Game-Theoretic Design of Optimal Two-Sided Rating Protocols for Service Exchange Dilemma in Crowdsourcing

Jianfeng Lu, Member, IEEE, Yun Xin, Zhao Zhang, Xinwang Liu, Kenli Li

Abstract—Despite the increasing popularity and successful examples of crowdsourcing, its openness overshadows important episodes when elaborate sabotage derailed or severely hindered collective efforts. A service exchange dilemma arises when non-cooperation among self-interested users, and zero social welfare is obtained at myopic equilibrium. Traditional rating protocols are not effective to overcome the inefficiency of the socially undesirable equilibrium due to specific features of crowdsourcing: a large number of anonymous users having asymmetric service requirements, different service capabilities, and dynamically joining/leaving a crowdsourcing platform with imperfect monitoring. In this paper, we develop the first game-theoretic design of the two-sided rating protocol to stimulate cooperation among self-interested users, which consists of a recommended strategy and a rating update rule. The recommended strategy recommends a desirable behavior from three predefined plans according to intrinsic parameters, while the rating update rule involves the update of ratings of both users, and uses differential punishments that punish users with different ratings differently. By quantifying necessary and sufficient conditions for a sustainable social norm, we formulate the problem of designing an optimal two-sided rating protocol that maximizes the social welfare among all sustainable protocols, provide design guidelines for optimal two-sided rating protocols and a low-complexity algorithm to select optimal design parameters in an alternate manner. Finally, illustrative results show the validity and effectiveness of our proposed protocol designed for service exchange dilemma in crowdsourcing.

Index Terms—service exchange, incentive mechanism, rating protocol, differential punishment, sustainable social norm, game theory

I. INTRODUCTION

CROWDSOURCING has emerged in recent years as a paradigm for leveraging human intelligence and activity at a large scale, it offers a distributed and cost-effective approach to obtain needed content, information or services by soliciting contributions from an undefined set of people, instead of assigning a job to designated employees. Over the past decade, numerous successful crowdsourcing platforms, such as Amazon Mechanical Turk (AMT), Yahoo! Answers, Upwork emerge. With the help of crowdsourcing platforms and with the power of a crowd, crowdsourcing is becoming increasingly popular as it provides an efficient and cheap method of obtaining solutions to complex tasks that are currently beyond computational capabilities but possible for humans.

Over the past decade, techniques for securing crowdsourcing operations have been expanding steadily, so has the number of applications of crowdsourcing. However, users in a crowdsourcing platform have the opportunity to exhibit antisocial behaviors due to the openness of crowdsourcing. Meanwhile, crowdsourcing has overshadowed important episodes when elaborate sabotage derailed or severely hindered collective efforts. As part of crowdsourcing, service exchange applications have proliferated as the medium that allows users to exchange valuable services. In a typical service exchange application, a user plays a dual role: as a client (sometimes also called a requester) who submits requested services to a crowdsourcing platform, and as a server (or worker) who chooses to devote a high/low level of efforts to work on a job and provides solutions to the client in exchange for rewards. Since providing services incurs costs to servers in terms of power, time, bandwidth, privacy leakage, etc., rational and self-interested users would be more inclined to devote low level efforts when being a server, and seek for services from others as a client, rather than providing services as a server. Under such circumstances, non-cooperative behaviors among self-interested users decrease their social welfare, which is a social dilemma. Therefore, an increased level of cooperation is considered to be socially desirable for service exchange in crowdsourcing platforms.

The main reason why users in the above service exchange game have the incentive not to cooperate with each other is the absence of punishments for such malicious behaviors. Self-interested users always adjust their strategies over time to maximize their own utilities, however, they cannot receive a direct and immediate benefit by choosing to be a server and devoting a high-level effort to provide high-quality services to other users (as clients). Such a conflict leads to an inevitable fact that, many users would be apt to be a client to request services, or be apt to be a server but devote a low-level effort to provide low-quality services. Thus, an important functionality of the crowdsourcing platform is to provide a good incentive mechanism for service exchange. And there is an urgent need to stimulate cooperation among self-interested users in crowdsourcing, under which self-interested users will be compelled to follow the social norm such that the inefficiency of the socially undesirable equilibrium will be overcome, i.e., if a user chooses to be a server in the first stage, and provides...
Incentives are key to the success of crowdsourcing as it heavily depends on the level of cooperation among self-interested users. There are two types of incentives, monetary and non-monetary. Incentive mechanisms based on monetary incentivize individuals to provide high-quality services relying on monetary or matching rewards in the form of micropayments, which in principle can achieve the social optimum by internalizing external effects of self-interested individuals. Although monetary incentives, in some sense, are the best and easiest way to motivate people [14], several challenges prevent monetary incentives from success in service exchange applications. Firstly, it is difficult to price small services (e.g., answer, knowledge, resources etc.) being exchanged between users as these are not real goods [15]. Deploying auctions to set the price may reduce the price to a certain degree, while it may cause implementation complexity, high delay, and currency inflation [16]. Secondly, as pointed out by [17], [18] and [19], “free-riding” may happen when rewards are paid before providing services, a server always has the incentive to take the reward without devoting enough effort, whereas if rewards are paid after the service exchange is completed, “false-reporting” may arise since the client has an incentive to lower or refuse rewards to servers by lying about the outcome of the task. Thirdly, although a monetary scheme is simple to be designed, it often requires a complex accounting infrastructure, which introduces computation overheads and substantial communication, and thus difficult to be implemented in reality [20], [21].

In addition to monetary incentives, some applications are endowed with different non-monetary incentive types, such as natural incentives, personal development, solidary incentives, material incentives, etc. [14]. Among these non-monetary incentives, rating protocols (as a form of solidary incentives) originally proposed by Kandori [22] have been shown to work effectively as incentive mechanisms to force cooperation in crowdsourcing platforms [12], [13], [15], [23], [24]. Generally speaking, a rating protocol labels each user by a rating label based on his past behaviors indicating his social status in the system. And users with different ratings are treated differently by the other users they interact with, and the rating of a user who complies with (resp. deviates from) the social norm goes up (resp. down). Hence, a user with high/low rating can be rewarded/punished by other users in a crowdsourcing platform who have not had past interactions with him. Furthermore, the use of ratings as a summary record of a user requires significantly less amount of information being maintained [25]. Hence, the rating protocol has a potential to form a basis for successful incentive mechanisms for service exchange in crowdsourcing platforms. Motivated by the above considerations, this paper is devoted to the study of incentive mechanisms based on rating protocol.

However, there are several major reasons that prevent existing works on the rating protocol to be directly implemented for incentive provision for service exchange in crowdsourcing: (i) Users have asymmetric service requirements and they can freely and frequently change their partners they interact with in most crowdsourcing platforms, which results in asymmetric interactions among those users, and it is more difficult to model and analyze [26], [27]; (ii) Taking into account the service capability of users and the spatial/temporal requirements of tasks, using the framework of anonymous random matching games in which each user is repeatedly matched with different partners over time for service exchange is inappropriate [12], [15]; (iii) User population is large, users are anonymous and not sufficiently patient, especially when those users with bad ratings may attempt to leave and rejoin the system as new members to avoid punishments (i.e., whitewashing) [23], [28]; (iv) In the presence of imperfect monitoring, a user’s rating may be wrongly updated, which will impact on rating protocol design, as well as social welfare loss [13], [24].

In this paper, we take into account the above features of service exchange in crowdsourcing into consideration, and propose a game-theoretic framework for designing and analyzing a class of rating protocols based incentive mechanisms, in order to stimulate cooperation among self-interested users and maximize the social welfare. To the best of our knowledge, update of ratings of both users (we name it as a two-sided rating) matched in the service exchange game is rarely tackled by other works. Using game theory to analyze how cooperation can be enforced and how to maximize the social welfare under the designed two-sided rating protocol, we rigorously analyze how users’ behaviors are influenced by intrinsic parameters and design parameters as well as users’ evaluation of their individual long-term utilities, in order to characterize the optimal design that maximizes users’ utilities and enforces cooperation among them. The main contributions of this paper are summarized as follows:

(i) We model the service exchange problem as an asymmetric game model with two stages, and show that inefficient outcome arises when no user cooperates with each other, and thus zero social welfare is obtained at myopic equilibrium, which is a social dilemma.
(ii) We develop the first game-theoretic design of two-sided rating protocols to stimulate cooperation among self-interested users, which consists of a recommended strategy and a rating update rule. The recommended strategy recommends a desirable behavior chosen from three predefined recommended plans according to intrinsic parameters, while the rating update rule involves the update of ratings of both users, and uses differential punishments that punish users with different ratings differently.
(iii) We formulate the problem of designing an optimal two-sided rating protocol that maximizes the social welfare among all sustainable rating protocols, provide design guidelines for determining whether there exists a sustainable two-sided rating protocol under a given recommended strategy, and design an algorithm achieving low-complexity computation via a two-stage procedure, each stage consists of two steps (we call this a two-stage two-step procedure), in an alternate manner.
(iv) We use simulation results to demonstrate how intrinsic parameters (i.e., costs, imperfect monitoring, user’s patience) impact on optimal recommended strategies, the
The remainder of this article is organized as follows. In section II, we describe the service exchange dilemma game with two-sided rating protocols. In section III, we formulate the problem of designing an optimal two-sided rating protocol. Then we provide the optimal design of two-sided rating protocols in Section IV. Section V presents simulation results to illustrate key features of the designed protocol. Finally, conclusions are drawn in Section VI.

