Exchanging Reputation Information Between Communities: A Payment-Function Approach

Georgia Kastidou
David R. Cheriton School of Computer Science
University of Waterloo, Canada
gkastidou@cs.uwaterloo.ca

Kate Larson
David R. Cheriton School of Computer Science
University of Waterloo, Canada
klarson@cs.uwaterloo.ca

Robin Cohen
David R. Cheriton School of Computer Science
University of Waterloo, Canada
rcohen@uwaterloo.ca

Abstract

We introduce a framework so that communities can exchange reputation information about agents in environments where agents are migrating between communities. We view the acquisition of the reputation information as a purchase and focus on the design of a payment function to facilitate the payment for information in a way that motivates communities to truthfully report reputation information for agents. We prove that in our proposed framework, honesty is the optimal policy and demonstrate the value of using a payment-function approach for the exchange of reputation information about agents between communities in multiagent environments. Using our payment function, each community is strengthened: it is able to reason more effectively about which agents to accept and can enjoy agents that are motivated to contribute strongly to the benefit of the community.

1 Introduction

Several researchers modeling trustworthiness of agents in multiagent systems have proposed the exchange of ratings of agents between peers, in a social network. This is especially useful in settings such as e-marketplaces, where a buying agent having little experience with a selling agent may ask other buyers for advice. One approach that has emerged is creating incentives for honesty when reputation ratings of agents are exchanged, in an effort to provide recipients with reliable information (e.g. [Jurca and Faltings, 2007; Miller *et al.*, 2005]).

In this paper, we examine a similar but distinct problem: how to promote honest exchange of information about the reputation of agents, between communities. By community, we mean a collection of agents that co-exist for a specific purpose. A primary example would be a file sharing community, where agents representing users make decisions about what to upload or download for the community and the reputation of an agent in the community is a reflection of how valuable its contributions are, to the other members. Another example of communities include auction settings where agents are brought together to perform several business transactions over time. Our perspective is that the quality of services that a community offers is based on the collective contributions of its agents. It is therefore important for communities to make careful decisions when accepting agents; it is likewise beneficial to encourage agents to be good contributors, in their communities.

Since agents can migrate between, or be members of, multiple communities, how an agent behaves in one community is of interest to others. This is particularly true in situations where communities can decide whether or not to allow a particular agent to join. Thus mechanisms that allow communities to exchange *reputation information* about agents can improve communities in two important ways. First, community A can use information obtained from community B in order to decide whether to allow an agent to join. Second, if agents are aware that their reputation will follow them from one community to another, then they may have incentive to be better community citizens.

Since communities have their own interests, which may not align with others, they may not always be willing to exchange information about their members freely. There are a variety of reasons for this. For example, consider an agent which participates in an online auction community A in which it has an outstanding contribution. Assume now that this agent is interested in joining another online auction community B. If the agent joins community B it might have to provide part of its contribution that so far was exclusively provided to community A to the new community B. Given this potential loss of contribution, the community A might be reluctant to truthfully reveal to B the contribution of the agent.

We leverage the fact that communities may be both buyers and sellers of information and construct a payment function that can be used to effectively provide sellers with appropriate payment for providing information.

The two main issues we are interested in addressing are:

1. how a community can be motivated to truthfully report its ratings and
2. how we can value the quality of the rating a community provides in order to compensate it with a fair payment.

As will be seen, we set our payment function to maximize the payment of a community only when it provides a truthful rating (for 1.) and introduce a set of properties a payment function should follow in order to promote honesty and fairness (for 2.), thus, providing an effective proposal for the exchange of reputation information between communities.

A key distinction that we make is between what we refer to as a good or a poor contributor. These are labels that correspond to the desirability for a community to accept or to reject the agent (the key decision-making that each community must undergo). The reputation information that is shared consists of both a rating and a type (good, poor). The community receiving the agent reputation explicitly evaluates the importance of this information. As will be shown, reputation information that leads to the correct decision about accepting or rejecting an agent leads to more lucrative payments to the sellers, thus promoting both honest reporting and fair payments. In Section 5 we return to clarify the distinction between communities sharing ratings of agents and current research where agents share ratings amongst themselves.

2 Model

Let C_i denote community i and let a_j denote agent j. We assume that if a_j is a member of community C_i then C_i can observe and judge the quality of agent a_j. In particular, we assume that community C_i maintains a *reputation model* for all member agents, and is able to assign a *reputation rating* r_j^i to agent a_j, where r_j^i is some real number from the interval $[\alpha, \beta]$, $0 \leq \alpha \leq \beta$.

If agent a_j wishes to join community C_k, then before welcoming a_j, community C_k will contact the communities in which a_j is currently, or was previously, a member. We denote the set of these communities as $S(a_j)$. The communities in $S(a_j)$ are asked to provide two bits of information. First, each community C_i is asked to report r_j^i, its reputation rating for agent a_j. Since communities may use different reputation models, or may interpret ratings differently, communities are also requested to provide *type information*, θ_j^i, for agent j. We assume that $\theta_j^i \in \{\text{good}, \text{poor}\}$, and that this is the *interpretation* that community C_i makes concerning its reputation rating r_j^i for agent a_j.[1]

In exchange for information, community C_i receives a payment from community C_k. This payment, P, depends on how useful or *important* community C_k finds the information provided by C_i. If community C_k is interested in possibly welcoming agent a_j, it will contact the communities in $S(a_j)$ to request information about the agent a_j. Each community $C_i \in S(a_j)$ reports its reputation rating (r_j^i) and type information (θ_i^j) for agent a_j, possibly misreporting the information.

