International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:3, 2007

Towards an Effective Reputation Assessment Process in Peer-to-Peer Systems

Farag Azzedin and Ahmad Ridha

Abstract**—** The need for reputation assessment is particularly strong in peer-to-peer (P2P) systems because the peers' personal site autonomy is amplified by the inherent technological decentralization of the environment. However, the decentralization notion makes the problem of designing a peer-to-peer based reputation assessment substantially harder in P2P networks than in centralized settings. Existing reputation systems tackle the reputation assessment process in an ad-hoc manner. There is no systematic and coherent way to derive measures and analyze the current reputation systems. In this paper, we propose a reputation assessment process and use it to classify the existing reputation systems. Simulation experiments are conducted and focused on the different methods in selecting the recommendation sources and retrieving the recommendations. These two phases can contribute significantly to the overall performance due to communication cost and coverage.

Keywords**—** P2P Systems; Trust, Reputation, Performance.

I. INTRODUCTION

PEER -to-Peer (P2P) systems have flourished in the Internet and have become the number one application in Internet bandwidth usage. P2P systems are used in different applications, ranging from file sharing, real-time communication, and up to searching for extraterrestrial intelligence.

Along with various benefits such as scalability, there are several top issues in P2P systems, amongst which are reputation issues [13]. Reputation issues are important as they affect the confidence of peers on the network [5]. Users must be given the ability to measure the trustworthiness of a transaction partner [2].

Many popular P2P applications have not implemented a mechanism to assess peers' reputation. Therefore, it is difficult to prevent malicious transactions. An example of malicious transaction in P2P file sharing applications is the pollution of popular files [12]. In distributed computing applications, this can have a major effect on the invalid results due to improper computation by malicious peers. In incorporating reputation mechanisms in P2P systems, we need to consider the effectiveness and efficiency of those mechanisms. P2P systems have taken a lot of bandwidth so that we should balance between the overhead of reputation assessment and the capability to eliminate malicious transactions.

Manuscript received March 27, 2007. This work was supported in part by the Deanship of Scientific Research (DSR) at King Fahd University of Petroleum and Minerals (KFUPM) under Grant JFRG-2006-16.

Dr. Farag Azzedin is an Assisstant Professor with KFUPM, P.O. Box 1290, Dhahran, Saudi Arabia. Dr. Farag Azzedin is the corresponding author. Phone: 03-860-3431; fax: 03-860-2174; e-mail: fazzedin@kfupm.edu.sa.

Ahmad Ridha is a Graduate Student with the Department of Information and Computer Science at KFUPM; e-mail:aridha@kfupm.edu.sa.

The need for reputation assessment is particularly strong in P2P systems because the peers' personal site autonomy is amplified by the inherent technological decentralization of the environment. However, the decentralization notion makes the problem of designing a P2P based reputation assessment substantially harder in P2P networks than in centralized settings. In a P2P setting, the reputation data retrieval aspect plays an important role in P2P networks and only a small fraction of the entire available feedback can be used. In addition, P2P environments have no central trusted authority to aggregate feedback in a particular way to enable the decision making process.

A. Motivation

The field of reputation online has spawned the interest of a number of scholars in technical as well as non-technical fields. For example, groups such as [15] aim to explore the role of trust and reputation and to explore effective ways to implement them. Effective reputation modeling is believed to be an enabler for a range of new computing services including enhanced e-commerce, ubiquitous computing, Grid computing, P2P computing, and probably a variety of collaborative and cooperative online activities. Many researchers have proposed reputation systems [2], [5], [3], [13]to assess the trustworthiness of a peer based on the recommendations from other peers in the P2P environment.

From a practical point of view, reputation (referred to as second-hand information, referrals, or ratings) schemes are already being used in many successful commercial systems such as eBay, BizRate, and Amazon. These online commercial systems allow their clients to give recommendations on other members or other resources such as books.

Various problems exist in practical as well as academic reputation systems. These problems are tackled in an ad-hoc manner and there is no systematic and coherent way to derive measures, analyze the current reputation systems [5], [10]. This motivated us to propose a reputation assessment process and evaluate its effectiveness to accurately predict the reputation of a target peer while minimizing the overhead of collecting, filtering, adjusting, and aggregating the recommendation requests.