II. SYSTEM MODELS

A. Service Exchange Dilemma Game

We consider a crowdsourcing where each user can offer a valuable service to other users. Examples of services are sensing task, expert knowledge, information resource, computing power, storage space, etc. In each period, a client generates a service request, which is sent to a server that can provide the requested service. We model such a process using uniform random matching, that is each user in the community is involved in two matches in every period, one as a client and the other as a server, each user is equally likely to receive exactly one request in every period, and the matching is independent in different periods. Note that the user with whom a user interacts as a client can be different from that with whom he interacts as a server, reflecting asymmetric interests between a pair of users in a given instant. Such a model well approximates the matching process between users in large-scale crowdsourcing systems where users interact with others in an ad-hoc fashion and the interactions between users are constructed randomly over time.

In this model, a user decides whether or not to request service (choosing to be a client/server), if the user chooses to be a server, he strategically determines his service quality (devoting a high/low level of efforts). Note that the decisions are sequential: the decision on role selection is made first, and then the decision on service quality is made next. We model this interaction as a sequential game. Formally, we have a two-stage game. In the first stage, a user’s action is chosen from a set \{C, S\}, where C stands for “choosing to be a client” (request service), whereas S stands for “choosing to be a server” (offer service). In the second stage, the server has a binary choice of whether being whole-hearted or being half-hearted in providing the service, while the client has no choice. The set of actions for the server is denoted by \{H, L\}, where H stands for “high level of effort” and L for “low level of effort”.

We assume that any C strategy is costly (consumes a cost c for choosing C). If the server devotes a high level of efforts to fulfill the client’s request, the client receives a service benefit of \(b > 0\), while the server suffers a service cost of \(s > 0\). If the server devotes a low level of efforts to the request, both users receive zero payoffs. Obviously, the server’s action determines the payoffs of both users. After a server takes an action, the client sends a report about the action of the server to the third-party device or infrastructure that manages the rating scores of users. However, the report is inaccurate, either by the client’s incapability of accurate assessment or by some system error with a small probabilities \(\varepsilon\). That is, \(L\) is reported when the server takes action \(H\) with probability \(\varepsilon\), and vice versa. Assuming a binary set of reports, it is without loss of generality to restrict \(\varepsilon \in (0, \frac{1}{2})\), because when \(\varepsilon = \frac{1}{2}\), reports are completely random and do not contain any meaningful information about the actions of users.

We find the subgame perfect equilibrium of the two-stage game. Each pair of users’ decisions made in the first stage (C or S) result in a different second-stage game (H or L). We first compute expected utilities in the second-stage game, and then turn back to compute expected utilities when both users choose their actions in the first-stage before knowing their productivities. When a user requests services as a client, a matching rule is used to determine corresponding server. We model the interaction between a pair of matched users in the second stage as a gift-giving game [29], and the payoff matrix of the gift-giving game between a client and a server is presented in Table I. We assume that \(b > \frac{c + s}{2}\) so that the service of a user creates a positive net social benefit, and social welfare is maximized when all servers choose action \(H\) in the gift-giving games they play, which yields payoff \((1 - \varepsilon)b - c - s\) to every user. On the contrary, action \(L\) is the dominant strategy for the server, which constitutes a Nash equilibrium of the gift-giving game.

We now take a step back and compute expected utilities for such a case. When the client consumes a cost \(c\) for choosing \(C\) in the first stage, and receives \(\varepsilon b - c\) payoff in the second stage, while the server will choose \(L\) and suffer a low cost, which is approximated by 0 here, also results in zero payoff. We summarize this in the \(CS\) and \(SC\) cells of the pay-off matrix in Table I. When both users choose to request services as clients, each user consumes a cost \(c\), but receives zero service benefit as there is no user offering service. We note this in the \(CC\) cell of the pay-off matrix in Table I. The case when both users choose to be servers is very similar, the only difference is that each user suffers no cost. The expected utility of each user is zero as no user requests service. The \(SS\) cell of the pay-off matrix in Table I describes such a case.

In summary, for any choice of parameters, only \(SS\) can be a Nash equilibrium of the service exchange game. When every user chooses his action to maximize his current payoff

**Table I**

Payoff Matrix of a Gift-Giving Game Between a Client and a Server

| Server | \(H\) | \(L\) |
|--------|-------|-------|
| Client | \((1 - \varepsilon)b, -s\) | \(\varepsilon b, 0\) |

**Table II**

The Expected Payoff Matrix for the First Stage Game

| \(C\) | \(-c, -c\) | \(\varepsilon b - c, 0\) |
| \(S\) | \(b, \varepsilon b - c\) | \(0, 0\) |
myopically, an inefficient outcome arises where every user receives zero payoff, which is a social dilemma. Under the current framework, nobody will take the initiative to help others, and do not expect to get help from others.

B. Two-sided Rating Protocols

We consider a two-sided rating protocol that consists of a recommended strategy and a rating update rule. The recommended strategy prescribes the contingent plan according to intrinsic parameters that the server should take based on ratings of both his own and his client’s. Here, we focus on one plan, while two other plans will be introduced in the later half of this article. The rating update rule involves the update of ratings of both users depending on their past actions as a server or a client, and uses differential punishments that punish users with different ratings differently. To the best of our knowledge, two-sided rating protocol in crowdsourcing is rarely tackled by other works. In the following, we give a formal definition of a two-sided rating protocol.

Definition 1. A two-sided rating protocol \( \mathcal{P} \) is represented as a 5-tuple \((\Theta, \sigma, \rho, \pi, \tau)\), i.e., a set of binary rating labels \( \Theta \), a social strategy \( \sigma \), a client/server ratio \( \rho \), a recommended strategy \( \pi \), and a rating update rule \( \tau \).

- \( \Theta = \{0, 1\} \) denotes the set of binary rating labels, where 0 is the bad rating, and 1 is the good rating.
- \( \sigma : \Theta \rightarrow A \) represents the adopted social strategy for a user with rating \( \theta \), where \( \sigma(\theta | \theta \in \Theta) \in \{\{C, S\} \times \{H, L\}\} \).
- \( \rho : \Theta \rightarrow R^+ \) keeps a record of the client/server ratio for a user with rating \( \theta \), which contains his history and current choices of \( \sigma(\theta | \theta \in \Theta) \in \{\{C, S\} \}
- \( \pi : \Theta \times \Theta \rightarrow A \) defines the strategy \( \pi(\theta_S, \theta_C) \in A \) which the server with rating \( \theta_S \) should select when faced with the client with rating \( \theta_C \).

\[
\pi(\theta_S, \theta_C) = \begin{cases} 
1, & \text{if } \theta_S \leq \theta_C \\
0, & \text{otherwise} 
\end{cases}
\]  

(1)

- \( \tau : \Theta \times \Theta \times A \rightarrow \Delta(\Theta \times \Theta) \) specifies how a user’s rating should be updated based on its adopted strategies and current rating as follows:

\[
\tau(\theta'_S, \theta'_C | \theta_S, \theta_C, r, \rho) = \begin{cases} 
\alpha_{\theta_S}, & \theta'_S = 1, r \geq \pi(\theta_S, \theta_C) \\
1 - \alpha_{\theta_S}, & \theta'_S = 0, r \geq \pi(\theta_S, \theta_C) \\
\beta_{\theta_S}, & \theta'_S = 0, r < \pi(\theta_S, \theta_C) \\
1 - \beta_{\theta_S}, & \theta'_S = 1, r < \pi(\theta_S, \theta_C) \\
\gamma_{\theta_C}, & \theta'_C = 1, \rho(\theta_C) \leq 1 \\
1 - \gamma_{\theta_C}, & \theta'_C = 0, \rho(\theta_C) \leq 1 \\
\delta_{\theta_C}, & \theta'_C = 0, \rho(\theta_C) > 1 \\
1 - \delta_{\theta_C}, & \theta'_C = 1, \rho(\theta_C) > 1 
\end{cases}
\]  

(2)

We characterize the erroneous report by a mapping \( R : \{0, 1\} \rightarrow \Delta(\{0, 1\}) \), where 0 and 1 represent “L” and “H”, respectively. \( \Delta(\{0, 1\}) \) is the probability distribution over \{0, 1\}, and \( R(r|q) \) is the probability that the client reports “r” given the server’s actual service quality “q”.

\[
R(r|q) = \begin{cases} 
1 - \varepsilon, & r = q \\
\varepsilon, & r \neq q 
\end{cases} \quad \forall r, q \in \{0, 1\}. 
\]  

(3)

