Efficient Working and Shirking in Information Sharing Networks

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

***Abstract*—In many systems, agents interact repeatedly with each other over an exogenously determined network and need to cooperate with each other by producing and sharing valuable knowledge or information with the agents with which they are connected. However, producing and sharing information can be costly for the agents themselves, while providing no direct immediate benefit to them. Hence, there are incentives for individual agents to shirk rather than to work - to free ride on the information production and sharing of other agents rather than to produce information themselves. In this paper, we develop a systematic framework for designing rating systems aimed at promoting efficient production and sharing in these networks, thereby significantly improving the social welfare (i.e. sum utility of agents) of such networks. The schemes proposed operated effectively even in settings where monitoring of agent behavior is subject to significant errors. In many scenarios our schemes achieve maximum social welfare; in others, we prove that optimal schemes necessarily fall short of maximum social welfare due to imperfect monitoring. The distinction between these scenarios arises from the tension between the social value of producing for others and the strategic value of withholding production. In some scenarios, the optimal scheme allows that less-productive agents shirk (not produce); this creates the largest incentives for more-productive agents to work at the socially-desired level. We establish conditions under which recommending "work" to all agents is the optimal strategy and develop low-complexity algorithms to determine the optimal strategy in general settings for arbitrary information sharing networks.**

***Index Terms*—Repeated games, rating systems, imperfect monitoring, efficient shirking**

I. Introduction

In recent years, extensive research efforts have been devoted to studying complex networks (e.g. social networks, ad hoc/sensor networks, biological networks and etc.), where agents interact with each other repeatedly over an exogenously-determined topology by cooperatively producing and sharing information or knowledge (e.g. beliefs and opinions, measurement results). The collaboration among agents is fundamental for many problems such as social learning in multi-agent signal processing networks [1], information diffusion for agent adaptation and learning in biological networks [2], belief consensus in social networks [3] and etc. However, undertaking such cooperative actions is costly and leads to no immediate benefits for the agents themselves. For networks where agents are strategic and self-interested, meaning that they aim to maximize their own utilities by strategically choosing their actions, the agents will participate in the collaborative process (i.e. produce and share information with others) only if this is beneficial to them. Hence, absent incentives for collaboration, these systems will work inefficiently or even collapse [4]. Therefore, a key challenge to ensure the survivability and efficient operation of systems of interconnected selfish agents is the design of efficient incentive schemes that promote the cooperation of agents.

J. Xu and M. van der Schaar are with the Dept. of Electrical Engineering, University of California Los Angeles (UCLA). Emails: jiexu@ucla.edu, mihaela@ee.ucla.edu.

In this paper, we design incentive schemes that promote the cooperation of self-interested agents interacting over information sharing networks, thereby improving the social welfare of the system (i.e. sum utilities of the participating agents). Since agents stay in the system for a long time and interact with other agents repeatedly, our solution exploits the ongoing nature of the agents' interaction and constructs incentive schemes in which the agents' current interactions depend on the past history of interactions within the network, in particular, on the extent to which the past behavior of an agent has been in accordance with the recommended behavior by the network administrator. For this, we build on the general theory of repeated games with imperfect monitoring, but with many necessary innovations.

Our approach is based on rating systems. The rating system is implemented and operated by the service or system administrator. Specifically, the administrator selects a social strategy which recommends to agents different actions (i.e. work or shirk) depending on their connectivity. For instance, a social strategy may recommend "work" (i.e. produce information) to all agents or it may recommend "work" to only a subset of agents, while allowing the rest of agents to shirk (i.e. not produce information). Depending on each agent's compliance to the recommended social strategy, the rating of each agent is updated over time: agents that behaved as directed enjoy high ratings - and hence greater access to information produced by others; agents that have not behaved as directed suffer low ratings - and hence less access to information produced by others. This low information access level which is designed by the network administrator for low-rated agents thus serves as a punishment device for agents who did not follow the social strategy. For instance, low-rated agents may have only limited access to the information (e.g. measurements, news, and other content) produced by the agents with which they are connected. However, we note that even though the maximal amount of information that can be obtained by an agent is limited by its information access level, the actual amount of information which the agent receives still depends on whether its neighboring agents work or shirk. (See an illustrative information sharing networks with the rating system in Figure 1.)

When designing the rating systems for networks where agents interact over networks, there are many challenges that

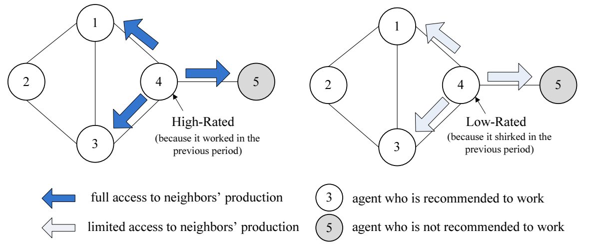

Fig. 1. An illustrative information sharing network with the rating system.

need to be considered. The first challenge is due to the imperfect monitoring of agents' actions. In practical systems, actions taken by the agents are usually imperfectly monitored due to environment noise or mistakes made by receiving agents. For example, even though an agent worked, a receiving agent could report its cooperation but this report may be lost or, alternatively, the receiving agent may mistakenly believe that the information provided by the providing agent is not of adequate quality and thus, it may mistakenly report that the providing agent was uncooperative. If there are no reporting errors, then the administrator can safely block a low-rated agent from accessing information created by other agents with which it is connected. This will provide the strongest incentives for agents to cooperate. However, because monitoring is imperfect, the frequency that a punishment is triggered (thereby resulting in undesirable social welfare loss) largely depends on the accuracy of the monitoring. Therefore, the incentive scheme design needs to explicitly take this into consideration in order to maximize the social welfare.

The second challenge comes from agents' heterogeneity in terms of connectivity, benefits and costs. One rating system design may provide sufficient incentives for some agents to work but not for others. In certain scenarios, in order to maximize the social welfare, the optimal rating system may compromise the shirking behavior of some agents, given that their contribution to the social welfare is minimal while a milder punishment can be adopted to sustain the incentive-compatibility of the remaining agents. Due to this tension between social value of producing for others and the strategic value of withholding production, shirking can also be "efficient" in information sharing networks with imperfect monitoring. With this insight, the optimal social strategy which recommends to agents specific levels of information production becomes significantly more challenging to determine (as compared to all existing related works [5][6][7] where to all agents "work" is recommended). This challenge stems from the heterogeneous incentives of agents in information sharing networks since they enjoy different levels of benefits and is further amplified by the fact that agents' incentives are coupled and interdependent in a (possibly extremely) complex way. In the worst case, the shirking behavior of one agent may cause a "chain reaction" which leads to all agents shirking and hence, to a network collapse.

Another important feature of the information sharing network we consider is that agents broadcast the *same* information to all their neighbors. Withholding information therefore punishes *all* neighbors. In the existing works [8][9] agents provide *different* services to each of their neighbors and withholding service therefore punishes only the *single* neighbor from whom service is withheld. Thus, in the present scenario it is much more difficult for an agent to threaten or use punishments in order to provide incentives for those with whom it interacts.

In this paper, we solve the optimal rating system design problem for the information sharing networks. In particular, our contributions include:

- We determine conditions under which the intuitive strategy where all agents should work is indeed optimal;
- We develop low-complexity algorithms to determine the optimal rating system for arbitrary information sharing networks. The optimality of the algorithms is proved under certain conditions;
- We study the impact of network structure on the emergence of "efficient" shirking (i.e. the optimal strategy allows less-productive agents to shirk).

Besides these rigorous analytical characterizations of the optimal rating system, our simulations also indicate that the proposed rating system is able to create incentives for agents to work at socially-desired level and significantly outperforms other widely-used incentive schemes such as Tit-for-Tat and Trigger strategies.

The proposed rating system is applicable to many applications where self-interested agents are inter-connected and need to exchange information/goods or services/favors with their neighbors. For instance, it can be applied to crowdsourcing, user-generated content networks, peer-to-peer networks etc.

The rest of this paper is organized as follows. Section II discusses the related works. Section III formalizes the information production and sharing game over networks. Section IV introduces the rating system. Section V designs the optimal incentive-compatible rating system. Section VI provides simulation results to highlight the features of the rating system design. Section VII concludes the paper.

II. Related Works

Information sharing networks, where agents interact repeatedly over network to share valuable information, content and other resources range from communication networks [1] to biological networks [2], social and economic networks [3][10]. A large research effort was dedicated in the past years to studying the emergent behavior arising from the interaction of individual agents in such networks as well as to designing efficient in-network signal processing methods. Most of these works assume that agents are cooperative and follow the strategy or protocol prescribed by the designer at all times. However, free riding behavior of agents has been observed in many types of information sharing networks such as online forums, microblogging websites and other online social networks [11][12][13][14]. When agents are self-interested, incentives for agents to cooperate need to be designed to ensure the survivability and efficient operation of these systems.