B. Contributions

This paper contributes to the reputation issue in P2P environments by proposing a reputation assessment process with

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:3, 2007

the following features: (a) The proposed reputation assessment entails various functions and traditionally these functions have been tightly coupled, which is not only an obstacle in identifying the common vulnerabilities of reputation systems but also makes the analysis and comparison of reputation systems hindering to identify emerging trends and open research issues, (b) The proposed reputation assessment enables us to address the focus of each reputation system and either inject strengths or remedies into the reputation system. To the best of our knowledge, this is the first attempt at developing a reputation assessment process that can be used as a classification, comparison, and analysis tool for reputation systems.

II. Related Work

A trust model for P2P computing environments is presented in [5]. This model uses a determined set of recommenders. The reputation is retrieved from the set of recommenders. The recommenders will provide recommendations based on their direct experience with the target peer. An honesty check mechanism is also presented to filter the recommendation set. The recommendations are then manipulated based on the recommenders' accuracy.

Gupta et al. [7] propose a reputation system for unstructured P2P file-sharing systems. The reputation is computed with a debit and credit concept with two methods of reputation tracking differing only in the trade-offs between reliability and overhead. Their approach uses a special node, Reputation Computation Agent (RCA), to which peers can enroll to participate in the reputation system.

Another model for unstructured P2P systems is presented in [3]. A peer assigns a Boolean satisfaction value for a transaction to indicate the outcome of the transaction. If previous experience exists, the satisfaction value is incorporated with past experiences using freshness value, based on accuracy of current trust compared to latest transaction. To check the target peer's reputation, the source peer sends a broadcast message to ask for recommendations.

To detect dishonest peers, the authors in [9] introduce the concept of suspicious transactions. A suspicious transaction is a transaction whose feedback is different from the one expected based on the target peer's reputation. To enable peers to move to different supernodes, each peer keeps a copy of his reputation. The data is encrypted using the supernode's secret key.

EigenTrust is proposed in [11]. It assigns a global trust value to every peer in a structured P2P file-sharing system. The system uses the value to choose the peers to download from. This way, untrustworthy peers are isolated. To get the target peer's global trust value, the source peer contacts the target peer's score managers. If they return different values, the majority value is chosen.

III. Proposed Reputation Assessment Process

A reputation based system enables peer x to gather recommendations from other peers regarding the reputation of peer y. The reputation of y might be polluted. Pollution happens due to dishonest recommenders trying to damage the reputation of y. The issues that we have to investigate are: (a) how much pollution or distortion a reputation system can tolerate? and (b) what steps a reputation system should take to minimize this pollution or distortion?

If x wants to enquire effectively about the reputation of y , x has to resolve the following issues: (a) Recommender selection: who should become the recommenders of x (i.e., from whom should x get the recommendations to compute the reputation of y), (b) Recommendation retrieval (i.e., how will the recommendations be retrieved), (c) Recommender filtering (which recommenders are useful towards computing the reputation of y), (d) Recommendation evaluation (how important are the recommendations), and (e) Recommendation manipulation (how will x process and combine the recommendations). This process of reputation assessment is illustrated in Figure 1.

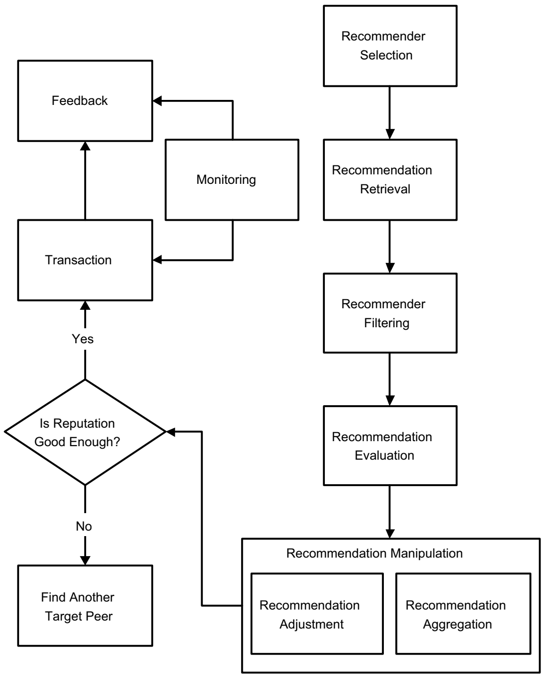

Fig. 1. Reputation assessment process.