A schematic representation of a rating update rule \( \tau \) is provided in Figure 1. Given a rating protocol \( \mathcal{P} \), each user \( i \) is tagged with a binary rating label \( \theta_i \in \Theta \triangleq \{0, 1\} \) representing its social status. Obviously, the higher \( \theta_i \) is, the better the social status the user \( i \) has. Ratings of users are stored and updated by the system administrator based on strategies adopted by the user in the transactions that he is engaged in. The rating scheme \( \tau \) can update a user’s rating at the end of each transaction or at the beginning of the next transaction. Under the rating update rule \( \tau \), a \( \theta_S \)-server (i.e., a server with rating \( \theta_S \)) will have rating 1 with probability \( \alpha_{\theta_S} \), and have rating 0 with probability \( 1 - \alpha_{\theta_S} \). We consider the cases where the server is rated by another user. Let \( \theta_C \) be the rating received by the server from the client and have rating \( \theta_C \) with probability \( \beta_{\theta_C} \) and have rating 0 with probability \( 1 - \beta_{\theta_C} \). While a \( \theta_C \)-client will have rating 1 with probability \( \gamma_{\theta_C} \), and have rating 0 with probability \( 1 - \gamma_{\theta_C} \), if the client/server ratio \( \rho \leq 1 \); otherwise, it will have rating 1 with probability \( \delta_{\theta_C} \) and have rating 0 with probability \( 1 - \delta_{\theta_C} \). Obviously, \( \alpha_{\theta_S} \) and \( \gamma_{\theta_C} \) can be referred to as the strength of reward imposed on servers and clients when they cooperate with each other, respectively, while \( \beta_{\theta_S} \) and \( \delta_{\theta_C} \) can be referred to as the strength of punishment imposed on clients when they expect to get excessive service from others rather than to serve others.

III. PROBLEM FORMULATION

A. Stationary Rating Distribution

Given a two-sided rating protocol \( \mathcal{P} \), suppose that each user always follows a given recommended strategy \( \pi \) and keep \( \rho \leq 1 \) in any period. As time passes, ratings of users are updated, and transition probabilities of a user’s rating can be expressed as follows. Then the transition from \( \eta_P^{(t)}(\theta) \) to \( \eta_P^{(t+1)}(\theta) \) is determined by the rating update rule \( \tau \), taking into account the rate \( \lambda \) for a user choosing to be a client and the error probability \( \varepsilon \), as shown in the following expressions:
distribution that users should follow, the stationary distribution is also user to follow the recommended strategy \( \theta \) follows a recommended strategy \( \theta \) matched with a \( \rho \) distribution, because we consider a continuum of users. Any unilateral deviation from an individual user would not affect the evolution of rating scores and thus the stationary distribution \( \{\eta_\pi(t)\}^t_{t=0} \) can be derived as follows.

\[
\begin{align*}
\eta_\pi(0) &= \frac{\phi_1}{1 + \phi_1 - \phi_0} \\
\eta_\pi(1) &= \frac{1 - \phi_0}{1 + \phi_1 - \phi_0}
\end{align*}
\]

Since the coefficients in the equations that define a stationary distribution are independent of the recommended strategy that users should follow, the stationary distribution is also independent of the recommended strategy, as can be seen from Eq. (5). Thus, we will write the stationary distribution as \( \{\eta_\pi(\theta)\} \) to emphasize its dependence on \( \pi \).

### B. Sustainable Conditions

The purpose of designing a social norm is to enforce a user to follow the recommended strategy \( \pi(\theta, \theta_C) \) and keep \( \rho \leq 1 \) in any period. We call a user who complies with such a social norm as a “compliant user”, otherwise, the user who deviates from the social norm is called as a “non-compliant user”. The compliant user will be rewarded, on the contrary, a non-compliant user will be punished in order to regulate his behavior. Since we consider a non-cooperative scenario, it is important to check whether a user can improve his long-term payoff by a unilateral deviation. Note that any unilateral deviation from an individual user would not affect the evolution of rating scores and thus the stationary distribution, because we consider a continuum of users.

Let \( c_\pi(\theta, \theta_C) \) be the cost paid by a \( \theta_C \)-server who is matched with a \( \theta_C \)-client and follows a recommended strategy \( \pi \), that is, \( c_\pi(\theta, \theta_C) = s \) if \( \pi(\theta, \theta_C) = 1 \), and \( c_\pi(\theta, \theta_C) = 0 \) if \( \pi(\theta, \theta_C) = 0 \). Similarly, let \( b_\pi(\theta, \theta_C) \) be the benefit received by a \( \theta \)-client who is matched with a \( \theta \)-server who follows a recommended strategy \( \pi \), that is, \( b_\pi(\theta, \theta_C) = b - c \) if \( \pi(\theta, \theta_C) = 1 \) and \( b_\pi(\theta, \theta_C) = -c \) if \( \pi(\theta, \theta_C) = 0 \). Since we consider uniform random matching, the expected period payoff of a \( \theta \)-user under a rating protocol \( \pi \) and a chosen rate \( \lambda \) before he is matched is given by

\[
v_{\pi, \lambda}(\theta) = \lambda \sum_{\theta' \in \Theta} \eta(\theta') b_\pi(\theta', \theta) - (1 - \lambda) \sum_{\theta' \in \Theta} \eta(\theta') c_\pi(\theta', \theta)
\]

To evaluate the long-term payoff of a compliant user, we use the discounted sum criterion in which the long-term payoff of a user is given by the expected value of the sum of discounted period payoffs starting from the current period. Let \( p_{\pi, \lambda}(\theta') | \theta \) be the transition probability that a \( \theta \)-user becomes a \( \theta' \)-user in the next period under a rating protocol \( \pi \) when he follows the recommended strategy and selects the chosen rate \( \lambda \), which can be expressed as

\[
p_{\pi, \lambda}(\theta' | \theta) = \begin{cases} 
\lambda \eta(\theta) + (1 - \lambda)(1 - \varepsilon)\alpha_\theta + 
\varepsilon(1 - \beta_\theta), & \text{if } \theta' = 1 \\
\lambda(1 - \eta(\theta)) + (1 - \lambda) \left[(1 - \varepsilon)(1 - \alpha_\theta) + \varepsilon\beta_\theta\right], & \text{if } \theta' = 0 
\end{cases}
\]

The expected long-term utility of a user in the repeated game starting from the current period is the infinite-horizon discounted sum of his expected one-period utility with a discount factor \( \omega \in (0, 1) \)

\[
v_{\pi, \lambda, \omega}(\theta) = v_{\pi, \lambda}(\theta) + \omega \sum_{\theta' \in \Theta} p_{\pi, \lambda}(\theta' | \theta) v_{\pi, \lambda, \omega}(\theta')
\]

Where \( \omega \) is the rate at which a user discounts his future payoff, and reflects his patience. It is obvious that larger discount factor reflects a more patience user. With a simple manipulation based on Eq. (6), Eq. (7) and Eq. (8), we have

\[
\begin{align*}
v_{\pi, \lambda}(0) &= \lambda \left[\eta(\theta)(1 - \varepsilon) + \eta(\theta)(1 - \epsilon)b - c \right] - (1 - \lambda)s \\
v_{\pi, \lambda}(1) &= \lambda [(1 - \epsilon)b - c] - (1 - \lambda) \eta(\theta) s \\
v_{\pi}^{\infty}(1) - v_{\pi}^{\infty}(0) &= \frac{\lambda \eta(\theta)(1 - 2\epsilon)b + (1 - \lambda)\eta(\theta)s}{1 + \omega (\phi_1 - \phi_0)}
\end{align*}
\]

Since users always aim to strategically maximize their own benefits, they will find in their own self-interest to comply with the social norm under a given two-sided rating protocol, if and only if they cannot benefit in terms of their long-term utilities upon deviations. We call such a protocol as a sustainable two-sided rating protocol, and give its formal definition as follows:

**Definition 2. (Sustainable Two-Sided Rating Protocols)** A two-sided rating protocol \( \pi \) is sustainable if and only if \( v_{\pi, \lambda, \omega}^{\infty}(\theta) \geq v_{\pi', \lambda', \omega}^{\infty}(\theta) \) for all \( \pi', \lambda' \) and \( \omega \).

In other words, a sustainable two-sided rating protocol \( \pi \) should maximize a user’s expected long-term utility at any period, such that no user can gain from a unilateral deviation regardless of the rating of his matched partner when every other user follows the recommended strategy \( \pi \) and selects \( \lambda = \frac{1}{2} \). It is obvious that the social welfare will be maximized when compliant users keep \( \rho \leq 1 \) (i.e., \( \lambda \leq \frac{1}{2} \)). Checking whether a rating protocol is sustainable in the second stage using the preceding definition requires computing deviation gains from all possible recommended strategies. By employing the criterion of unimprovability in Markov decision theory [30], a user’s strategic decision problem can be formulated as a Markov decision process under a two-sided rating protocol \( \pi \), where the state is the user’s rating \( \theta \), and the action is his chosen strategy \( \sigma(\theta) \). We thus establish the one-shot deviation principle for sustainable two-sided rating protocols, which provides simpler conditions.