A large body of research has been dedicated to the design of incentive mechanisms which encourage cooperation among agents. Some popular mechanisms rely on pricing mechanisms or reputation mechanisms. Pricing mechanisms [15][16] are appropriate in many settings, but they are not adequate incentive schemes for many networks where much of the appeal is that the information is free (e.g. online social networks). Under a reputation mechanism, an agent is assigned with a global reputation [17] 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 such as effective information gathering techniques [18] or determining the impact of reputation on a seller's prices and sales [19][20]. The few works providing theoretical results consider either one (or a few) long-lived agent(s) interacting with many short-lived agents [21][22][23] or anonymous, homogenous and unconnected agents selected to interact with each other using random matching [5][6][7][24][25].

This paper significantly differs from existing incentive mechanisms in the following aspects. First, few existing works consider incentive design in a network interaction environment where agents repeatedly interact over networks/graphs. The network constraint makes incentive design much more challenging since actions and incentives are coupled in a complicated manner. Second, the agent interaction cannot be formulated using a standard repeated game in which two interacting agents have double coincidence of wants. Due to the asymmetric interests of the agents, conventional incentive mechanisms used in repeated games based on direct reciprocity (such as Tit-for-Tat strategy [8]) fail to work. Our rating system exploits social/indirect reciprocity to provide agents with incentives to cooperate. Third, our work is unique because it carefully considers the impact of imperfect monitoring and the heterogeneity of agents on incentives. No such prior work considering all these issues exists. Last but not least, we derive non-trivial and surprising results: network settings exist where we can prove that shirking by certain agents is beneficial.

Few research results exist which study the interaction of self-interested agents repeatedly interacting over graphs. For instance, [26] studies how the topology influences the cooperation level of networked agents through the use of community enforcement. In [27], network patterns that sustain favor exchanges are characterized. However, system design is missing in these papers. In [28], the authors show that heterogeneity may preclude the emergence of cooperative behavior. Therefore, partitioning a group into more homogeneous subgroups can enable cooperative behavior which may not be feasible otherwise. However, different from our work, [28] assumes that the designer can determine the connectivity between agents (i.e. which agent can interact with which agent). In our setting, agents are connected over an exogenously determined topology and therefore, the design is restricted by the underlying topology, thereby complicating the rating system design.

Our prior work [9] studies networks in which each agent provides *different* services to each of its neighbors. For example, in wireless relay networks, a node can forward different packets to different destination nodes; the destination node only benefits by receiving the packets addressed to itself but not packets addressed to other agents. In that work, we use distributed optimization techniques to design incentive-compatible rating systems. The systems that we design exploit the fact that an agent can withhold service for a specific neighbor but continue the relaying service for other neighbors: the punishments can be made differential. In the current paper, the focus is on information sharing networks where an agent broadcasts information to all its neighbors. For example, an agent may post a review or a recommendation that all its neighbors can see. However, since this sharing is public, if an agent does not post, none of its neighbors obtain the benefit. Hence, not posting punishes all neighbors; it is not possible to punish one neighbor and not another. The current setting requires completely different analysis and yields different rating schemes with surprisingly different properties, for instance, shirking may become efficient in certain networks.

Table I highlights the significant differences of our rating system from existing works.

III. System Model

A. Information Production Game over a Network

We consider a network of N agents, represented by a set $\mathcal{N} = \{1, 2, ..., N\}$. Agents are connected according to a topology matrix $\mathcal{G}$. If $g_{ij} = 1$, then there is a directed link from agent i to agent j. If $g_{ij} = 0$, then there is not such a link. Note that links are directed and hence, g_{ij} does not necessarily equal g_{ji}. We assume that this underlying topology is predetermined for the following analysis and we do not consider the network formation process. We refer readers interested in network formation problems to [27].

We consider a discrete time system. In each period, each agent i chooses an action $a_i \in \mathcal{A} = \{0, 1\}$ where $a_i = 1$ stands for "work" and $a_i = 0$ stands for "shirk". If an agent works, the agents that connect to it receive benefits. Denote $b_{ij} \in \mathbb{R}^+$ as the benefit that agent i obtains from agent j. If there is no link from agent i to agent j, then it must be $b_{ij} = 0$. Note that b_{ij} could be different from $b_{kj}, (k \neq i)$ even though both agent i and agent k connect to the same agent j. Working is costly and we denote $c_i > 0$ as the cost for agent i to work [1]. We assume that shirking incurs no cost. Agents interact repeatedly with each other and discount the next period utility by a discount factor $\delta \in [0, 1)$.

Given the action profile $\boldsymbol{a}(t)$ of all agents, the utility of agent i at any period t is

$$u_i(\boldsymbol{a}) = \sum_{j:g_{ij}=1} b_{ij} a_j - c_i a_i \tag{1}$$

If agents are compliant, then the network administrator can assign actions to agents to maximize the social welfare, which is defined as the sum utility of all agents, by solving the following optimization problem,

$$\begin{aligned} \underset{\boldsymbol{a}}{\text{maximize}} \quad & V = \sum_i u_i(\boldsymbol{a}) \\ \text{subject to} \quad & a_i \in \{0, 1\}, \forall i \end{aligned} \tag{2}$$

[1] Here we assume that working does not bring benefit to the agent itself. However, if working also brings some benefit $b_{ii} > 0$ to agent i itself, we can consider the net cost $c_i - b_{ii}$ in the system model.

TABLE I
COMPARISONS WITH EXISTING WORKS.

	[Kandori 1992; Ellison 1995]	[Dellarocas 2005]	[Zhang et al. 2013]	[Xu et al 2014]	This work
Incentive mechanism	Differential services	Monetary rewards	Differential services	Differential services	Differential services
Asymmetry of interests	No	No	Yes	Yes	Yes
Number of long-lived players	Multiple	One	Multiple	Multiple	Multiple
Player interactions	One-to-one random matching	One-to-one random arrival	One-to-one random matching	One-to-many link-based over a network	One-to-many node-based over a network
Action space	Discrete	Discrete	Discrete	Continuous	Discrete
Benefit and cost	Homogeneous	Homogeneous	Homogeneous	Heterogeneous	Heterogeneous
Optimization criterion	Individual long-term utility	Individual long-term utility	Sum utility of all players	Sum utility of all players	Sum utility of all players
Resulting optimal strategy	Cooperate	Cooperate	Cooperate	Cooperate	Not necessarily all cooperate

The optimal solution of the above problem can be characterized as follows

$$a_i = \begin{cases} 1, \text{if } c_i < \sum\limits_{j:g_{ij}=1} b_{ji} \\ 0, \text{otherwise} \end{cases} \tag{3}$$

Basically, if an agent i's contribution to the social welfare (i.e. $\sum\limits_{j:g_{ji}=1} b_{ji}$) is greater than its production cost, then it is optimal for agent i to work. In the following analysis, we focus on the case that $c_i < \sum\limits_{j:g_{ji}=1} b_{ji}, \forall i$ and hence, the optimal action profile is $\boldsymbol{a}^* = \mathbf{1}$. That is, the social welfare is maximized when all agents work. However, if agents are self-interested, then the only Nash equilibrium (NE) of this information sharing game is $\boldsymbol{a}^{NE} = 0$ because information production only incurs a cost but no direct benefit to the agents. Therefore, the social welfare at NE is 0: $V^{NE} = \sum\limits_i u_i(\boldsymbol{a}^{NE}) = 0$.

B. Problem Formulation

The objective of the network administrator is to design an incentive mechanism to provide self-interested agents with incentives to work in order to maximize the social welfare. Let π be an incentive mechanism and V^π be the social welfare achieved by π if it is incentive-compatible, meaning that all agents follow the strategy given by the incentive mechanism in their own self-interest. The design problem is therefore,

$$\begin{aligned} &\underset{\pi}{\text{maximize}} && V^\pi \\ &\text{subject to} && \pi \text{ is Incentive-Compatible} \end{aligned} \tag{4}$$

IV. The Rating System

In this paper, we design a simple but practical rating system aiming to maximize the social welfare of the agent network. The reasons why we choose such a rating system are two-fold. First, the proposed rating system is simple to implement and easy to maintain in practical systems. The implementation and maintenance complexity is of significant importance especially when the system is of large-scale. Second, adopting such a simple rating structure allows us to better focus on the impact of agents' heterogeneity induced by the connectivity on the rating system performance in the imperfect monitoring scenarios.

The rating system has three components.