Based on the information received from communities in $S(a_j)$, C_k decides whether to accept agent a_j or to reject it. If C_k accepts agent a_j then it gets to *observe* and *evaluate* a_j. By doing so, C_k is able to assign both a rating, $\hat{r}$, and a type, θ_j^k to the agent. If community C_k decides to *reject* agent a_j, then the payment procedure is a little more complicated since the community never gets the opportunity to directly observe and evaluate agent a_j. Instead, community C_k sets the type of agent a_j to be $\theta_j^k = \text{poor}$, and computes $\hat{r}$ to be the average reputation rating of all communities in $S(a_j)$ that also assigned type $\theta_j^i = \text{poor}$. That is, if $B = \{i | i \in S(a_j) \text{ and } \theta_j^i = \text{poor}\}$ then:

$$\hat{r} = \frac{\sum_{i \in B} r_j^i}{|B|}.$$

Only after this decision is made do the communities in $S(a_j)$ get paid for the information that they provided. In particular, for each community $C_i \in S(a_j)$, the payment it receives for its information (r_j^i and θ_j^i) is

$$P(r_j^i, \hat{r}, \theta_j^k) = \bar{\alpha} \cdot I(r_j^i, \hat{r}, \theta_j^k) + \bar{\beta} \quad (1)$$

where $\bar{\alpha} \in \mathbb{R}^+$ and $\bar{\beta} \in \mathbb{R}$ are constants set by community C_k, and $I(r_j^i, \hat{r}, \theta_j^k)$ is the *Importance Function* used by community C_k to determine the usefulness of the reputation information provided by community C_i.

While the parameters $\bar{\alpha}$ and $\bar{\beta}$ are interesting in that community C_k can use them to scale the payments made to the other communities, the key part of the payment function is the *Importance Function*. The rest of this paper focuses on what this function should be, and what properties it should exhibit in order to create an effective payment scheme for exchanging reputation information between communities of agents.

3 The Importance Function

In this section we describe the basic desirable properties of the Importance Function (*IMF*). In the previous section we informally introduced *IMF* as $I(r_j^i, \hat{r}, \theta_j^k)$ where r_j^i was the reputation rating given by community C_i, $\hat{r}$ was community C_k's (the community making the payment) reputation rating of the agent, and θ_j^k was C_k's assigned type to the agent. We define I as $I : \mathbb{R} \times \mathbb{R} \times \{\text{good}, \text{poor}\} \to \mathbb{R}$ such that $I(x, \hat{r}, \theta) = -\infty$ if $x < \alpha$ or $x > \beta$ for some predefined $\alpha, \beta \geq 0$. Because of this, a community C_i, reporting on agent a_j is best off revealing a *legal* rating.

Our first desired property is that I is continuous when a legal reputation rating is given.

Property 1. *Let $\alpha \geq 0$. In the restricted domain $[\alpha, \beta] \times [\alpha, \beta] \times \{\text{good}, \text{poor}\}$, I is continuous.*

If the community accepts the agent, then it is able to observe and evaluate the agent, and thus determine its own reputation rating, $\hat{r}$. When determining payments we desire that the communities that provided the *more accurate* information are rewarded. Consider the following example.

Example 1. *Assume that community C_k requests information about agent a_j. Assume that four communities submitted reputation ratings $r_j^1 = 0.6$, $r_j^2 = 0.65$, $r_j^3 = 0.75$ and $r_j^4 = 0.8$. After observing the agent, community C_k set $\hat{r} = 0.7$.*

We would like our importance function to reward the communities that submitted reputation ratings of 0.65 and 0.75 more than those that submitted 0.6 and 0.8, respectively. That is, we would like to capture the property that as x approaches $\hat{r}$, the importance of x increases. Property 2 captures this.

[1] For example, a rating 0.55 in one community could indicate an agent which is a poor contributor while in another community, which might be more strict in providing high ratings, the same rating might indicate an agent which is a good contributor.

Property 2. *For any* $\hat{r} \in [\alpha, \beta]$ *and for any* $\theta \in \{\text{good}, \text{poor}\}$, *if* $x \in [\alpha, \hat{r}]$ *then* $I(x, \hat{r}, \theta)$ *is strictly monotonically increasing, and if* $x \in [\hat{r}, \beta]$ *then* $I(x, \hat{r}, \theta)$ *is strictly monotonically decreasing.*

Given Property 2, the *IMF* should reward communities C_2 and C_3 in Example 1 more than communities C_1 and C_4, respectively. We may want, however, to further distinguish between the information received from communities C_2 and C_3 since even though the reputation values were equally far from $\hat{r}$, C_2 stated that a_j had a lower reputation than observed, while community C_3 stated a_j had a higher reputation than observed.

