Social Norm Design for Information Exchange Systems with Limited Observations

Jie Xu, *Student Member, IEEE,* Mihaela van der Schaar, *Fellow, IEEE,*

***Abstract*—Information exchange systems, such as BitTorrent, Yahoo Answers, Yelp, Amazon Mechanical Turk, differ in many ways, but all share a common vulnerability to selfish behavior and free-riding. In this paper, we build incentives schemes based on social norms. Social norms prescribe a social strategy for the agents in the system to follow and deploy reputation schemes to reward or penalize agents depending on whether they follow or deviate from the prescribed strategy when selecting actions. Because agents in these systems often have only limited capability to observe the global system information, e.g. the reputation distribution of the agents participating in the system, their beliefs about the reputation distribution are heterogeneous and biased. Such belief heterogeneity causes a positive fraction of agents to not follow the social strategy. In such practical scenarios, the standard equilibrium analysis deployed in the economics literature is no longer directly applicable and hence, the system design needs to consider these differences. To investigate how the system designs need to change, we focus on a simple social norm with binary reputation labels but allow adjusting the punishment severity through randomization. First, we model the belief heterogeneity using a suitable Bayesian belief function. Next, we formalize the agents' optimal decision problems and derive in which scenarios they follow the prescribed social strategy. Then we study how the system state is determined by the agents' strategic behavior. We are particularly interested in the robust equilibrium where the system state becomes invariant when all agents strategically optimize their decisions. By rigorously studying two specific cases where agents' belief distribution is constant or is linearly influenced by the true reputation distribution, we prove that the optimal reputation update rule is to choose the mildest possible punishment. This result is further confirmed for more sophisticated belief influences in simulations. In conclusion, our proposed design framework enables the development of optimal social norms for various deployment scenarios with limited observations.**

***Index Terms*—game theory, limited observations, reputation.**

I. INTRODUCTION

AS THE WEB has evolved, it has become increasingly social. People turn to the web to exchange ideas, data and services, as evidenced by the popularity of sites like Wikipedia, Bit-Torrent, Yahoo Answers, Yelp and Amazon Mechanical Turk (AMT). While these systems, which we refer to as *information exchange systems*, differ in many ways, they share a common vulnerability to selfish behavior and free-riding. For example, a worker on AMT may attempt to complete jobs with as little effort as possible while still being paid; an agent in a peer-to-peer system may wish to download files without using bandwidth to upload files for others. In order for these sites to thrive, participants must be properly motivated to contribute.

Distributed optimization techniques have been applied extensively in engineering to enable the efficient usage of resources by obedient or cooperative agents. Only in recent years have engineers started to investigate incentive issues in systems formed by self-interested agents. Many of the existing mechanisms to combat free-riding problems rely on game-theoretic approaches and can be classified as either *pricing* mechanisms or *reciprocity* mechanisms. Pricing mechanisms are appropriate in some settings, but do not make sense for applications like Yahoo Answers, Wikipedia, or Yelp, where much of the appeal is that the information is free.

Under a reciprocity mechanism, an agent is rewarded or punished based on its behavior in the system according to a differential service scheme [1] [2]. This preferential treatment provides an incentive for agents to cooperate, and can be implemented using either virtual currency [3] [4] [5] or reputation. However, prior work shows that even optimal designs based on virtual currency cannot achieve optimal performance [6]. Depending on how an agent's reputation is generated, reciprocity-based protocols can be classified as direct reciprocity mechanisms [7], or indirect reciprocity mechanisms [8]. Direct reciprocity implies that the interaction between two agents is influenced only by the history of their mutual interactions, and not by their interactions with other agents. Though easy to implement, direct reciprocity requires frequent interactions between two agents in order to establish accurate mutual ratings. This is restrictive in systems characterized by high churn, asymmetry of interests, or infrequent interactions between any pair of agents, such as most peer production systems, online labor markets, and review sites.

Protocols that are based on indirect reciprocity typically assign to each agent a global reputation [9] based on its past interactions with all other agents in the system. A differential service scheme recommends actions based only on the reputations of agents, and not on their entire history of interactions. Much of the existing work on reputation mechanisms is concerned with practical implementation details. Some focuses on effective information gathering techniques [10] [11]. Many empirical studies have been done on the impact of reputation on a seller's prices and sales [12] [13]. The few works [14] [15] [16] providing theoretical results typically consider one (or a few) long-lived seller(s) interacting with many short-lived buyers. This model and the corresponding analysis are not appropriate for information exchange systems where there are many interacting agents playing the role of buyer or seller or both, contributing and seeking information.

Manuscript received 15 December 2011; revised 1 June 2012.

The authors are with the Department of Electrical Engineering, University of California Los Angeles (UCLA), Los Angeles, CA, 90095 USA (e-mail: jiexu@ucla.edu, mihaela@ee.ucla.edu; phone: (310)825-5843; fax: (310)206-4685).

Digital Object Identifier 10.1109/JSAC.2012.121205.

To rigorously capture the impact of various strategy and protocol design choices, a framework based on social norms was proposed in [17] which was originally designed to sustain cooperation in a community with a large population of individuals participating in anonymous random matching games [18] [19] [20] [22]. In an incentive scheme based on a social norm, each individual is assigned a dynamic label indicating its reputation or status based on past behavior, and individuals with different labels are treated differently by others in the system. Hence, a social norm can be adopted easily in social communities with an infrastructure that collects, processes, and delivers information about individuals' behavior. Most of the existing works on reputation systems assume that agents have complete knowledge of the system, i.e. the reputation distribution in particular. However, this assumption is unrealistic in practice. First, none of the major existing online information exchange systems (e.g. Yahoo! Answers, Yelp) provide reputation statistics to agents in real-time. Second, even if the reputation statistics are provided online and are accessible by the agents, the agents may not believe their accuracy and may not rely on them to make decisions. People tend to form their beliefs based on their own experience or experiences of "trusted" people. Hence, agents have only incomplete information about the system and their beliefs are heterogeneous.

In this paper, we build incentive schemes for information exchange systems based on social norms, which explicitly consider these scenarios. The main contributions of this paper are summarized as follows:

- We design simple social norms for information exchange systems using binary reputation and adjust the punishment severity through randomization. This class of social norms is simple and easy to implement while being close to the optimal strategy in the unlimited observations case [17].
- We model the agents' heterogeneous beliefs of the system state (i.e. the reputation distribution) based on their limited observations using a Bayesian belief model. We require the equilibrium to be robust to small disturbances and prove the existence of robust equilibrium and we study what are the conditions under which the equilibrium exists, as well as how its efficiency depends on the agents' observations.
- We prove that agents follow the social strategy only if their beliefs about the system state are above certain thresholds, i.e., they need to have sufficient "trust" in the society. These thresholds are analytically computed and depend on the system characteristics (e.g. discount factor, benefit and cost). Using this result, we rigorously study the system design problem and prove that, in most scenarios, the optimal punishment design is to use the mildest possible punishment, thereby leading to a different social norm design than in the complete information cases.

