A Point-Based Incentive Scheme for P2P Reputation Management Systems

Takuya NISHIKAWA and Satoshi FUJITA

Department of Information Engineering, Hiroshima University
Higashi-Hiroshima, 739-8527, Japan

Abstract—This paper proposes an incentive scheme for P2P resource management systems which encourages the users to evaluate the “trustworthiness of evaluations” given by the other users. More concretely, in the proposed scheme, each user earns a reward from other users by evaluating their evaluations, and a user which acquires a large number of evaluations will be granted a right to access high quality services. To this end, we introduce the notion of evaluation points which mediates the “evaluation of services” and the “evaluation of evaluations.” The performance of the proposed incentive scheme is evaluated by simulation. The simulation results indicate that: 1) the proposed scheme certainly encourages the users to conduct an evaluation of evaluations, 2) it encourages users to provide high quality evaluations of the services, and 3) a rational strategy for the users is to repeat evaluation of evaluations after conducting a certain number of evaluations of services.

Keywords: Peer-to-Peer resource management, reputation, incentive scheme.

1. Introduction

In recent years, Peer-to-Peer systems (P2Ps) have attracted considerable attention in the field of distributed computing and network applications. A P2P consists of a large number of autonomous computers called peers which can play the role of a client and a service provider at the same time. P2Ps have many advantages such as the fault-tolerance and the high scalability compared with traditional client/server systems, so that they are used in many network applications such as file sharing, live streaming, and IP phone. However, due to the anonymity of participants and the openness of the underlying network, P2Ps involve several critical issues such as the lack of authenticity and the low reliability of transactions which should be conducted with anonymous clients and/or service providers.

To overcome such a weakness of P2Ps, the use of trust management systems has also attracted attention in recent years. So far, many types of trust management systems have been proposed in the literature [17], [2], [15], [6], [7]. The main objective of the trust management in P2Ps is to make assessments and decisions regarding the dependability of potential transactions, and to allow users and the system managers to increase and correctly represent the reliability of themselves and their systems [5]. It is also commonly recognized that a key challenge in designing a trust management system in P2P environment is how to overcome the lack of central authorities to conduct assessments and necessary decisions (i.e., each user must make a decision about the trust of other users on her own responsibility while she could refer to the decision of other users collected to a central server), and how to overcome the lack of prior information concerned with other participants, since they are distributed systems consisting of many anonymous users.

In general trust management systems, the “reputation” of customers is commonly used as the source of information concerned with the trustfulness of service providers. In fact, online auction systems such as eBay and customer review services in shopping sites such as Amazon, try to increase their reliability by allowing customers to make an assessment of their past counterparty or of their product reviews, and by disclosing the result of such assessments to all users. Even in the P2P environments, the outcome of assessments on past transactions could be effectively used to identify (and sometimes to exclude) malicious users who involve potential risks, such as the upload of inauthentic files, a long delay of transactions and a sudden cancellation of ongoing transactions.

A critical issue in the reputation-based trust management in P2Ps with no central authorities, is that the reputation given by a malicious user may not be reliable. Although it would be possible to omit “all” reputations given by suspected users, we cannot identify a sufficient number of reliable users under such a pessimistic approach if the number of unsuspected users is not large. We need to carefully take into account the reputation given by every user (including suspected ones) to identify as large number of reliable users as possible. Another critical issue in actual reputation management systems is that it tends to lack the number of evaluations provided by the participants. For example, in KaZaA file sharing system, only 1% of shared files are evaluated by the users [8]. Too small number of evaluations would degrade the accuracy of the resultant reputation, even if we could effectively eliminate less reliable evaluations by using a scheme proposed in [11], for example.
In this paper, we propose an incentive scheme which encourages the users to evaluate the trustworthiness of evaluations given by the other users. More concretely, in the proposed scheme, each user earns a reward from other users by evaluating the evaluation given by them, and a user which acquired a large number of evaluations will be granted a right to access high quality services. To this end, we introduce a formal model of the evaluation cost as well as the definition of evaluation point which mediates the “evaluation of services” and the “evaluation of evaluations.” The performance of the proposed incentive scheme is evaluated by simulation. The simulation results indicate that: 1) the proposed scheme certainly encourages the participants to conduct an evaluation of evaluations, 2) it encourages to provide high quality evaluations of the services, and 3) a rational strategy for the participants is to repeat evaluation of evaluations after conducting a certain number of evaluations of services.

