappa,\phi}(\theta,\mu) = \frac{1}{N-1}\sum_{\tilde{\theta}\neq 0} m_\theta(\tilde{\theta})\left[b(\tilde{\theta},\theta,\phi(\tilde{\theta},\theta)) - c(\theta,\tilde{\theta},\sigma(\tilde{\theta}))\right] - \frac{1}{N-1} m_\theta(0)c(\theta,0,\sigma(0)) \qquad (3)$$

Here $b(\tilde{\theta},\theta,\phi(\tilde{\theta},\theta))$ is the one-period benefit which this user

[1] The analysis can be extended to a general case when the service request rate per period is a positive real number. It is assumed to be 1 in this manuscript only for illustration purposes.

[2] The framework in this manuscript is also applicable to the case when the action space of servers is not binary.

can receive when its matched server has a reputation $\tilde{\theta}$ and follows ϕ; $c(\theta,\tilde{\theta},\sigma(\tilde{\theta}))$ is the one-period cost of this user when its matched client has a reputation $\tilde{\theta}$. The expected utility if a user following ϕ is compactly denoted as $v_\kappa(\theta,\mu)$.

A user's long-term utility in the repeated game is evaluated with the infinite-horizon discounted sum criterion as

$$v_\kappa^\infty(\sigma^{t_0},\theta^{t_0},\mu^{t_0}) = E\{\sum_{t=t_0}^{\infty}\delta^{t-t_0}v_\kappa(\sigma^t,\theta^t,\mu^t)\}, \qquad (4)$$

where $\delta \in [0,1)$ is a user-defined discount factor representing the user's preference to its future utility.

Under a social norm κ, the best response for a self-interested user in any period t_0 is

$$\sigma^* = \arg\max_\sigma v_\kappa^\infty(\sigma,\theta^{t_0},\mu^{t_0}). \qquad (5)$$

Given the best response dynamics, we are interested in whether this community configuration may converge in the long-run, i.e. at equilibrium, each self-interested user holds a fixed reputation and plays a fixed strategy after a sufficiently long time. This provides the protocol designer with guidelines for selecting the correct social norm based on the community characteristics, e.g. the utility structure (b,c) and the discount factor δ, in order to optimize the sharing efficiency in the community. Particularly, we focus on the *stochastic stability* of the community as defined in [8] with $\varepsilon \to 0$, which approximates the situation where the operation error happens infrequently in the community. Generally speaking, a community configuration μ is a *stochastically stable equilibrium* if, in the long-run, it is nearly certain that the community lies within every small neighbourhood of μ as the operation error ε approaches slowly to 0 [8].

III. User Adaptation and The Community's Long-run Evolution

In this section, we analyze the evolution of the community. Each user has a probability $\gamma \in (0,1)$, which is called as its adaptation rate, to adapt its strategy at the beginning of each period. The optimization can be formalized as a Markov Decision Process (MDP) [15].

A. The MDP Formulation and The Optimal Policy

The MDP to be solved by a user is formalized below.

State: The state of a user is its reputation and opponent configuration, which is defined as $s = (\theta,\eta) \in \mathcal{S}$.

Action: The action of a user is the service threshold of its serving strategy $\sigma \in \Gamma$, as $a \in \mathcal{A} = \{0,1,\cdots,L+1\}$.

Reward function: The one-period reward function $r(s,a)$ is defined as the expected one-period utility as in (3). Similarly, the long-term reward function $R(s)$ of a user is defined as the discounted sum $R(s) = \sum_{t=0}^{\infty}\delta^t r(s^t,a^t)$.

Policy and Value function: The solution of the MDP is a policy $\pi : \mathcal{S} \to \{0,1,\cdots,L+1\}$, which maps each state to a service threshold. The value function is thus defined as the expected long-term reward under a policy π,

$$V^\pi(s) = E\{r(s,a)\big|\pi\} + \delta\sum_{s'} p(s' \mid s,\pi)V^\pi(s'). \qquad (6)$$

The above MDP can be solved using common computation methods such as value iteration [15], with the resulting optimal policy and value function being $\{\pi^*(s)\}$ and $\{V^*(s)\}$.

First, it can be shown that the best response of a user in any period is always above the service threshold that is specified in the social strategy, i.e. $\{h_\theta\}$, regardless of the user's reputation θ and the opponent configuration η. This is proved in the lemma below by contradiction with the basic idea as follows. If $\pi^*(\theta,\eta) < h_\theta$ for some θ, it implies that when the reputation is θ, the user chooses to provide more services to the community than what is required by the social strategy and thus consumes more service cost in the current period. On the contrary, the user also gets a higher probability to be punished by the social norm for deviating from ϕ [3], and hence a lower expected future utility compared to what it can receive by following ϕ. It follows that the user will receive a lower long-term utility in this case than by following ϕ, which contradicts the fact that $\pi^*(\theta,\eta)$ is the best response to maximize a user's expected long-term utility.