(1) **Recommended Strategy**. The administrator recommends different actions to different agents according to a mapping $\sigma : \mathcal{N} \to \{1, \text{null}\}$ where $\sigma(i)$ is the action recommended to agent i. If $\sigma(i) = 1$, then "work" is recommended to agent i; if $\sigma(i) = \text{null}$, then there is no recommendation to agent i. Effectively, a recommended strategy partitions the agent set into two complementary subsets P and $\bar{P}$ of agents such that "work" is recommended to only agents in P. Recommended strategies can be partially ordered. We say $\sigma_1 \prec \sigma_2$ if they satisfy the following condition: $\forall i \in \mathcal{N}$, if $\sigma_1(i) = 1$, then $\sigma_2(i) = 1$ and there exists at least one agent j such that $\sigma_1(j) = \text{null}$ while $\sigma_2(j) = 1$. Given the recommended strategy, we can also compute the maximum benefit that can be obtained by an agent i if all agents follow the strategy, denoted by $b_i(\sigma) = \sum\limits_{j:g_{ij}=1} \sigma(j) b_{ij}$.

(2) **Rating Set and Information Access Levels**. We consider a simple binary rating set $\{0, 1\}$ where $\theta_i(t) = 1$ means that the rating of agent i at time t is high while $\theta_i(t) = 0$ means that its rating is low. High-rated agents have full access to the information produced by its neighbors, and hence if $\theta_i(t) = 1$, then agent i's benefit is $b_i(\sigma)$. Low-rated agents have less access to the information produced by its neighbors limited by the network administrator. Therefore, if $\theta_i(t) = 0$, then agent i's benefit is $p_i b_i(\sigma)$ where $0 \leq p_i < 1$ is a punishment parameter determined by the network administrator. In general p_i is not necessarily the same for all agents. We call these *discriminative policies* (DPs). However, in many systems,

access policies often do not discriminate among agents so it is also important to study the case that $p_i = p, \forall i$ in which the punishment level only depends on the ratings but not agents' identities. We call these *non-discriminative policies* (NDPs). We note that such a punishment device can be easily implemented in practice, either by the administrator or the interacting agents themselves. For instance, on Facebook, only a fraction of content produced by agents is displayed to low-rated agents.

(3) **Rating Update Rule**. The rating update rule, which is executed at the end of each period, decides how the ratings of the agents should be updated according to their production actions and the strategies that are recommended to them. It is a mapping $\tau : \{1, \text{null}\} \times \mathcal{A} \to [0, 1]$ which decides the probability that the next period rating of an agent i is high (i.e. $\theta_i(t+1) = 1$) according to the recommended strategy $\sigma(i)$ and the current period action $a_i(t)$. In particular, the following simple update rule is adopted in this paper:

$$\tau(1,1) = 1, \tau(1,0) = 0, \tau(\text{null}, \cdot) = 1 \tag{5}$$

Specifically, the rating update depends on whether it followed the recommendation or not. The agents that have behaved as directed enjoy high ratings while the agents that have not behaved as directed suffer low ratings. If there is no recommendation to an agent, then the agent's rating is constantly high. Note, however, that the above update rule applies to a perfect monitoring scenario where production actions are perfectly monitored without errors. In practice, monitoring is never perfect and hence, we explicitly take into account the imperfect monitoring of agent behavior in our rating system design. Therefore, the rating update rule is modified to be

$$\tau(1,1) = 1 - \epsilon^{\text{MD}}, \tau(1,0) = \epsilon^{\text{FA}}, \tau(\text{null}, \cdot) = 1 \tag{6}$$

where $\epsilon^{\text{MD}} \in [0, 0.5]$, $\epsilon^{\text{FA}} \in [0, 0.5]$ are the miss detection and false alarm error probabilities, respectively.

Under the above proposed rating system, the design parameters are summarized by the recommended strategy σ and the information access levels $\boldsymbol{p} = \{p_i, \forall i : \sigma(i) = 1\}$. Hence, we concisely write a rating system as $\pi = (\sigma, \boldsymbol{p})$. We note again that in the NDP case, $p_i = p, \forall i$. Next we compute the achievable social welfare of a rating system by assuming that it is incentive-compatible.

If the rating system is incentive-compatible, then agents to whom "work" is recommended will always work and agents to whom "work" is not recommended will always shirk. Hence, for an agent i, its utility at period t is

$$u_i(t) = \begin{cases} b_i(\sigma) - c_i \mathbf{1}_{\sigma(i)=1}, \text{if } \theta_i(t) = 1 \\ p_i b_i(\sigma) - c_i \mathbf{1}_{\sigma(i)=1}, \text{if } \theta_i(t) = 0 \end{cases} \tag{7}$$

where $\mathbf{1}$ is the indicator function. For an agent i to whom "work" is recommended, the stationary distribution of its rating can be computed by solving a two-state Markov chain:

$$\text{Prob}(\theta_i = 1) = 1 - \epsilon^{\text{MD}}, \text{Prob}(\theta_i = 0) = \epsilon^{\text{MD}} \tag{8}$$

Therefore, the expected utility of agent i over the states is

$$\mathbb{E}\mu_i = [(1 - \epsilon^{\text{MD}}) + \epsilon^{\text{MD}} p_i] b_i(\sigma) - c_i \tag{9}$$

If "work" is not recommended to an agent i, the expected utility of agent i is

$$\mathbb{E}\mu_i = b_i(\sigma) \tag{10}$$

Therefore, the achievable social welfare of a rating system $\pi = (\sigma, \boldsymbol{p})$ (if all agents follow) is

$$\begin{aligned} & V(\sigma, \boldsymbol{p}) \\ = & \sum_{i:\sigma(i)=1} \{[(1 - \epsilon^{\text{MD}}) + \epsilon^{\text{MD}} p_i] b_i(\sigma) - c_i\} + \sum_{i:\sigma(i)=\text{null}} b_i(\sigma) \\ = & \sum_i b_i(\sigma) - \sum_{i:\sigma(i)=1} [\epsilon^{\text{MD}}(1 - p_i) b_i(\sigma) + c_i] \end{aligned} \tag{11}$$

As we can see from (11), in order to maximize the social welfare, the administrator should choose a high p_i (i.e. a milder punishment) and recommend "work" to as many agents as possible (since the optimal social welfare is achieved at $\boldsymbol{a}^* = \mathbf{1}$). However, p_i cannot be chosen as arbitrarily high since that will not provide sufficient incentives for agents to follow the recommended strategy. Instead, the choice of $\boldsymbol{p}$ should be tailored to the specific choice of the strategy σ as will be shown in the next section.

V. Optimal Incentive-Compatible Rating System

In this section, we first study the feasible region and the optimal choice of $\boldsymbol{p}$ given the recommended strategy σ such that the rating system is incentive-compatible. Next, we jointly optimize σ and $\boldsymbol{p}$ to maximize the social welfare.

A. Optimal $\boldsymbol{p}$ for a given σ

As mentioned in the last section, the information access levels $\boldsymbol{p}$ for low-rated agents cannot be chosen as arbitrarily high since the punishment will not be strong enough to deter agents from deviating from the recommended strategy. Agents will follow the recommended strategy if and only if the utility by following outweighs the utility by deviating. Since agents are long-lived in the network, we consider their long-term utilities when evaluating agents' incentives. Consider an agent i such that $\sigma(i) = 1$, let $U_i^\infty(\theta_i)$ be its long-term utility when its rating is θ_i. $U_i^\infty(\theta_i)$ can be computed recursively as follows:

$$\begin{aligned} U_i^\infty(1) &= b_i(\sigma) - c_i + \delta[(1 - \epsilon^{\text{MD}}) U_i^\infty(1) + \epsilon^{\text{MD}} U_i^\infty(0)] \\ U_i^\infty(0) &= p_i b_i(\sigma) - c_i + \delta[(1 - \epsilon^{\text{MD}}) U_i^\infty(1) + \epsilon^{\text{MD}} U_i^\infty(0)] \end{aligned} \tag{12}$$

In the right-hand side of either of the above two equations, the first term is the utility that agent i obtains in the current period depending on its rating and the second term is the discounted continuation utility in the subsequent periods.

Alternatively, if agent i unilaterally deviates from the recommended strategy only in the current period, then its long-term utility becomes,

$$\begin{aligned} \tilde{U}_i^\infty(1) &= b_i(\sigma) - c_i + \delta[\epsilon^{\text{FA}} U_i^\infty(1) + (1 - \epsilon^{\text{FA}}) U_i^\infty(0)] \\ \tilde{U}_i^\infty(0) &= p_i b_i(\sigma) - c_i + \delta[\epsilon^{\text{FA}} U_i^\infty(1) + (1 - \epsilon^{\text{FA}}) U_i^\infty(0)] \end{aligned} \tag{13}$$

The one-shot deviation principle [29] in the repeated games theory allows us to only compare the values of $U_i^\infty(\theta)$ and $\tilde{U}_i^\infty(\theta)$ in order to determine whether the incentive-compatible constraint is satisfied for agent i. Basically the one-shot deviation principle says that if an agent cannot gain by unilaterally deviating from the recommended strategy only in the current period and following the recommended strategy afterwards, then it cannot gain by switching to any other strategies (possibly multiple-shot deviations) either, and vice versa.

Before we study the optimal $\boldsymbol{p}$, we introduce here an important network characteristic given the recommended strategy σ.