**Lemma 1. (One-Shot Deviation Principles)** A two-sided rating protocol \( \pi \) satisfies the one-shot deviation principle if
and only if
\[ \frac{1}{2} [\eta_P(1)(1 - 2\varepsilon)b + \eta_P(0)s] \geq \max \left\{ \frac{\eta_P(1)s}{\omega(\phi_1 - \phi_0)}, \frac{s}{\omega(1 - 2\varepsilon)(\alpha_0 + \beta_0 - 1)}, \frac{\eta_P(1)s}{\omega(1 - 2\varepsilon)(\alpha_1 + \beta_1 - 1)} \right\} \]
(10)

Proof: For the “if” part: A user’s expected long-term utility when he adopts the recommend strategy \( \pi \) for all \( \theta \in \Theta \), can be expressed as \( v^\infty_{p,\pi}(\theta) \) in Eq. (8) (here, we fix \( \lambda = \frac{1}{2} \)). If the user unilaterally deviates from \( \pi \) to \( \pi' \) at rating \( \theta \), his expected long-term utility becomes
\[ v^\infty_{p,\pi',\theta}(\theta) = v_{p,\pi',\theta}(\theta) + \omega \sum_{\theta' \in \Theta} p_{p,\pi',\theta}(\theta', \alpha)v^\infty_{p,\pi'}(\theta') \] (11)
Where \( p_{p,\pi',\theta}(\theta') \) is the transition probability that a non-compliant \( \theta \)-server becomes a \( \theta' \)-server in the next period under \( \mathcal{P} \), which is expressed as
\[ p_{p,\pi',\theta}(\theta') = \begin{cases} \frac{1}{2} \gamma + \frac{1}{2} (1 - \gamma), & \text{if } \theta' = 1 \\ \frac{1}{2} (1 - \gamma) + \frac{1}{2} (1 - \gamma), & \text{if } \theta' = 0 \end{cases} \] (12)
By comparing these two payoffs \( v^\infty_{p,\pi}(\theta) \) and \( v^\infty_{p,\pi',\theta}(\theta) \), and solving the following inequality:
\[ v^\infty_{p,\pi}(\theta) - v^\infty_{p,\pi',\theta}(\theta) = \frac{1}{2} \sum_{\theta_1 \in \Theta} \eta_P(\theta_1)(c_{\pi}(\theta, \theta_1) - c_{\pi}(\theta, \theta_1)) + \omega \sum_{\theta' \in \Theta} [p_{p,\pi}(\theta') - p_{p,\pi',\theta}(\theta')]v^\infty_{p,\pi'}(\theta') \geq 0. \] (13)
If \( \theta = 0 \), then for each \( \tilde{\theta} \in \Theta \), \( \pi(\theta, \tilde{\theta}) = 1 \), \( c_{\pi}(\theta, \tilde{\theta}) = s \) and \( c_{\pi}(\theta, \tilde{\theta}) = 0 \), we have
\[ v^\infty_{p,\pi}(1) - v^\infty_{p,\pi}(0) \geq \frac{s}{\omega(1 - 2\varepsilon)(\alpha_0 + \beta_0 - 1)} \] (14)
While if \( \theta = 1 \) and \( \tilde{\theta} = 1 \), then \( \pi(\theta, \tilde{\theta}) = 1 \), \( c_{\pi}(\theta, \tilde{\theta}) = s \) and \( c_{\pi}(\theta, \tilde{\theta}) = 0 \). Else if \( \theta = 0 \), then \( \pi(\theta, \tilde{\theta}) = 0 \) for each \( \tilde{\theta} \in \Theta \), self-interested users have no incentive to deviate from \( \pi(\theta, \tilde{\theta}) = 0 \). Hence, we have
\[ v^\infty_{p,\pi}(1) - v^\infty_{p,\pi}(0) \geq \frac{\eta_P(1)s}{\omega(1 - 2\varepsilon)(\alpha_1 + \beta_1 - 1)} \] (15)
We have inequality in Eq. (10) by substituting \( v^\infty_{p}(1) - v^\infty_{p}(0) = \frac{1}{2} [\eta_P(1)(1 - 2\varepsilon)b + \eta_P(0)s] \) into the LHS of Eq. (14) and Eq. (15). Hence, the two-sided rating protocol \( \mathcal{P} \) is satisfied with the one-shot deviation principle if Eq. (10) holds.

For the “only if” part: Suppose the rating protocol \( \mathcal{P} \) is satisfied with the one-shot deviation principle, then clearly there are no profitable one-shot deviations. We can prove the converse by showing that if \( \mathcal{P} \) is not satisfied with the one-shot deviation principle, there is at least one profitable one-shot deviation. Since \( c_{\pi}(\theta, \tilde{\theta}) \) and \( c_{\pi}(\theta, \tilde{\theta}) \) are bounded, this is true by the unimprovability property in Markov decision theory.

Lemma 1 shows that if a user cannot gain by unilaterally deviating from \( \pi \) only in the current period and following \( \pi \) afterwards, he can neither gain by switching to any other recommended strategy \( \pi' \), and vice versa. \( \frac{1}{2} \sum_{\theta \in \Theta} \eta_P(\theta)(c_{\pi}(\theta, \tilde{\theta}) - c_{\pi}(\theta, \tilde{\theta})) \) of Eq. (13) can be interpreted as the current gain from choosing \( \pi' \) in the second stage, while \( \omega \sum_{\theta \in \Theta} [p_{p,\pi}(\theta') - p_{p,\pi}(\theta')]v^\infty_{p,\pi'}(\theta') \) of Eq. (13) represents the discounted expected future loss due to the different transition probabilities incurred by choosing \( \pi' \).

After analyzing sustainable conditions in the second-stage, we then step back to analyze sustainable conditions in the first stage when both users choose their strategies in the first-stage before knowing their productivities. In the first stage, users decide the optimal chosen rate \( \lambda \), and follow the recommended strategy \( \pi \) in their self-interest. Under the service exchange dilemma game, a \( \theta \)-user will find it optimal to choose to be a client in the first stage, as his revenue is maximized when his matched \( \theta \)-server chooses to follow the recommended strategy \( \pi \) in the second stage, which yields payoff \( b_{\pi}(\theta, \tilde{\theta}) \) for him. On the contrary, choosing to be a server will suffer a cost \( c_{\pi}(\theta, \tilde{\theta}) \). However, social welfare is maximized if and only if every user chooses to be a server or a client with the same probability \( \lambda = \frac{1}{2} \), which we name it as the principle of fairness inspired by [31], and derive incentive constraints that characterize sustainable conditions in the first stage as shown in Lemma 2.

**Lemma 2. (The Principle of Fairness)** A two-sided rating protocol \( \mathcal{P} \) satisfies the principle of fairness if and only if
\[ \frac{1}{2} [\eta_P(1)(1 - 2\varepsilon)b + \eta_P(0)s] \geq \max \left\{ \frac{(1 - \varepsilon)b - c + \eta_P(1)s}{\omega(1 - 2\varepsilon)(\alpha_1 + \beta_1 + \gamma_1 - 2(1 - \delta_1))}, \frac{\eta_P(0)(1 - \varepsilon)b - c + \eta_P(1)s}{\omega(1 - 2\varepsilon)(\alpha_0 + \beta_0)} \right\} \]
(16)

Proof: For the “if” part: Assume that each user selects \( \lambda = \frac{1}{2} \) in the first stage, and adopts the recommend strategy \( \pi \) in the second stage, then his expected long-term utility can be expressed as
\[ v^\infty_{p,\pi,\lambda}(\theta) = v_{p,\pi,\lambda}(\theta) + \omega \sum_{\theta' \in \Theta} p_{p,\pi,\lambda}(\theta', \alpha)v^\infty_{p,\pi,\lambda}(\theta') \] (17)
Where \( p_{p,\pi,\lambda}(\theta)(\theta') \) is the transition probability that a complaint \( \theta \)-user becomes a \( \theta' \)-user in the next period when he selects \( \lambda = \frac{1}{2} \) in the first stage under the rating protocol \( \mathcal{P} \), which can be found in Eq. (7).