Lemma 1. $\pi^*(\theta,\eta) \ge h_\theta$, for any $\theta \in \Theta$ and η.

Proof: See [13]. ■

Moreover, it can be shown that under a constant opponent configuration, the service threshold of a bad (good) user's best response monotonically decreases against the reputation. Formally, this monotonicity can be represented as $\pi^*(\theta_1,\eta) \ge \pi^*(\theta_2,\eta)$, if $\theta_1 < \theta_2 < h$ or $h \le \theta_1 < \theta_2$. The average service that has to be provided by a bad (good) user in one period is constant regardless of its reputation, and only depends on the current opponent configuration and its selection of the sharing strategy. On the contrary, the average service that a bad (good) user expects to receive monotonically increases with its reputation. Therefore, given a fixed opponent configuration and a fixed strategy, a bad (good) user will always obtain a higher long-term utility with a higher reputation and thus will have less incentive to deviate from this strategy.

[3] The social norm does not only punish users who do not provide services as required. If a user provides service to another user who is supposed to be punished by the social norm, this user itself will also be punished.

Lemma 2. $\pi^*(\theta_1,\eta) \geq \pi^*(\theta_2,\eta)$ if $\theta_1 < \theta_2 < h$ or $h \leq \theta_1 < \theta_2$.

Proof: See [13]. ■

Thus, it can be concluded from Lemma 1 and 2 that users who play the best response are always harmful to the community.

B. The Community's Long-run Evolution and Stochastic Stability

In this section, we examine the evolution of the community under the best response dynamics. We define the strategy configuration of the community as $\{\pi_\mu^*(\theta)\}_{\theta=0}^L$. We first show the community configuration μ cannot be stabilized in the long-run with a positive operation error ε. A Markov chain analysis is employed to illustrate this result. It has been proved in [13] that the community configuration evolves as a Markov chain on the finite space

$$\mathcal{U} = \left\{ \begin{array}{l} \mu = (n(0), n(1), \ldots, n(L)) \\ \mid n(\theta) \in \mathbb{N} \ for\ 0 \leq \theta \leq L, and\ \sum_{\theta=0}^{L} n(\theta) = N \end{array} \right\} \quad (7)$$

whose size is $|\mathcal{U}| = \begin{pmatrix} N+L \\ L \end{pmatrix}$. The transition probabilities are given by $p_{\mu\mu'} = p(\mu' \mid \mu)$ between any two configurations μ, $\mu' \in \mathcal{U}$ and $P = [p_{\mu\mu'}]$ is the transition matrix. Note that with $\varepsilon > 0$, all entries in P are positive. It is well-known then that this Markov chain is irreducible and aperiodic, which introduces a unique stationary distribution. Let

$$\Delta_{|\mathcal{U}|} \equiv \left\{ \begin{array}{l} \boldsymbol{q} \in \mathbb{R}^{|\mathcal{U}|+1} \mid \\ q_i \geq 0 \ for\ i = 0,1,2,\cdots,|\mathcal{U}|\text{-}1 \ and\ \sum_i q_i = 1 \end{array} \right\} \quad (8)$$

be the $|\mathcal{U}|$-dimensional simplex. Any $\boldsymbol{q} \in \Delta_{|\mathcal{U}|}$ represents a probability distribution on the set of configurations. Indexing all configurations in $\mathcal{U}$ from 0 to $|\mathcal{U}|$, q_i then represents the frequency that the community stays at the i-th configuration in the long-run. A stationary distribution is a row vector $\boldsymbol{\omega}(\varepsilon) = (\omega_0, \omega_1, \ldots, \omega_{|\mathcal{U}|}) \in \Delta_{|\mathcal{U}|}$ satisfying $\boldsymbol{\omega}(\varepsilon) = \boldsymbol{\omega}(\varepsilon)P$. When P is strictly positive, not only $\boldsymbol{\omega}(\varepsilon)$ exists and is unique, it also preserves stability and ergodicity.

Moreover, as all entries in P are positive, we have $q_i > 0$ for all $0 \leq i \leq |\mathcal{U}|$, whose value only depends on ε and κ. Therefore, when time goes to infinity, the community will spend a positive fraction of time in all possible configurations in $\mathcal{U}$. We can thus conclude that the community will oscillate between different configurations and will never converge to a unique one in the long-run when $\varepsilon > 0$.