Define

$$\delta(\hat{r}, \epsilon, \theta) = I(\hat{r} + \epsilon, \hat{r}, \theta) - I(\hat{r} - \epsilon, \hat{r}, \theta)$$

for any $\epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$. This measures the difference in the *IMF* value when communities over-report and under-report by the same amount. If $\delta(\hat{r}, \epsilon, \theta) = 0$ for all ϵ then the *IMF* would treat over- and under-reported reputation ratings equally. We believe that the *IMF* should be used to reward communities that provide ratings which deviate towards the *correct direction* (i.e. good or poor) higher than those communities who provided ratings of equal deviation but towards the wrong direction.

Property 3. *For any* $\hat{r}$ *and any* $\epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$ *let* $I(x, \hat{r}, \theta)$ *be such that*

$$\delta(\hat{r}, \epsilon, \text{good}) > 0$$

and

$$\delta(\hat{r}, \epsilon, \text{poor}) < 0.$$

In words, Property 3 says that if C_k determines that the agent is a good agent, then communities who reported higher reputation ratings should be rewarded more than communities who reported lower reputation ratings, assuming that the difference from $\hat{r}$ is the same. A similar property should hold if C_k determined that the agent was poor. Referring back to Example 1, if C_k determined that $\theta = \text{good}$, then community C_3 should have a higher *IMF* value than C_2, and C_4 should have a higher *IMF* value than C_1. If $\theta = \text{poor}$, then C_2 should have a higher *IMF* value than C_3 and C_1 should have a higher *IMF* value than C_4.

Our last two desired properties describe how $\delta(\hat{r}, \epsilon, \theta)$ should behave.

Property 4. *For any* $\hat{r}$

- $\delta(\hat{r}, \epsilon, \text{good})$ *is strictly monotonically increasing in* ϵ,
- $\delta(\hat{r}, \epsilon, \text{poor})$ *is strictly monotonically decreasing in* ϵ.

We interpret Property 4 in that if an agent is judged to be good, then communities who submitted high reputation ratings for the agent deserve higher *IMF* values, since they were offering support for the agent (and vice-versa for the case when an agent is judged to be poor). Referring back to Example 1, if the agent was judged to be good, then community C_4 would receive the highest *IMF* value, whereas if the agent was judged to be poor, then community C_1 would receive the highest *IMF* value.

Property 5. *For any* $\epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$ *and for any* $\theta \in \{poor, good\}$, $\delta(\hat{r}, \epsilon, \theta)$ *is monotonically decreasing in* $\hat{r}$.

Property 5 states that a given deviation ϵ has different significance for different values of $\hat{r}$. For instance, if the agent's type is judged to be good then the significance of a deviation ϵ increases as the reported rating $\hat{r}$ decreases. This is due to the fact that as $\hat{r}$ decreases the rating $\hat{r} - \epsilon$ might be crossing the cutoff $\bar{r}$ value that a community considers in order to accept an agent or not. In particular, the further the rating $\hat{r} - \epsilon$ crosses $\bar{r}$ the more in doubt it can put a community regarding the agent's real value. Analogous, is the case where the agent's type is judged to be poor.

3.1 Importance Function and Payments

The Importance function (*IMF*) forms the foundation of our payment system, and thus the properties of the *IMF* have a profound influence on the properties of the payments, and the incentives for communities to report their reputation ratings and type information when requested.

We first note that if the *IMF* satisfies Properties 1 and 2, then it is uniquely maximized when communities report $\hat{r}$ (*i.e.* $I(\hat{r}, \hat{r}, \theta)$ is the global maximum). Since the payment a community receives is an affine transformation of the *IMF*, a community has incentive to report the reputation rating that it truly believes C_k will experience if C_k accepts the agent.

We introduce Properties 3 through 5 so as to ensure a certain level of *fairness* in the system. While the communities who present the most accurate information to community C_k benefit the most from the *IMF*, communities, which provide information that tried to convince community C_k to make the appropriate decision with respect to the agent, are also well rewarded payment-wise.

4 A Class of Importance Functions

In the previous section we outlined the desireable properties for an *IMF*. The obvious question is then *Does there exist any functions which could be used as* IMF*s?* In this section we introduce a class of functions that satisfy Properties 1 to 4, and that contains a subclass which satisfies Property 5.

Let $\phi : [\alpha, \beta] \rightarrow \mathbb{R}^+$ and $\psi : [\alpha, \beta] \rightarrow \mathbb{R}^+$ be arbitrary continuous functions on $[\alpha, \beta]$. Let ϕ be strictly monotonically decreasing and let ψ by strictly monotonically increasing. Now, define *IMF* $I(x, \hat{r}, \theta)$ as

$$I(x, \hat{r}, \theta) = \begin{cases} \int_\alpha^\beta \phi(y)dy - |\int_{\hat{r}}^x \phi(y)dy|, & \text{if } \theta = \text{good} \\ \\ \int_\alpha^\beta \psi(y)dy - |\int_{\hat{r}}^x \psi(y)dy|, & \text{if } \theta = \text{poor} \end{cases} \quad (2)$$

We now show that for any choice of ψ and ϕ, $I(x, \hat{r}, \theta)$ satisfies Properties 1 to 4. Later we show that for particular choices of ϕ and ψ it is possible to also satisfy Property 5.