As a \( \theta \)-user can receive the benefit \( b_{\pi}(\theta, \tilde{\theta}) \) if and only if he chooses to be a client in the current period under the recommended strategy \( \pi \), otherwise, he will suffer a cost \( c_{\pi}(\theta, \tilde{\theta}) \). Without loss of generality, we now suppose that a user derives from \( \lambda = \frac{1}{2} \) to \( \lambda = 1 \) in the current period, and following \( \lambda = \frac{1}{2} \) afterwards, then his expected long-term utilities can be expressed as
\[ v^\infty_{p,\pi,\lambda}(\theta) = v_{p,\pi,\lambda}(\theta) + \omega \sum_{\theta' \in \Theta} p_{p,\pi,\lambda}(\theta', \alpha)v^\infty_{p,\pi,\lambda}(\theta') \] (18)
Where \( p_{P,\lambda=1}(\theta'|\theta) \) can be computed based on Eq. (2).
\[
p_{P,\lambda=1}(\theta'|\theta) = \left\{ \begin{array}{ll}
1 - \delta_{\theta}, & \text{if } \theta' = 1 \\
\delta_{\theta}, & \text{if } \theta' = 0
\end{array} \right.
\]
(19)

By comparing Eq. (17) with Eq. (18), we can check whether a \( \theta \)-user has an incentive to deviate from \( \lambda = \frac{1}{2} \) as follows
\[
v_{P,\lambda=1}(1) - v_{P,\lambda=1}(1) = v_{P,\lambda=1}(\theta) - v_{P,\lambda=1}(\theta) + \omega \left[(1 - \varepsilon)\alpha_0 + \varepsilon(1 - \beta_0) + \gamma_0 - 2(1 - \delta_0)\right] + (v_{P,\lambda=1}(1) - v_{P,\lambda=1}(1)) \geq 0
\]
(20)

With simple manipulation based on Eq. (4), we obtain that \( v_{P,\lambda=1}(1) = (1 - \varepsilon)b - c + \eta_P(1)s \) and \( v_{P,\lambda=1}(0) = [\eta_P(0)(1 - \varepsilon) + \eta_P(1)\varepsilon]b - c \), together with Eq. (9), we have
\[
v_{P,\lambda=1}(1) - v_{P,\lambda=1}(1) = \frac{1}{2}[(1 - \varepsilon)b - c + \eta_P(1)s]
\]
(21)
\[
v_{P,\lambda=1}(0) - v_{P,\lambda=1}(0) = \frac{1}{2}\left\{[\eta_P(0)(1 - \varepsilon) + \eta_P(1)\varepsilon]b - c + s\right\}
\]
(22)

If \( \theta = 1 \), Eq. (20) can be rewritten as
\[
v_{P,\lambda=1}(1) - v_{P,\lambda=1}(1) \geq \frac{(1 - \varepsilon)b - c + \eta_P(1)s}{\omega[(1 - \varepsilon)\alpha_1 + \varepsilon(1 - \beta_1) + \gamma_1 - 2(1 - \delta_1)]}
\]
(23)

While if \( \theta = 0 \), Eq. (20) can be rewritten as
\[
v_{P,\lambda=1}(1) - v_{P,\lambda=1}(1) \geq \frac{[\eta_P(0)(1 - \varepsilon) + \eta_P(1)\varepsilon]b - c + s}{\omega[(1 - \varepsilon)\alpha_0 + \varepsilon(1 - \beta_0) + \gamma_0 - 2(1 - \delta_0)]}
\]
(24)

Combining Eq. (23) and Eq. (24), sufficient conditions that a two-sided rating protocol \( P \) is satisfied with the principle of fairness can be obtained, as shown in inequality (16).

For the “only if” part: Suppose \( P \) is satisfied with the principle of fairness, then clearly there are no profitable deviations (i.e., \( \rho > 1 \) or \( \lambda > \frac{1}{2} \)) in the first stage. We can prove the converse by showing that if \( P \) is not satisfied with the principle of fairness, there is at least one profitable deviation. Since the RHS of Eq. (16) is bounded, this is true by the unimprovability property in Markov decision theory.

Using one-shot deviation principle and the principle of fairness, we can derive incentive constraints that characterize necessary and sufficient conditions for a two-sided rating protocol to be sustainable, which is formalized in the next theorem.

**Theorem 1.** A two-sided rating protocol \( P \) is sustainable if and only if it is satisfied with both of the one-shot deviation principle and the principle of fairness.

**Proof:** This proof can be directly obtained from Lemma 1 and 2, and is omitted here.

---

**IV. OPTIMAL DESIGN OF TWO-SIDED RATING PROTOCOLS**

In this section we investigate the design of an optimal two-sided rating protocol that solves the two-sided rating protocol design problem under a given recommended strategy \( \pi \), i.e., selecting the optimal rating update rule \( \tau \), which are determined by design parameters \( (\alpha_\theta, \beta_\theta, \gamma_\theta, \delta_\theta) \). In order to characterize an optimal design which is denoted as \( (\alpha_\theta, \beta_\theta, \gamma_\theta, \delta_\theta) \), \( \forall \theta \in \Theta \), we investigate the impacts of design parameters on the social welfare \( U_P \triangleq \sum_{\theta \in \Theta} \eta_P(\theta)v_P(\theta) \), and the incentive for satisfying constraints in Eq. (25).

**A. Existence of a Sustainable Two-sided Rating Protocol**

We first investigate whether there exists a sustainable two-sided rating protocol under \( \pi \), i.e., determining whether there exists a feasible solution for the design problem of Eq. (25).

**Theorem 2.** A sustainable two-sided rating protocol \( P \) under the recommended strategy \( \pi \) exists if and only if
\[
\omega \geq \max \left\{ \frac{2s}{(1 - 2\varepsilon)[(1 - \frac{5}{2}\varepsilon + \varepsilon^2)b + \frac{5}{2}s]}}, \frac{(2 - 2\varepsilon)b - 2c + (2 - \varepsilon)s}{[(1 - \frac{5}{2}\varepsilon + \varepsilon^2)b + \frac{5}{2}s](2 - \varepsilon)}, \frac{(3\varepsilon - 2\varepsilon^2)b - 2c + 2s}{[(1 - \frac{5}{2}\varepsilon + \varepsilon^2)b + \frac{5}{2}s](2 - \varepsilon)} \right\}
\]
(26)
Proof: For the “if” part: Among the eight design parameters, \(\alpha_1, \alpha_0, \gamma_1\) and \(\gamma_0\) can be referred to as reward factors imposed on compliant users, while \(\beta_1, \beta_0, \delta_1\) and \(\delta_0\) can be referred to as punishment factors imposed on non-compliant users. The incentive for self-interested users to be a compliant user is maximized when we maximize all of reward factors and punishment factors, i.e., \(\alpha_0 = \beta_0 = \gamma_0 = \delta_0 = 1, \forall \theta \in \Theta\). Then, Eq. (25) can be transformed into

\[
\frac{1}{2}(1 - \frac{\varepsilon}{2})(1 - 2\varepsilon) b + \frac{\varepsilon}{2}s \geq \max \left\{ \frac{1}{2}(1 - \frac{\varepsilon}{2})c + (1 - \frac{\varepsilon}{2})s, \frac{1}{2}\frac{(\varepsilon - \varepsilon^2) b - c + s}{\omega(2 - \varepsilon)} \right\}
\]

(27)

It is obvious that \(\frac{1}{2}(1 - \frac{\varepsilon}{2})b + \frac{\varepsilon}{2}s\) can be revised as follows

\[
\frac{1}{2}(1 - \frac{\varepsilon^2}{2})b + \frac{\varepsilon}{2}s \geq \max \left\{ \frac{1}{2}\frac{\varepsilon}{\omega(2 - \varepsilon)}, \frac{1}{2}\frac{(\varepsilon - \varepsilon^2)b - c + s}{\omega(2 - \varepsilon)} \right\}
\]

(28)

By solving Eq. (28), we can obtain Eq. (26), that is, there exists a feasible solution for the design problem. It shows that Eq. (25) always has a feasible solution if users have sufficient patience (i.e., when the discount factor \(\omega\) is large). We assume that \(\omega < 1\) as no one can be 100% patient. Therefore, there exists a sustainable two-sided rating protocol if Eq. (26) holds.

For the “only if” part: Suppose Eq. (26) hold, it is easy to determine whether constraints in the design problem of Eq. (25) are satisfied, similar as the above, the “only if” part can be proved.

B. Optimal Values of the Rating Update Rule

In this section, we assume that Eq. (26) holds, that is, there exists a feasible solution for the two-sided rating protocol design problem of Eq. (25). Our goal is to select \((\alpha^*_0, \beta^*_0, \gamma^*_0, \delta^*_0), \forall \theta \in \Theta\) to maximize the social welfare \(U_\theta\), that is, maximizing reward factors \(\alpha_0, \gamma_0, \forall \theta \in \Theta\), and minimizing punishment factors \(\beta_0, \delta_0, \forall \theta \in \Theta\). With this idea, Theorem 3 gives the optimal value of reward/punishment factors except \(\beta_0, \forall \theta \in \Theta\).

Theorem 3. Given a sustainable two-sided rating protocol \(\mathcal{P}\), \(\alpha^*_0 = \gamma^*_0 = \delta^*_0 = 1, \forall \theta \in \Theta\) is always the optimal solution to Eq. (25).

Proof: Social welfare \(U_\theta\) monotonically increases with reward factors \(\alpha_0, \gamma_0, \forall \theta \in \Theta\), and the upper bound of them is 1, with which the incentive constraints in Eq. (25) are satisfied. As \(U_\theta\) monotonically decreases with punishment factors \(\beta_0, \forall \theta \in \Theta\), and given \(\alpha^*_0 = \gamma^*_0 = 1, \forall \theta \in \Theta\), the design problem of Eq. (25) is transformed into the selection of the smallest \(\beta_0, \forall \theta \in \Theta\), with which the incentive constraints Eq. (16) are satisfied. It is obvious that the smallest \(\beta_0, \forall \theta \in \Theta\) can be obtained when we select the largest \(\delta_0, \forall \theta \in \Theta\), and the upper bound of \(\delta_0, \forall \theta \in \Theta\) is 1. Since \(U_\theta\) is only determined by \(\beta_0, \forall \theta \in \Theta\), rather than \(\delta_0, \forall \theta \in \Theta\), in order to provide sufficient incentive and get as less \(\beta_0, \forall \theta \in \Theta\) as possible, we have \(\delta_0 = 1, \forall \theta \in \Theta\). Hence, this statement follows.