Definition 1. *The critical benefit-to-cost ratio of a network given the recommended strategy σ is* $\gamma(\sigma) = \min_{i:\sigma(i)=1}\{b_i(\sigma)/c\}$.

Theorem 1 below establishes the feasible and the optimal values of $\boldsymbol{p}$ given a recommended strategy σ.

Theorem 1. *Given a recommended strategy σ, an incentive-compatible rating system can be constructed if and only if $\gamma(\sigma) \geq \frac{1}{\delta(1-\epsilon^{MD}-\epsilon^{FA})}$ and*

- *In DP systems, $p_i \leq 1 - \frac{1}{\delta(1-\epsilon^{MD}-\epsilon^{FA})}\frac{c_i}{b_i(\sigma)}, \forall i$.*
- *In NDP systems, $p \leq 1 - \frac{1}{\delta(1-\epsilon^{MD}-\epsilon^{FA})\gamma(\sigma)}$*

Moreover, the optimal $\boldsymbol{p}^*$ is the binding value of the above inequalities.

Proof: Consider the DP case and consider an agent i such that $\sigma(i) = 1$, we apply the one-shot deviation principle,

$$\begin{aligned} & U_i^\infty(\theta) - \tilde{U}_i^\infty(\theta) \\ = \quad & -c_i + \delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})(U_i^\infty(1) - U_i^\infty(0)) \geq 0 \end{aligned} \tag{14}$$

According to (13), we have $U_i^\infty(1) - U_i^\infty(0) = (1-p_i)b_i(\sigma)$. Substituting this into the above equation, we obtain the condition on p_i, i.e. $p_i \leq 1 - \frac{1}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}\frac{c_i}{b_i(\sigma)}$. Obviously, for p_i to have a valid value in $[0,1)$, it must be true that $\frac{b_i(\sigma)}{c_i} \geq \frac{1}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}$. For the rating system to be incentive compatible, this must hold for all agents. Since the utility of agent i is increasing in p_i, the optimal p_i is the binding value.

A similar analysis can be conducted for the NDP case. The major difference is that the common p has to be chosen as the maximum among all $p_i^*, \forall i$ to satisfy all agents' incentive. ■

Theorem 1 has two implications. First, in order to design an incentive-compatible rating system, the benefit-to-cost ratio of agents should be larger than a threshold which is determined by the environment parameters. In particular, as agents value less the future utilities (smaller δ) and the monitoring becomes less accurate (larger $\epsilon^{\text{MD}}, \epsilon^{\text{FA}}$), the higher threshold is required to ensure that agents can obtain a sufficiently large benefit by working at a cost. Second, the optimal access levels for low-rated agents depend on both the environment parameters as well as the recommended strategy. The optimal $\boldsymbol{p}^*$ (or p^*) in fact represents the tradeoff of incentive-compatibility and the social welfare in the imperfect monitoring scenario. If $\boldsymbol{p} > \boldsymbol{p}^*$ (or $p > p^*$), then the rating system is not incentive-compatible and there is at least one agent who will not follow the recommended strategy; if $\boldsymbol{p} < \boldsymbol{p}^*$ (or $p < p^*$), much social welfare loss will incur due to the undesired rating drop in the imperfect monitoring scenario.

Importantly, the optimal $\boldsymbol{p}^*$ (or p^*) is a function of the recommended strategy σ. Hence, the optimal social welfare by an incentive-compatible rating system given the recommended strategy σ is

$$V(\sigma) = \begin{cases} \sum_i b_i(\sigma) - \sum_{i:\sigma(i)=1}\left(1 + \frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}\right)c_i \\ \qquad \text{(DP)} \\ \sum_i b_i(\sigma) - \sum_{i:\sigma(i)=1}\left(\frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}b_i(\sigma) + c_i\right) \\ \qquad \text{(NDP)} \end{cases} \tag{15}$$

Therefore, according to Theorem 1, the rating system design problem is refined to determine the optimal recommended strategy that maximizes the social welfare,

$$\begin{aligned} \underset{\sigma}{\text{maximize}} \quad & V(\sigma) \\ \text{subject to} \quad & \gamma(\sigma) \geq \frac{1}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})} \end{aligned} \tag{16}$$

In general, the above maximization problem is difficult to solve since the strategy space is combinatorial and huge (i.e. 2^N) and moreover, the objective function depends on σ in an extremely complex way (especially for the non-discriminative policy case). In the next subsections, we determine the optimal recommended strategy (hence, the optimal rating system) in various scenarios. In some scenarios, our rating system achieves maximum social welfare; in others, they necessarily fall short of maximum social welfare due to the imperfect monitoring but we determine the optimal design that achieves the highest social welfare among all rating systems under consideration.

B. Optimality of "All agents work"

As noted before, the strategy space is huge. Thus, searching in this space for the optimal solution is thus a demanding task. Intuitively, it seems that the optimal recommended strategy should be "all agents should work", denoted by σ^{ALL}. If this intuition was correct, then the administrator could simply recommend σ^{ALL}. In this subsection, we determine the conditions under which this strategy is indeed optimal.

Theorem 2. *For both DP and NDP, if $\delta\gamma(\sigma^{ALL}) > 1$, then there exist ϵ_1, ϵ_2 (possibly different for DP and NDP) such that for all ϵ^{MD}, ϵ^{FA} that satisfy $\frac{\epsilon^{MD}}{1-\epsilon^{MD}-\epsilon^{FA}} < \epsilon_1$ and $\epsilon^{MD}+\epsilon^{FA} < \epsilon_2$, the optimal strategy is σ^{ALL}.*

Proof: (1) We first study DP. Consider any recommended strategy σ, we can compute $V(\sigma^{\text{ALL}}) - V(\sigma)$ as follows,

$$\begin{aligned} & V(\sigma^{\text{ALL}}) - V(\sigma) \\ = \quad & \sum_i b_i(\sigma^{\text{ALL}}) - \sum_i \left(1 + \frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}\right)c_i \\ & -\sum_i b_i(\sigma) + \sum_{i:\sigma(i)=1}\left(1 + \frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}\right)c_i \\ = \quad & \sum_i b_i(\sigma^{\text{ALL}}) - \sum_i c_i - \sum_i b_i(\sigma) \\ & + \sum_{i:\sigma(i)=1} c_i - \frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}\sum_{i:\sigma(i)=\text{null}} c_i \end{aligned} \tag{17}$$

In this equation, $\sum_i b_i(\sigma^{\text{ALL}}) - \sum_i c_i - \sum_i b_i(\sigma) \geq 0$ due to the fact that the social welfare is maximized when all agents work. Therefore, there must exist a value ϵ_σ such that if

$$\frac{\epsilon^{\text{MD}}}{1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}}} < \epsilon_\sigma \triangleq \delta \frac{\sum_i b_i(\sigma^{\text{ALL}}) - \sum_i c_i - \sum_i b_i(\sigma)}{\sum_{i:\sigma(i)=\text{null}} c_i} \tag{18}$$

then $V(\sigma^{\text{ALL}})-V(\sigma) \geq 0$ holds. Since the recommended strategy space is finite, there must exists an value $\epsilon_1 = \min_{\sigma \neq \sigma^{\text{ALL}}} \epsilon_\sigma$ such that σ^{ALL} maximizes the achievable social welfare.

(2) We then study the non-discriminative policy. Consider any recommended strategy σ, we compute $V(\sigma^{\text{ALL}}) - V(\sigma)$ as follows,

$$\begin{aligned} & V(\sigma^{\text{ALL}}) - V(\sigma) \\ = & \sum_i b_i(\sigma^{\text{ALL}}) - \sum_i \left(\frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})\gamma(\sigma^{\text{ALL}})} b_i(\sigma^{\text{ALL}}) + c_i \right) \\ & - \sum_i b_i(\sigma) + \sum_{i:\sigma(i)=1} \left(\frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})\gamma(\sigma)} b_i(\sigma) + c_i \right) \\ = & \sum_i b_i(\sigma^{\text{ALL}}) - \sum_i c_i - \sum_i b_i(\sigma) + \sum_{i:\sigma(i)=1} c_i \\ & - \frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})} \left(\sum_i \frac{b_i(\sigma^{\text{ALL}})}{\gamma(\sigma^{\text{ALL}})} - \sum_{i:\sigma(i)=1} \frac{b_i(\sigma)}{\gamma(\sigma)} \right) \end{aligned} \tag{19}$$

In the above equation, $\sum_i b_i(\sigma^{\text{ALL}}) - \sum_i c_i - \sum_i b_i(\sigma) + \sum_{i:\sigma(i)=1} c_i \geq 0$. If $\sum_i \frac{b_i(\sigma^{\text{ALL}})}{\gamma(\sigma^{\text{ALL}})} - \sum_{i:\sigma(i)=1} \frac{b_i(\sigma)}{\gamma(\sigma)} \leq 0$, then $V(\sigma^{\text{ALL}})-V(\sigma) \geq 0$ trivially holds regardless of the monitoring errors. If $\sum_i \frac{b_i(\sigma^{\text{ALL}})}{\gamma(\sigma^{\text{ALL}})} - \sum_{i:\sigma(i)=1} \frac{b_i(\sigma)}{\gamma(\sigma)} > 0$, then there must exist an ϵ_σ such that if

$$\frac{\epsilon^{\text{MD}}}{1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}}} < \epsilon_\sigma \triangleq \delta \frac{\sum_i b_i(\sigma^{\text{ALL}}) - \sum_i c_i - \sum_i b_i(\sigma) + \sum_{i:\sigma(i)=1} c_i}{\sum_i \frac{b_i(\sigma^{\text{ALL}})}{\gamma(\sigma^{\text{ALL}})} - \sum_{i:\sigma(i)=1} \frac{b_i(\sigma)}{\gamma(\sigma)}} \tag{20}$$

then $V(\sigma^{\text{ALL}})-V(\sigma) \geq 0$ holds. Since the recommended strategy space is finite, there must exists an value $\epsilon_1 = \min_{\sigma \neq \sigma^{\text{ALL}}} \epsilon_\sigma$ such that σ^{ALL} maximizes the achievable social welfare.