First, since both ϕ and ψ are continuous on $[\alpha, \beta]$, then $I(x, \hat{r}, \theta)$ is also continuous on the restricted domain $[\alpha, \beta] \times [\alpha, \beta] \times \{\text{good}, \text{poor}\}$. That is, $I(x, \hat{r}, \theta)$ satisfies Property 1.

As the x approaches $\hat{r}$ from the left the area that is defined by $|\int_x^{\hat{r}} \phi(y)dy|$ is strictly decreasing. Consequently, $I(x, \hat{r}, \theta)$ is strictly increasing in $[a, \hat{r}]$. As the x goes away from $\hat{r}$ then the area that is defined by the $|\int_{\hat{r}}^x \phi(y)dy|$ is strictly increasing. Thus, $I(x, \hat{r}, \theta)$ is strictly decreasing in $[\hat{r}, \beta]$. Similarly, it can be proved that the function ψ satisfies Property 2 as well.

Property 3 is also satisfied. We can rewrite $\delta(\hat{r}, \epsilon, \theta)$ as:

$$\delta(x, \hat{r}, \theta) = \begin{cases} -\int_{\hat{r}}^{\hat{r}+\epsilon} \phi(y)dy + \int_{\hat{r}-\epsilon}^{\hat{r}} \phi(y)dy, & \text{if } \theta = \text{good} \\ -\int_{\hat{r}}^{\hat{r}+\epsilon} \psi(y)dy + \int_{\hat{r}-\epsilon}^{\hat{r}} \psi(y)dy, & \text{if } \theta = \text{poor} \end{cases}$$

We show the case for $\theta = \text{good}$ since the case when $\theta = \text{poor}$ is analogous. Given that $\hat{r} \in [\alpha, \beta]$ and ϕ is strictly monotonically decreasing and positive in $[\alpha, \beta]$,:

$$\int_{\hat{r}}^{\hat{r}+\epsilon} \phi(y)dy < \int_{\hat{r}}^{\hat{r}+\epsilon} \phi(\hat{r})dy \text{ and } \int_{\hat{r}}^{\hat{r}+\epsilon} \phi(\hat{r})dy = \int_{\hat{r}-\epsilon}^{\hat{r}} \phi(\hat{r})dy$$

and

$$\int_{\hat{r}-\epsilon}^{\hat{r}} \phi(\hat{r})dy < \int_{\hat{r}-\epsilon}^{\hat{r}} \phi(y)dy$$

Thus:

$$\int_{\hat{r}}^{\hat{r}+\epsilon} \phi(y)dy < \int_{\hat{r}-\epsilon}^{\hat{r}} \phi(y)dy \Leftrightarrow \delta(\hat{r}, \epsilon, \text{good}) > 0$$

Regarding Property 4 we need to show that the partial derivative of $\delta(\hat{r}, \epsilon, \theta)$ with respect to ϵ is greater than zero for $\theta = \text{good}$ and less than zero for $\theta = \text{poor}$, where

$$\frac{\partial \delta(\hat{r}, \epsilon, \theta)}{\partial \epsilon} = \begin{cases} \phi(\hat{r} - \epsilon) - \phi(\hat{r} + \epsilon), & \text{if } \theta = \text{good} \\ \psi(\hat{r} - \epsilon) - \psi(\hat{r} + \epsilon), & \text{if } \theta = \text{poor} \end{cases}$$

Given that ϕ and ψ are strictly monotonically increasing and strictly monotonically decreasing, respectively, and $\hat{r} \in [\alpha, \beta]$, where $\beta > \alpha \geq 0$, and $\epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$: $\phi(\hat{r} + \epsilon) < \phi(\hat{r} - \epsilon)$ and $\psi(\hat{r} - \epsilon) < \psi(\hat{r} + \epsilon)$. Thus:

$$\begin{cases} \frac{\partial \delta(\hat{r}, \epsilon, \theta)}{\partial \epsilon} > 0, & \text{if } \theta = \text{good} \\ \frac{\partial \delta(\hat{r}, \epsilon, \theta)}{\partial \epsilon} < 0, & \text{if } \theta = \text{poor} \end{cases} \tag{3}$$