The remainder of this paper is organized as follows. Section 2 overviews related works. Section 3 describes a model of P2P reputation systems. Section 4 describes an incentive scheme proposed in this paper. Section 5 shows the simulation results. Finally, Section 6 concludes the paper with future works.

2. Related Work

How to give an incentive to the users to evaluate received services has been a main concern in realizing practical P2P reputation systems. Miura and Kawaura [10] focused on a knowledge-sharing community called Yahoo!Chiebukuro3 and analyzed the motivation of users to provide answers to given queries using a questionnaire survey. In the analysis, they classified the motivation of users into four types, i.e., 1) assistance motivation, 2) reciprocal motivation, 3) social motivation and 4) reward motivation, where each type respectively means: 1) the tendency of helping a questioner, 2) repaying a kindness in the past and expecting future benefits, 3) social meaning of answering behavior and 4) rewards acquired by answering.

Different from knowledge-sharing communities, in P2P reputation systems, the assistance motivation and the reciprocal motivation do not work well, since the target of assistance is not clear in P2Ps and it has a strong anonymity in nature. In fact, existing incentive mechanisms designed for P2Ps promote the reward motivation to encourage peers to reciprocally cooperate. For example, in eMule [4], a higher priority is given to a downloder if it shares a large number of files with the other peers, and BitTorrent [1] adopts the Tit-for-Tat strategy in such a way that a peer which uploaded chunks to another peer will be granted a right to download chunks from that peer.

In this paper, we will focus on the reward motivation similar to existing works. However, unlike conventional schemes which use the provisioning of resources such as an upload bandwidth and shared files as a concrete contribution, in our scheme, we will use the “evaluation of evaluations” as the source of contributions. This enables the participants to make a contribution to the system much easier than conventional schemes which are merely based on the resource provisionings. Note that the notion of evaluation of evaluations has already been proposed in the literature in a slightly different context [13]. The main difference to our scheme is that in the previous scheme, the evaluation of evaluations called feedback reputation is linked with the evaluation of services so that the feedback reputation concerned with a service is automatically notified to all peers which evaluated the service in the past, when (and only when) a peer makes the evaluation of the service. In contrast, in our scheme, we explicitly separate the evaluation of evaluations and the evaluation of service so as to increase the chance of contributing to the system as an evaluator.

3. Model

3.1 Model of P2P

In this paper, we consider a model of P2P consisting of a tracker and a set of peers. The set of peers might contain a malicious peer, but it must follow the protocol described below (more concretely, the only parameter controlled by a malicious peer is services and evaluations). The set of peers is divided into two subsets, i.e., a set of Service Providers (SPs) and a set of Client Peers (CPs), where the intersection of those subsets is not empty in general. Each CP can receive a service from an SP, whereas each SP can provide several services to CPs. The tracker issues (virtual) points to CPs in reward for an evaluation. Such a behavior of the tracker is realized either by using a virtual server as in hybrid P2Ps or by using a secure authentication chain proposed in the literature [3].

The quality of a service is evaluated by CPs by describing a survey report, where the evaluator of a service must be a recipient of the service. In the following, we call such an evaluation of services a qualification. Each qualification is associated with a real number in [0, 1] called qualification value, where value 1 indicates that the quality of the service is absolutely high and value 0 indicates that the quality of the service is absolutely low. We use symbol \( Q_{p,s} \) to denote the value of a qualification of service \( s \) evaluated by CP \( p \).

In addition to the qualification of services, in our model, each CP can evaluate qualifications given by other CPs by designating a real number in [0, 1], where similar to the case of qualifications, the evaluator of a qualification must be a recipient of the corresponding service. In the following, we call such an evaluation of qualifications a vote. In a vote, value 1 indicates that the evaluator completely agrees.
with the qualification. We use symbol \( V_{p,q} \) to denote the value associated with a vote which is given by CP \( p \) on a qualification of service \( s \) conducted by CP \( q \).