Proposition 1. When the operation error $\varepsilon > 0$, the community never converges to a unique community configuration, regardless of the initial configuration it starts with.

Proof: See [13]. ■

Next, we analyse the stochastic stability of the community when $\varepsilon \to 0$. To facilitate the analysis, we first define the concept of the limiting distribution.

Definition 1. The limiting configuration distribution of the community is defined by $\bar{\boldsymbol{\omega}} = \lim_{\varepsilon \to 0} \boldsymbol{\omega}(\varepsilon)$.

The existence and the uniqueness $\bar{\boldsymbol{\omega}}$ can be proved as in [6], which is omitted here. Hence, a community configuration indexed as i in $\mathcal{U}$ is a stochastically stable equilibrium if and only if $\bar{\omega}_i > 0$.The following proposition proves that a stochastically stable equilibrium should be composed only by the population of reputations 0 and L.

Proposition 2. When the community converges to a stochastically stable equilibrium in the long-run, it does not have positive populations of reputations other than 0 and L. That is, if $\bar{\mu} = \{\bar{n}(0), \cdots, \bar{n}(L)\}$ is a stochastically stable equilibrium, it should preserve the following property

$$\bar{n}(\theta) = 0 \text{, for all } \theta \in \{1, \cdots L-1\}. \quad (9)$$

Proof: See [13]. ■

Proposition 2 characterizes the property that a stochastically stable equilibrium preserves. However, since there are usually multiple configurations that satisfy (9) as well as the incentive constraints, which configuration the system converges to is still left undetermined. In the following proposition, we further refine the set of stochastically stable equilibrium using the idea of "basin of attraction" [6], and provide the condition with which there is a unique stochastically stable equilibrium. To prove this, we need to characterize the structure of the strategy configuration $\{\pi_\mu^*(\theta)\}$ first.

Lemma 3. $\pi_\mu^*(\theta_1) \geq \pi_\mu^*(\theta_2)$ if $\theta_1 < \theta_2 < h$ or $h \leq \theta_1 < \theta_2$.

Proof: See [13]. ■

Using the result of Lemma 3, we are now able to prove the following proposition, which characterizes the community configuration in the long-run.

Proposition 3. We have $\bar{n}(L) = N$ in the long-run if and only if $\delta > \frac{c}{b}$ and $h < H$, where H is the solution of the following equation

$$\delta^H b - \delta^{H-1} c = (1-\delta^H) cH \quad (10)$$

Proof: See [13]. ■

We briefly explain the above three conditions. Condition (1) is intuitive. By deviating from the social norm, the maximum gain on a user's one-period utility is the saving of its immediate service cost, which is proportional to c, at the loss of its future utility which is proportional to b. Hence, when c/b decreases, the current gain of a user also decreases with its future loss increasing at the same time, which provides stronger incentives for a user to comply with the social rule. A similar analysis can be applied to Condition (2). The discount factor δ adjusts the weights that a user places on its current and future utilities. With a larger δ, a user puts a higher weight on its future utility, and thus becomes more interested in increasing its reputation and obtaining a higher future utility rather than deviating to save its immediate service cost. As a result, the incentive for a user to comply with the social rule increases. Condition (3) contradicts the traditional opinion that a user's incentive will increase when the punishment in the protocol becomes more severe. In our framework, this statement is correct only for good users. As outlined by the proof, when the punishment is too severe, which is represented here by a high social threshold in the social rule, a bad user has to wait a long period of time to recover its reputation (i.e. becoming a good user), which harms its incentive to comply with the social rule. This prohibits the protocol designer from increasing h arbitrarily.

IV. EXPERIMENTS

In our experiments, we simulate the adaptation behavior of $N = 1000$ users. The setting are as follows: $L = 3$, $c = 1$, $\delta = 0.5$, $h = 2$. We run the experiment for 10^8 periods and measure the average reputation distribution over every 2.5×10^7 periods. This can be used as an approximation on the community configuration in the long-run and the results are illustrated in Table 1, which shows the community configurations when $\varepsilon = 0.2$ and 0.05, respectively. The community configuration oscillates when $\varepsilon = 0.2$ and cannot converge to a unique stationary point. Hence, the average reputation distribution changes after every periods, with each reputation taking a positive fraction in the population in the long-run. When $\varepsilon = 0.05$, the community configuration converges to the stochastically stable equilibrium. As a result, almost all users are of reputations 0 and L in the long-run, with users of reputation 0 providing no service and users of reputation L acting in a mutually cooperative manner with each other.

