Supporting Privacy in Decentralized Additive Reputation Systems

Elan Pavlov, Jeffrey S. Rosenschein, and Zvi Topol

Hebrew University, Givat Ram, Jerusalem 91904, Israel,
{elan, jeff, zvit}@cs.huji.ac.il,
WWW home pages: http://www.cs.huji.ac.il/~{elan, jeff, zvit}

Abstract. Previous studies have been suggestive of the fact that reputation ratings may be provided in a strategic manner for reasons of reciprocation and retaliation, and therefore may not properly reflect the trustworthiness of rated parties. It thus appears that supporting privacy of feedback providers could improve the quality of their ratings. We argue that supporting perfect privacy in decentralized reputation systems is impossible, but as an alternative present three probabilistic schemes that support partial privacy. On the basis of these schemes, we offer three protocols that allow ratings to be privately provided with high probability in decentralized additive reputation systems.

1 Introduction

In recent years, reputation systems have emerged as a way to reduce the risk entailed in interactions among total strangers in electronic marketplaces. Such systems collect and aggregate feedback about past behavior of participants in electronic transactions, so as to derive reputation scores assumed to predict likely future behavior.

Centralized reputation systems, such as the system in use by the electronic auction site eBay [1], collect and store reputation ratings from feedback providers in a centralized reputation database. These ratings are then processed to produce a publicly available reputation measure that can be obtained by querying the database. In eBay, for example, both buyers and sellers participating in a transaction may provide one of three possible feedbacks: positive (+1), neutral (0), and negative (-1). The reputation score of a user is simply the sum of his accumulated ratings over a period of six months.

Decentralized reputation systems, on the other hand, do not make use of a central repository to collect and report reputation ratings [2]. In this type of system, participants help one another with the provision of reputation ratings in order to evaluate the trustworthiness of potential transaction partners. Each participant is responsible for his own local repository of reputation through the collection and propagation of feedback when needed.

One concern about reputation systems (which has received relatively little attention in the trust and reputation management literature), is that of feedback providers' privacy. An empirical study conducted by Resnick *et al.* [3] on data sets extracted from eBay's reputation system reported a high correlation between buyer and seller ratings. Moreover, more than 99% of the feedback provided was positive.

This might be due to the fact that mutually satisfying transactions are simply the (overwhelming) norm. However, it might also be the case that when feedback providers' identities are publicly known, reputation ratings can be provided in a strategic manner for reasons of reciprocation and retaliation, not properly reflecting the trustworthiness of the rated parties. For example, a user may have an incentive to provide a high rating because he expects the user he rates to reciprocate, and provide a high rating for either the current interaction or possible future ones.

This type of strategic manipulation in the process of feedback provision is likely to occur also in decentralized reputation systems. There too, agents providing feedback would like to ensure that the ratings they provide cannot be abused by malicious agents in a way that can affect them negatively in the future. An example of such malicious behavior might occur if individual ratings were first reported to the rated agent, who can then retaliate or reciprocate on his turn (when he is given an opportunity to rate the feedback providers).

The logic of anonymous feedback to a reputation system is thus analogous to the logic of anonymous voting in a political system. It potentially encourages truthfulness by guaranteeing secrecy and freedom from explicit or implicit influence. Although this freedom might be exploited by dishonest feedback providers, who tend to report exaggerated feedbacks, it seems highly beneficial for honest ones, protecting the latter from being influenced by strategic manipulation issues as described above.

1.1 Structure of Paper

The rest of the paper is organized as follows. Section 2 describes the problem setting with which we are dealing, while Section 3 presents the notion of Decentralized Additive Reputation Systems and gives an example of one — the Beta Reputation system. Section 4 proves an impossibility result and suggests methods of partially circumventing it. Section 5 then suggests three protocols achieving probabilistic privacy in decentralized additive reputation systems. Section 6 surveys related work, and Section 7 concludes by summarizing our results and suggesting directions for future research.

2 Problem Setting

We assume that each user in the system is represented by an *agent*, which performs necessary computations and communication activities with other agents, on behalf of the user. We also assume authenticated, secure channels between every two users. Such channels can be achieved via standard technologies such as SSL (Secure Sockets Layer).