Finally, Property 5 is also satisfied by (2) if:

$$\begin{cases} 2\phi(\hat{r}) - \phi(\hat{r} + \epsilon) - \phi(\hat{r} + \epsilon) \leq 0 \\ 2\psi(\hat{r}) - \psi(\hat{r} + \epsilon) - \psi(\hat{r} + \epsilon) \leq 0 \end{cases} \tag{4}$$

for $\forall \hat{r} \in [\alpha, \beta]$, where $\beta > \alpha \geq 0$, and $\forall \epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$. More specifically, we need to prove that the partial derivatives of $\delta(\hat{r}, \epsilon, \theta)$ with respect to $\hat{r}$ are less than or equal to zero. This is true if (4) holds, given that:

$$\frac{\partial \delta(\hat{r}, \epsilon, \theta)}{\partial \hat{r}} = \begin{cases} 2\phi(\hat{r}) - \phi(\hat{r} + \epsilon) - \phi(\hat{r} - \epsilon) \leq 0, & \text{if } \theta = \text{good} \\ 2\psi(\hat{r}) - \psi(\hat{r} + \epsilon) - \psi(\hat{r} - \epsilon) \leq 0, & \text{if } \theta = \text{poor} \end{cases}$$

Thus, if ϕ and ψ satisfy inequalities (4) then Property 5 is also satisfied.

Proposition 1 states that certain linear transformations of the functions ϕ and ψ can be also used to create an *IMF*.

Proposition 1. *Let $\phi : [\alpha, \beta] \to \mathbb{R}^+$ and $\psi : [\alpha, \beta] \to \mathbb{R}^+$ be arbitrary continuous functions on $[\alpha, \beta]$, $\beta > \alpha \geq 0$. Let ϕ be strictly monotonically decreasing and let ψ by strictly monotonically increasing in $[\alpha, \beta]$, and:*

$$\begin{cases} 2\phi(\hat{r}) - \phi(\hat{r} + \epsilon) - \phi(\hat{r} + \epsilon) \leq 0 \\ 2\psi(\hat{r}) - \psi(\hat{r} + \epsilon) - \psi(\hat{r} + \epsilon) \leq 0 \end{cases}$$

for $\forall \hat{r} \in [\alpha, \beta]$ and $\forall \epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$. The function:

$$I(x, \hat{r}, \theta) = \begin{cases} \int_{\alpha}^{\beta} \Phi(y)dy - |\int_{\hat{r}}^{x} \Phi(y)dy|, & \text{if } \theta = \text{good} \\ \int_{\alpha}^{\beta} \Psi(y)dy - |\int_{\hat{r}}^{x} \Psi(y)dy|, & \text{if } \theta = \text{poor} \end{cases}$$

where $\Phi(y) = \lambda_1 * \phi(y) + \lambda_2$, $\Psi(y) = \kappa_1 * \psi(y) + \kappa_2$, $\lambda_1, \kappa_1 \in \mathbb{R}^+$, $\lambda_2, \kappa_2 \in \mathbb{R}$ and $\Phi(y), \Psi(y) \geq 0$ in $[\alpha, \beta]$, is a valid IMF *which satisfies Properties 1 to 5.*

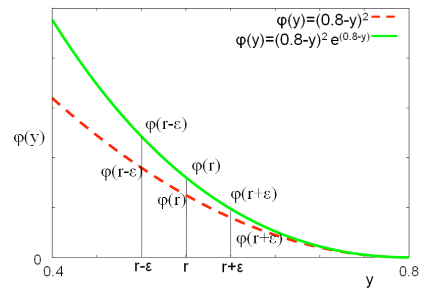

Figure 1: Examples of ϕ

Due to its simplicity and space limitations the proof is omitted.

4.1 Examples

In this section we introduce two examples of an *IMF*. Let $\phi(y) = (\beta - y)^n$ and $\psi(y) = y^m$. Then:

$$I(x, \hat{r}, \theta) = \begin{cases} \int_{\alpha}^{\beta} (\beta - y)^n dy - |\int_{\hat{r}}^{x} (\beta - y)^n dy| & \theta = \text{good} \\ \int_{\alpha}^{\beta} y^m dy - |\int_{\hat{r}}^{x} y^m dy| & \theta = \text{poor} \end{cases}$$

where $x, \hat{r} \in [\alpha, \beta]$, $\beta > \alpha \geq 0$, and $n, m \in \mathbb{N}^+$. An example of $\phi(y)$ with $\beta = 0.8$ and $n = 2$ is depicted in Figure 1.

Since ϕ and ψ are continuous, positive and strictly monotonically decreasing and strictly monotonically increasing, respectively (as can be seen by a simple check of the first and second derivatives), $I(x, \hat{r}, \theta)$ satisfies Properties 1 to Property 4. In order to prove that $I(x, \hat{r}, \theta)$ also satisfies Property 5 it is sufficient to prove that

$$2\phi(\hat{r}) - \phi(\hat{r}+\epsilon) - \phi(\hat{r}+\epsilon) = 2(\beta - \hat{r})^n - (\beta - (\hat{r}+\epsilon))^n - (\beta - (\hat{r}-\epsilon))^n \leq 0 \tag{5}$$

and

$$2\psi(\hat{r}) - \psi(\hat{r} + \epsilon) - \psi(\hat{r} + \epsilon) = 2\hat{r}^m - (\hat{r} + \epsilon)^m - (\hat{r} - \epsilon)^m \leq 0 \tag{6}$$

for $\forall \hat{r} \in [\alpha, \beta]$ and $\forall \epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$. Inequality (5) can be proved by induction. More specifically:

For $n = 0$, we have

$$2(\beta - \hat{r})^0 - (\beta - (\hat{r} + \epsilon))^0 - (\beta - (\hat{r} - \epsilon))^0 = 2 - 2 = 0$$

For $n = 1$, we have

$$2\beta - 2\hat{r} - \beta + \hat{r} - \epsilon - \beta + \hat{r} - \epsilon = -2\epsilon \leq 0$$

Assume that it is true for $n = k$:

$$2(\beta - \hat{r})^k - (\beta - (\hat{r} + \epsilon))^k - (\beta - (\hat{r} - \epsilon))^k \leq 0 \tag{7}$$

we will prove that it is also true for $n = k + 1$:

$$\begin{aligned} & 2(\beta - \hat{r})^{k+1} - (\beta - (\hat{r} + \epsilon))^{k+1} - (\beta + (\hat{r} - \epsilon))^{k+1} \\ = \; & (\beta - \hat{r})(2(\beta - \hat{r})^k - (\beta - (\hat{r} + \epsilon))^k - (\beta - (\hat{r} - \epsilon))^k) + \\ + \; & \epsilon((\beta - (\hat{r} + \epsilon))^k - (\beta - (\hat{r} - \epsilon))^k) \leq 0 \end{aligned} \tag{8}$$

Inequality (8) is true as it is the summation of two negative numbers. Similarly, inequality (6) can be proved.

In the second example we consider $\phi(y) = (\beta - y)^n e^{(\beta - y)}$ and $\psi(y) = y^m e^y$. In this case the *IMF* (2) will become:

$$I(x, \hat{r}, \hat{\theta}) = \begin{cases} \int_{\alpha}^{\beta} (\beta - y)^n e^{(\beta - y)} dy - |\int_{\hat{r}}^{x} (\beta - y)^n e^{(\beta - y)} dy| & \theta = \text{good} \\ \int_{\alpha}^{\beta} y^m e^y dy - |\int_{\hat{r}}^{x} y^m e^y dy| & \theta = \text{poor} \end{cases} \tag{9}$$

An example of ϕ for $\beta = 0.8$ and $n = 2$ is depicted in Figure 1.

Similarly with the first example, since ϕ and ψ are continuous, positive, and strictly monotonically decreasing and strictly monotonically increasing respectively (as can be seen by a simple check of the first and second derivatives), $I(x, \hat{r}, \theta)$ satisfies Properties 1 to 4.

As in the case of the first family, in order to prove that *IMF* (9) also satisfies Property 5 it is sufficient to show that $\forall n, m \in \mathbb{N}, \forall \hat{r} \in [\alpha, \beta]$ and $\forall \epsilon \in (0, \min[\hat{r} - \alpha, \beta - \hat{r}]]$:

$$\begin{aligned} & 2\phi(\hat{r}) - \phi(\hat{r}+\epsilon) - \phi(\hat{r}+\epsilon) = \\ = \; & 2(\beta - \hat{r})^n e^{\beta - \hat{r}} - (\beta - (\hat{r}+\epsilon))^n e^{(\beta - (\hat{r}+\epsilon))} + \\ - \; & (\beta - (\hat{r}-\epsilon))^n e^{(\beta - (\hat{r}-\epsilon))} \leq 0 \end{aligned} \tag{10}$$

and

$$2\psi(\hat{r}) - \psi(\hat{r}+\epsilon) - \psi(\hat{r}+\epsilon) = 2\hat{r}^m e^{\hat{r}} - (\hat{r}+\epsilon)^m e^{\hat{r}+\epsilon} - (\hat{r}-\epsilon)^m e^{\hat{r}-\epsilon} \leq 0 \tag{11}$$

Consider the case where $\theta = \text{poor}$. We will prove that inequality (11) is true by induction (similarly inequality (10) can be proved).[2] For $m = 0$ we have: $2 - e^{\epsilon} - e^{-\epsilon}$ which is clearly less or equal to zero. For $m = 1$ we have:

$$2r - e^{\epsilon}(\hat{r}+\epsilon) - e^{-\epsilon}(\hat{r}-\epsilon) = \epsilon(e^{-\epsilon} - e^{\epsilon}) + (2 - e^{\epsilon} - e^{-\epsilon}) \leq 0 \tag{12}$$

which is true since it is the summation of two negative numbers. Assume now that (11) is true for $m = k$:

$$2r^k - e^{\epsilon}(\hat{r}+\epsilon)^k - e^{-\epsilon}(\hat{r}-\epsilon)^k \leq 0$$

We will prove that it is also true for $m = k + 1$.

$$\begin{aligned} 2r^{k+1} - e^{\epsilon}(\hat{r}+\epsilon)^{k+1} - e^{-\epsilon}(\hat{r}-\epsilon)^{k+1} \leq 0 \Leftrightarrow \\ -e^{\epsilon}(\hat{r}+\epsilon)^k(\hat{r}+\epsilon) + 2r^k r - e^{-\epsilon}(\hat{r}-\epsilon)^k(\hat{r}-\epsilon) \leq 0 \Leftrightarrow \\ (-e^{\epsilon}(\hat{r}+\epsilon)^k + 2r^k - e^{-\epsilon}(\hat{r}-\epsilon)^k)r + \epsilon(e^{-\epsilon}(\hat{r}-\epsilon) - e^{\epsilon}(\hat{r}+\epsilon)) \leq 0 \end{aligned}$$

which is true since it is the summation of two negative numbers. Consequently, the inequality (11) holds for $\forall m \in \mathbb{N}$.

