Decoupling Service and Feedback Trust in a Peer-to-Peer Reputation System

Gayatri Swamynathan, Ben Y. Zhao and Kevin C. Almeroth

Department of Computer Science, UC Santa Barbara
{gayatri, ravenben, almeroth}@cs.ucsb.edu

Abstract. Reputation systems help peers decide whom to trust before undertaking a transaction. Conventional approaches to reputation-based trust modeling assume that peers reputed to provide trustworthy *service* are also likely to provide trustworthy *feedback*. By basing the credibility of a peer's feedback on its reputation as a transactor, these models become vulnerable to malicious nodes that provide good service to badmouth targeted nodes. We propose to decouple a peer's reputation as a *service provider* from its reputation as a *service recommender*, making the reputation more robust to malicious peers. We show via simulations that a decoupled approach greatly enhances the accuracy of reputations generated, resulting in fewer malicious transactions, false positives, and false negatives.

1 Introduction

The explosive growth in the Internet in the last decade has resulted in an increase in the use and popularity of online peer-to-peer (P2P) communities. P2P file sharing communities like Gnutella [9] involve millions of users who interact daily to transfer files among each other free of cost. The success of this type of a P2P community relies on cooperation amongst all the peers in the community. However, peers are anonymous and can act in their self-interests. This open and anonymous nature makes the network difficult to police and vulnerable to a variety of attacks.

A number of attacks can interfere with the operation of a P2P system. One common attack is the "whitewashing attack" where a free-riding node repeatedly joins the network under a new identity in order to avoid the penalties imposed on free-riders [8]. A more serious type of attack is when malicious peers exploit file sharing networks to distribute viruses and Trojan horses. The *VBS.Gnutella* worm, for example, stores trojan executables in network nodes. *Mandragore*, a Gnutella worm, registers itself as an active peer in the network, and in response to intercepted queries, provides a renamed copy of itself for download [5]. Peers also need to detect inauthentic file attacks, in which corrupted or blank files are passed off as legitimate files. Hence, it is necessary for P2P communities to combat these threats by motivating cooperation and honest participation within their network. Reputation systems help address this need by establishing a trust mechanism that helps peers decide whom to trust before undertaking a transaction.

A number of reputation systems have been proposed or deployed in practice. While systems like eBay use a centralized approach [7], a number of decentralized reputation systems encourage cooperation and punish malicious behavior. These systems, within the bounds of their assumptions, demonstrate the ability to significantly reduce the number of malicious transactions in a P2P system [5, 1, 3, 6, 10, 12].

A central challenge in building a reputation system is to make it robust to misleading or unfair feedback. Malicious peers can subvert the reputation system by assigning poor reputation ratings to honest peers and good ratings to other malicious peers. To cope with malicious feedback, most existing reputation systems incorporate into their trust model the notion of *correlated trust*: peers reputed to provide trustworthy service, in general, will likely provide trustworthy feedback. Consequently, in these models the credibility of a peer's *feedback* is weighed by its reputation as a *service provider*.

While useful as a simple defense against malicious ratings, the correlated trust assumption can easily fail or be manipulated. A peer providing honest service can be incentivized to give false feedback about other peers' service. Similarly, colluding malicious nodes can offer honest service for the express purpose of boosting their reputations so they can badmouth the peers they are attacking.

This paper offers three key contributions. First, we propose a peer-to-peer reputation system that increases robustness against fake and misleading feedback by decoupling service and feedback reputations. Second, we show via simulation how our reputation system drastically reduces the rate of malicious transactions in a P2P system. Finally, we compare our scheme against correlated trust models in existing reputation systems. Our simulations show that strategic peers can exploit correlated trust models to increase malicious transactions, false positives and false negatives in the system. Our decoupled reputation system significantly reduces all of these behaviors.

The remainder of the paper is organized as follows. Related work is discussed in Section 2. In Section 3, we discuss our decoupled trust model and present our reputation system. In Section 4, we present our simulation settings and performance evaluation. Finally, Section 5 concludes the paper with suggested future work.

2 Related Work

Reputation management involves several components, including trust modeling, data storage, communication and reputation safeguards. Most research efforts have focused on solving only specific reputation management issues such as reputation storage, communication or attack safeguards [5, 6, 12].

eBay, the largest person-to-person auction site, uses a reputation-based trust scheme where, after each transaction, buyers and sellers rate each other using the *Feedback Forum* [7]. Reputation profiles are designed to predict future performance and help users decide whom to transact with [13]. eBay, however, uses a central authority to manage all communication and coordination between peers, essentially eliminating much of the complexity present in decentralized systems.