By substituting \(\alpha_0 = \gamma_0 = \delta_0 = 1, \forall \theta \in \Theta\) into Eq. (5), we have

\[
\begin{align*}
\eta_0(0) &= \frac{\varepsilon \beta_1}{2 + \varepsilon \beta_1 - \varepsilon \beta_0} \\
\eta_0(1) &= \frac{2 - \varepsilon \beta_1 - \varepsilon \beta_0}{2 + \varepsilon \beta_1 - \varepsilon \beta_0}
\end{align*}
\]

(29)

It is obvious that \(0 < \eta_0(0) < \frac{1}{2} < \eta_0(1)\) as \(\beta_1, \beta_0 \in [0, 1]\). Hence, the social welfare \(U_\theta\) monotonically decreases with \(\eta_0(0)\), while \(\eta_0(0)\) monotonically increases with both \(\beta_1\) and \(\beta_0\), and should be sufficiently small in order to increase \(U_\theta\). However, \(\beta_1\) and \(\beta_0\) also cannot be too small since it cannot provide sufficient punishment for self-interested users to comply with the social norm.

Given \(\alpha_0 = \gamma_0 = \delta_0 = 1, \forall \theta \in \Theta\), the design problem in Eq. (25) w.r.t \(\beta_0\) and \(\beta_1\) can be rewritten as

\[
\begin{align*}
\min_{(\beta_0, \beta_1)}& \quad \frac{1}{1 + \frac{\varepsilon \beta_0}{2 - \varepsilon \beta_1}} \\
\text{s.t.}& \quad (2 - \varepsilon \beta_1)(1 - 2\varepsilon) b + s \varepsilon \beta_1 \geq \max \left\{ \frac{(1 - \varepsilon)(2 + \varepsilon \beta_1 - \varepsilon \beta_0)s}{\omega(2 - \varepsilon)}, \frac{(1 - \varepsilon)(2 + \varepsilon \beta_1 - \varepsilon \beta_0)s}{\omega(2 - \varepsilon)} \right\} \\
& \quad \max \left\{ \frac{1}{2}\frac{\varepsilon}{\omega(2 - \varepsilon)}, \frac{1}{2}\frac{(\varepsilon - \varepsilon^2)b - c + s}{\omega(2 - \varepsilon)} \right\}
\end{align*}
\]

(30)

We now design an algorithm to solve the non-convex optimization problem in Eq. (30) inspired by [32, 33], it achieves low-complexity computation via two-stage two-step in an alternate manner. The proposed algorithm is outlined in Algorithm 1, where \(obj^t\) denotes value of \(\frac{1}{1 + \frac{\varepsilon \beta_0}{2 - \varepsilon \beta_1}}\) at the \(t\)-th iteration. It should be noted that the value of \(obj^t\) of Algorithm 1 is guaranteed to be monotonically decreasing when optimizing one variable with another fixed in each iteration [34]. Meanwhile, \(\frac{1}{1 + \frac{\varepsilon \beta_0}{2 - \varepsilon \beta_1}}\) is lower-bounded by \(\frac{\varepsilon}{2}\) with the presence of imperfect monitoring (i.e., \(\varepsilon > 0\)). Therefore, Algorithm 1 is guaranteed to converge. The detailed explanation of Algorithm 1 can be found in the proof of Theorem 4.

Theorem 4. Given \(\alpha^*_0 = \gamma^*_0 = \delta^*_0 = 1, \forall \theta \in \Theta\) and a residual \(\varepsilon\), the output of \(\beta^*_0, \forall \theta \in \Theta\) by Algorithm 1 is the optimal solution to Eq. (30).

Proof: Algorithm 1 consists of two stages, and each stage consists of two steps. In stage (i), we first fix \(\beta_0 = 1\), and then update both \(\beta_1\) and \(\beta_0\). Where stage (ii) is symmetric with stage (i), the only difference is that we first fix \(\beta_1 = 1\) and then update both \(\beta_0\) and \(\beta_1\). Step (i): Optimizing with fixed \(\beta_0\). Given \(\beta_0\), the optimization problem in Eq. (30) w.r.t \(\beta_1\) can be rewritten as

\[
\begin{align*}
\min_{\beta_1}& \quad \frac{\varepsilon \beta_1^2 + y_1\beta_1 + z_1}{2} \leq 0, i \in \{1, 3, 4\} \\
\text{s.t.}& \quad x_j \beta_1^2 + y_j \beta_1 + z_j \geq 0, j \in \{2\}
\end{align*}
\]

(31)
**Algorithm 1** Alternate Optimal Design of Punishment factors $\beta_0$ and $\beta_1$

**Input:** $b, c, s, \varepsilon, \omega$ and $\epsilon$.

**Output:** $\beta_0^*$ and $\beta_1^*$.

1. Initialize $\beta_0^* = 1$ and $t = 1$.
2. **repeat**
   3. Update $\beta_1^t$ by solving Eq. (34) with given $\beta_0^{t-1}$.
   4. Update $\beta_0^t$ by solving Eq. (37) with given $\beta_1^t$.
   5. $t = t + 1$.
6. **until** $(obj_{t-1} - obj_t)/obj_t \leq \epsilon$
7. $obj^* = obj_t$
8. Set $\beta_0^t = 1$ and $t = 1$.
9. **repeat**
10. Update $\beta_0^t$ by solving Eq. (34) with given $\beta_1^{t-1}$.
11. Update $\beta_1^t$ by solving Eq. (37) with given $\beta_0^t$.
12. $t = t + 1$.
13. **until** $(obj_{t-1} - obj_t)/obj_t \leq \epsilon$
14. $(\beta_0^*, \beta_1^*) = \arg \min \{obj^*, obj_t\}$.

**Where**

$x_1 = s \omega^2 \omega, y_1 = s \omega(2 + 2 \omega - \beta_0), z_1 = (2 - \varepsilon_0)[s(2 + 2 \varepsilon_0) - \omega b(1 - 2 \varepsilon_0)^2 \beta_0], x_2 = s \omega \omega(1 - 2 \varepsilon_0), y_2 = \omega(2 + \omega - \beta_0)(b(1 - 2 \varepsilon_0)^2 - \varepsilon_0), z_2 = s(2 + 2 \omega - \beta_0)(2 - \varepsilon_0), x_3 = \omega^2 \omega(b - b - c + s), y_3 = \omega^2 \omega(2 - \varepsilon_0)(s + c + 2b - 3c) + \varepsilon(b - b - c)(2 - \omega \varepsilon_0) - 2 \omega \varepsilon_0, z_3 = (2 - \varepsilon_0)[(2 - \varepsilon_0)(b - b - c - 2b \omega(1 \varepsilon_0 - 2 \varepsilon_0))] + \varepsilon^2 \omega^2(b - b - c) + \varepsilon(b - b - c)(2 - \omega \varepsilon_0)(s + c + 2b - 3c), z_4 = (2 - \varepsilon_0)[(2 - \omega \varepsilon_0)(s + c + 2b - 3c) - \omega b(1 \varepsilon_0 - 2 \varepsilon_0)].$

By solving inequalities in Eq. (31), we have

$\beta_1 \in \left\{ - \frac{y_1}{2x_1} - \sqrt{\left(\frac{y_1}{2x_1}\right)^2 - z_1}, - \frac{y_1}{2x_1} + \sqrt{\left(\frac{y_1}{2x_1}\right)^2 - z_1} \right\}, \forall i \in \{1, 3, 4\}$

and

$\beta_1 \in \left\{ 0, \frac{y_2}{2x_2} - \sqrt{\left(\frac{y_2}{2x_2}\right)^2 - z_2}, \frac{y_2}{2x_2} + \sqrt{\left(\frac{y_2}{2x_2}\right)^2 - z_2} \right\}$

Let $\psi_y = \frac{y_2}{2x_2} - \sqrt{\left(\frac{y_2}{2x_2}\right)^2 - z_2}, \psi_1^\varepsilon = \min_{i \in \{1, 3, 4\}} \left\{ - \frac{y_i}{2x_i} + \sqrt{\left(\frac{y_i}{2x_i}\right)^2 - z_i}, \frac{y_i}{2x_i} - \sqrt{\left(\frac{y_i}{2x_i}\right)^2 - z_i} \right\}, \psi_2^\varepsilon = \frac{y_2}{2x_2} + \sqrt{\left(\frac{y_2}{2x_2}\right)^2 - z_2}, \Psi_1 = [\psi_1^\varepsilon, \psi_1^\varepsilon], \text{and} \Psi_2 = [\psi_2^\varepsilon, \psi_2^\varepsilon].$ Eq. (32) can be rewritten as

$\beta_1 \in \Psi_1 \cup \Psi_2$ (33)

The optimal value of $\beta_1$ for Eq. (31) can be conducted as follows

$\min_{s.t. \beta_1 \in [0, 1]} \bigcup_{i \in \{1, 2\}} \Psi_i$ (34)