Recommender Selection: In this phase, the source of recommendations (i.e., the recommenders) is determined. In reputation systems such as [3], the recommenders are every peer (except the source peer). In other reputation systems such as [5], [7] a specific set of recommenders are selected. Depending on the selection scheme used, there is an impact on the P2P network bandwidth usage, the filtering mechanism on the set of recommenders, and on the number of opportunities that a dishonest recommender is given to penetrate through the recommendation network and pollute the reputation of the target peer.

Recommendation Retrieval: Recommendation retrieval is then performed by sending recommendation requests. A source peer can send recommendation request to all other peers through flooding limited by a time-to-live value such as in [3].

If the network diameter is low, the request can reach most or all peers. This means that will cover most of the recommenders who have first-hand information about the target peer. But, of course, this is not a good utilization of the P2P network bandwidth.

Recommender Filtering: In this phase, peer x wants to know the most useful recommendation. For example, if the recommender is dishonest and gave a wrong recommendation, peer x needs to filter this recommender out. Please note that in this phase, the objective to do filtering is based on recommender's attributes. We can also do filtering based on the recommendation's attributes, but this is left for the recommendation filtering phase.

Recommendation Evaluation: In the evaluation phase, we decide how important a certain recommendation in relation to the other recommendations. Also in this phase, we will assign weights to the various recommendations received regarding the reputation of peer y. The decision making and the weight assignment can be done by the source peer [5], the supernodes [10], or the score managers [8]. Reputation systems can assign transaction-based weights, recommender-based weights, or use a combination scheme to assign the weights. Transaction-based weights are assigned based on the transaction value, transaction frequency, transaction importance, or the time of transaction. For example, this approach is done in [14]. Recommender-based weights are related to the attributes of the recommender such as his trustworthiness or the frequency of transactions between him and the source peer. This approach is used, for example, in [4].

Recommendation Manipulation: This phase calculates the reputation of the target peer based on the recommendation. The calculation involves two parts: adjustment and aggregation. Recommendation adjustment is performed to deal with subjectivity and accuracy problem. Once they are adjusted, there are several aggregation methods available such as average, weighted average, OWA, and so on. The weights from the recommendation evaluation phase are used in the aggregation part.

Recommendation adjustment has been dealt with differently by different researcher groups. Many reputation systems ignore the need of adjustment for recommendation [2], [3], [14]. The recommendations received from the recommenders need to be adjusted before being used. This is because trust is not completely transitive. If peer z trusts peer y with a trust level TL, then peer x who has z as a recommender, may not have the same TL for y. Hence, x needs to gauge the discrepancy between z's TL regarding y.

Recommendation Aggregation is a process that takes as input the individual recommendations from recommenders and their assigned weights of importance from the previous phase of recommendation evaluation. Its output is a single reputation value for the target peer based on the recommendations. There are two main directions of research in this recommendation aggregation process: (a) Which aggregation algorithm should be used and (b) Where the recommendations should be aggregated.

Transaction Monitoring: In any transaction between peer x and peer y, there is a trust concern from x as well as from y. Peer x can monitor the transaction using offline, online, or combining offline and online mechanisms. Peer x observes the transaction or the transaction records to determine whether any abuses have taken place by the target peer. While the source peer should have special configuration to define what conditions exactly cause a breach in the transaction contract.

Transaction Feedback: Using transaction feedback, peer x can update its own local experiences with peer y that it interacted with. This feedback process is done for future references. For example, if peer x wants to interact with peer y that it had interacted with before, then peer x can evaluate its past experience with peer y. The feedback process can be extended to a complaint mechanism in the reputation assessment process. A complaint mechanism is very crucial to post the reputation of target peers in a place where other peers can access. Having a complaint mechanism is motivated by this example. If there is no complaint mechanism, then each peer keeps to its direct experiences, with other peers that it interacted with them, to itself. This set of peers will not propagate the information about the untrustworthy peer unless they were asked for recommendations. This way, the untrustworthy peer might be globally undetected for sometime and be given chance to harm other peers.