We are concerned with the following problem: a *querying agent* A_q has to decide whether to interact with a potential partner, the *target agent* A_t. A_q has incomplete information about A_t. It either has no prior knowledge about A_t's past behavior at all, since both agents do not have a common history of interactions, or its experience with A_t is too limited or outdated, so that it cannot derive a meaningful reputation measure regarding the trustworthiness of the target agent.

In a decentralized reputation system, A_q consults a group of agents, or *witnesses*, $\{W_1, W_2, ..., W_n\}$, considered to have a reputation score regarding A_t. One way to obtain such a set of witnesses is through a series of referrals from agents residing in the same social network of A_t (see [2] for further details about how to obtain such a set of witnesses). We denote the reputation rating of witness i by r_i. Although r_i is generally represented by a vector of finite dimension (measuring reputation over different contexts of interest), we will assume without loss of generality throughout the paper that r_i is a scalar. We are interested in a method assuring that whenever feedbacks received from the witnesses are combined in an additive manner, their privacy is properly maintained, i.e., feedbacks are not revealed to any other agent in the system, nor to possible third parties.

We divide agents participating in the feedback provision process into two types: *curious but non-malicious agents* (which we call "curious agents") and *malicious agents*. Curious agents follow the protocol; that is, curious witnesses provide honest feedback about the target agent, and do not try to interfere with the correct flow of the protocol in order to change or corrupt the result obtained at the end of the process (the combined reputation rating). The main concern about such agents is that they might try to reveal reputation ratings in different ways, including collusion with other agents.

Malicious agents, on the other hand, might try to actually tamper with the protocols, provide dishonest feedback in order to bias the combined reputation rating according to their interests, or even render the resulting rating unusable.

In our scenario, the querying agent can act only as a curious agent. Clearly, it would not be in its interest to interfere with the rating calculation in any way. An example of a querying agent acting curiously would be if the target agent itself masquerades as a querying agent in order to reveal the reputation ratings of witnesses.

3 Decentralized Additive Reputation Systems

We here define Decentralized Additive Reputation Systems, and follow with an example of such a reputation system, the Beta Reputation system.[1]

Definition 1. *Reputation System R is said to be a Decentralized Additive Reputation System if it satisfies two requirements:*

1. *Feedback collection, combination, and propagation are implemented in a decentralized way.*
2. *Combination of feedbacks provided by agents is calculated in an additive manner.*

The Beta Reputation system presented in [4] and described in the next subsection is an example of a reputation system satisfying both requirements. eBay's reputation system, on the other hand, satisfies only the second requirement, i.e., it is additive but centralized.

[1] Our approach in this paper is broadly applicable to Decentralized Additive Reputation Systems, but we specifically present the Beta Reputation system as one example.

3.1 The Beta Reputation System

The Beta Reputation system is based on the beta-family of probability density functions which are typically used to represent a posteriori probability distributions of binary events. The beta functions are continuous functions of the form $f(p|a,b)$ which can be expressed as:

$$f(p|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{(a-1)}(1-p)^{(b-1)} \tag{1}$$

where Γ is the gamma function, a generalization of the factorial function to real values, $0 \leq p \leq 1$, $a > 0$, $b > 0$, $p \neq 0$ if $a < 1$ and $p \neq 1$ if $b < 1$. The expectation of the beta distribution can be shown to be:

$$E(p) = \frac{a}{a+b} \tag{2}$$

Given a binary stochastic process with two possible outcomes $\{o_1, o_2\}$, the probability p of observing o_1 in the future as a function of past observations of r_1 instances of o_1 and r_2 instances of o_2 is given by: a = r_1 + 1, b = r_2 + 1, where $r_1 \geq 0$ and $r_2 \geq 0$. The expectation can now be written as:

$$E(p) = \frac{r_1 + 1}{r_1 + r_2 + 2} \tag{3}$$

Letting o_1 be a positive outcome of an interaction between two agents and o_2 be a negative one from the point of view of the rating agent, r_1 and r_2 could be seen as the degree of satisfaction and dissatisfaction respectively. Since the agent's satisfaction after a transaction is not necessarily binary, (r_1, r_2) is represented as a pair of continuous values. The expectation value is then defined to be the reputation rating about the target agent:

$$Rep(r_1, r_2) = \frac{r_1 + 1}{r_1 + r_2 + 2} \tag{4}$$

Let A_t be the target agent and let A_1 and A_2 be two agents that interacted with A_t in the past. Let $Rep^1(r_1^1, r_2^1)$ be A_1's reputation rating about A_t and let $Rep^2(r_1^2, r_2^2)$ be the reputation rating of A_2. The combined reputation value is then obtained by calculating:

$$r_1^* = r_1^1 + r_1^2 \tag{5}$$

$$r_2^* = r_2^1 + r_2^2 \tag{6}$$

and plugging the results into (4), to obtain $Rep^*(r_1^*, r_2^*)$. This additive property of the Beta Reputation system, which is both commutative and associative, could be generalized to any number of agents.