We highlight the differences from existing works in the following. First, the information based on which agents make their self-interested decisions is fundamentally different from all the existing works. In the existing works on social reciprocation, agents have complete and accurate information of the system state. Alternatively, in this work, we consider more realistic scenarios where agents only acquire incomplete and inaccurate information about the system state. Second, since the information based on which agents make decisions is incomplete, the equilibrium concept is different. For example, in [17], the equilibrium is a public perfect equilibrium. However, in this paper, agents' decisions are based on their own observations and beliefs, the equilibrium concept that we use is Bayes-Nash equilibrium. We also require the equilibrium to be robust to smaller disturbances. Third, the agents' strategic behaviors at equilibrium differ significantly from the existing results where all agents follow the recommended protocol in equilibrium. In the considered model in this paper, agents have heterogeneous beliefs under any system state and hence, not all the agents will follow the recommended strategy. To better illustrate the differences, we categorize the existing works in Fig. 1.

Empirical Study	Theoretical Study		
	One long-lived agents vs. many short-lived agents	Many long-lived agents (random matching)	
		Complete information of the system state	Incomplete information of the system state
[12][13]	[14][22]	[17][18][19]	This paper

Fig. 1. Comparisons with existing works.

The rest of this paper is organized as follows. Section II describes the basic model, the structure of the social norms and the belief model based on which the agents make decisions. Section III investigates the agents' decision problem. System dynamics and the equilibrium are then studied. In Section IV, the impact of punishment on the equilibrium performance is investigated. The optimal design is derived for two specific Bayesian belief functions. Simulations are conducted in Section V followed by conclusions in Section VI.

II. System Model

A. Setup

We consider an information exchange system where agents request and provide information or resources. We utilize the widely-used continuum model (mass 1), implicitly assuming that the agent population is large and static. The system is modeled as a discrete-time system where time is divided into periods. When a requester generates a task, it is posted on the website and a provider is assigned to solve the task. We assume (as in Yelp, Yahoo Answers and etc.) that there is no price associated with the task and that the provider is the only strategic entity that needs to decide whether or not to solve the task. Upon accepting, the provider incurs a cost c to fulfill the task while the requester receives a benefit b . We assume that $b > c > 0$ to make providing the service socially valuable and denote $\gamma = b/c$ as the benefit-to-cost ratio. This is a simple gift-giving game (see Fig. 2) in which the dominant strategy for the provider is not to provide service. Incentives for providers to contribute their services can be constructed if the provider is long-lived and will also become a requester in

	Provider	
	Provide service	Not provide service
Requester	*b, -c*	*0, 0*

Fig. 2. The utility matrix of the gift-giving game.

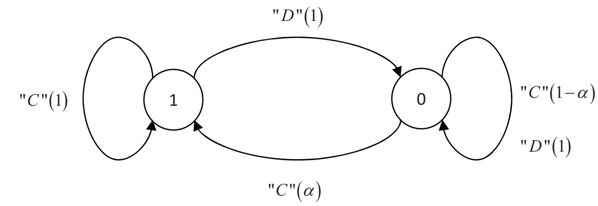

Fig. 3. Reputation update rule.

the future. We assume that agents discount the future utility by $\beta \in (0, 1)$. We assume that in each period, each agent requests a task to be solved and another agent is randomly assigned to solve this task. This random matching model is common in the economics literature [18] [19] [20]. Nevertheless, our analysis also applies in deployment scenarios where only a fraction $\lambda \in [0, 1]$ of the population generates tasks in each period and it is omitted here because our analysis will be very similar. In the considered case when $\lambda = 1$, each agent is a requester as well as a provider.

B. Punishment adjustable social norm

In the considered information exchange systems, the protocol (system) designer will design a social norm κ, which is composed of a social strategy σ, a reputation update rule τ, and a reputation set Θ. Each agent is tagged with a reputation θ representing its social status. We consider only two available reputation labels for the agents $\Theta = \{0, 1\}$ with $\theta = 1$ indicating a good status and $\theta = 0$ indicating a bad status. Denote the social strategy by the mapping $\sigma : \Theta \to \mathcal{A}$, where Θ is the reputation set of the requester and $\mathcal{A} = \{0, 1\}$ stands for the action set[1] of the provider. The action $a = 1$ represents the case where the provider offers the service while $a = 0$ when it does not provide service. Hence, the social strategy is $\sigma(1) = 1, \sigma(0) = 0$. The social strategy favors good agents because it suggests to the providers to only provide service to good requesters but not to provide service to bad requesters. At a first glance, this strategy may seem similar to the well-known Tit-for-Tat (TfT) strategy which rewards agents for cooperative behaviors and punishes them for non-cooperative behaviors. However, for the TfT strategy to work successfully it needs to be based on the history of past reciprocation of the same agents which are repeatedly interacting over time. Hence, it requires direct reciprocity between interacting agents while the social strategy proposed here is based on indirect reciprocity and it is applicable in systems where agents have infrequent interactions and are anonymous.

The social norm provides incentives to providers to adhere to the social strategy by affecting their reputation based on the action they take. Intuitively, agents who follow the social strategy should receive good reputations and those who do not should receive bad reputations. Denote the reputation update rule by the mapping $\tau : \Theta \times \Theta \times \mathcal{A} \to [0, 1]$, where $\tau(\theta, \theta', a)$ indicates the probability that the provider has a good reputation in the next period when the provider's reputation is θ, the requester's reputation is θ' and the provider takes action a. The deployed update rule is: $\tau(\theta, 0, a) = \theta, \forall a$ and

$$\tau(1, 1, a) = \begin{cases} 1 & \text{if} a = \sigma(1) \\ 0, & \text{if} a \neq \sigma(1) \end{cases}$$

$$\tau(0, 1, a) = \begin{cases} \alpha & \text{if} a = \sigma(1) \\ 0, & \text{if} a \neq \sigma(1) \end{cases}$$

Essentially, if the provider deviates from the prescribed social strategy when meeting a good requester [2], its reputation drops to 0; if the bad provider follows the prescribed social strategy, it restores a good reputation with probability $\alpha \in [0.1]$. Hence, for an agent to receive service when it becomes a requester in the future, it needs to follow the social strategy as a provider in the current period. The parameter α adjusts the severity of punishment of the social norm which needs to be designed by the system designer. Such randomization can be easily implemented by a central entity that maintains and processes agents' reputations. For $\alpha = 1$, the punishment is the mildest, allowing the bad provider to restore its reputation after a single cooperation with a good requester; for $\alpha = 0$, the punishment is the harshest, preventing the bad provider from having a good reputation again in the future no matter how this behaves; for $\alpha \in (0, 1)$, the expected time periods for which the agent remains in the bad reputation is at least $1/\alpha$. Even though we focus on a system using only binary reputation labels, randomization affects the punishment severity similar to a system using multiple (more than 2) reputation labels. We portray the aforementioned reputation update rule in Fig. 3.