Note that we have not yet considered the constraint in (16). However it is easy to see that if $\delta\gamma(\sigma^{\text{ALL}}) > 1$ and if we let $\epsilon^{\text{MD}} + \epsilon^{\text{FA}} \leq 1 - \frac{1}{\delta\gamma(\sigma^{\text{ALL}})} \triangleq \epsilon_2$, then we ensure that σ^{ALL} is an incentive compatible strategy. This proves the theorem. ■

Theorem 2 is an important result regarding the optimality of "efficient working". It states that if monitoring errors are sufficiently small, then the optimal recommended strategy is σ^{ALL}. Therefore, the administrator can simply adopt a strategy that recommends "work" to all agents. We note that Theorem 2 is not as straightforward as it may appear because it provides a much stronger result than the intuitive asymptotic optimality of σ^{ALL} as $\epsilon^{\text{MD}} \to 0$ and $\epsilon^{\text{FA}} \to 0$. Instead, it states that there exists a continuous non-degenerate interval of ϵ^{MD} and ϵ^{FA} such that σ^{ALL} is indeed the optimal recommended strategy.

C. Determining the optimal strategy

We have shown that when monitoring is sufficiently perfect, recommending "work" to all agents is optimal. In general, it is difficult to construct the optimal recommended strategy for arbitrary networks. Since the strategy space is huge (i.e. 2^N), relying on exhaustive search thus is extremely time-consuming (i.e. $O(2^N)$). In this subsection, we propose efficient algorithms, namely the Iterative Deletion (ID) algorithms, to determine a "good" recommended strategy whose complexity is $O(N)$. We note that the ID algorithm may not be able to find the global optimal strategy in all cases but we prove that under certain conditions that the strategy determined by ID algorithm corresponds to the optimal strategy. Nevertheless, our simulations show that the ID algorithm performs extremely well - it finds the optimal rating system in almost all cases. We present the ID algorithm below.

Algorithm Iterative Deletion Algorithm

Input: $\mathcal{N}$, G, $b_{ij}, \forall i, j$ and $c_i, \forall i$.
Output: The optimal rating system $\pi^* = (\sigma^*, \boldsymbol{p}^*)$
$t = 1$, $\sigma^t = \sigma^{\text{ALL}}$.
while $\sigma^t(i) \neq$ null for some i **do**
  Compute $\gamma(\sigma^t)$
  if $\gamma(\sigma^t) \geq \frac{1}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}$ **then**
    Compute $V(\sigma^t)$
  else $V(\sigma^t) = 0$
  end if
  Set $i^* \leftarrow \arg \min_{i:\sigma^t(i)=1} \frac{b_i(\sigma^t)}{c_i}$
  Construct new σ^{t+1} to be

$$\sigma^{t+1}(i) = \begin{cases} \sigma^t(i), \text{if } i \neq i^* \\ \text{null}, \text{if } i = i^* \end{cases} \tag{21}$$

end while
Let $\sigma^* = \arg\max_t V(\sigma^t)$ and $\boldsymbol{p} = \boldsymbol{p}(\sigma^*)$.

The main idea of the ID algorithm is that we iteratively try to use a milder punishment by being tolerant of the shirking behavior of the agents with the least benefit-to-cost ratio in the last iteration. Since using a milder punishment can reduce the social welfare loss caused by the undesired rating drop in the imperfect monitoring scenario, the social welfare may be improved. Since at most N strategies are visited, the complexity of the ID algorithm is $O(N)$. In Theorem 3, we prove that the resulting strategy by running the ID algorithm is the optimal strategy when certain conditions hold, for both discriminative and non-discriminative policies.

Theorem 3. *For both DP and NDP, there exists ϵ_3 (possibly different for DP and NDP) such that for all $\epsilon^{MD}, \epsilon^{FA}$ that satisfy $\frac{\epsilon^{MD}}{1-\epsilon^{MD}-\epsilon^{FA}} < \epsilon_3$, the strategy determined by the ID algorithm is optimal.*

Before we prove Theorem 3, we present and prove a technical lemma.

Lemma 1. *(1) For DP, there exists ϵ_3 such that for all $\epsilon^{MD}, \epsilon^{FA}$ that satisfy $\frac{\epsilon^{MD}}{1-\epsilon^{MD}-\epsilon^{FA}} < \epsilon_3$, then $\forall \sigma_1 \succ \sigma_2$, $V(\sigma_1) > V(\sigma_2)$.*

(2) For NDP, there exists ϵ_3 such that for all $\epsilon^{MD}, \epsilon^{FA}$ that satisfy $\frac{\epsilon^{MD}}{1-\epsilon^{MD}-\epsilon^{FA}} < \epsilon_3$, then $\forall \sigma_1 \succ \sigma_2$, if $\gamma(\sigma_1) > \gamma(\sigma_2)$, then $V(\sigma_1) > V(\sigma_2)$.

Proof: (1) Consider any $\sigma_1 \succ \sigma_2$, we can compute the difference $V(\sigma_1) - V(\sigma_2)$ to be

$$V(\sigma_1) - V(\sigma_2) = \sum_{i:\sigma_1(i)=1,\sigma_2(i)=\text{null}} \left(\sum_j b_{ji} - \left(1 + \frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}\right) c_i \right) \quad (22)$$

Hence, if

$$\frac{\epsilon^{\text{MD}}}{1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}}} < \epsilon_3 \triangleq \delta \frac{\sum_j b_{ji} - c_i}{c_i} \quad (23)$$

then $V(\sigma_1) - V(\sigma_2) \geq 0$.

(2) Consider any $\sigma_1 \succ \sigma_2$, we can compute the difference $V(\sigma_1) - V(\sigma_2)$ to be

$$\begin{aligned} &V(\sigma_1) - V(\sigma_2) = \\ &\sum_i b_i(\sigma_1) - \sum_{i:\sigma_1(i)=1} c_i - \sum_i b_i(\sigma_2) + \sum_{i:\sigma_2(i)=1} c_i \\ &-\tfrac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})} \left(\sum_{i:\sigma_1(i)=1} \tfrac{b_i(\sigma_1)}{\gamma(\sigma_1)} - \sum_{i:\sigma_2(i)=1} \tfrac{b_i(\sigma_2)}{\gamma(\sigma_2)} \right) \\ &> \sum_i b_i(\sigma_1) - \sum_{i:\sigma_1(i)=1} c_i - \sum_i b_i(\sigma_2) + \sum_{i:\sigma_2(i)=1} c_i \\ &-\tfrac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})\gamma(\sigma_1)} \left(\sum_{i:\sigma_1(i)=1} b_i(\sigma_1) - \sum_{i:\sigma_2(i)=1} b_i(\sigma_2) \right) \end{aligned} \quad (24)$$

The last inequality is due to $\gamma(\sigma_1) > \gamma(\sigma_2)$. Because $\boldsymbol{a}^* = \mathbf{1}$ is optimal we have $\sum_i b_i(\sigma_1) - \sum_{i:\sigma_1(i)=1} c_i - \sum_i b_i(\sigma_2) + \sum_{i:\sigma_2(i)=1} c_i \geq 0$. Moreover, because $\sigma_1 \succ \sigma_2$, $\sum_{i:\sigma_1(i)=1} b_i(\sigma_1) - \sum_{i:\sigma_2(i)=1} b_i(\sigma_2) > 0$. Therefore, there must exist ϵ_3 small enough such that $V(\sigma_1) - V(\sigma_2) > 0$. Since the space of ordered strategy pair is finite, we let $\epsilon_3 = \min_{\sigma_1 \succ \sigma_2} \epsilon_{\sigma_1,\sigma_2}$. This proves the second part of this lemma. ■

Now we are ready to prove Theorem 3.

proof of Theorem 3: (1) We prove for the discriminative policy first. Without loss of generality, we assume that, in each iteration, exactly one agent is deleted. We further relabel the indices of the agents according to the deletion sequence. Therefore σ^t recommends "work" to agents $\{t, ..., N\}$.