4 Witness Selection

An inherent problem with decentralized reputation systems is the collusion of $n - 1$ witnesses along with a dishonest (either curious or malicious) querying agent in order to reveal the reputation information of an honest witness. The querying agent can *choose*

$n-1$ dishonest agents and a single honest agent. If the function calculating reputation is reversible, then there is no protocol that can anonymously calculate reputation. This yields the following lemma:

Lemma 1. *For a reversible reputation function F that accepts n witnesses and outputs a reputation, if there are $n-1$ dishonest witnesses, there is no protocol that deterministically anonymously calculates reputation.*

Proof. For any protocol there might be $n-1$ dishonest witnesses and one honest one. If the querying agent is malicious then he can create such a set deterministically. Thus, collusion between the n dishonest agents would expose the reputation score of the honest witness.

To circumvent this inherent limitation, we look at probabilistic methods of ensuring that there is a large number of honest witnesses.

Lemma 2. *Let $N > 1$ be the number of potential witnesses and let $n > 0$, $n < N$ be the number of witnesses participating in the process. Let $b < N$ be the number of dishonest agents in N. If honest agents are uniformly distributed over N, then there exists a witness selection scheme that guarantees at least two honest witnesses with probability greater than $(1-\frac{1}{n})(\frac{N-b-1}{N-1})$.*

Proof. Consider the following witness selection scheme: A_q chooses the first witness W_1. Each witness chosen, with probability $1-\frac{1}{n}$, chooses another witness to participate in the feedback collection process and with probability $\frac{1}{n}$ does not invite additional witnesses. At some point, an honest witness is chosen. Let W_h be the first honest witness to be chosen. If b' dishonest witnesses were chosen before W_h, then W_h chooses another honest witness with probability $P_r \geq (1-\frac{1}{n})(\frac{N-b-1}{N-b'-1}) \geq (1-\frac{1}{n})(\frac{N-b-1}{N-1})$.

Similar witness selection schemes can be implemented using protocols for leader selection resilient to linear size coalitions, such as the one described in [5]. Witness selection is equivalent to leader selection; thus, n witnesses are selected by n activations of the leader selection protocol. It is also possible to use the same instance of the protocol to select more than one witness.

Sometimes it is not enough to ensure that there is a large number of honest witnesses in the group; we might also need to make sure that there is a predefined proportion between the size of the group and the number of honest witnesses in it, as in the case of Section 5.3. This is achieved by the following lemma, provided that A_q is honest.

Lemma 3. *Let $N > 0$ be the number of potential witnesses and let $n > 0$, $n < N$ be the number of witnesses participating in the process. Let $b < N$ be the number of dishonest agents in N. If honest agents are uniformly distributed over N, then there exists a witness selection scheme that guarantees at least $n(\frac{N-b-n}{N})$ honest witnesses in the group of witnesses participating in the process, with high probability.*

Proof. Consider the following witness selection scheme: A_q chooses the first witness W_1. At this point, the size of the group of witnesses participating in the process k is 2. Given a group of size k, the agents in the group collectively flip a weighted coin in order to decide whether to extend the group. With probability $1 - \frac{1}{n}$ they choose at random another agent from N to join the group, and with probability $\frac{1}{n}$ they stop. The expected number of coin tosses until the group stops is n. At each coin toss, the probability of choosing an honest witness to join the group is greater than $\frac{N-b-n}{N}$; thus, the expected number of honest witnesses in the group is greater than $n(\frac{N-b-n}{N})$. If we denote $\mu = n(\frac{N-b-n}{N})$, then by Chernoff bounds (see for example [6]), the probability that the number of honest witnesses is substantially smaller than μ, namely $(1-\delta)\mu$, is less than $e^{-\frac{\mu\delta^2}{2}}$.