C. Belief heterogeneity and trust

In this subsection, we model the agents' belief heterogeneity. Because agents are far-sighted, their decisions depend on how they evaluate the system state, i.e., the reputation distribution of the system. Since we are considering a binary reputation system, the reputation distribution can be fully described by the fraction of agents with good reputations, which we define as the *social reputation* ρ_s and use in the reminder of this paper. We consider the scenarios where agents have incomplete information of this social reputation. Agents only have inaccurate and heterogeneous beliefs about the social reputation ρ_s. In practical systems, the accurate value of ρ_s is difficult to obtain unless agents have full access to all agents' reputation information. Agents form beliefs about the social reputation according to their observations of a limited number of reputations of other agents or their limited memory

[1]The social norm can be easily extended to scenarios where actions are not binary, e.g. multiple effort levels. In that case, the reputation update rule is agent who exerts a certain level of effort drops to low reputation with a probability.

[2]We will use in the remainder of this paper the term "good/bad" users to refer to the users with good/bad reputations.

of agents' reputations with whom they have interacted in the past. In both cases, their information of the social reputation ρ_s is inaccurate and heterogeneous.

We assume that agents' belief ρ of the social reputation ρ_s follows a conditional distribution $f_{\rho|\rho_s}(\cdot)$ of the true ρ_s. This agent-specific belief can also be interpreted as the agents' own "*trust*" in the system. The belief distribution function satisfies the following properties:

1) $f_{\rho|\rho_s}(\cdot)$ has full support on [0,1]. That is, all beliefs can emerge for a given social reputation.
2) $f_{\rho|\rho_s}(\cdot)$ is continuous in ρ_s. The true ρ_s has a continuous impact on agents' beliefs.

III. System Dynamics and Equilibrium

In this section, we discuss the system dynamics and formally define the Bayes-Nash equilibrium. For this, we first need to formalize the provider's decision problem and characterize the social reputation that arises in the steady-state in our model.

A. Agent's decision problem

We begin by investigating a typical provider's decision problem based on its belief ρ. The provider's decision will be based on its own reputation θ, the requester's reputation θ' and its belief ρ toward the social reputation. The provider chooses an action $a(\theta,\theta'|\rho) \in \mathcal{A}$ to maximize its total expected discounted utility. Depending on which action the provider takes, the reputation transition follows the reputation update rule. The provider will follow the social strategy if the long-run payoff is larger than the payoff obtained by deviating and will deviate otherwise. The long-term payoff is calculated as

$$U(a|\theta,\theta',\rho) = u(a|\theta,\theta') + \beta V(\tau(\theta,\theta',a),\rho)$$

where $u(\alpha|\theta,\theta')$ is the stage payoff and $V(\tau(\theta,\theta',a),\rho)$ is the expected discounted future payoff. Note that $V(\tau(\theta,\theta',a),\rho)$ only depends on the reputation in the next period and is

$$V(\theta_0,\rho) = \mathbb{E}_\rho\{\sum_{t=0}^{\infty}\beta^t u(a_t|\theta_t,\theta'_t)\}$$

Note that the expectation of future payoffs is different for agents who have different "trust" ρ. It is well-known in the game-theoretic literature that reasoning in incomplete information scenarios becomes very complex for the players, if not impossible. Specifically, rational behavior not only depends on the agent's belief about the environment but also involves agents forming higher-order beliefs. Therefore, most game-theoretic models dealing with incomplete information scenarios assume that agents have only limited reasoning capabilities [23] [24]. Similarly, in this paper, we assume that agents possess simple reasoning capabilities about the play of agents with whom they interact based on their reputations: agents' subjective beliefs are that agents with high reputation follow the social strategy and agents with low reputation defect. Under this assumption, the forward trajectory of ρ remains constant because agents with low reputations defect and hence, they maintain low reputations while agents with high reputations follow the recommended strategy and hence, they maintain high reputations. The next proposition shows that the provider needs to have sufficient "trust" in the society in order for it to be willing to follow the social strategy.

Proposition 1. *The optimal action $a^*(\theta,\theta'|\rho)$ for the provider with a belief ρ to follow the prescribed social strategy has a threshold property, i.e.,*

$$a^*(1,1|\rho) = \begin{cases} 1, \text{if } \rho \geq \rho_G \\ 0, \text{if } \rho < \rho_G \end{cases};\ a^*(0,1|\rho) = \begin{cases} 1, \text{if } \rho \geq \rho_B \\ 0, \text{if } \rho < \rho_B \end{cases}$$

and $a^*(\theta,\theta'|\rho) = \sigma(\theta')$ for all other cases where,

$$\rho_G = \frac{1-\beta}{\beta(\gamma-\alpha)}, \rho_B = \frac{1-\beta}{\beta\alpha(\gamma-1)}$$

Proof: See Appendix A. ■

The above proposition proofs that agents will only follow the social strategy if they believe that the society is in a sufficiently good state. Moreover, it also provides several subtle insights about the agents' behaviors: (1) Since both ρ_G, ρ_B are strictly positive, there is always a positive fraction of agents who will deviate because of their (heterogeneous) beliefs. (2) Since $\rho_G \leq \rho_B$, incentives of good agents to cooperate is always at least as large as those of bad agents. (3) Since ρ_G is increasing with α and ρ_B is decreasing with α, punishment has opposite effects on agents' incentives: harsher punishment increases the incentives of good agents to cooperate but also increases the incentives of bad agents to deviate.

B. Dynamics and equilibrium

We assume initially that the social reputation is ρ_S. The limited observations of agents induce heterogeneous beliefs. Agents optimize their strategies a^* according to Proposition 1, and these strategies induce dynamics in the new social reputation $\Phi(\rho_s, a^*)$. The equilibrium requires a consistency check: the steady state social reputation remains invariant, i.e. $\rho_s = \Phi(\rho_s, a^*)$.

Definition 1. *(Bayes-Nash equilibrium in the information exchange system.) Given that an information exchange system is characterized by β,γ and a punishment design α, let ρ_s be a social reputation, $f_{\rho|\rho_s}(\rho)$ be the induced belief distribution due to limited observations, and a^* be the optimal strategy for the agents given the beliefs. We say that $(\rho_s, f_{\rho|\rho_s}, a^*)$ constitutes an equilibrium if*

1) *Agents adopt the optimal strategy a^* to maximize their expected utilities (as in Proposition 1).*
2) *The invariant property holds $\rho_s = \Phi(\rho_s, a^*)$.*

It is worth noting that the agents' optimal strategy does not rely on the current social reputation ρ_s since the threshold beliefs are only functions of β,γ,α but not ρ_s. However, because the belief distribution is induced by ρ_s, the fraction of agents who follow the social strategy is thus influenced by ρ_s, which in turn determines the social reputation in the next period. The new social reputation in the next period can be calculated as follows

$$\begin{aligned}\Phi(\rho_s, a^*) =& \rho_s^2 F_{\rho|\rho_s}(\rho \geq \rho_G) \\ &+ \alpha(1-\rho_s)\rho_s F_{\rho|\rho_s}(\rho \geq \rho_B) + \rho_s(1-\rho_s)\end{aligned}$$