Now consider any strategy σ' that does not emerge on the iteration path. Let j be the smallest agent index (after relabeling as above) such that $\sigma'(j) = 1$. It is easy to see that $\sigma' \prec \sigma^j$. Therefore, for all i, we have $b_i(\sigma') \leq b_i(\sigma^j)$. Hence, if σ' is incentive-compatible, then σ^j is also incentive-compatible. Moreover, the difference $V(\sigma^j) - V(\sigma')$ satisfies

$$V(\sigma^j) - V(\sigma') \geq \sum_{i:\sigma^j(i)=1,\sigma'(i)=\text{null}} \left(\sum_k b_{ki} - c_i - \frac{\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})} c_i \right) \quad (25)$$

According to Lemma 1, $V(\sigma^j) - V(\sigma') \geq 0$. Since σ' is arbitrary, we now conclude that the optimal σ must be among the strategies emerging on the iteration path.

(2) We then prove for the non-discriminative policy. Similarly, without loss of generality, we assume that, in each iteration, exactly one agent is deleted. We further relabel the indices of the agents according to the deletion sequence. Therefore σ^t recommends "work" to agents $\{t, ..., N\}$. Moreover, among all N strategies emerging on the iteration path, let $\sigma^{t_1}, ..., \sigma^{t_N}$, where $t_1 < ... < t_N$, be the set of strategies such that $\gamma(\sigma^{t_m}) > \gamma(\sigma^1)$. Note that $\sigma^1 = \sigma^{\text{ALL}}$.

Now consider any strategy σ' that does not emerge on the iteration path. Let j be the smallest agent index (after relabeling as above) such that $\sigma'(j) = 1$. We discuss two possibilities below.

First, suppose there exists $t_m = j$. Then it must be $\sigma' \prec \sigma^{t_m}$ and $\gamma(\sigma') \leq \gamma(\sigma^{t_m})$. According to Lemma 1 part 2, $V(\sigma') < V(\sigma^{t_m})$. Note that if σ' is incentive-compatible then σ^{t_m} is also incentive-compatible.

Second, suppose there does not exist $t_m = j$. In this case, $\gamma(\sigma') < \gamma(\sigma^j) < \gamma(\sigma^1)$. Moreover, $\sigma' \prec \sigma^1$. According to Lemma 1 part 2, $V(\sigma') < V(\sigma^1)$. Note that if σ' is incentive-compatible then σ^1 is also incentive-compatible.

Since σ' is arbitrary, we now conclude that the optimal σ must be among the strategies emerging on the iteration path. ■

D. When is "shirking" efficient?

In this subsection, we study the impact of network topology on the optimal design, assuming that the administrator adopts a non-discriminative policy and the benefits and costs are identical for all agents, i.e. $b_{ij} = b, \forall i, j : g_{ij} = 1$ and $c_i = c, \forall i$. Therefore, the benefit that an agent can obtain only depends on the number of neighbors that it connects with. We are interested in determining the network connectivity conditions under which shirking becomes more "efficient" than working. We show this by studying a particular network topology - the "core-periphery" topology. "Core-periphery" topology plays an important role in networking research since a lot of social networks exhibit this topology or consists of many "core-periphery" components [30]. We first give a formal definition of the "core-periphery" topology.

Definition 2. *In a core-periphery topology, agents are divided into two sets: a core set and a periphery set. The connectivity satisfies: (1) every agent in the periphery set connects to some agents in the core set; (2) every agent in the core set connects to some other agents in the core set; (3) there is no edge between any two agents in the periphery set.*

Assume that there are K core agents, each connecting to $K_1 < K$ other core agents, and LK periphery agents, each connecting to $K_2 < K_1$ core agents. Hence, L represents the density of the periphery agents. We denote the strategy that recommends "work" to only the core agents by σ^{CORE}.

Proposition 1. *If $bK_2\delta(1-\epsilon^{MD}-\epsilon^{FA}) > c\epsilon^{MD}$, then there exists L_{min} such that*

- *for all $L \leq L_{min}$, the optimal strategy is σ^{CORE}.*
- *for all $L > L_{min}$, the optimal strategy is σ^{ALL}.*

Proof: To simplify notations, let $\alpha = \frac{c\epsilon^{\text{MD}}}{\delta(1-\epsilon^{\text{MD}}-\epsilon^{\text{FA}})}$. The social welfare of σ^{ALL} is

$$\begin{aligned} &V(\sigma^{\text{ALL}}) = Kb(K_1+LK_2) + LKbK_2 \\ &-\tfrac{\alpha}{bK_2}(Kb(K_1+LK_2) + LKbK_2 + (1+L)Kc) \end{aligned} \quad (26)$$

The social welfare of σ^{CORE} is

$$V(\sigma^{\text{CORE}}) = KbK_1 + LKbK_2 - \frac{\alpha}{bK_1}(KbK_1 + Kc) \quad (27)$$

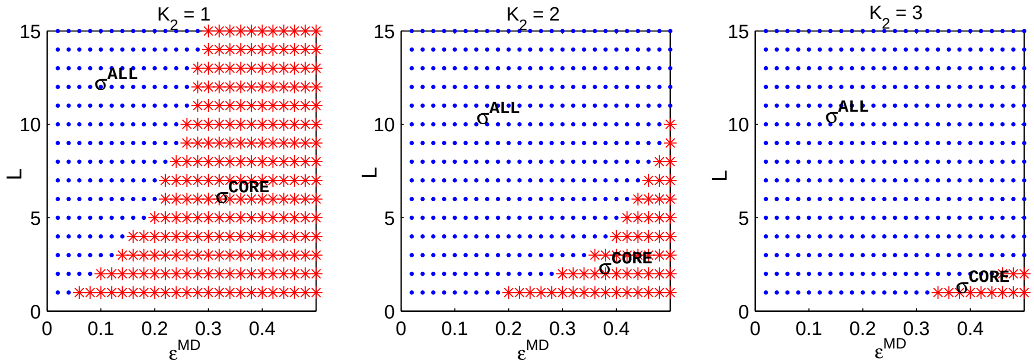

Fig. 2. Optimal recommended strategies under various scenarios.

Their difference $V(\sigma^{\text{ALL}}) - V(\sigma^{\text{CORE}})$ is

$$K\left((bK_1 - \alpha - \frac{\alpha}{bK_2}c)L - \alpha(\frac{K_1}{K_2} - 1)(1 + \frac{c}{bK_1})\right) \quad (28)$$

If $bK_1 - \alpha - \frac{\alpha}{bK_2}c \leq 0$, then $V(\sigma^{\text{ALL}}) - V(\sigma^{\text{CORE}}) < 0$ for all L. In this case, we let $L_{min} = \infty$. If $bK_1 - \alpha - \frac{\alpha}{bK_2}c > 0$, then there must exist a $L_{min} < \infty$ such that $\forall L < L_{min}$, $V(\sigma^{\text{ALL}}) - V(\sigma^{\text{CORE}}) < 0$ and $\forall L \geq L_{min}$, $V(\sigma^{\text{ALL}}) - V(\sigma^{\text{CORE}}) \geq 0$. This completes the proof. ■

Proposition 1 shows that shirking can also become more "efficient" than working in some scenarios. This is because in order to incentivize all agents (in particular, the periphery agents) to work, a very strong punishment needs to be imposed. However, in the imperfect monitoring scenario, strong punishment causes much undesired social welfare loss, thereby reducing the social welfare. When the density of periphery agents is small, this social welfare loss outweighs the social welfare contribution of the periphery agents and hence, higher social welfare can be achieved by allowing the shirking behavior of the periphery agents. We note that "efficient" shirking can only emerge in networks where agents are heterogeneous. In networks where agents are homogeneous, agents' incentives are either satisfied or not satisfied at the same time. Hence, "shirking" requires all agents to shirk at the same time so it is inefficient in any case. The "efficient" shirking effect causes a fairness problem between a core agent and its connected periphery agents in the sense that only the core agent works while the periphery agents shirk. One way to address this fairness issue is for the rating system to adopt a discriminative policy that uses different punishment levels for core and periphery agents. Specifically, to incentivize the periphery agents to work, a stronger punishment is used for the periphery agents than the core agents since periphery agents are more difficult to be incentivized.

VI. Simulation Results

In this section, we provide numerical results to illustrate the efficacy of our proposed rating system. We focus on only the non-discriminative policies (i.e. common punishment for all agents) to avoid redundant results and due to the fact that non-discriminative policies are easier to implement in practice.