Choosing ϕ and ψ

In order to find the most suitable instances of the above families of functions a number of different criteria could be used. For example, one criterion is bounding the maximum value of $\delta(\hat{r}, \epsilon, \theta)$.

More specifically, consider the case of $\theta = \text{good}$, a criterion could be to choose the function ϕ in such a way that the maximum value of the $\delta(\hat{r}, \epsilon, \theta)$ (which essentially defines the maximum difference of the *IMF* value two ratings of equal deviation can have) is bounded by U and L, where $U, L \in \mathbb{R}^+$, in order to avoid over-penalizing communities that under-reported the rating of an agent. A simple way of finding ϕ and ψ that satisfy the following bounds:

$$L \leq max[\delta(\hat{r}, \epsilon, good)] \leq U \textbf{ and } L' \leq min[\delta(\hat{r}, \epsilon, poor)] \leq U'$$

in an interval $[w_1, w_2]$, where $U, L, w_1, w_2 \in \mathbb{R}^+$ and $U', L' \in \mathbb{R}^-$ is presented in Proposition 2.

Proposition 2. *Given an upper bound U and a lower bound L finding the function ϕ such that for $\forall \epsilon \in (0, min[w_2 - \hat{r}, \hat{r} - w_1]]$ and $\forall \hat{r} \in [w_1, w_2]$:*

$$L \leq max[\delta(\hat{r}, \epsilon, good)] \leq U$$

is equivalent to finding the function ϕ such that:

$$L \leq \int_{\hat{r}-\bar{\epsilon}}^{x} \phi(y)dy \leq U$$

where $\bar{\epsilon} = min[w_2 - \hat{r}, \hat{r} - w_1]$, and x is the solution to: $I(\hat{r} + \bar{\epsilon}, \hat{r}, good) = I(x, \hat{r}, good)$. While finding the function ψ that for $\forall \epsilon \in (0, min[w_2 - \hat{r}, \hat{r} - w_1]]$ and $\forall \hat{r} \in [w_1, w_2]$

$$L' \leq min[\delta(\hat{r}, \epsilon, poor)] \leq U'$$

is equivalent to finding the function ψ such that:

$$L' \leq \int_{x}^{\hat{r}+\bar{\epsilon}} \phi(y)dy \leq U'$$

where $\bar{\epsilon} = min[w_2 - \hat{r}, \hat{r} - w_1]$, and x is the solution to: $I(\hat{r} - \bar{\epsilon}, \hat{r}, poor) = I(x, \hat{r}, poor)$.

Proof. Consider the case where $\theta = good$.[3] Given that $\delta(\hat{r}, \epsilon, good)$ is monotonically increasing in ϵ: i) it is maximized when $\epsilon = min[\hat{r} - w_1, w_2 - \hat{r}]$, and ii) there will be a $x \in (\hat{r} - \epsilon, \hat{r})$ such that $I(x, \hat{r}, good) = I(\hat{r} + \epsilon, \hat{r}, good)$. Thus, we can write $\delta(\hat{r}, \epsilon, \theta)$ as:

$$\delta(\hat{r}, \epsilon, \theta) = \int_{\hat{r}-\bar{\epsilon}}^{x} \phi(y)dy$$

□

Essentially, Proposition 2 states that in order to find a ϕ that leads to an $\delta(\hat{r}, \epsilon, \theta)$ whose maximum value is between two bounds L and U, we just have to find a ϕ such that:

$$L \leq \int_{\hat{r}-\bar{\epsilon}}^{x} \phi(y)dy \leq U$$

where $\bar{\epsilon} = min[w_2 - \hat{r}, \hat{r} - w_1]$, and x is the solution to: $I(\hat{r} + \bar{\epsilon}, \hat{r}, good) = I(x, \hat{r}, good)$. Analogous is the case for the function ψ.

5 Discussion & Future Work

In this paper we present a payment function approach to promote honest exchange of information about the reputation of agents, between communities. Our payment function is motivated by the work on scoring rules [Savage, 1971]: a framework for eliciting probabilistic information from agents. We differ from the formal definition of scoring rules in that we determine a value based on the community's declarations, whether the information was judged to be good or not (a binary event) and how it relates to the judging agent's valuation (a continuous variable), while scoring rules typically are either concerned with binary events or continuous settings, but not both [Zohar and Rosenschein, 2008]. We still retain the salient features of scoring rules, such as incentive compatibility.

Other authors have investigated the use of incentive mechanisms in order to promote trustworthiness when reporting about agents in multiagent systems [Jurca and Faltings, 2003; 2007; Miller *et al.*, 2005]. Jurca and Faltings offer a side payment mechanism and proves that rational software agents under this mechanism will truthfully share their reputation information [Jurca and Faltings, 2003]. They further examine

[2] Given that $e^r > 0$ we can omit it.