We denote $\Delta(\rho_s) = \Phi(\rho_s, a^*) - \rho_s$ as the change in the social reputation and thus,

$$\Delta\left(\rho_s\right) = \underbrace{\alpha\left(1-\rho_s\right)\rho_s F_{\rho|\rho_s}\left(\rho \geq \rho_B\right)}_{\text{bad to good}} - \underbrace{\rho_s^2 F_{\rho|\rho_s}\left(\rho \leq \rho_G\right)}_{\text{good to bad}} \quad (1)$$

with

$$F_{\rho|\rho_s}\left(\rho \geq \rho_B\right) = \int_{\rho=\rho_B}^{1} f_{\rho|\rho_s}\left(\rho\right) d\rho$$

$$F_{\rho|\rho_s}\left(\rho \leq \rho_G\right) = \int_{\rho=0}^{\rho_G} f_{\rho|\rho_s}\left(\rho\right) d\rho$$

The first part in (1) represents the fraction of agents whose reputations change from bad to good and the second part is the fraction of agents whose reputations change from good to bad. To constitute an equilibrium, it is sufficient and necessary that $\Delta(\rho_s) = 0$. However, in information exchange systems it is of paramount importance that the resulting equilibrium is robust to small disturbances (e.g. small reputation update errors).

Definition 2. *(Robust equilibrium) The equilibrium with ρ_s is robust if and only if*

$$\Delta\left(\rho_s\right) = 0 \quad \text{and} \quad \frac{d\Delta\left(\rho_s\right)}{d\rho_s} < 0$$

Now we study the conditions under which robust equilibrium exists.

Proposition 2. *Given that an information exchange system is characterized by β, γ and the punishment design is α, the existence of the robust equilibrium depends on ρ_B.*

1) *If $\rho_B > 1$, $\rho_s = 0$ is the unique robust equilibrium.*
2) *If $\rho_B \leq 1$, there exists at least one robust equilibrium $\rho_s \in (0, 1)$.*

Proof: Omitted proofs can be found in [26]. ■

Proposition 2 proves that neither full efficiency nor zero efficiency will occur in the robust equilibrium in the limited observations case. As we will see later, the actual efficiency will depend on the punishment severity which needs to be carefully designed. Before proceeding to that, we compare the achievable efficiency for the limited observations case with that for the unlimited observations case to illustrate the different design aspects.

C. Unlimited observations

In this subsection, we investigate how the system evolves if agents make unlimited observations (i.e. $f_{\rho|\rho_s}(\rho) = I(\rho - \rho_s)$) to illustrate why the system design should be different than in the limited observations case. Suppose that the system starts with an initial social reputation $\rho_s^0 \in [0, 1]$, and we examine in which long-run state $\rho_s^{t\to\infty}$ that the system will be trapped in.

Proposition 3. *With unlimited observations, the long-run system state is (1) If $\rho_s^0 \geq \rho_B, \rho_s^{t\to\infty} = 1$. (2) If $\rho_s^0 \leq \rho_G, \rho_s^{t\to\infty} = 0$. (3) If $\rho_G < \rho_s^0 < \rho_B, \rho_s^{t\to\infty} = \rho_s^0$.*

Proof: Omitted proofs can be found in [26]. ■

We see that, in the unlimited observations case, appropriately choosing the initial social reputation (e.g. by influencing the experiences of sufficient people) can lead to full efficiency while starting from the wrong initial social reputation leads to zero efficiency regardless of the choice of α. This differs from the limited observations case where neither the full efficiency (cooperation) nor the zero efficiency (no cooperation) systems can occur in a robust equilibrium. The achievable efficiency depends on the punishment severity of the social norm and hence, this needs to be carefully designed as discussed in the next section.

IV. Optimal Punishment Design

The minimum social reputation beliefs ρ_G, ρ_B that sustain cooperation are determined by the punishment. The harsher the punishment is (smaller α), fewer good providers deviate while also fewer bad providers cooperate to restore their reputations. Hence, when designing the punishment, the tension between providing incentives to good versus bad agents needs to be considered. In this section, we characterize the impact of punishments on the achievable system efficiency. Our focus is on maximizing the cooperation among the agents and hence, we use the social reputation, i.e. the fraction of good agents in the system, as the efficiency metric.

The objective of the system designer in our model is to choose the optimal punishment α, given the network environment parameters β, b, c such that the social reputation is maximized (hence, the probability that agents cooperate is also maximized which leads to the maximized social welfare). Formally, the design problem is to solve

$$\begin{aligned} &\underset{\alpha}{\text{maximize}} \quad \rho_s \\ &\text{subject to} \quad \text{the system is in a robust equilibrium} \end{aligned}$$

A. Belief distribution function based on observations

In this subsection, we describe a specific belief distribution function based on the agents' observations for further analysis. Suppose that in each period an agent is able to observe M reputations of other agents. The agent's belief ρ is formed based on these observation results. We define M as the observation granularity which quantifies how much agents are able to observe the system and define m as the observation result which indicates the number of good reputations observed. The observation granularity represents how well informed the agents in the system are and hence, we treat it as an intrinsic characteristic of the information exchange system. Hence, we assume that M is exogenously determined[3] and the designer selects the social norm that maximizes the system's efficiency. Note that even if the observation results of agents are the same, it is still possible for them to form different beliefs. We therefore use the widely-adopted representative agent model [25] to determine how heterogeneous beliefs are formed.

First we consider the case where all agents have the same observation granularity, i.e. all agents make M observations. Therefore, there are $M + 1$ representative agents, each representing agents having the same observation results (i.e. $m = 0, 1, ..., M$ respectively). In Bayesian statistics, the beta

[3] Alternatively, the observation granularity could also be a design parameter. In that case, there might be associated with making observations and the designer needs to optimize M given costs.

distribution $B(m+1, M-m+1)$ can be seen as the posterior probability of the parameter of a binominal distribution after observing m successes and $M-m$ failures. Therefore, we use the beta distribution to model a representative agent's belief when the observation result is m high reputations out of M other agents (similar belief function is used in [21] to model agents' posterior beliefs after observations).

$$\begin{aligned} f_{m,M}(\rho) &= B(m+1, M-m+1) \\ &= \frac{\Gamma(M+2)}{\Gamma(m+1)\Gamma(M-m+1)} \rho^m (1-\rho)^{M-m} \end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function. For an individual agent who has the same observation result of this representative agent, its belief is then a realization according to $f_{m,M}(\rho)$. Since the fractions of agents who have different observation results follow a binominal distribution with the parameter ρ_s (which is the true social reputation), the belief distribution function is given by

$$\begin{aligned} f^M_{\rho|\rho_s}(\rho) &= \mathbb{E}_m\{f_{m,M}\} \\ &= \sum_{m=0}^{M} \binom{M}{m} \rho_s^m (1-\rho_s)^{M-m} f_{m,M}(\rho) \end{aligned}$$

Next, we consider the case where the observation granularity is a random variable to model the scenario where agents may make different numbers of observations. Suppose the observation granularity follows an independent identical distribution $\Omega(\cdot)$ over non-negative integer numbers. Because we consider a continuum model, a fraction $\Omega(M)$ of the agents makes M observations. Therefore, the belief distribution of all agents is

$$f_{\rho|\rho_s}(\rho) = \sum_{M\in\mathbb{N}} \Omega(M) f^M_{\rho|\rho_s}(\rho)$$

The observation granularity determines the knowledge agents can obtain regarding the social reputation of the system. A larger will lead to more accurate beliefs for the agents about the social reputation. We make a few discussions about the impact of observation granularity.