The reader should note that in the above framework, the target of evaluation is either a service or a qualification, and it does not directly measure up the SP who provided the service nor the CP who provided a qualification. This reflects a natural insight such that a good player might not always exhibit a good performance.

### 3.2 Model of Evaluation

Next, we describe the model of evaluation conducted by each node. The parameter controlled by each evaluator is the cost of evaluation as well as its opinion on the target, and the resulting quality of evaluations is determined by the cost and several parameters while we assume that the evaluator is not aware of those parameters. Let \( c_s \) be a local variable representing the cost of qualification. Before conducting a qualification, each CP sets a value in range \([c_{\min}, c_{\max}]\) to its local variable \( c_s \), which indicates that “how much effort will be necessary to complete the qualification.” The quality of qualification monotonically increases as \( c_s \) increases, and increases in proportion to the skill of the evaluator represented by a real number in range \((0, 1]\) (the skill of evaluator is not disclosed to any CP including the evaluator itself). More precisely, we assume that the quality of qualification is defined as follows:

\[
\Phi(c_s, \sigma) := \frac{c_s}{c_{\max}}^k \times \sigma \tag{1}
\]

where \( \sigma \) denotes the skill of the evaluator and \( k \) is an appropriate parameter greater than one. The reader should note that \( \Phi \) is a convex function with respect to \( c_s \), which is intended to model a situation in which less effort leads to much worse quality. In addition, the reader should remind that the quality of qualification is independent of the qualification value \( Q_{p,s} \) given by the evaluator.

Let \( c_r \) denote the cost of voting. In the following, we assume that \( c_r \) takes a fixed value smaller than or equal to \( c_{\min} \). This definition reflects an intuition such that a vote simply judges whether a given qualification is useful and match its own opinion concerned with the corresponding service. More specifically, the value of \( V_{p,q} \) is determined by the evaluation of the qualification value \( Q_{p,s} \) and the evaluation of the quality of qualification given by Equation (1). A detailed model of such evaluations used in simulations is given in the Appendix. The key idea behind the model is that each user in the real world cannot be completely objective in providing a vote, since her vote on a qualification should be biased by the closeness of the qualification to her opinion.

A list of parameters used in our model is summarized in Table 1.

### 4. Proposed Scheme

#### 4.1 Incentive Scheme

A collection of qualifications “approximates” the actual quality of the corresponding services. However, it might contain malicious qualifications which intentionally provide wrong values to illegally control the “reputation” of the corresponding services. Although the impact of such malicious qualifications could be reduced by increasing the number of collected qualifications, it is difficult to collect many qualifications since the cost of qualification is generally large. In order to overcome such a problem, our model adopted the notion of voting. In other words, if we collected a sufficient number of votes for each qualification, we could evaluate the trustworthiness of qualifications and accurately evaluate the quality of services without increasing the number of qualifications.

In this paper, we propose a point-based scheme to encourage voting and qualification. The proposed incentive scheme, which is based on the notion of **evaluation point** and **contribution point**, is described as follows (see Figure 1 for illustration):

- By conducting a qualification, the CP receives \( P_s \) evaluation points from the tracker, and
- When CP \( p \) votes for a qualification given by CP \( q \), 1) \( p \) receives \( P_r \) evaluation points from \( q \) and 2) \( q \) receives \( P_c \) contribution points from the tracker, where
  - if the evaluation points possessed by \( q \) are less than \( P_r \), then any CP cannot vote for the qualifications given by \( q \),
  - \( P_r \) is fixed to satisfy inequality \( P_r \leq \left( \frac{c_e}{c_{\max}} \right) \times P_s \) in order to encourage qualification rather than voting, and
  - \( P_c \) equals to \( V_{p,q} \in [0, 1] \), i.e., as the value associated with the vote increases, the contribution point received from the tracker increases.

#### 4.2 Service Differentiation

Service differentiation is a common technique used in many incentive schemes. In the proposed scheme, we realize
a service differentiation using contribution points. More precisely, by paying contribution point to an SP, the CP is granted to receive a high quality service from the SP. Since it can receive contribution points from the tracker equivalent to the value associated to the acquired votes, it works as an incentive to give a high quality qualification which attracts many votes with a high evaluation value. In addition, the notion of evaluation points works as another incentive to give a vote for existing qualifications, although it is not directly connected to the service differentiation. In fact, in order to acquire many votes from evaluators, it must have enough evaluation points which can be earned only through: 1) the issue of a qualification on a service, or 2) a vote for a qualification given by another CP.