A. Efficient Working versus Efficient Shirking

First we study when the optimal scheme requires less productive agents to shirk. We simulate the core-periphery networks studied in Section V.D and show how the monitoring accuracy and the network connectivity influence the resulting optimal rating system design. In this set of simulations, we fix $K = 50$, $K_1 = 40$, $\epsilon^{\text{FA}} = 0.05$, $\delta = 0.99$, $b = 2$ and $c = 1$. We vary the miss detection error probability ϵ^{MD}, the periphery degree K_2 and the density of periphery agents L. Figure 2 shows the optimal strategy for various values of ϵ^{MD} and L. As predicted by Proposition 1, for a fixed ϵ^{MD} and K_2, there is a cutoff value of L above which the optimal strategy is that all agents work and below which the optimal strategy is that only core agents work. This is because as there are more periphery agents in the network, their contribution to the social welfare becomes more significant. Therefore, the optimal strategy should incentivize all agents to work. For a fixed K_2, this cutoff value increases with ϵ^{MD}, suggesting that the impact of undesired rating drop due to imperfect monitoring on the social welfare becomes more significant. Therefore, it is better to use a milder punishment by allowing periphery agents to shirk. For a fixed ϵ^{MD}, the cutoff value of L decreases with K_2. Since K_2 represents the productivity of periphery agents, this suggests that if agents become more productive, "work" should be recommended to them because their contribution to the social welfare dominates over the social welfare loss caused by the undesired punishment.

B. Effectiveness of the ID algorithm

Next, we study the effectiveness of the propose ID algorithm. First we provide an example to illustrate how the ID algorithm works. We consider a network with $N = 6$ agents on

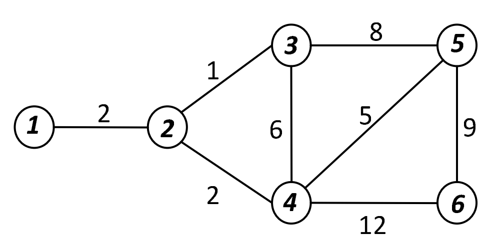

Fig. 3. An illustrative information sharing network.

the topology shown in Figure 3. For illustrative purposes, the costs for all agents are the same, i.e. $c_i = 1, \forall i$ and the benefits between connected agents are symmetric, i.e. $b_{ij} = b_{ji}$. The exact values of benefits are shown on the edges in the figure. The discount factor is $\delta = 0.9$. We perform the ID algorithm for the monitoring error probabilities being $\epsilon^{\text{MD}} = 0.1$ and $\epsilon^{\text{MD}} = 0.3$.

The recommended strategies emerging on the algorithm path and their corresponding social welfare are reported in Table II. The emerging recommended strategies are the same for both cases since the deletion process does not depend on the monitoring errors. However, the optimal social welfare is different for different strategies and for different error probabilities. For $\epsilon^{\text{MD}} = 0.1$, the optimal recommended strategy recommends "work" to agents 2, 3, 4, 5, 6. For $\epsilon^{\text{MD}} = 0.3$, the optimal recommended strategy recommends "work" to agents 3, 4, 5, 6. As we can see, when the monitoring error probability is large, allowing some agents to shirk significantly outperforms the strategy that recommends "work" to all agents.

We further evaluate the performance of the ID algorithm for random graphs. In this set of simulations, we consider a random network of $N = 10$ agents (we use a small number of agents in order to enable the exhaustive search approach for comparison purpose). The network is constructed in the following way: each agent i is associated with a value $q_i = i^{-0.2}$; agent i and agent j are connected with each other with probability $q_i q_j$. We fix $b = 1.5$, $c = 1$, $\delta = 0.99$ and $\epsilon^{\text{FA}} = 0$. We vary ϵ^{MD} to investigate its impact on the resulting optimal rating system. Figure 4 illustrates the optimal social welfare when agents are compliant, the social welfare by adopting the optimal rating systems determined by exhaustive search, the ID algorithm and simply recommending "work" to all agents. Each simulation result is obtained by averaging the social welfare of 500 randomly generated networks. This simulation has several implications:

- The social welfare achieved by the rating system decreases with the increase of the monitoring error. This is because imperfect monitoring inevitably introduces undesired rating drop which leads to social welfare loss.
- Even though we prove (in Theorem 3) the optimality of the ID algorithm only under certain conditions, the ID algorithm actually is able to find the optimal rating system in almost all scenarios. We note again that the complexity of the exhaustive search is $O(2^N)$ while the complexity of the ID algorithm is significantly lower, i.e $O(N)$.
- As predicted by Theorem 2, the simple rating system design σ^{ALL} which recommends "work" to all agents is optimal when the monitoring is sufficiently accurate (in the simulated networks, at least for $\epsilon^{\text{MD}} < 0.05$). However, σ^{ALL} can be very inefficient when the monitoring error is large. There are two forces that lead to this result. First, in order to incentivize less-productive agents to work, a strong punishment is required and hence, much social welfare is induced by undesired rating drop. Second, when the monitoring error is high, even the strongest punishment (i.e. $p = 0$) may not be able to sustain the incentive-compatibility of σ^{ALL}.

TABLE II
RECOMMENDED STRATEGIES AND THEIR SOCIAL WELFARE EMERGING ON THE ALGORITHM PATH.

t	1	2	3	4	5	6	$V(\sigma^t)$ ($\epsilon^{MD} = 0.1$)	$V(\sigma^t)$ ($\epsilon^{MD} = 0.3$)
1	**2**	5	15	25	22	21	78.44	62.57
2	--	**3**	15	25	22	21	**79.46**	69.35
3	--	--	**14**	23	22	21	78.29	**76.28**
4	--	--	--	17	**14**	21	64.54	63.23
5	--	--	--	**12**	--	12	43.75	43.04
6	--	--	--	--	--	**0**	0	0

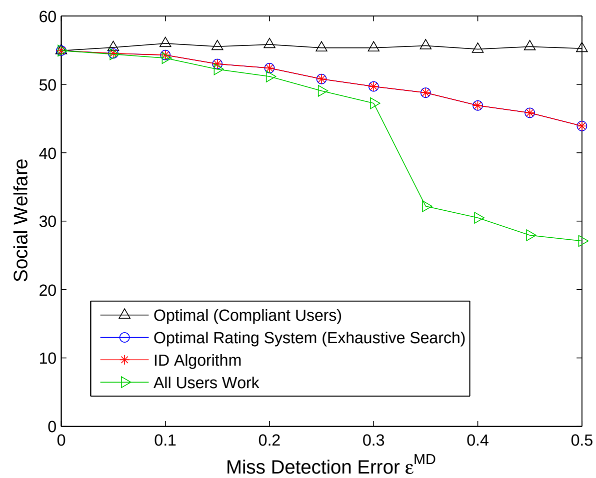

Fig. 4. Performance of the ID algorithm.

C. Comparison with other incentive mechanisms

In this experiment, we compare the performance of the proposed rating system with two benchmark incentive mechanisms, i.e. the widely-studied Grim-Trigger strategy (e.g. [29]) and Tit-for-Tat strategy (e.g. [8]). In the deployment under consideration, these strategies are described below:

- *Grim-Trigger strategy*: If an agent is observed to be not working in a period, the network administrator permanently blocks it from accessing the network (hence, it does not have access to its neighbors' information in all following periods).
- *Tit-for-Tat strategy*: If an agent is observed to be not working in a period, the network administrator blocks it from accessing to its neighbors' content in the next period. If an agent is observed to be working in a period, the network administrator grants it full access to its neighbors' information in the next period.

Figure 5 shows the social welfare achieved by these two benchmark schemes and our proposed rating system for $N = 30, 40, 50$ and various miss detection error probabilities. In each simulation, a random graph is generated in a similar manner as in Section VI.B. We fix parameters to be $b_0 = 1.5$, $c = 1$, $\delta = 0.99$, $\epsilon^{\text{FA}} = 0$. Each point of the curve is obtained by running 500 simulations. As we can see, Trigger strategy only achieves efficiency when the monitoring is perfect. In the imperfect monitoring scenario ($\epsilon^{\text{MD}} > 0$), all agents will eventually be blocked from the network, leading to the

TABLE III
PERFORMANCE IMPROVEMENT AGAINST TFT FOR VARIOUS NETWORK SIZES. (%)

N	*100*	*200*	*300*	*400*	*500*
$\epsilon^{MD} = 0.1$	7.06	7.10	7.11	7.12	7.13
$\epsilon^{MD} = 0.2$	15.18	15.28	15.31	15.33	15.34
$\epsilon^{MD} = 0.3$	24.58	24.78	24.85	24.89	24.91

network collapse. Tit-for-Tat strategy is less vulnerable to monitoring errors. However, because 1) the punishment to the agents when they are observed to be not working is not tailored to the specific network and monitoring environment and 2) incentivizing all agents may not even be possible, Tit-for-Tat strategy induces much social welfare loss when the monitoring error becomes large. Instead, our rating system is able to use the "correct" punishment tailored to the specific network structure as well as the monitoring environment and recommend "work" to only the "correct" agents and hence, it achieves high social welfare even when the monitoring error is large. Our simulations indicate a performance improvement up to 80% against the Tit-for-Tat strategy by the proposed optimal rating system design.