Aberer and Despotovic propose a decentralized reputation system for P2P networks where data is stored on a P-Grid [1]. Their system assumes most network peers are honest, and reputations in the system are expressed as complaints. Though the method works well, it is not at all robust to dynamic peer personalities.

EigenTrust [10] is a reputation system for P2P networks that attempts to combat the spread of inauthentic files. Each peer is associated with a global trust value that reflects the experiences of all other peers in the network with the target peer. Peers use these trust values to choose who they download from, as a consequence, the community identifies and isolates malicious peers from the network. The limitation of EigenTrust is that it assumes the existence of pre-trusted peers in the network.

While the systems mentioned so far assume a correlation between service and feedback reputations, a few have actually developed separate metrics for evaluating *service* trust and *feedback* trust [14, 2]. PeerTrust [14] is a reputation framework that includes an adaptive trust model. To decouple feedback trust from service trust, peers use a *personalized similarity measure* to more heavily weigh opinions of peers who have provided similar ratings for a common set of past partners. In a large P2P system, however, finding a statistically significant set of such past partners is likely to be difficult. As a consequence, peers will often have to make choices among a set of candidates for which there is no information.

CONFIDANT [2] attacks the problem of false ratings using a Bayesian approach in a mobile ad-hoc network. They distinguish between reputation, how well a node behaves in routing and trust, and how well it behaves in the reputation system. A node distributes only first-hand information to other nodes, and only accepts other first-hand information if those opinions are similar (within a threshold) to its own opinion. Compared to this system where a node's referral is interpreted subjectively per node, our proposal produces a system-wide referrer rating per node. Our proposal is also generalizable to any environment using a reputation system.

Previous trust models do not provide a general model for decoupling service trust and feedback trust. In this paper, we propose a reputation system in which each peer is associated with two trust values: one for its role as a service provider in the P2P network, and the other for its role as a service recommender in the reputation system.

3 The Trust Model

Our reputation system associates with each peer two sets of reputation ratings: an aggregated service rating (*s-rating*) and an aggregated feedback rating (*f-rating*). Additionally, the system maintains for each peer a list of peers that has rated it and its rating. Service ratings are normalized values ranging from -1.0 to 1.0 with 0 indicating a neutral rating. Feedback ratings are normalized values that range from 0 to 1.0 with 1.0 indicating a good rater. Initially, the s-rating is set to 0, and the f-rating is set to 1.0 for all peers.

Consider a peer, A, that queries for a file. In order to make a decision on which responding peer to transact with, A chooses the peer with the highest aggregated service rating. While this can result in an unbalanced load distribution in the network, a probabilistic approach can be employed to distribute load [10]. After finishing a transaction with a service provider, B, A provides to B either a rating of -1 (unsatisfactory) or 1 (satisfactory) depending on the outcome. This rating is weighed by *f-rating(A)*, *i.e.* the feedback rating of A. This implies that A needs to be well-reputed as a feedback provider in order for its opinions to have an effect on B's service reputation. That is, the feedback from those peers with higher feedback trust ratings will be weighed more than those with lower feedback ratings.

At the end of the transaction, A also needs to send feedback rating updates to all peers that had rated B earlier. If A received good (or bad) service from B, it provides a rating of 1 to all the peers that rated B as good (or bad) prior to the transaction.

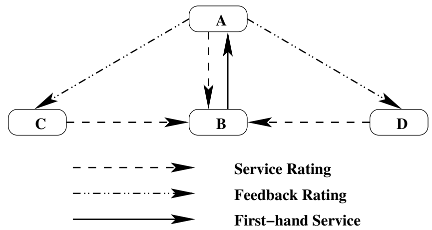

Fig. 1. Decoupling service and feedback reputation: after interacting with B, peer A modifies B's service reputation, but also modifies the feedback reputations of B's previous raters C and D.

This rating is in turn weighed by A's feedback rating. In the case that the outcome of A's transaction with B did not match with a prior service rating, A assigns a feedback rating of 0 to the originator of the rating. This process is shown in Figure 1, where peer A interacts with B, updates B's service reputation, and updates the feedback ratings of C and D, who contributed to B's service reputation.