- $M = 0$: $f_{m,M}(\rho) = f_{0,0}(\rho)$ is constant and hence, $f_{\rho|\rho_s}(\rho)$ is constant and independent of ρ_s, thereby implying that the agents' beliefs of the social reputation are uniformly random.
- $M \to \infty$: $f_{\rho|\rho_s}(\rho) \to I(\rho - \rho_s)$, where $I(\cdot)$ is the indicator function, thereby implying that as the observation granularity becomes infinite, agents have perfect knowledge of the social reputation.

In the remainder of this section, we will consider the constructed belief distribution function and assume that all agents have the same observation granularity. Simulations are conducted for the case when agents have different observation granularities in Section V.

B. *Efficiency bounds*

In this subsection, we rigorously investigate how the observation granularity affects the system performance.

Proposition 4. *Given that an information exchange system is characterized by β, γ, M, for a design parameter α, the robust equilibrium ρ_s^* is bounded as follows*

$$\frac{\alpha(1-\rho_B)^{M+1}}{\alpha(1-\rho_B)^{M+1} + \left(1-(1-\rho_G)^{M+1}\right)} \leq \rho_s^* \leq \frac{\alpha\left(1-\rho_B^{M+1}\right)}{\alpha\left(1-\rho_B^{M+1}\right) + \rho_G^{M+1}}$$

Proof: Omitted proofs can be found in [26]. ∎

Corollary 1. *Fix β, γ, M, for large γ, the robust equilibrium ρ_s^* is bounded away from 1,*

$$\rho_s^* \leq 1 - \left(\frac{1-\beta}{\beta(\gamma-1)}\right)^{M+1}, \forall \alpha \in [0,1]$$

Proof: Omitted proofs can be found in [26]. ∎

The above result shows that the upper bound depends on the granularity of observations. If the system designer wants to achieve a higher efficiency, it is necessary that agents are able to make more observations to acquire more accurate reputation distribution information. In some systems, the number of observations can be designed by the designer. For example, the website designer may only allow agents to access the reputations of a limited number of other agents due to privacy and security concerns. Therefore, the tradeoff between efficiency and privacy needs to be carefully considered. However, in this paper, we assume that the observation granularity is exogenously determined and characterize the information availability of the system.

In the following we analyze several specific cases of limited observations which induce different agent belief distributions.

C. *Example 1: $M = 0$ (constant belief distribution)*

We first consider the simplest case: $M = 0$, i.e. agents have no observation. In the belief model that we use, $M = 0$ corresponds to the case that agents have a (constant) uniform belief over all possible social reputations, namely $f_{\rho|\rho_s}(\rho) = 1$ and

$$F_{\rho|\rho_s}(\rho \geq \rho_B) = 1 - \rho_B, F_{\rho|\rho_s}(\rho \leq \rho_G) = \rho_G$$

For this simple case, we are able to explicitly determine the unique robust equilibrium:

$$\rho_s^* = \frac{\alpha(1-\rho_B)}{\alpha(1-\rho_B) + \rho_G} = \frac{\alpha - \frac{1-\beta}{\beta(\gamma-1)}}{\alpha - \frac{1-\beta}{\beta(\gamma-1)} + \frac{1-\beta}{\beta(\gamma-\alpha)}}$$

It is equivalent to consider the maximization problem,

$$\max_{\alpha} \left(\alpha - \frac{1-\beta}{\beta(\gamma-1)}\right)(\gamma - \alpha)$$

The objective function is a quadratic function. The maximum is achieved at

$$\alpha^* = \min\left\{\frac{\gamma\beta(\gamma-1) + 1 - \beta}{2\beta(\gamma-1)}, 1\right\}$$

The above equation therefore provides the optimal α which is designed to maximize the efficiency when $M = 0$.

Proposition 5. *Given that an information exchange system is characterized by β, γ and for $M = 0$, the optimal punishment parameter is*

$$\alpha^* = \min\left\{\frac{\gamma\beta\left(\gamma-1\right)+1-\beta}{2\beta\left(\gamma-1\right)}, 1\right\}$$

and the induced robust equilibrium is

$$\rho_s^* = \begin{cases} \frac{\frac{1}{4}\left(\gamma-\frac{1-\beta}{\beta(\gamma-1)}\right)^2}{\frac{1}{4}\left(\gamma-\frac{1-\beta}{\beta(\gamma-1)}\right)^2+\frac{1-\beta}{\beta}}, \text{if } \frac{\gamma\beta(\gamma-1)+1-\beta}{2\beta(\gamma-1)} < 1 \\ 1-\frac{1-\beta}{\beta(\gamma-1)}, \text{if } \frac{\gamma\beta(\gamma-1)+1-\beta}{2\beta(\gamma-1)} \geq 1 \end{cases}$$

Proof: Omitted proofs can be found in [26]. ■

D. Example 2: $M = 1$ (linear belief distribution)

In this subsection we consider the case $M = 1$. For example, agents observe the reputation of one other agent by sampling the system. This can be interpreted as agents having a linear belief distribution regarding the true social reputation. The belief function thus is given by

$$\begin{aligned} f_{\rho|\rho_s}(\rho) &= \rho_s f_{1,1}(\rho) + (1-\rho_s) f_{0,1}(\rho) \\ &= 2\left(1-\rho_s + (2\rho_s-1)\rho\right) \end{aligned}$$

Note $f_{1,1}(\rho) = \rho, f_{0,1}(\rho) = 1-\rho$. Hence, the cumulative belief functions are linear in ρ,

$$F_{\rho|\rho_s}(\rho \geq \rho_B) = (1-\rho_B)^2 + 2\rho_B(1-\rho_B)\rho_s$$

$$F_{\rho|\rho_s}(\rho \leq \rho_G) = \rho_G(2-\rho_G) - 2\rho_G(1-\rho_G)\rho_s$$

To solve $\Delta(\rho_s) = 0$, it is equivalent to solve $\Delta(\rho_s)/\rho_s = 0$ for $\rho_s \neq 0$. Let $g(\rho_s) = \Delta(\rho_s)/\rho_s$.

$$\begin{aligned} g(\rho_s) = \alpha\Big((1-\rho_B)^2 + 2\rho_B(1-\rho_B)\rho_s\Big)(1-\rho_s) \\ - \left(\rho_G(2-\rho_G) - 2\rho_G(1-\rho_G)\rho_s\right)\rho_s \end{aligned}$$

The above function is a quadratic function regarding ρ_s. It is difficult to determine the robust equilibrium and even more difficult to analyze the impact of punishment directly. In the following, we instead first establish tighter upper and lower bounds of the efficiency in the robust equilibrium than the general bounds given in Proposition 4 when γ is large. In most information exchange systems, service cost is much smaller than the benefit that can be obtained from receiving this service, therefore we focus on the systems where γ is large. Using the new bounds we can derive the optimal punishment based on which optimal social norms can be designed.