Next we discuss the communication overhead of the proposed rating system. After each interaction, agents report the observed actions of all their neighbors to the network administrator who will synthesize this knowledge to compute the agents' next period ratings. Specifically, each agent i sends one message (i.e. a vector of size $\sum_j g_{ij}$) about the observed actions of its neighbors which is to the network administrator. The total number of messages sent from agents to the network administrator is N since there are N agents. Based on these messages, the administrator updates each agent's rating and then informs each agent of its own rating and neighbors' ratings of the next period. For agent i, the message sent by the administrator is a vector of size $\sum_j g_{ij} + 1$. This requires another N message exchanges since there are N agents. Therefore, the total number of messages exchanged in one period is $2N$ (note that a message is a vector). The benchmark schemes (i.e. Grim-Trigger strategy and Tit-for-Tat strategy) can be implemented in a fully decentralized way, requiring no information exchange/communication overheads among agents. However, the communication overhead of the control messages of our proposed rating system is negligible compared to the information/traffic exchanged among agents when they exchange services. With only limited message exchanges, our proposed scheme can achieve significantly better performance than the benchmark schemes.

Finally, we study the scalability of the proposed rating system in terms of the network size. The performance improvement (in %) against the Tit-for-Tat scheme is reported in Table III. As we can see, the proposed rating system scales with the network size and even improves slightly when the network becomes larger.

VII. CONCLUSION

In this paper, we studied how to design rating systems aimed at maximizing the social welfare of information sharing networks where agents are interconnected according to an exogenously determined topology. We showed that it is possible to exploit the ongoing nature of agents' interactions to design rating systems to incentivize agents to work (i.e. produce information) in order to improve the social welfare of the network. Our analysis showed that the imperfect monitoring and the agent's heterogeneity in terms of benefits and specific connectivity can strongly influence the agents' self-interested decisions and incentives. Surprisingly, in some scenarios, allowing a certain level of shirking behavior of agents can achieve even higher social welfare than incentivizing all agents to work. This significantly differs from existing works on ratings and social norms in which incentivizing all agents to work is always optimal.

In this paper, we assumed the existence of a central network administrator who operates the rating system. In many networks and systems, such as crowdsourcing, peer-to-peer, cybersecurity, a central agency tracking the reputation of agents does often exist in practice [31] and can be used to implement our proposed strategies. However, developing decentralized rating systems is important for many networks and systems that do not have such agencies or in which connecting with such agencies is prohibitively expensive. This is an important topic for future research.

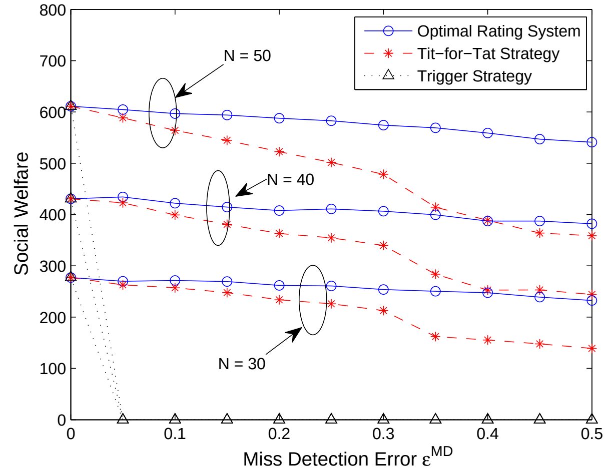

Fig. 5. Performance comparison with Tit-for-Tat and Trigger strategies.

REFERENCES

[1] V. Krishnamurthy and H. V. Poor, "Social learning and bayesian games in multiagent signal processing: How do local and global decision makers interact?" *Signal Processing Magazine, IEEE*, vol. 30, no. 3, pp. 43–57, 2013.

[2] F. S. Cattivelli and A. H. Sayed, "Modeling bird flight formations using diffusion adaptation," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2038–2051, 2011.

[3] A. Nedic and A. Ozdaglar, "Cooperative distributed multi-agent," *Convex Optimization in Signal Processing and Communications*, p. 340, 2010.

[4] B. Q. Zhao, J. C. Lui, and D.-M. Chiu, "Analysis of adaptive incentive protocols for p2p networks," in *INFOCOM 2009, IEEE*. IEEE, 2009, pp. 325–333.

[5] Y. Zhang, J. Park, and M. van der Schaar, "Rating protocols for online communities," *ACM Transactions on Economics and Computation*, 2013.

[6] J. Xu and M. van der Schaar, "Social norm design for information exchange systems with limited observations," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 11, pp. 2126–2135, 2012.
[7] M. Kandori, "Social norms and community enforcement," *The Review of Economic Studies*, vol. 59, no. 1, pp. 63–80, 1992.
[8] V. Srinivasan, P. Nuggehalli, C.-F. Chiasserini, and R. R. Rao, "Cooperation in wireless ad hoc networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 2. IEEE, 2003, pp. 808–817.
[9] J. Xu, Y. Song, and M. van der Schaar, "Sharing in networks of strategic agents," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 4, pp. 717–731, Aug 2014.
[10] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, vol. 1, no. 1, pp. 3–49, 2011.
[11] J. Park and M. van der Schaar, "A game theoretic analysis of incentives in content production and sharing over peer-to-peer networks," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 4, pp. 704–717, 2010.
[12] C. Lampe, R. Wash, A. Velasquez, and E. Ozkaya, "Motivations to participate in online communities," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2010, pp. 1927–1936.
[13] S. Jain, Y. Chen, and D. C. Parkes, "Designing incentives for online question and answer forums," in *Proceedings of the 10th ACM conference on Electronic commerce*. ACM, 2009, pp. 129–138.
[14] V. K. Singh, R. Jain, and M. S. Kankanhalli, "Motivating contributors in social media networks," in *Proceedings of the first SIGMM workshop on Social media*. ACM, 2009, pp. 11–18.
[15] D. Bergemann and D. Ozmen, "Optimal pricing with recommender systems," in *Proceedings of the 7th ACM conference on Electronic commerce*. ACM, 2006, pp. 43–51.
[16] J. K. MacKie-Mason and H. R. Varian, "Pricing congestible network resources," *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 7, pp. 1141–1149, 1995.
[17] H. Masum and Y.-C. Zhang, "Manifesto for the reputation society," *First Monday*, vol. 9, no. 7, 2004.
[18] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The eigentrust algorithm for reputation management in p2p networks," in *Proceedings of the 12th international conference on World Wide Web*. ACM, 2003, pp. 640–651.
[19] S. Ba and P. A. Pavlou, "Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior," *MIS quarterly*, pp. 243–268, 2002.
[20] P. Resnick and R. Zeckhauser, "Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system," *Advances in applied microeconomics*, vol. 11, pp. 127–157, 2002.
[21] C. Dellarocas, "Reputation mechanism design in online trading environments with pure moral hazard," *Information Systems Research*, vol. 16, no. 2, pp. 209–230, 2005.
[22] M. Fan, Y. Tan, and A. B. Whinston, "Evaluation and design of online cooperative feedback mechanisms for reputation management," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 17, no. 2, pp. 244–254, 2005.
[23] G. Zacharia, A. Moukas, and P. Maes, "Collaborative reputation mechanisms for electronic marketplaces," *Decision Support Systems*, vol. 29, no. 4, pp. 371–388, 2000.
[24] M. Okuno-Fujiwara and A. Postlewaite, "Social norms and random matching games," *Games and Economic behavior*, vol. 9, no. 1, pp. 79–109, 1995.
[25] G. Ellison, "Cooperation in the prisoner's dilemma with anonymous random matching," *The Review of Economic Studies*, vol. 61, no. 3, pp. 567–588, 1994.
[26] S. N. Ali and D. Miller, "Enforcing cooperation in networked societies," *Unpublished manuscript, University of California at San Diego*, 2009.
[27] M. O. Jackson, *Social and economic networks*. Princeton University Press, 2010.
[28] M. Haag and R. Lagunoff, "Social norms, local interaction, and neighborhood planning," *International Economic Review*, vol. 47, no. 1, pp. 265–296, 2006.
[29] G. J. Mailath and L. Samuelson, "Repeated games and reputations: long-run relationships," *OUP Catalogue*, 2006.
[30] E. M. Airoldi and K. M. Carley, "Sampling algorithms for pure network topologies: a study on the stability and the separability of metric embeddings," *ACM SIGKDD Explorations Newsletter*, vol. 7, no. 2, pp. 13–22, 2005.
[31] AmazonMechanicalTurk, https://www.mturk.com.