Consequently, the service trust value and feedback trust value of a peer, u, denoted by *s-rating(u)* and *f-rating(u)*, are defined as:

$$\textit{s-rating(u)} = \alpha * \textit{s-rating(u)} + \beta * (r_u * \textit{f-rating(i)})$$

$$\textit{f-rating(u)} = \frac{1}{n_u} * \sum_{i=1}^{n_u} f_u * \textit{f-rating(i)}$$

where r_u indicates a service rating of -1 or 1; f_u is the feedback rating which can be 0 or 1 depending on malicious feedback or helpful feedback; n_u represents the total number of transactions that have made use of u's feedback; and α and β are normalized weight factors, between 0 and 1, used to exponentially decay reputation ratings.

Peers can exhibit dynamic personalities, *i.e.* they are honest at times and dishonest at others. For example, once a peer has established a good reputation in the network, it can abuse it. Also, honest peers can be subverted at any time and begin behaving badly. Hence, peer reputations must be representative of more recent behavior rather than old ratings. Our model follows previous models in exponentially decaying reputation to weigh recent feedback more heavily than older feedback. This allows reputations to become negative if a node becomes malicious, or recover if a formerly malicious node becomes honest. Furthermore, a dynamic system also allows honest nodes to recover from poor ratings given by malicious nodes.

In our model, we do not explicitly define how reputations and records of ratings are stored. The issue of reputation storage is orthogonal to our problem of decoupling reputation. Different storage models would not impact our reputation accuracy. In a self-storing model, peers can compute and maintain their own reputations, storing them along with ratings signed by raters. Another option is to store each peer's reputations away from the peer. For example, Eigentrust [10] and P-Grid [1] use distributed hash tables to determine where individual reputations are stored in the P2P system.

4 Performance Evaluation

We first evaluate the effectiveness of our method for limiting malicious behavior, then compare our approach to conventional correlated trust approach. Our results show not only a decrease in the number of malicious transactions, but also a significant reduction in the number of false positives and negatives reported.

To limit storage and communication overhead, we use a time window so that only records of a peer's transactions within the window are stored. Only the most recent service ratings are stored and feedback rating updates are only applied to those peers who rated a node recently. This reduces the communication costs associated with updating feedback ratings. The storage and communication costs of our reputation system are reasonable and justified given its significant benefits.

4.1 Simulation Environment

We implement our simulations in C using tools built on the Stanford Graph Base (SGB) [11]. The SGB platform represents a peer community and takes a peer model and topology graphs generated from the GT-ITM Topology Generator [4]. Table 1 summarizes the main parameters used and their default values.

	Parameter	Value Range	Nominal Value
Peer Model	Number of peers in the network	50-1000	500
	Percentage of honest peers	0-100	60
	Percentage of malicious peers	0-100	40
	Number of strategic peers	0-100	0
	Percentage of peers responding to a query request	0-20	10
Simulation	Number of query cycles in one experiment	50-1000	500
	Number of experiments over which results are averaged	5	5

Table 1. Simulation Parameters

Our network simulation proceeds in cycles. For simplicity, we assume that every peer in the network makes one transaction in each query cycle. We model the distribution of files and query responses using a Zipf distribution. Finally, P2P file sharing networks are often clustered by content categories. We assume only one content category in our implementation with file popularities defined to follow a Zipf distribution.

Our peer model involves three types of behavior patterns in the network, namely, *honest*, *dishonest* and *strategic*. Honest peers are truthful in providing service and feedback while dishonest peers provide incorrect service and incorrect feedback. Strategic peers are opportunistic peers that try to exploit the correlated trust model to spread bad information. They either provide good service and dishonest feedback or bad service and honest feedback. We vary the percentage of strategic peers in our experiments to illustrate the benefits of using our approach in scenarios where honest peers could report dishonest feedback about others, and vice versa.

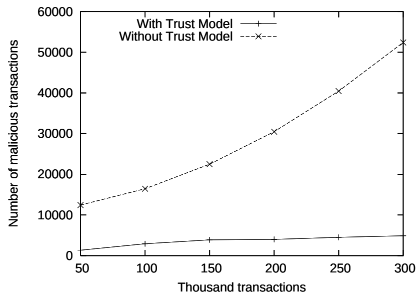

Fig. 2. Measuring malicious transactions in a network with and without our reputation model (40% nodes are malicious).