Table 1 also shows how the service benefit b impacts the stochastically stable equilibrium. As expected, there are only users of reputations 0 and L in the community after sufficiently long time. When $b = 1.5$, the instant saving of the service cost outweighs the future benefit of obtaining a high reputation and hence, a majority of users will maintain the lowest reputation 0. When $b = 5$, it is more attractive to obtain a high reputation so as to receive higher future service benefit, and thus all users converge to reputation L with $n(L) \approx N$ in the community. Hence, our result verifies the conclusion in Proposition 3 that $\overline{n}(L) = N$ in the long-run.

TABLE I. THE EVOLUTION OF THE COMMUNITY IN 10^8 PERIODS

Periods	2.5×10^7	5×10^7	7.5×10^7	1×10^8
$\varepsilon = 0.05$ and $b = 3$				
$\theta = 0$	34%	25%	23.3%	23%
$\theta = 1$	7%	2.5%	1.2%	1%
$\theta = 2$	13%	2.5%	1.2%	1%
$\theta = 3$	46%	70%	74.3%	75%
$\varepsilon = 0.2$ and $b = 3$				
$\theta = 0$	18%	26%	41%	18%
$\theta = 1$	47%	19%	28%	23%
$\theta = 2$	25%	14%	7%	32%
$\theta = 3$	10%	55%	24%	27%

Periods	2.5×10^7	5×10^7	7.5×10^7	1×10^8
$\varepsilon = 0.05$ and $b = 1.5$				
$\theta = 0$	52%	49%	54.7%	56%
$\theta = 1$	7%	3%	2%	0.5%
$\theta = 2$	9%	3%	2%	0.5%
$\theta = 3$	32%	45%	41.3%	43%
$\varepsilon = 0.2$ and $b = 1.5$				
$\theta = 0$	11%	7%	2%	2%
$\theta = 1$	2%	1.5%	0.5%	0.5%
$\theta = 2$	2%	1.5%	0.5%	0.5%
$\theta = 3$	85%	90%	97%	27%

Since the fraction of users of reputation L in the stochastically stable equilibrium depends on the ratio of $\frac{\delta^h b}{c}$, similar results as those in Table 1 can be expected when we change δ and c.

In the final part of the experiment, we assume that the community characteristics, e.g. the utility structure (b,c) and the discount factor δ are not fixed but vary over time and consider how such variation will impact the long-run evolution of the community. As an example, we use b as the representative variable to plot the result and assume it varying over time following a Gaussian distribution with the mean $\overline{b} = 3$ and the variance $\sigma^2 = 0.01$. Fig. 1 depicts the social welfare of the community over time. We consider two selections on the social strategy as $h = 1$ and $h = 2$. In both cases, the social welfare when b is variable (solid lines) is smaller than the social welfare when b is constant (dotted lines). This is due to the fact that our designed protocol only guarantees users sufficient incentive to comply with the social norm when b is at its mean value 3. When b deviates from its mean value, users might have incentive to adapt to the best response and deviate from the social norm. In addition, since most users maintain reputation L in the stochastically stable equilibrium when $\overline{b} = 3$ and $\delta = 0.5$, the social norm with $h = 2$ can provide larger incentives for users of reputation L to follow the social strategy. Therefore, it is more robust against the variation on b, which maintains $n(L)$ at a higher level than the social norm with $h = 1$. As a

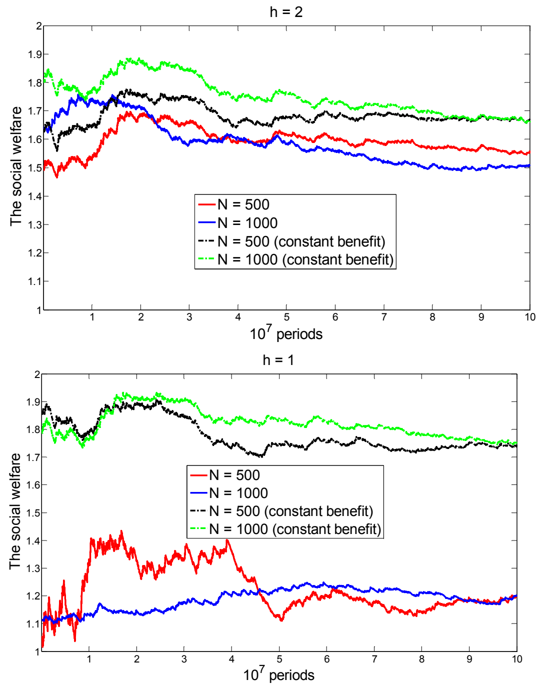

Fig. 1. The evolution of the social welfare in the long-run when the stage game benefit varies over time ($L = 3, \delta = 0.5, \varepsilon = 0.05$, $c = 1$)

result, the social norm with $h = 2$ delivers higher social welfare for both $N = 500$ and $N = 1000$ when b is variable.