This type of collective coin flipping scheme can be implemented as follows: the agents agree on value v, $0 \leq v \leq n$. Every agent i chooses at random and independently $log_2(n)$ bits, x_i, and sends them to the other agents in the group. Each agent calculates $x = x_1 \oplus x_2 \oplus \ldots \oplus x_n$. If $x = v$ the agents stop, otherwise the agents continue. The decision about which new witness is to join the group could be rendered random in a similar way. Note that if at least one honest witness is present, then the value of x is guaranteed to be random. This scheme requires $\sum_{k=2}^{n+1} k^2 = O(n^3)$ messages among the agents.

5 Privacy in Decentralized Additive Reputation Systems

In this section, we present three different protocols achieving privacy in Decentralized Additive Reputation Systems. The basic idea behind the protocols is to consider the feedback provided by each witness to be his private information, or *secret*. The sum of secrets represents the combined reputation rating, and should be constructed without revealing the secrets.

5.1 Towards Achieving Privacy

One protocol achieving privacy in the presence of curious but non-malicious agents is the following:

1. Initialization Step: the querying agent, A_q, orders the agents in a circle: $A_q \rightarrow W_1 \rightarrow W_2 \rightarrow \ldots \rightarrow W_n \rightarrow A_q$ and sends each witness i the identity of his successor in the circle, i.e., witness $i+1$. W_n is sent the identity of A_q.
2. A_q chooses $r_q \neq 0$ at random and sends it to W_1.
3. Upon reception of r_p from his predecessor in the circle, each agent W_i $i = 1 \ldots n$ calculates $r_p + r_i$, where r_i is the reputation score of W_i about the target agent, and sends it to his successor in the circle.
4. Upon reception of the feedback from W_n, A_q subtracts r_q from it and plugs the result into the additive reputation system engine, that calculates the combined reputation rating.

Lemma 4. *If agents do not collude, then at the end of the protocol the querying agent obtains the sum of the feedbacks, such that feedbacks are not revealed to any of the agents.*

Proof. Every witness i adds in stage 3 his reputation rating to the number he previously received from his predecessor in the circle, so W_n sends to A_q the sum $\sum_{i=1}^{n}(r_i) + r_q$. Therefore, in stage 4, when A_q subtracts from this sum his random number r_q, he obtains the sum of the feedbacks. The random number r_q that A_q contributes at stage 2 masks the feedback provided by W_1, as it is different from zero, so W_2 doesn't reveal it. From this point in the protocol, no agent can guess any of the feedbacks provided by his predecessors.

If we consider transmissions of r_p between two adjacent agents in the circle as a single message, we can see that in this scheme $O(n)$ messages are passed among the agents.

A prominent drawback of this approach is its lack of resilience to collusion among agents. Two witnesses, W_{i-1} and W_{i+1}, $i = 2 \ldots n-1$, separated by a single link in the circle, namely W_i, could collude against W_i and reveal its private information, i.e., his feedback, by subtracting the rating transmitted by W_{i-1} from the one transmitted to W_{i+1}.

In the following subsections we will provide a way to overcome this vulnerability through the description of two protocols resilient to collusion of up to $n-1$ witnesses with high probability.

5.2 Privacy Through Secret Splitting

In this subsection, we present a simple protocol that provides privacy for curious agents, yet is resilient with high probability to collusion of up to $n-1$ agents, if witnesses are selected as described in the first witness selection scheme proposed in Section 4.

1. Initialization Step: A_q sends to the witnesses $\{W_1, ..., W_n\}$ the details of all agents participating in the process, i.e., identities of the n witnesses and itself, and chooses r_q at random.
2. Each of the $n+1$ agents participating in the protocol splits its secret, i.e., its reputation score, into $n+1$ shares in the following way: agent i chooses n random numbers $s_{i,1}, ..., s_{i,n}$, and calculates $s_i = r_i - \sum_{k=1}^{n}(s_{i,k})$. He keeps s_i and sends $s_{i,1}, ..., s_{i,n}$ to the n other agents, such that each agent j receives share $s_{i,j}$.
3. Each agent j calculates $val_j = \sum_{i=1}^{n}(s_{i,j}) + s_j$, and sends val_j to the querying agent.
4. The querying agent calculates, upon reception of val_i $i = 1 \ldots n$ from the n witnesses, $r = \sum_{j=1}^{n+1}(val_j) - r_q$ and provides r to the reputation engine.