4.3 Lifetime of Evaluation Points

In the above point-based scheme, evaluation points can be infinitely provided to the system by the tracker. Thus, to avoid an “inflation” of the evaluation points which reduces the relative value of the evaluation points, we introduce the notion of lifetime to the evaluation points, so that a point is expired from the system if the lifetime of the point becomes zero. The lifetime is set to $L$ at the time of provisioning by the tracker and is linearly decreased as the elapsed time increases, while it is “reset” to $L$ when it is transferred to another CP as a reward of voting. The notion of lifetime has an important side effect such that CPs should continuously conduct evaluations to keep a sufficient amount of evaluation points. Such an effect of lifetime will be evaluated by simulation in the next section.

5. Simulation

5.1 Setup

We conducted simulations to evaluate the performance of the proposed scheme. In the simulation, we consider a P2P consisting of a single SP and several CPs, where the number of services provided by the SP is fixed to 150. The reader should note that the SP models a collection of independent SPs, i.e., we do not consider a situation in which several SPs simultaneously provide services to a CP in a collaborative manner. The simulation time is divided into 500 intervals called time steps, and in each time step, each CP conducts one of the following three actions:

QUAL: Randomly select a service and receive it. After that, conduct a qualification of the service by spending a cost selected from $[c_{\text{min}}, c_{\text{max}}] := [1, 5]$.

VOTE: Randomly select a qualification of a service which has been received by the CP, and vote for it.

NONE: Randomly select a service and receive it, but no qualification is conducted.

With respect to the way of contribution to the system, we assume that CPs are classified into the following four types: A CP is said to be of Type 1 (resp. 2 and 3), if it prefers to action QUAL (resp. VOTE and NONE) and the credibility of the CP is high (i.e., 0.9), where the credibility of each service is randomly selected from range $[0, 1]$ and a model of evaluations which takes into account the credibility of evaluations is given in the Appendix. A CP is said to be of Type 4 if it prefers to action QUAL but the credibility of the CP is low (i.e., 0.0). More detailed specification of the setting

| Table 2: The probability of selecting each action. |
|----------------|----------------|----------------|
| QUAL           | VOTE           | NONE           |
| Type1          | 80%            | 10%            | 10%            |
| Type2          | 10%            | 80%            | 10%            |
| Type3          | 10%            | 10%            | 80%            |
| Type4          | 80%            | 10%            | 10%            |

When $P < 30$.

<table>
<thead>
<tr>
<th>QUAL</th>
<th>VOTE</th>
<th>NONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type1</td>
<td>40%</td>
<td>10%</td>
</tr>
<tr>
<td>Type2</td>
<td>10%</td>
<td>80%</td>
</tr>
<tr>
<td>Type3</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Type4</td>
<td>40%</td>
<td>10%</td>
</tr>
</tbody>
</table>

When $P \geq 30$. 
used in the simulations is summarized in Table 2. As shown in the table, we assume that the behavior of a CP depends on the evaluation points $P$ possessed by the CP. Namely, if $P$ is smaller than a threshold, which is fixed to 30 in the simulation, it follows probabilities shown in the left table and otherwise, it follows probabilities shown in the right table. It should be noted that the probability of choosing action $QUAL$ by a CP of Type 1 or 4 becomes small if it possesses enough evaluation points. This reflects a natural intuition such that the incentive to conduct qualifications should become weak if it possesses enough evaluation points.

Under the above settings, we simulated the behavior of CPs and observed earned contribution points and the cost required for the evaluations. The other common parameters are fixed as follows: 1) the number of CPs is 30 for each type, 2) the amount of evaluation points earned through evaluations are $P_s = 5$ and $P_r = 1$, and 3) the lifetime of each evaluation point is 20 time steps.

### 5.2 Comparison by Types

At first, we evaluate the behavior of CPs by their types. Parameters used in Equation (1) are fixed as $c_s = 5$, $\sigma = 1.0$, and $k = 2.0$.