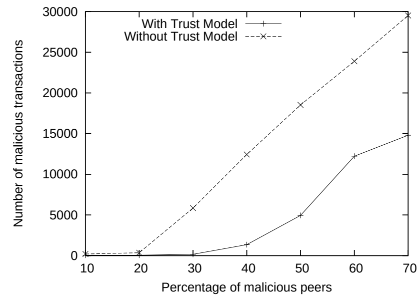

Fig. 3. Measuring malicious transactions in a network with and without our reputation model (Number of transactions is 50,000).

Our first set of experiments evaluates the effectiveness of our reputation system at detecting malicious behavior compared to conventional correlated trust. Each result presented is an average of five randomized runs, and the standard deviation is less than 2%.

4.2 Effectiveness against Malicious Behavior

We set the number of malicious peers to 40% in a network of 500 peers. On the x-axis, the number of transactions ranges from 50,000 to 300,000. As seen in Figure 2, without a reputation system, an increase in the number of transactions results in a corresponding increase in the number of malicious transactions. However, our trust model results in a significant reduction in the number of bad transactions in the system. After about 100,000 transactions, the number of malicious transactions is close to constant.

Figure 3 shows results for a similar experiment, but instead of varying the total number of transactions, varies the number of malicious peers in the network. We perform the test for 50,000 transactions over 500 peers, and vary the percentage of malicious peers from 10% to 70%. As seen in the figure, the number of malicious transactions is substantially lower when a trust model is employed. When a small percentage of network peers are malicious, they are easily detected and avoided, resulting in a very low number of malicious transactions. However, as malicious nodes become the majority ($> 50\%$) they begin to overwhelm honest nodes, resulting in a significant increase in malicious transactions. This result demonstrates the natural collusion between dishonest nodes that form a network majority.

4.3 Benefits of Decoupling Service and Feedback Trust

In our second set of experiments, we evaluate the benefits of our approach compared to the conventional approach of correlating service trust and feedback trust. In the correlated approach, ratings assigned to the service provider at the end of a transaction are weighed only by the service rating of the rater. That is, the feedback from those peers with higher service ratings will be weighed more than those with lower service ratings.

We set the number of transactions to 50,000 for 500 peers. We first evaluate the number of malicious transactions using both approaches in a network with only static peer personalities. Honest peers always provide honest service and feedback, and dishonest peers always provide malicious service and feedback. We vary the percentage of

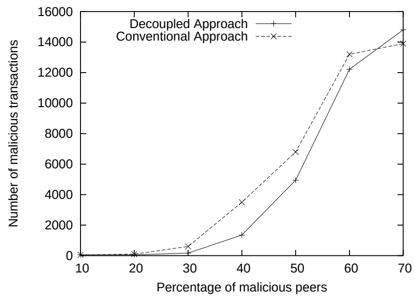

Fig. 4. Malicious transactions in networks with a conventional trust model and our decoupled model (50,000 transactions).

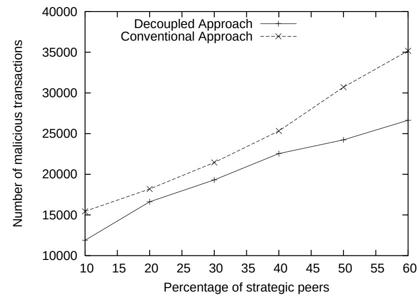

Fig. 5. Malicious transactions in networks with a conventional trust model and our decoupled model (40% malicious nodes, the percentage of strategic nodes varies).

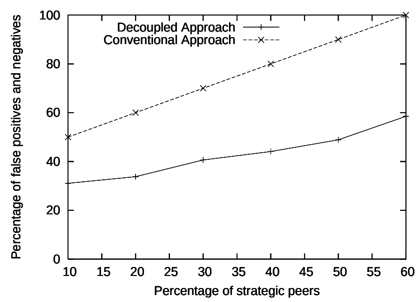

Fig. 6. False positives and negatives in a network with a conventional trust model and our decoupled model (40% malicious nodes, the percentage of strategic nodes varies).

malicious peers in the network from 10% to 70%. As seen in Figure 4, both approaches perform well in reducing the total number of malicious transactions, with our model generally being more accurate. When peers exhibit static personalities, the assumption that a honest peer will provide honest feedback holds true. Hence, correlated trust-based reputation models work as well as our decoupled model.