Proposition 6. *Given that an information exchange system is characterized by β, γ and $M = 1$, for a design parameter α, the robust equilibrium ρ_s^* is bounded by*

$$\frac{\alpha(1-\rho_B)^2}{\alpha(1-\rho_B)(1-3\rho_B)+\rho_G(2-\rho_G)} \leq \rho_s^* \leq \frac{\alpha(1-\rho_B)^2}{\alpha(1-\rho_B)^2+\rho_G^2}$$

Proof: Omitted proofs can be found in [26]. ■

The upper bound in the above proposition has significant implications for the optimal punishment design. In order to maximize the upper bound, it is equivalent to consider the following maximization problem

$$\max_\alpha \alpha(1-\rho_B)^2 \text{ or } \max_\alpha \alpha\left(1-\frac{1-\beta}{\alpha\beta\gamma}\right)^2$$

Expanding the above objective function, we get

$$\alpha\left(1-\frac{1-\beta}{\alpha\beta\gamma}\right)^2 = \alpha + \frac{1}{\alpha}\left(\frac{1-\beta}{\beta\gamma}\right)^2 - 2\frac{1-\beta}{\beta\gamma}.$$

Recall that we need to ensure that $\rho_B < 1$ since otherwise the only robust equilibrium is 0 according to Proposition 3 and hence, the feasible α needs to satisfy

$$\alpha \geq \frac{1-\beta}{\beta\gamma}.$$

Therefore, choosing $\alpha = 1$ maximizes the objective function and hence, it maximizes the upper bound for all feasible α. Note that for $\alpha = 1$, the upper bound is indeed the actual efficiency because the upper and lower bounds are identical. Therefore, $\alpha = 1$ maximizes the efficiency of the robust equilibrium. The following proposition restates this result and also determines the social reputation in equilibrium.

Proposition 7. *Fix $\beta, \gamma, M = 1$, for large γ, the robust equilibrium ρ_s^* is maximized by choosing $\alpha^* = 1$, and the optimal solution is*

$$\rho_s^* = \frac{\left(1-\frac{1-\beta}{\beta\gamma}\right)^2}{\left(1-\frac{1-\beta}{\beta\gamma}\right)^2+\left(\frac{1-\beta}{\beta\gamma}\right)^2}$$

Note that this robust equilibrium efficiency is close to 1 when γ is large or β is close to 1. We will numerically illustrate the higher order cases, i.e. values of M other than 0 and 1, in the simulation section. However, from the above analysis of the two specific examples, several key design insights can be drawn: when the benefit-to-cost ratio is large, (1) it is optimal to choose $\alpha = 1$, which is the mildest punishment possible, and (2) larger M leads to a higher social reputation for $\alpha = 1$.

V. Simulations and Discussions

In this section, we provide some simulation results. In Fig. 4, the social reputation evolutions for different observation granularities (i.e. $M = 0, 1, 2$) are illustrated. For each observation granularity, we start the system from two different initial states $\rho_s^{t=0} = 0.1$ and $\rho_s^{t=0} = 0.99$ to illustrate how the system state evolves from very different initial states. $\rho_s^{t=0} = 0.1$ represents a system where the majority of agents have low reputations and $\rho_s^{t=0} = 0.99$ is a system where almost all agents have high reputations. The simulations show that the system indeed converges to the same robust equilibrium where the social reputation does not change. For the simulated cases, the robust equilibrium is unique. However, there could be multiple robust equilibria in which case different initial states would converge to different robust equilibria.

Fig. 5 illustrates the impact of the punishment probability α on the system performance for various observation granularities. For a given mild punishment (large α), more observations have higher efficiency. For a given harsh punishment (small α), it is possible that more observations can lead to lower

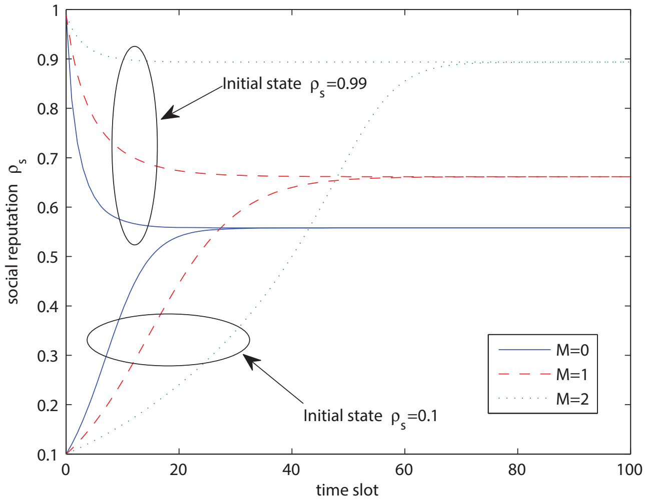

Fig. 4. Stable state of the system for $\beta = 0.25, \gamma = 8$ and $\alpha = 0.9$.

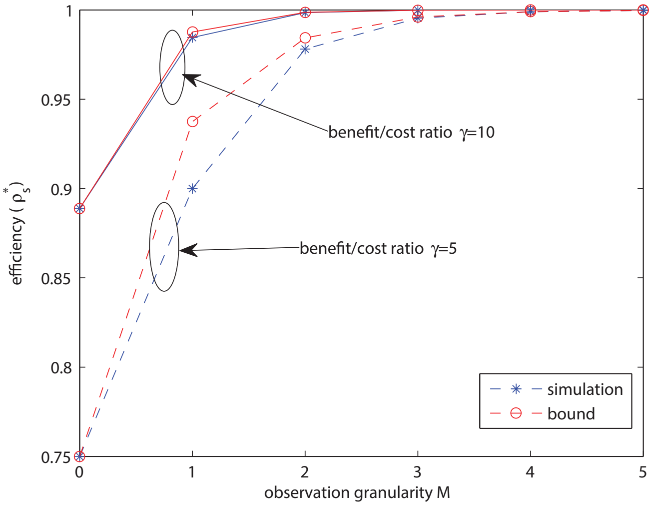

Fig. 6. Bounds on the efficiency for various observation granularities. ($\beta = 0.5$)