Figure 2 illustrates the time transition of the average contribution points earned from the tracker, where the horizontal axis is the elapsed time and each curve is associated with a type of CPs. CPs of Types 1 and 4 earn many contribution points over time, since they conduct more $QUAL$’s than the other two types. Although there is a difference between Types 1 and 4 which is due to the difference of the credibility of CPs, the influence of the credibility is limited. Such a trend slightly changes if we consider the evaluation cost. Figure 3 illustrates the time transition of contribution points per cost. While CPs of Type 3 earn only few contribution points, the contribution point per cost gradually approaches to Type 1 and eventually overtakes it. Such a high efficiency of Type 3 is due to the expiration of the evaluation points. See Table 3. This table summarizes that: 1) how many points are earned (the second column), 2) how many points are paid as a reward (the third column), 3) how many points are expired (the fourth column), and 4) how many points are possessed on average (the first column). Although CPs of Type 3 earn a small amount of evaluation points, the ratio of expired points to the earned points is 78% which is smaller than 87% of Type 2. Thus, although the number of qualifications is small and they are rarely chosen as the target for a vote in an early stage of the simulation, as the number of (accumulated) qualifications increases, the low probability of selecting $QUAL$ improves the efficiency of Type 3 with respect to the earned contribution points per cost.

The reader might think that CPs of Type 2 are “useless” since they exhibit the worst performance with respect to both of the above two metrics, but it is not true. They actually play a crucial role in the proposed scheme. Figure 4 shows the result of simulation without CPs of Type 2. The amount of contribution points earned by the CPs (of
Types 1, 3 and 4) significantly decreases compared with Figure 2, which indicates that under the proposed scheme, VOTE is necessary for all CPs to earn a sufficient amount of contribution points.

By the above observations, we can conclude that a rational strategy for CPs to quickly, efficiently earn contribution points is to repeat voting to keep the amount of evaluation points after conducting a certain number of qualifications. In practical situations, however, such a simple strategy is not enough since older qualification becomes less attractive to earn many votes with a high evaluation value. Thus, CPs should repeat such a strategy with an appropriate interval.

5.3 Impact of Qualification Cost

In the second simulation, we evaluated the impact of qualification cost $c_s$ to the contribution points, by considering a P2P consisting of a single SP and CPs of Types 1 and 4. Parameters are fixed as in the first simulation except for the qualification cost of CPs of Type 1. More specifically, we partition Type 1 into three Types such that a CP is said to be of Type 1A (resp. 1B and 1C) if it sets $c_s$ to 5 (resp. 3 and 1). We prepare 30 CPs for each of the above four types.

Figure 5 illustrates the time transition of contribution points earned by the CPs. The amount of contribution points monotonically increases as $c_s$ increases, since we are assuming that the quality of qualifications is a monotonic function of $c_s$ in Equation (1). In addition, CPs of Type 4 earn more contribution points than Types 1B and 1C, which indicates that the amount of contribution points is more sensitive to the spent qualification cost rather than the credibility of evaluators. This encourages CPs to provide high quality qualifications by spending more cost.

Figure 6 illustrates the time transition of the contribution points per cost. We can see that CPs of Type 1C earn contribution points much more efficiently than the other three types. The badness of Types 1A and 1B compared with Type 1C could be explained as follows. The first reason is that in the simulation, qualifications to be evaluated are randomly selected without considering the quality of qualifications. In other words, the increase of $c_s$ does not always increase the chance to be evaluated. The second reason is that the amount of earned contribution points was a concave function of the qualification cost. In other words, although CPs of Type 1A spent qualification cost which is five times as large as the qualification cost spent by Type 1C, the amount of earned contribution points is only the twice of Type 1C. To overcome such an inefficiency for CPs which try to provide high quality qualifications, we should introduce another mechanism so that the earned contribution point increases according to a concave function of the acquired vote values. This important and interesting matter is left as a future work.

6. Concluding Remarks

In this paper, we propose an incentive scheme for P2P reputation management systems focusing on the cost of evaluations. The proposed scheme is based on the notion of evaluation points which mediates between the evaluation of services and the evaluation of evaluations, and the notion of contribution points which is used for the service differentiation. The performance of the scheme is experimentally evaluated by simulation.