V. Conclusion

We have studied the problem of designing social norm based protocols for online communities and analyzed the adaptation behavior of users under such protocols. Knowledge on the evolution of the community in the long-run can facilitate the protocol designers to design protocols which achieve efficient social welfare. Our framework can be extended in several directions, among which we mention four. First, users in the community do not necessarily need to be homogeneous as discussed in Section III. Different users can have different benefits and costs for the service received/provided. Also, they can choose different discount factors δ when evaluating the long-term utility. The discount factor that a user chooses can be dynamically adjusted over time depending on its own expected lifetime in the community. It is an interesting problem of analyzing how the user heterogeneity impacts the design of efficient protocols. Second, clients can use more complicated decision rules while reporting the servers' actions to the community manager in order to maximize their own long-term utility, instead of always reporting truthfully. Third, online communities may be subject to practical constraints such as topological constraints, in which users can only observe the local information and different users at different locations do not necessarily share the same community information. Hence, the analysis in this paper needs to be extended to scenarios where users adapt based on partial and heterogeneous information. Finally, users adopt a simple belief model. However, a more sophisticated belief model can be introduced into our framework such that users can update their beliefs on others based on their observation. For example, the formation of user beliefs and opinions in social networks are extensively studied in [16] and [17]. Understanding how the evolutions of user belief and user strategy will impact each other will be an appealing direction.

References

[1] E. Adar, B. A. Huberman, "Free Riding on Gnutella," *First Monday*, vol. 5, No. 10, Oct. 2000.

[2] S. Guo, Y. Gu, B. Jiang, and T. He, "Opportunistic Flooding in Low-Duty-Cycle Wireless Sensor Networks with Unreliable Links," *15th Annual Int'l Conf. on Mobile Computing and Networking (MobiCom)*, 2009

[3] B. Cohen, "Incentives Building Robustness in BitTorrent," *Proc. P2P Econ. Workshop*, Berkeley, CU, 2003.

[4] M. Feldman, K. Lai, I. Stoica, and J. Chuang. "Robust Incentive Techniques for User-to-User Networks," *Proc. ACM Conf. on Elec. Commerce*, no. 4, pp. 102 – 111, 2004.

[5] P. Johnson, D. Levine, W. Pesendorfer, "Evolution and Information in a Gift-Giving Game," *Journal of Econ. Theory*, vol. 100, no. 1, pp 1-21, 2001.

[6] M. Kandori, G. Mailath, and R. Rob, "Learning, Mutation, and Long-run Equilibria in Games," *Econometrica*, 1993, vol. 61, no. 1, pp. 29 – 56, 1993.

[7] M. Kandori, R. Rob, "Evolution of Equilibria in the Long-run: A General Theory and Applications," *Journal of Economic Theory*, vol. 65, no. 2, pp. 383 – 414, 1995.

[8] D. Foster, P. Young, "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, vol. 38, no. 2, pp. 219 – 232, 1990.

[9] Y. Zhang, J. Park, M. van der Schaar, "Social Norm for Online communities," *arXiv:1101.0272 [cs.SI]*.

[10] KaZaA User-to-User File Sharing CU, [Online]. Available: http://www.kazaa.com/, [Online]. Available

[11] G. J. Mailath, L. Samuelson, "Repeated Games and Reputations: Long-Run Relationships," *Oxford University Press*, 2006.

[12] M. Nowak, K. Sigmund, "Evolution of Indirect Reciprocity," *Nature* 437, pp. 1291 – 1298, 2005.

[13] Y. Zhang, M. van der Schaar, "Strategic Learning and Robust Protocol Design for Online Communities with Selfish *Users," UCLA Technical Report*, http://medianetlab.ee.ucla.edu/.

[14] B. Skyrms, R. Pemantle, "A Dynamic Model of Social Network Formation," *Adaptive Networks (Understanding Complex Systems)*, pp. 231 – 251, 2009.

[15] R. Sutton, A. Barto, "Reinforcement Learning: An Introduction," *MIT Press*, Cambridge, MA, 1998.

[16] D. Acemoglu, A. Ozdaglar, "Opinion Dynamics and Learning in Social Networks," MIT LIDS report 2851.

[17] V. Krishnamurthy, "Quickest Time Herding and Detection for Optimal Social Learning," *arXiv:1003.4972 [cs.IT]*.