We introduce strategic behavior in our second experiment. Malicious peers may try to "rig the system" by providing honest service and feedback in some cases but dishonest feedback in others. Similarly, honest peers may, at times, give malicious feedback or service to some peers due to jealousy or competition. We use a network with 40% malicious peers who will provide both bad service and bad feedback. We vary the percentage of strategic peers from 10% to 60%, with the remaining nodes being totally honest. Half of the strategic peers provide good service as a service provider and malicious feedback as a service recommender. The other half provide bad service but honest feedback. Figure 5 demonstrates that our decoupled approach significantly outperforms the conventional approach in reducing the number of malicious transactions. While strategic peers take advantage of the correlated trust assumption in conventional systems to spread incorrect ratings, our decoupled model correctly identifies nodes as malicious service providers or sources of malicious feedback.

In our last experiment, we demonstrate how our decoupled trust model reduces the number of false positives and negatives reported in the P2P network. False positives and negatives represent the amount of false information fed into the reputation system. Such disinformation are the source of malicious transactions and are difficult to remove, once inserted. Again, we use a network with 40% malicious nodes, and vary the percentage of strategic peers in the network from 10% to 60%. As seen in Figure 6, our decoupled approach results in significantly fewer false positives and negatives than the conventional model. We note that the relatively high numbers of false reports are due to the high number (40%) of initial malicious nodes in these network setups.

5 Conclusions and Future Work

We have proposed a reputation-based trust model that improves accuracy by removing the assumption of correlation between service quality and feedback quality. The model decouples trust associated with each peer based on the role it plays, both as a service provider and as a service recommender. This decoupled approach incorporates reputations of both the service provider and the requester in the computation of trust values and, in this way, makes our model more robust to peer maliciousness. Our results report fewer false positives and negatives in the system as compared to the conventional approach of correlating the trust values. As ongoing work, we are building a more sophisticated trust model and working towards safeguarding our system from collusion.

References

1. ABERER, K., AND DESPOTOVIC, Z. Managing trust in a Peer-2-Peer information system. In *Proc. of CIKM* (Atlanta, GA, USA, Nov. 2001).
2. BUCHEGGER, S., AND BOUDEC, J. L. A robust reputation system for P2P and mobile ad-hoc networks. In *Proc. of the 2nd P2PEcon Workshop* (June 2004).
3. BURTON, K. Design of the openprivacy distributed reputation system, May 2002. http://www.peerfear.org/papers/openprivacy-reputation.pdf.
4. CALVERT, K. L., DOAR, M. B., AND ZEGURA, E. W. Modeling internet topology. *IEEE Communications Magazine 35*, 6 (June 1997), 160–163.
5. DAMIANI, E., DI VIMERCATI, D. C., PARABOSCHI, S., SAMARATI, P., AND VIOLANTE, F. A reputation-based approach for choosing reliable resources in peer-to-peer networks. In *Proc. of CCS* (Nov. 2002).
6. DEWAN, P., AND DASGUPTA, P. Pride: Peer-to-Peer reputation infrastructure for decentralized environments. In *Proc. of WWW* (May 2004).
7. EBAY. ebay home page, http://www.ebay.com, 2005.
8. FELDMAN, M., PAPADIMITRIOU, C., CHUANG, J., AND STOICA, I. Free-riding and whitewashing in peer-to-peer systems. In *Proc. of WEIS* (May 2004).
9. GNUTELLA. The gnutella protocol specification v0.4, 2001.
10. KAMVAR, S. D., SCHLOSSER, M. T., AND GARCIA-MOLINA, H. The eigentrust algorithm for reputation management in P2P networks. In *Proc. of WWW* (May 2003).
11. KNUTH, D. E. *The Stanford GraphBase: A Platform for Combinatorial Computing*. 1993.
12. OOI, B. C., LIAU, C. Y., AND TAN, K.-L. Managing trust in peer-to-peer systems using reputation-based techniques. In *Proc. of WAIM* (August 2003).
13. RESNICK, P., AND ZECKHAUSER, R. Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. *Advances in Applied Microeconomics 11* (Jan. 2001).
14. XIONG, L., AND LIU, L. Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE Trans. on Knowledge and Data Engineering 16*, 7 (2004).