Lemma 5. *If the agents participating in the protocol are curious, then at the end of the last stage, the querying agent obtains the sum of the feedbacks, such that feedbacks are not revealed to any of the agents with probability greater than* $(1 - \frac{1}{n})(\frac{N-b-1}{N-1})$.

Proof. At stage 2 of the protocol, each agent i distributes n random shares, but keeps in private a share s_i, that along with the distributed shares uniquely defines his secret. At stage 3, each agent sums his private share along with n random numbers he receives from the other agents, masking his private share, such that when he sends this sum to the querying agent, his private share cannot be revealed, unless the other $n - 1$ witnesses and the querying agent form a coalition against him. The latter case occurs with probability less than $1 - (1 - \frac{1}{n})(\frac{N-b-1}{N-1})$, if agents are self-ordered as suggested in the first witness selection scheme in Section 4. At stage 4, the querying agent calculates: $r = \sum_{j=1}^{n+1}(val_j) - r_q = \sum_{j=1}^{n+1}(\sum_{i=1}^{n}(s_{i,j}) + s_j) - r_q = \sum_{j=1}^{n+1}(r_j - s_j + s_j) - r_q = \sum_{j=1}^{n} r_j$ and thus obtains the sum of feedbacks.

This protocol requires $O(n^2)$ messages among the agents participating in the process, as opposed to $O(n)$ messages in the protocol from the previous subsection. On the other hand, the current protocol is resilient against collusion of up to $n - 1$ agents with high probability.

This protocol works well in the presence of curious agents, but malicious agents can tamper with it in various ways. A simple yet effective attack is the provision of reputation ratings out of range, such that the resulting reputation score is affected in an extreme way or is even rendered unusable. For example, if the reputation ratings should be positive integers in the range [1, 100] and there are 5 witnesses, one of the witnesses providing a reputation rating of 500 renders the resulting sum greater than 500, hence unusable. The following subsection presents another protocol that ensures that the provided reputation ratings lie within a predefined range.

5.3 Achieving Privacy Using Verifiable Secret Sharing

In this subsection, we suggest a protocol that achieves privacy in Decentralized Additive Reputation Systems, resilient with high probability to collusion of up to $n - 1$ curious agents participating in the process, and supports validity checking of the feedback provided. We use the Pederson Verifiable Secret Sharing scheme [7], which is based on Shamir Secret Sharing [8] and a discrete-log commitment method, in a manner similar to what is described in [9]. Both the Shamir Secret Sharing scheme and the discrete logarithm commitment are homomorphic in nature, making them suitable building blocks to use with additive reputation systems.

One of the properties of the Shamir Secret Sharing scheme is its resilience to up to $n/2$ malicious agents. Thus, the presence of more than $n/2$ such agents might be problematic for an honest querying-agent. If witnesses are selected as described in the second witness selection scheme proposed in Section 4 and if $b < \frac{N}{2} - n$, then with high probability, there are less than $n/2$ malicious agents.

For the purpose of this protocol, we assume that the reputation rating provided by W_i, r_i, is an integer in the group G_q of prime order q. The protocol is as follows:

1. Initialization Step: A_q selects a group G_q of a large prime order q with generators g and h, where $log_g h$ is hard to find. He sends to the witnesses $\{W_1, \ldots, W_n\}$, g and h and the details of all agents participating in the process, i.e., the n witnesses and itself.
2. Witness i chooses two polynomials of degree n: $p^i(x) = p_0^i + p_1^i x + p_2^i x^2 + \ldots + p_n^i x^n$ and $q^i(x) = q_0^i + q_1^i x + q_2^i x^2 + \ldots + q_n^i x^n$. The witness then sets r_i as p_0^i. The other coefficients of the polynomials are chosen at random uniformly from G_q.