**Step (ii): Optimizing with fixed $\beta_1$.** Similar to step (1), given $\beta_1$, the optimization problem in Eq. (30) w.r.t $\beta_0$ can be rewritten as

$\min_{s.t. x_j^\varepsilon \beta_1^2 + y_j^\varepsilon \beta_1 + z_j \leq 0, j \in \{1, 2, 3, 4\}} \beta_1$ (35)

Where $x_1' = \omega(2 + \omega - \beta_0), y_1' = -2 \omega(2 + \omega - \beta_0) - \omega(2 + 2 \omega - \beta_0) - \omega(2 + b \varepsilon_1)(2 + 2 \varepsilon_0), x_2' = s(2 + \varepsilon_0) - \omega b(1 - 2 \varepsilon_0)^2 \beta_1, y_2' = \omega b(1 - 2 \varepsilon_0)^2 \beta_1 - 2 \omega b(1 - 2 \varepsilon_0)^2 \beta_1, z_2' = 2s(2 + \omega - \beta_0) - \omega b(1 - 2 \varepsilon_0)^2 \beta_1, x_3' = \varepsilon^2 \omega(b - b - c + s), y_3' = \omega(2 - \varepsilon_0)(s - c + 2b - 3c) + \varepsilon(b - b - c)(2 - \omega - \beta_0)(s - c + 2b - 3c), z_4' = (2 - \varepsilon_0)[(2 - \omega - \beta_0)(s - c + 2b - 3c) - \omega b(1 - 2 \varepsilon_0)].$

Eq. (35) can be solved as follows

$\beta_0 \in \left\{ - \frac{y_j'}{2x_j'} - \sqrt{\left(\frac{y_j'}{2x_j'}\right)^2 - z_j'}, - \frac{y_j'}{2x_j'} + \sqrt{\left(\frac{y_j'}{2x_j'}\right)^2 - z_j'} \right\}, \forall j \in \{1, 2, 3, 4\}$ (36)

Let $\varphi_u = \min_{j \in \{1, 2, 3, 4\}} \left\{ - \frac{y_j'}{2x_j'} - \sqrt{\left(\frac{y_j'}{2x_j'}\right)^2 - z_j'}, - \frac{y_j'}{2x_j'} + \sqrt{\left(\frac{y_j'}{2x_j'}\right)^2 - z_j'} \right\},$ Eq. (36) can be rewritten as

$\min \beta_0 \left\{ s.t. \beta_0 \in [0, 1] \cap \varphi_1, \varphi_u \right\}$ (37)

Given $\beta_1 = 1$, we can obtain the smallest $\beta_1$, with which the incentive constraints Eq. (31) are satisfied, where $\beta_1$ takes its value from the domain $\Phi_1$. Where if we fix $\beta_1 = 1$, the output of Eq. (35) is the smallest value of $\beta_0$, and next we update $\beta_1$ by solving Eq. (37) with given $\beta_0$, where $\beta_1$ takes its value from the domain $\Phi_2$. It is obvious that the global optimal solution will be derived from these two domains as there does not exist another domain for $\beta_1^t$ with given $\beta_0^t$. And hence, this statement follows.

**C. Optimal Values of the Rating Update Rules with a Stricter Recommended Strategy**

Eq. (25) shows that there exists a feasible solution for the design problem of Eq. (25) under the condition when the user is sufficiently patient with his discount factor $\omega$. However, such a condition may not hold with a small $\omega$ or max $\left\{ - \frac{y_j}{2x_j} - \sqrt{\left(\frac{y_j}{2x_j}\right)^2 - z_j}, - \frac{y_j}{2x_j} + \sqrt{\left(\frac{y_j}{2x_j}\right)^2 - z_j} \right\}, \psi_1^\varepsilon \in \{1, 3, 4\}$ (32) + \varepsilon(2 - \omega - \beta_0) - \omega b(1 - 2 \varepsilon_0)].$ Eq. (32) can be rewritten as

The optimal value of $\beta_1$ for Eq. (31) can be conducted as follows

$\min_{s.t. \beta_1 \in [0, 1]} \bigcup_{i \in \{1, 2\}} \Psi_i$ (34)

$\text{Step (ii): Optimizing with fixed } \beta_1.$ Similar to step (1), given $\beta_1$, the optimization problem in Eq. (30) w.r.t $\beta_0$ can be rewritten as

$\min \beta_1 \left\{ s.t. x_j^\varepsilon \beta_1^2 + y_j^\varepsilon \beta_1 + z_j \leq 0, j \in \{1, 2, 3, 4\} \right\}$ (35)

$\varepsilon b - c - (1 - \lambda) \eta \varepsilon(1)s$, $v_{\varepsilon_1, \lambda}(e) = \lambda[(1 - \varepsilon)b - c - (1 - \lambda) \eta \varepsilon(1)s]$, $v_{\varepsilon_1, \lambda}(e) = \lambda(1 - \varepsilon)b - \frac{\lambda(1 - \varepsilon)b}{1 + \omega(\phi_1 - \phi_0)}$. (39)
The two-sided rating protocol design problem with a strict recommended strategy \( \pi_s \), that is, Eq.\((25)\) can be rewritten as follows:

\[
\max_{(\tau, \pi)} U_P \triangleq \sum_{\theta \in \Theta} \eta_P(\theta) v_P(\theta) = \frac{1}{2} (1-\varepsilon)b - c - s - \eta_P(0)(1-2\varepsilon)b + s)
\]

\[
\text{s.t. } \frac{(1-2\varepsilon)b}{2 + 2\omega(\phi_1 - \phi_0)} \geq \frac{\eta_P(1)s}{\omega(1-2\varepsilon)(\alpha_0 + \beta_0 - 1)} \cdot \frac{(1-\varepsilon)c + \varepsilon(1-\beta_0)}{\omega[1-\varepsilon]a_1 + 2(1-\beta_0) - 2(1-\delta_0)}
\]

Eq.\((40)\) w.r.t \( \beta_0 \) and \( \beta_1 \) can be rewritten as

\[
\min_{(\beta_0, \beta_1)} \frac{1}{1 + \frac{2 - \varepsilon \beta_0}{\varepsilon \beta_1}}
\]

\[
\text{s.t. } (1 - 2\varepsilon)b \geq \omega[1-\varepsilon]a_1 + \varepsilon(1-\beta_0) - 2(1-\delta_0)
\]

An important observation is that Algorithm 1 can be efficiently used to handle the design problem of Eq.\((42)\) in a similar manner, which is omitted here.

V. ILLUSTRATIVE RESULTS

In this section, we provide numerical results to illustrate the key features of our proposed two-sided rating protocol designed for service exchange dilemma in crowdsourcing. First of all, we show how to determine the optimal recommended strategy as intrinsic parameters vary. Secondly, we investigate the impact of intrinsic parameters on design parameters. Finally, we examine the performance of the optimal design of two-sided rating protocols. Throughout our experiments, the benefit \( b \) for unit service exchange is normalized to be 1, cost \( c \) and \( s \) are restricted to be smaller than \( b \).

A. Optimal Recommended Strategy Against Intrinsic Parameters

Figure 2 shows that the optimal recommended strategy is determined by intrinsic parameters \( c, s, \varepsilon \), and \( \omega \). When both of \( c \) and \( \omega \) are sufficiently large (given \( b=1, s = 0.2 \) and \( \varepsilon = 0.05 \)) as shown in Figure 2(a), the recommended strategy \( \pi \) can be sustained, and thus be selected to be the optimal choice among the other two candidate recommended strategies \( \pi_s \) and \( \pi_0 \) (The server will be recommended to provide
low-quality service regardless of his own and his client’s ratings under $\pi_0$.) As $c$ or $\omega$ decreases, $\pi$ cannot be sustained any more, hence, the optimal recommended strategy changes from $\pi$ to $\pi_s$. The main reason behind this phenomenon is that, smaller $c$ and $\omega$ introduce a higher probability for self-interested users to deviate from the principle of fairness, which needs a stricter recommended strategy $\pi_s$ to increase users’ incentive to comply with the social norm. When the region of $c$ and $\omega$ in which $\pi_s$ is not sustained, $\pi_0$ will be the unique sustainable recommended strategy, which yields zero social welfare for each user, since full cooperation cannot be achieved in such a scenario.

Figure 2(b) shows the optimal recommended strategy with $b = 1$, $c = 0.4$ and $\varepsilon = 0.05$ as $s$ and $\omega$ vary. As $s$ increases, a user with a higher $\omega$ has higher probability to be a compliant user. Otherwise, $\pi$ will not be sustained, and thus the stricter recommended strategy $\pi_s$ will be selected to be optimal. When there does not exist any sustainable rating protocol in the region of $s$ and $\omega$, in order to obtain maximal social welfare, the optimal recommended strategy changes from $\pi$ to $\pi_s$ and eventually to $\pi_0$. A similar phenomenon can be found in Figure 2(c), which plots the change of optimal recommended strategy versus $\varepsilon$ and $\omega$ with given $b = 1$, $c = 0.4$ and $s = 0.2$.