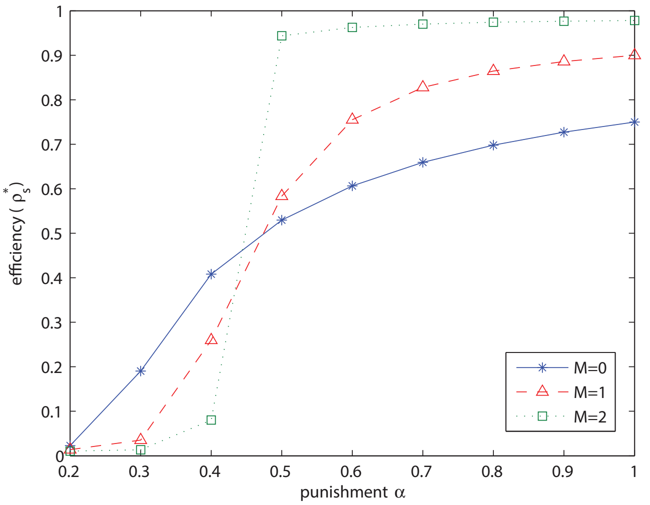

Fig. 5. Impact of the punishment parameter α for various observation granularities. ($\gamma = 5, \beta = 0.5$)

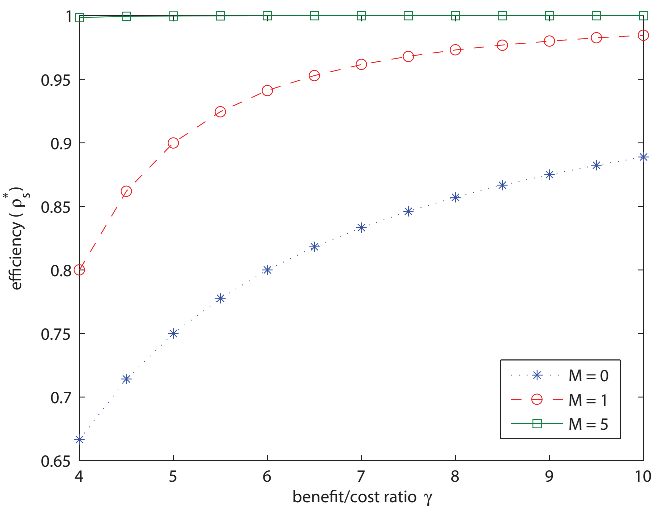

Fig. 7. Impact of the benefit-to-cost ratio γ on the system efficiency. ($\beta = 0.5$)

efficiency. However, since choosing a mild punishment is always better than choosing a harsh punishment, basically M should be larger to achieve higher efficiency. It suggests that in order to obtain the better performance, agents need to have more accurate information of the social reputation. Fig. 6 further compares the simulated optimal efficiency in equilibria with the upper bound established in Proposition 3. The established bound is close to the simulation points and the performance becomes close to full efficiency as M increases.

Fig. 7 and Fig. 8 illustrate the impact of the benefit-to-cost ratio γ and the discount factor β on the system efficiency. For a given punishment probability, the system performance improves with large γ. The discount factor β has a similar impact: larger β leads to better performance.

The case where agents have different observation granularities is investigated in Fig. 9. In this set of simulations, a fraction $1 - q$ of agents have observation granularity $M = 1$. For the other fraction q of agents, their observation granularities are uniformly randomly distributed between 0 and 2-6. Fig. 9 shows how q affects the system performance under various α. We see that the design does not change even if there is some randomness in the observation granularities. The optimal punishment design is still $\alpha = 1$ which is the mildest possible punishment.

Next, we discuss the optimal punishment design. The fact that the mildest punishment is optimal may seem counter-intuitive. However, this finding can be easily explained as follows. Punishment is often used to prevent agents from misbehaving. When agents have high reputations, harsher punishments impose greater threats on these agents if they would deviate. Hence, it may seem that harsher punishments are needed to obtain a better performance. However, this intuition is only valid when all agents are on the equilibrium path, i.e. they always follow the social strategy. For the limited observations scenario, there are always a positive fraction of

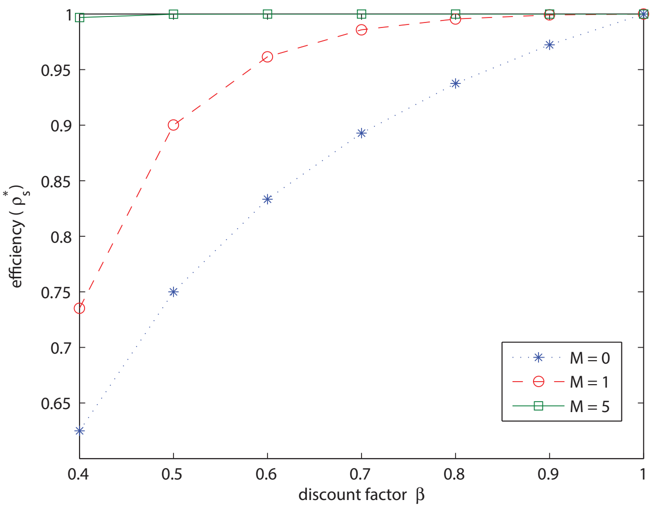

Fig. 8. Impact of the discount factor β on the system efficiency. ($\gamma = 5$).

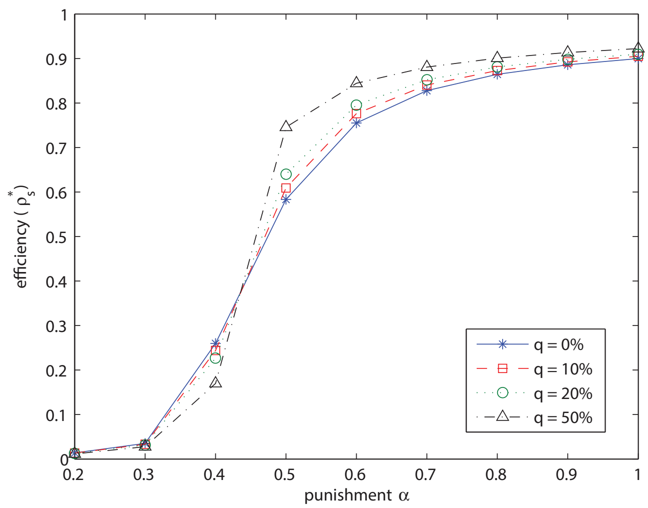

Fig. 9. Impact of randomness in the observation granularity on the system efficiency. ($\beta = 0.5, \gamma = 5$)

agents who deviate no matter what the punishment is. Once agents are in the punishment phase, harsher punishments become disincentives for them to restore their reputations. When the benefit-to-cost ratio is large, the punishment has greater impact on the disincentives for agents with low reputations than the incentives for agents with high reputations. Therefore, choosing the mildest punishment is the best design in those scenarios.

VI. Conclusion

In this paper, we design the optimal social norm protocol for information exchange systems in which agents have heterogeneous beliefs due to limited observations of the system. First, the optimal provider strategy is shown to have a threshold property: agents cooperate only when they have sufficient "trust" in the system (i.e. believe that sufficient agents are cooperating). Second, a Bayesian belief model is proposed to rigorously model the agents' belief heterogeneity and determine the impact of these beliefs on the system efficiency. Finally, the impact of the punishment severity on the robust equilibrium and the achievable system efficiency is rigorously studied. When agents can make unlimited observations, either full or no cooperation emerges in the robust equilibrium. However, in the more realistic limited observations scenario, full efficiency can never be achieved and different punishment strategies lead to different robust equilibria having different efficiencies. Hence, the system designer needs to select the optimal punishment to maximize the system efficiency given the agents' observation granularity. We show that choosing the mildest punishment is optimal for most systems and support this finding with both analytical and simulation results. While in this paper agents are assumed to possess simple reasoning capabilities, an important direction for future research will constitute considering more sophisticated reasoning and belief formation models for the agents. We have followed here the well-known adage "one has to start somewhere", but we are keenly aware that there is much more to be done.