The topics for future work are listed as follows:

1) We need to evaluate the effectiveness of the proposed scheme under a more practical setting. The application of the scheme to existing reputation management systems would be a perspective way to do so.

2) We need to carefully examine the influence of the anonymity to the incentive scheme, since in actual reputation systems, the opinion of a big name strongly affects the reputation given by other normal users.

3) We need to evaluate the robustness of the scheme against malicious attacks, and try to increase the robustness without incurring additional cost to the users.

References

The RBA model recently proposed by Nishikawa and Fujita to model such a variance of qualification value of the corresponding service depends on applications. For example, in the case of video streaming, the criteria would include download speed, interruption of playback, picture quality and others. In addition, since it is purely subjective, it differs for each evaluator, and of course, it is generally different from the actual (i.e., objective) quality of the corresponding service. To model such a variance of qualification value \( Q_{p,s} \), we use the RBA model recently proposed by Nishikawa and Fujita [11]. More concretely, we assume that \( Q_{p,s} \) is a random value selected from the following range:

\[
\left[ \text{max}\{0, Q_{s} + t_{p}^{*} - 1\}, \text{min}\{1, Q_{s} - t_{p}^{*} + 1\} \right]
\]

where \( t_{p}^{*} \) is the credibility of evaluator \( p \) and \( Q_{s} \) is the actual quality of service \( s \). The credibility of an evaluator is a real number in \([0,1]\) which means the accuracy of evaluations conducted by the evaluator. The reader should note that parameter \( t_{p}^{*} \) is introduced only for the modeling of the behavior of the evaluator and the value of \( t_{p}^{*} \) is not disclosed to any CP including the evaluator.

The evaluation of the quality of qualifications, which is the first part of a vote, is conducted in a similar way. More concretely, we assume that it is a random value selected from the following range:

\[
\left[ \text{max}\{0, \Psi + t_{p}^{*} - 1\}, \text{min}\{1, \Psi - t_{p}^{*} + 1\} \right],
\]

where \( t_{p}^{*} \) is the credibility of the evaluator \( p \) and \( \Psi \) is the quality of the qualification calculated by Equation (1). The evaluation of value \( Q_{q,s} \), which is the second part of a vote, is conducted as follows. Let \( p \) be the evaluator and let \( E_{p,q}^{s} \) denote the outcome of the evaluation. The key idea of our model is to focus on the proximity between \( Q_{q,s} \) and \( p \)'s opinion on service \( s \). Such a proximity is used by the evaluator in providing a vote consciously or unconsciously. Let \( Q_{p,s} \) be the qualification value on service \( s \) which is given by \( p \) if it was conducted by \( p \). Then, the value of \( E_{p,q}^{s} \) is calculated as follows:

\[
E_{p,q}^{s} := 1 - \left| Q_{q,s} - \hat{Q}_{p,s} \right|.
\]

Finally, the value \( V_{p,s}^{q} \) associated to a vote for qualification \( Q_{q,s} \) is calculated by taking an average of the above two values.

Appendix

Each qualification given by CP \( p \) is associated with a real number \( Q_{p,s} \) representing the quality of the corresponding service \( s \), where the criterion for determining the value of \( Q_{p,s} \) depends on applications. For example, in the case of video streaming, the criteria would include download speed, interruption of playback, picture quality and others. In addition, since it is purely subjective, it differs for each evaluator, and of course, it is generally different from the actual (i.e., objective) quality of the corresponding service. To model such a variance of qualification value \( Q_{p,s} \), we use the RBA model recently proposed by Nishikawa and Fujita [11]. More concretely, we assume that \( Q_{p,s} \) is a random value selected from the following range:

\[
\left[ \text{max}\{0, Q_{s} + t_{p}^{*} - 1\}, \text{min}\{1, Q_{s} - t_{p}^{*} + 1\} \right]
\]

where \( t_{p}^{*} \) is the credibility of evaluator \( p \) and \( Q_{s} \) is the actual quality of service \( s \). The credibility of an evaluator is a real number in \([0,1]\) which means the accuracy of evaluations conducted by the evaluator. The reader should note that parameter \( t_{p}^{*} \) is introduced only for the modeling of the behavior of the evaluator and the value of \( t_{p}^{*} \) is not disclosed to any CP including the evaluator.