3. W_i sends to each agent $j, j = 1, \ldots, i-1, i+1, \ldots, n+1$, from the set $\{W_1, \ldots, W_{i-1}, W_{i+1}, \ldots, W_n, A_q\}$ the point j on his polynomials, i.e., $p^i(j)$ and $q^i(j)$ along with commitments on the coefficients of its polynomials of the form: $g^{p^i_0} h^{q^i_0}, \ldots, g^{p^i_n} h^{q^i_n}$.
4. Witness m, upon reception of $p^1(m), p^2(m), \ldots, p^{m-1}(m), p^{m+1}(m), \ldots, p^n(m)$ and $q^1(m), q^2(m), \ldots, q^{m-1}(m), q^{m+1}(m), \ldots, q^n(m)$, calculates $p^m(m)$, $q^m(m)$, $s_m = \sum_{i=1}^{n} p^i(m)$ and $t_m = \sum_{i=1}^{n} q^i(m)$, and sends s_m and t_m to A_q. A_q calculates $s_{n+1} = \sum_{i=1}^{n} p^i(n+1)$ and $t_{n+1} = \sum_{i=1}^{n} q^i(n+1)$.
5. Upon reception of $s_1, \ldots, s_n$ and $t_1, \ldots t_n$, A_q obtains $s(0)$, the reputation rating, where $s(x) = \sum_{i=1}^{n} p^i(x)$ in the following manner: it computes $\sum_{i=1}^{n+1} s_i L_i(0)$, where $L_i(0)$ is the Lagrange polynomial at 0, and in this case could be expressed by: $L_i(0) = \mathit{\Pi}_{j=1, j \neq i}^{n+1} \frac{j}{j-i}$.

At the end of the last stage of the protocol, A_q holds the sum of the reputation ratings provided, as required. At stages 4 and 5, agents can verify that the shares they received from the other agents are valid using the homomorphic property of the commitments received at the end of stage 3. Complaints about invalid shares may be resolved by the accused agent sending the disputed point on the polynomial to A_q, since A_q cannot use it to reconstruct his secret.

For stage 3 we need a practical zero knowledge proof for the validity of the reputation ratings to be conducted between the witnesses and the querying agent; such a proof is provided, e.g., by [9].

This protocol requires $O(n^3)$ messages to be passed among the agents (due to the witness selection scheme) and does not reveal the reputation ratings of the witnesses involved since no less than $n+1$ different points on a polynomial of degree n are required for interpolation. It also requires linear work on the part of the agents.

6 Related Work

Much research concerning trust and reputation management has been conducted in recent years. Researchers have suggested different models of trust and reputation, both for centralized and decentralized systems. Most of the work on decentralized reputation systems, including [10–12], focus on efficient algorithms for distributed storage, collection and aggregation of feedbacks, but not on manipulative feedback provision.

Bin and Singh [2] propose a distributed reputation management system, where trust is modelled based on the Dempster-Shafer theory of evidence. In [13], they suggest a method for detection of deceptive feedback provision in their system, by applying a weighted majority algorithm adapted to belief functions. It is not clear, however, that their suggested scheme is efficient against wide-scale reciprocation and retaliation in the feedback provision process.

Dellarocas suggests in [14] a collaborative filtering-based method to deal with the problem of unfair ratings in reputation systems. His method is applicable to centralized reputation systems. It is not clear whether this method could be efficiently applied in the decentralized case.

There has been little work on privacy and anonymity concerns related to reputation management systems. Ismail *et al.* [15, 16] propose a security architecture based on electronic cash technology and designated verifier proofs. Their suggested architecture is targeted at centralized reputation systems and does not seem suitable for decentralized systems, on which we focus our attention.

Kinateder and Pearson [17] suggest a privacy-enhanced peer-to-peer reputation system on top of a *Trusted Computing Platform* (TCP); see [18] for more details on TCP. The platform's functionality along with the use of pseudonymous identities allow the platform to prove that it is a trusted platform, yet to conceal the real identity of the feedback provider. A possible privacy-breach in the IP layer is handled by the use of MIX cascades or anonymous web-posting. As opposed to our scheme, this approach is dependent on a specific platform, which is currently arousing controversy in the computing community. Further details on this issue can be found in [19].

7 Conclusions and Future Work

Decentralized reputation systems do not make use of a central repository to collect and report reputation ratings; participants help one another with the provision of reputation ratings in order to evaluate the trustworthiness of potential transaction partners. This kind of reputation system is a natural match for many kinds of distributed environments, including popular peer-to-peer systems. Systems are being used not only for content sharing (e.g., KaZaA, Gnutella), but for social and business interactions (e.g., Friendster, LinkedIn), classified advertising (e.g., Tribe Networks), and ecommerce (CraigsList), and while not all of these have a peer-to-peer architecture, they are all potentially modelled by peer-to-peer alternatives. Reliable distributed reputation systems in these settings would provide an important service to these communities.