Appendix A
Proof of Proposition 1

First we consider the decision problem when the provider has a good reputation. Obviously, if the requester's reputation is bad, it is optimal that the provider follows the social strategy and plays defect (not provide). If the requester's reputation is good, the provider may have incentives to deviate from the social strategy due to the immediate cost. For a provider with high reputation to provide service, it requires,

$$U(1|1,1,\rho) \geq U(0|1,1,\rho)$$

which further yields,

$$c \leq \beta(V(1,\rho) - V(0,\rho))$$

Next we consider a provider' incentive problem to provide service when it has a low reputation. Similarly, the only possible deviation occurs when the requester has a high reputation. For the provider to follow the social strategy, it requires,

$$U(1|0,1,\rho) \geq U(0|0,1,\rho)$$

In both cases, we need to determine the future payoff difference $V(1,\rho) - V(0,\rho)$. The discounted future payoff can be written in the following recursive form.

$$\begin{aligned} V(1,\rho) &= \rho(b-c) + \beta V(1,\rho) \\ V(0,\rho) &= -\rho c + \beta[\rho\alpha V(1,\rho) + (1-\rho\alpha)V(0,\rho)] \end{aligned}$$

and hence,

$$V(1,\rho) - V(0,\rho) = \frac{\rho b}{1 - \beta(1-\rho\alpha)}$$

Therefore, the agents' trust ρ needs to be above the following thresholds:

$$\rho \geq \frac{1-\beta}{\beta(\gamma-\alpha)}, \text{if } \theta = 1$$

$$\rho \geq \frac{1-\beta}{\beta\alpha(\gamma-1)}, \text{if } \theta = 0$$

REFERENCES

[1] O. Loginova, H. Lu, and X. H. Wang, "Incentive schemes in peer-to-peer networks," *The B.E. Journal of Theoretical Economics*, Oct. 2008.

[2] K. Ranganathan, M. Ripeanu, A. Sarin, and I. Foster, "Incentive mechanism for large collaborative resource sharing," *In Proc. IEEE International Symposium on Cluster Computing and the Grid*, 2004.

[3] P. Antoniadis, C. Courcoubetis, and B. Strulo, "Comparing economics incentives in peer-to-peer networks," *Computer Networks*, 46(1):133-146, 2004.

[4] E. J. Friedman, J. Y. Halpern, and I. A. Kash, "Efficiency and nash equilibria in a scrip system for P2P networks," *In Proc. Seventh ACM Conference on Electronic Commerce*, 2006.

[5] R. Landa, D. Griffin, R. Clegg, E. Mykoniati, and M. Rio, "A sybilproof indirect reciprocity mechanism for peer-to-peer networks," *In Proc. INFOCOM*, 2009.

[6] J. Xu, M. van der Schaar, and W. Zame, "Designing practiable distributed exchange for online communities," Technical report available at http://arxiv.org/abs/1108.5871.

[7] R. L. Trivers, "The evolution of reciprocal altruism," *Quarterly review of biology*, 46(1):35-57, 1971.

[8] R. D. Alexander, The Biology of Moral Systems, Aldine de Gruyter, New York, 1987.

[9] H. Masum and Y. Zhang, "Manifesto for the reputation society," *First Monday*, 9(7), 2004.

[10] S. Kamvar, M. T. Schlosser, and H. G. Molina, "The eigentrust algorithm for reputation management in P2P networks," *In Proc. 12th International Conf. on World Wide Web*, 2003.

[11] A. Ravoaja and E. Anceaume, "Storm: a secure overlay for P2P reputation management," *In Proc. 1st International Conf. on Self-Adaptive and Self-Organizing Systems*, 2007.

[12] S. Ba and P. Pavlou, "Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior," *MIS Quart*, 26(3):243-268, 2002.

[13] P. Resnick, R. Zeckhauser, "Trust among strangers in internet transactions: empirical analysis of eBay's reputation system," *Advances in Applied Microeconomics*, 11, 2002.

[14] C. Dellarocas, "Reputation mechanism design in online trading environments with pure moral hazard," *Information Systems Research*, 16(2):209-230, 2005.

[15] M. Fan, Y. Tan, and A. B. Whinston, "Evaluation and design of online cooperative feedback mechanism for reputation management," *IEEE Trans. Knowl. Data Eng.*, 17(3):244-254, 2005.

[16] G. Zacharia, A. Moukas, and P. Maes, "Collaborative reputation mechanisms in electronic marketplaces," *Decision Support Systems*, 29(4):371-388, 2000.

[17] Y. Zhang, J. Park, and M. van der Schaar, "Social norms based incentive mechanisms for peer-to-peer networks," *ICASSP*, 2011.

[18] G. Ellison, "Cooperation in the prisoner's dilemma with anonymous random matching," *Review of Economic Studies*, 61(3):567-588, 1994.

[19] M. Kandori, "Social norms and community enforcement," *Review of Economic Studies*, 59(1):63-80, 1992.

[20] M. Okuno-Fujiwara and A. Postlewaite. "Social norms and random matching games," *Games and Economic Behaviors*, 9(1):79-109, 1995.

[21] K. Iyer, R. Johari, M. Sundararajan, "Mean field equilibria of dynamic auctions with learning," *In Proc. 12th ACM conference on Electronic Commerce*, 2011.

[22] D. Fudenberg and P. Resnick, "The social cost of cheap pseudonyms," *J. Economics & Management Strategy*, Vol, 10, I. 2, pp 173-199, 2001.

[23] S. Morris and H. S. Shin, "Global games: theory and applications," book chapter in Advances in economics and econometrics: theory and applications, eighth world congress, Cambridge University Press, 2003.

[24] J. Kadane and P. Larkey, "Subjective probability and the theory of games," *Management Science*, 28, 113-120, 1982.

[25] J. E. Hartley, "The representative agent in macroeconomics," London New York: Routledge, 1997.

[26] omitted proofs, http://ee.ucla.edu/∼jiexu/documents/jsac_netecon.pdf.

Jie Xu (SM'09) received the B.S. and M.S. degrees in Electronic Engineering from Tsinghua University, Beijing, China, in 2008 and 2010, respectively. He is currently a Ph.D. student with the Electrical Engineering Department, UCLA. His primary research interests include game theory and strategic design in multi-agent networks.

Mihaela van der Schaar (F'10) is currently a Chancellor's Professor in Electrical Engineering Department, UCLA. Her research interests include engineering economics and game theory, strategic design, online reputation and social media, dynamic multi-user networks and system designs.