**B. The Impact of Intrinsic Parameters on Design Parameters**

Figure 3 plots how the optima design ($\beta_0$ and $\beta_1$) is influenced by intrinsic parameters $c$, $s$, $\varepsilon$, and $\omega$. There does not exist a sustainable two-sided rating protocol with the recommended strategy $\pi$ when $c$ is sufficiently small, as shown in Figure 3(a). This is because users will choose to be a client with a higher probability $\lambda > \frac{1}{2}$, that is they will deviate from the principle of fairness when $c$ is small (e.g., $c < 0.305000$). As $c$ increases, we can obtain a higher $\beta_0$ and a lower $\beta_1$, this is due to the fact that a higher $c$ introduces a higher $\eta_P(0)$, and hence it needs a higher $\beta_0$ to punish non-compliant users. When $c$ is sufficiently large, $\beta_0$ and $\beta_1$ will not change, since a large $c$ do not affect on the sustainable of rating protocols, and thus it is no necessary to enhance the punishment factors.

Figure 3(b) is very similar with Figure 3(c), these two figures plot the impact of $s$ and $\varepsilon$ on the design parameters $\beta_0$ and $\beta_1$, respectively. As $s$ and $\varepsilon$ increase, it becomes more difficult to incentivize users to comply with the social norm, and punishment factors $\beta_0$ and $\beta_1$ should be increased to sustain a rating protocol. Different from Figure 3(b) and 3(c), punishment factors $\beta_0$ and $\beta_1$ decrease as $\omega$ increases, this is because it is easier to give incentives to comply with the social norm and punishments through the designed two-sided rating protocol. It can be observed that $\beta_0$ is always larger than $\beta_1$ in all of the four subfigures in Figure 3, this is due to

---

**Figure 3.** The impact of design parameters $\beta_0$ and $\beta_1$ against intrinsic parameters: (a) $c$; (b) $s$; (c) $\varepsilon$; (d) $\omega$. 
the fact that \( \eta_P(0) < \eta_P(1) \), in order to sustain the designed two-sided rating protocol with sufficient punishment, a larger \( \beta_0 \) and a lower \( \beta_1 \) will receive a better performance.

C. Performance Efficiency

Figure 4 examines the performance of the optimal design two-sided rating protocol against intrinsic parameters \( c, s, \varepsilon \) and \( \omega \) (denoted as \( U_P \)). For the comparison, the social optimum \((1 - \varepsilon)b - c - s\) (denoted as \( U_C \)) is considered, which can be exactly achieved only by users with \( \rho = 1 \) (or \( \lambda = \frac{1}{2} \)), who provide high-quality service all the time when they are matched as servers, however, it is not an equilibrium. In other words, such a social optimum is impossible to be achieved. Our goal is to be as close as possible to this social optimum.

The social welfare \( U_P \) monotonically decreases with \( c, s \) and \( \varepsilon \), but increases with \( \omega \), where \( U_C \) is only determined by \( c, s \) and \( \varepsilon \), and is independent of \( \omega \) with given \( b = 1 \). In Figure 4(a), the performance gap between \( U_P \) and \( U_C \) is almost unchanged. The major reason is that the impact of \( c \) on punishment factors can almost be ignored, thus both \( U_P \) and \( U_C \) change only against the value of \( c \) with given \( b, s, \varepsilon \) and \( \omega \). While as shown in Figure 4(b) and 4(c), the performance gap between \( U_P \) and \( U_C \) becomes more significant as \( s \) and \( \varepsilon \) increase, respectively. This is because the incentive to comply with the social norm decreases and so is his one-period utility, and hence punishment factors \( \beta_0 \) and \( \beta_1 \) will be increased to provide sufficient incentive, thus reduces the social welfare \( U_P \). In particular, the gap between \( U_P \) and \( U_C \) is gradually narrowed as \( \omega \) increases, as shown in Figure 4(d). Since users have more patience when the discount factor \( \omega \) of users increases from 0 to 1, it becomes easier to sustain a two-sided rating protocol, thereby leading to a decrease in punishment factors and an increase in social welfare.

VI. CONCLUSION AND FUTURE WORK

In this paper, we proposed a service exchange dilemma in a two stage game, and developed a game-theoretic design of two-sided rating protocol to stimulate cooperation among self-interested users, and thus overcome the inefficiency of the socially undesirable equilibrium. By rigorously analyzing how users’ behaviors are influenced by intrinsic parameters, design parameters, as well as users’ valuation of their individual long-term utilities, we characterize the optimal design by selecting eight optimal design parameters \( (\alpha_\theta, \beta_\theta, \gamma_\theta, \delta_\theta), \forall \theta \in \Theta \), where we proved that \( \alpha_\theta^* = \gamma_\theta^* = \delta_\theta^* = 1, \forall \theta \in \Theta \) is always the optimal solution of Eq. \((125)\), and designed a two-stage two-step algorithm to select \( \delta_\theta^*, \forall \theta \in \Theta \) which achieves low-complexity computation in an alternate manner. The social welfare \( U_P \) obtained by our proposed two-sided rating protocol \( P \) can be very close to the social optimum, especially when users are
sufficiently patient and the monitoring and reporting error is small.

In the following, we summarize a few limitations of our current protocol and point out a few directions for future work. Firstly, we assume that the client is always truth-telling when reporting the service quality. However, a strategic client may be benefitted by falsely reporting the outcome in order to drop the server’s rating, and thus be served by other low rating servers with a higher probability. Hence, how to evaluate the quality of service in a fair and simple way is very urgent. Secondly, the optimal recommended strategy changes from $\pi$ to $\pi_s$ and eventually to $\pi_0$ if $\omega$ is sufficiently small, it is interesting to maximize the social welfare when anonymous users are non-long-lived in a crowdsourcing platform. Thirdly, our design did not consider imperfect monitoring in the second-stage of the service exchange dilemma game, as we believe that $C$ is reported when the user chooses to be a server with a very small probability, and vice versa. However, actions $C$ and $S$ may failure with a probability, and thus it is essential to design appropriate rating protocols to solve such a problem.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61402418, 11531011, 11771013, 61503342, 61672468, 61602418, and in part by the Social Development Project of Zhejiang Provincial Public Technology Research under Grant 2017C33054.

REFERENCES

[1] W. J. Howe, “The rise of crowdsourcing”, Wired magazine 14 (6), pp. 1-4, 2006.
[2] J. Howe, “Crowdsourcing: A Definition”, 2006. [Online]. Available: http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_2.html
[3] Amazon Mechanical Turk. [Online]. Available: http://www.mturk.com/
[4] Yahoo! Answers. [Online]. Available: https://answers.yahoo.com/
[5] Upwork. [Online]. Available: https://www.upwork.com/
[6] H. Hu, G. Li, Z. Bao, Y. Cui, and J. Feng, “Crowdsourcing-based real-time urban traffic speed estimation: From trends to speeds”, in Proc. of 32nd International Conference on Data Engineering (ICDE), Helsinki, Finland, 16-20 May, 2016, pp. 883-894.
[7] Y. Wu, Y. Wang, W. Hu, and G. Cao, “SmartPhoto: A Resource-Aware Crowdsourcing Approach for Image Sensing with Smartphones”, IEEE Transactions on Mobile Computing, vol. 15 , issue. 5, pp.1249-1263, May 2016.
[8] A. Tarable, A. Nordio, E. Leonardi, and M. A. Marsan, “The importance of being earnest in crowdsourcing systems”, in Proc. of 34th IEEE Conference on Computer Communications (INFOCOM), Hong Kong, China, 26 April-1 May, 2015, pp. 2821-2829.
[9] A. Kittur, J. V. Nickerson, M. S. Bernstein, E. M. Gerber, and et al., “The future of crowd work”, in Proc. of 16th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW), San Antonio, Texas, USA, 23-27 Feb, 2013, pp. 1301-1318.
[10] V. Naroditskiy, N. R. Jennings, H. P. Van, and M. Cebrian, “Crowdsourcing contest dilemma”, Journal of the Royal Society Interface, vol. 11, no. 99, pp. 1-8, Aug. 2014.
[11] K. Oishi, M. Cebrian, A. Abeleelu, and N. Masuda, “Iterated crowdsourcing dilemma game”, Scientific Reports, 4:4100, pp. 1-7, 2014.
[12] Y. Zhang, and M. van der Schaar, “Socially-optimal Design of Service Exchange Platforms with Imperfect Monitoring”, ACM Transactions on Economics and Computation, vol. 3, no. 4, pp. 25:1-25:25, Jul. 2015.
[13] J. Lu, C. Tang, X. Li, and Q. Wu, “Designing Socially-Optimal Rating Protocols for Crowdsourcing Contest Dilemma”, IEEE Transactions on Information Forensics and Security, vol.12, issue 6, pp.1330-1344, Jun. 2017.
[14] A. I. Chitlappilly, L. Chen, and S. Amer-Yahia, “A Survey of General-Purpose Crowdsourcing Techniques”, IEEE Transactions on Knowledge and Data engineering, vol. 28, no. 9, pp. 2246-2266, Sep. 2016.