Additive Reputation systems are those in which the combination of feedbacks provided by agents is calculated in an additive manner. They are a particular class of reputation systems with the attractive property of simplicity in the calculation of results.

In this paper, we have shown that there are limits to supporting perfect privacy in decentralized reputation systems. In particular, a scenario where $n-1$ dishonest witnesses collude with the querying agent to reveal the reputation rating of the remaining honest witness demonstrates that perfect privacy is not feasible. As an alternative, we have suggested a probabilistic scheme for witness selection to ensure that such a scenario occurs with small probability.

We have offered three protocols that allow ratings to be privately provided in decentralized additive reputation systems. The first protocol is not resilient against collusion of agents, yet is linear in communication and simple to implement, and might be used when dishonest witnesses are not an issue. The other two protocols are based on our probabilistic witness selection scheme, and are thus probabilistically resistant to collusion of up to $n-1$ witnesses. The second protocol achieves privacy through secret splitting and requires $O(n^2)$ messages among the agents. Its main drawback is its inability to ensure that ratings are provided correctly within the predefined range. The third protocol, based on Pederson Verifiable Secret Sharing, makes use of zero knowledge proofs to circumvent this vulnerability. It requires $O(n^3)$ messages among the agents and some computation on the part of the agents, compared to the second protocol.

In future work, we plan to study schemes and protocols achieving privacy in the general case, i.e., in decentralized reputation systems which are not necessarily additive. In addition, we plan to study other approaches to improve the feedback provided in reputation systems, such as through the design of mechanisms inducing agents to reveal their honest feedback. The combination of privacy and complementary mechanisms promoting truthful feedback revelation will make reputation systems more robust than ever. We believe that such reputation systems would provide solid ground for ecommerce to prosper.

References

1. eBay auction site: http://www.ebay.com (2003)
2. Yu, B., Singh, M.: Distributed reputation management for electronic commerce. Computational Intelligence **18** (2002) 535–549
3. Resnick, P., Zeckhauser, R.: Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. In: Working paper for the NBER Workshop on Emprical Studies of Electronic Commerce. (2000)
4. Josang, A., Ismail, R.: The beta reputation system. In: The Proceedings of the 15th Bled Conference on Electronic Commerce, Bled, Slovenia (2002)
5. Saks, M.: A robust noncryptographic protocol for collective coin flipping. SIAM Journal on Discrete Mathematics **2** (1989) 240–244
6. Motwani, R., Raghavan, P.: Randomized Algorithms. Cambridge University Press (1995)
7. Pederson, T.: Non-interactive and information secure veriable secret sharing. In: Advances in Cryptology - Crypto '91. (1991) 129–140
8. Shamir, A.: How to share a secret. Communications of the ACM **22** (1979) 612–613
9. R. Cramer, M. Franklin, L.S., Yung, M.: Multi-authority secret ballot elections with linear work. Technical Report CS-R9571, Centrum voor Wiskunde en Informatica (1995)
10. Aberer, K., Despotovic, Z.: Managing trust in a peer-2-peer information system. In: Proceedings of 9th International Conference on Information and Knowledge Management, Atlanta (2001)
11. Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: Proceedings of the 33rd Hawaii International Conference on System Sciences, Maui, Hawaii (2000)
12. Kinateder, M., Rothermel, K.: Architecture and algorithms for a distributed reputation system. In: Proceedings of the First International Conference on Trust Management, Crete, Greece (2003)
13. Yu, B., Singh, M.: Detecting deception in reputation management. In: Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems. (2003) 73–80
14. Dellarocas, C.: Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In: Proceedings of the 2nd ACM Conference on Electronic Commerce, Minneapolis, MN (2000)
15. Ismail, R., Boyd, C., Josang, A., Russel, S.: Strong privacy in reputation systems (preliminary version). In: the proceedings of WISA 2003. (2003)
16. Ismail, R., Boyd, C., Josang, A., Russel, S.: A security architecture for reputation systems. In: The Proceedings of EC-WEB 2003. (2003)
17. Kinateder, M., Pearson, S.: A privacy-enhanced peer-to-peer reputation system. In: Proceedings of the 4th International Conference on Electronic Commerce and Web Technologies (EC-Web 2003), Prague, Czech Republic (2003)
18. tcpa homepage: http://www.trustedcomputing.org (2003)
19. againsttcpa homepage: http://www.againsttcpa.com (2003)