[3] The proof for $\theta = \text{poor}$ is analogous.

how to discourage collusion among the reporting agents [Jurca and Faltings, 2007]. Miller et al. offer a somewhat different kind of incentive mechanism with scores kept in a central location, but is also able to demonstrate that this mechanism is incentive-compatible, creating an equilibrium under certain well-defined conditions [Miller *et al.*, 2005]. Our approach is to also create incentives for honesty but for the case of communities sharing information with other communities and doing so by carefully determining an appropriate payment for information. As a result, it is possible to reward reports that assist the buying communities most effectively (getting these communities closer to the end result of enjoying agents that are good contributors).

Unlike other proposed approaches for sharing reputation information that focus primarily on learning how to judge the trustworthiness of the sender, to then discount or disregard reputation ratings that are received [Jurca and Faltings, 2003; Teacy *et al.*, 2006], in our approach we require both type and rating information to be provided and are capable of coping with subjective differences in the way the ratings are calculated. Another approach that aims to explore subjective differences was proposed by Regan et al. [Regan *et al.*, 2006]. They presented a Bayesian-based approach to learn the evaluation function of an agent providing a rating, in order to make use of the information received. In our work we focus on determining the rewards that each community should receive with respect to the quality of the information it provided. A key contribution of our work therefore is that it promotes truthful reports and distributes the payments in a fair manner. Our next step is to take advantage of methods such as those proposed by Regan et al. [Regan *et al.*, 2006], as part of our effort to learn the behaviour of the selling communities and interpret the information that is received.

In our current approach the reputation of the agents is part of the information that it is being exchanged between the communities. Other multiagent researchers have also relied on exchanging reputation ratings [Dellarocas, 2001; Teacy *et al.*, 2006; Ismail and Josang, 2002], although in those cases the exchange of the ratings is between agents.[4] However, a limitation with a reputation rating for the case of communities is that it is highly related to the particular needs of the community. For example, a trustful agent might have low reputation inside a community because although it is willing to contribute, there is no current interest in the services it offers or simply the services it offers do not mesh with the needs of the community. On the other hand a malicious agent can temporarily create a good reputation and then become deceptive in the future, once it has been accepted in another community. An effective reputation system should be able to ensure minimal impact when the second problem occurs but it may not be able to cope with the case of a low reputation simply due to current lack of demand or due to incompatibility of services. Thus, there is a need to find a more comprehensive metric for evaluating an agent. One way we suggest is to consider the *trustworthiness* of an agent as the extent to which the agent is honest and contributes good quality information or services and is an active and consistent participant. Ultimately, our aim is to integrate a model for the trustworthiness of agents, as part of the buying community's reasoning about accepting agents.

In conclusion, in this paper we present the properties of a payment function which compensates the participant communities with respect to the importance of the provided information, and then we provide examples of specific families of functions that can be used. Finally, we give directions of possible criteria in selecting a specific instance of the latter families. We also offer an original proposal for exchanging reputation information, consisting of both a rating and a type, allowing the value of information received to affect the evaluation and hence the rewarding of the community providing the information. As such, we demonstrate the value of a payment function approach for exchanging reputation information between communities.

[4]Note that our approach could also be used for individual agents.

References

[Dellarocas, 2001] Chrysanthos Dellarocas. Analyzing the economic efficiency of ebay-like online reputation reporting mechanisms. In *EC '01: Proceedings of the 3rd ACM Conference on Electronic Commerce*, pages 171–179, New York, NY, USA, 2001. ACM Press.

[Ismail and Josang, 2002] R. Ismail and A. Josang. The beta reputation system. In *Proceedings of the 15th Bled Conference on Electronic Commerce*, 2002.

[Jurca and Faltings, 2003] Radu Jurca and Boi Faltings. An incentive compatible reputation mechanism. In *Proceedings of the IEEE Conference on E-Commerce*, pages 285–292, 2003.

[Jurca and Faltings, 2007] Radu Jurca and Boi Faltings. Collusion-resistant, incentive-compatible feedback payments. In *EC '07: Proceedings of the 8th ACM Conference on Electronic commerce*, pages 200–209, New York, NY, USA, 2007. ACM.

[Miller *et al.*, 2005] Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Manage. Sci.*, 51(9):1359–1373, 2005.

[Regan *et al.*, 2006] Kevin Regan, Pascal Poupart, and Robin Cohen. Bayesian reputation modeling in emarketplaces sensitive to subjectivity, deception and change. In *Proceedings of AAAI-06*, 2006.

[Savage, 1971] Leonard J. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

[Teacy *et al.*, 2006] W. T. Teacy, Jigar Patel, Nicholas R. Jennings, and Michael Luck. Travos: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006.

[Zohar and Rosenschein, 2008] Aviv Zohar and Jeffrey S. Rosenschein. Mechanisms for information elicitation. *Artif. Intell.*, 172(16-17):1917–1939, 2008.