IV. Performance Evaluation

To test the proposed reputation process, a simulation experiments were developed to compare the effectiveness and efficiency of different methods in recommender selection and recommendations retrieval phases. The simulations are set up based on three methods i.e.: (a) Using flooding (FL), (b) Using recommendation tree (RT). That is, peer x has specific recommender set, and (c) Using P-Grid [1]. The simulation experiments only include these two components of reputation assessment process namely, recommender selection and recommendations retrieval.

A. Goals of the Simulation

For each method (i.e., FL and RT), we run several schemes with different parameters. The parameters used in the simulation are shown in table I. We introduce the term acquaintance rate (AR) which is the probability that one peer knows another. Higher acquaintance rate means a peer has had transactions with more peers. The value ranges from 0 (every peer does have any transactions) to 1 (every peer has had transactions with all other peers). Performance indicators are calculated for each run. Recall and precision are used for FL and RT while P-Grid uses hop count.

B. Performance Metrics

The performance metrics for recommender selection and recommendation retrieval depend on the method used. For undetermined and specific set of recommender, the performance can be measured using recall and precision. These metrics are analogous to recall and precision in information retrieval[6].

TABLE I
EXOGENOUS AND DESIGN PARAMETERS USED IN THE SIMULATION.

Symbol	Definition	Values
N	# of peers	$N = (128, 256, 512, 1024)$
λ	Mean inter-arrival time	$\lambda = (0.1)$
AR	Acquaintance rate	$AR = (0.0 - 1.0)$
Num_{rec}	# of recommenders	4
TTL	Time to live	7
$Tran_{max}$	Maximum # of transactions	$5N$

A peer is regarded relevant when it has first hand information about the target peer. Recall, R, is calculated as follows:

$$R = \frac{C_R}{P_R} * 100\% \quad (1)$$

where C_R is the set of contacted peers that have first hand information about the target peer and P_R is the set of all peers that have first hand information about the target peer. The lowest value is 0 percent which means all peers having first hand information are not contacted. The highest value, 100%, is achieved if all peers having first hand information are contacted.

On the other hand, precision, P, is calculated as follows:

$$P = \frac{C_R}{C} * 100\% \quad (2)$$

where C is the set of contacted peers. The lowest value is 0% means none of the contacted peers have first hand information. The highest value, 100%, is achieved if all of the contacted peers have first hand information. Recall and precision are irrelevant for systems that store all information about a peer at one or more storage peers such as in structured peer-to-peer systems. If the system is stable and the network is secure and reliable, both recall and precision are 100%. For this type of systems, we can measure the performance using: (a) hop count: how many peers are contacted before reaching the storage nodes, and (b) structure overhead: the cost of keeping the structure stable (especially if the churn rate is high).

C. Simulation Results and Discussion

Simulation results for FL are shown in Figures 2 and 3. Naturally, precision improves as AR increases. We can see that the recall is relatively high ($> 75\%$) and stable. This is mainly because in FL recommendation requests are sent to all connected peer and the range is only limited by the TTL. The high recall is beneficial to closely approximate the global reputation. However, it depends on the number of dishonest peers in the network. Systems using FL must be able to verify the validity of the recommendations. Otherwise, it is susceptible to collusion between malicious peers.

Simulation results for RT are shown in Figures 4 and 5. Precision behavior is similar with FL; it is directly related with AR. As for recall, it seems to be inversely related with AR. This is due to the behavior of the recommendation tree. In systems with low AR, it is more likely that all or most of the recommenders do not have first hand information so the tree is expanded to the next level which means including more recommenders and increasing the chance of higher recall. On the other hand, if AR is high, the recommenders are more likely to have first hand information so that the tree is not expanded which results in low recall. Therefore, as the network size increases which usually means AR decreases, the recall should increase.

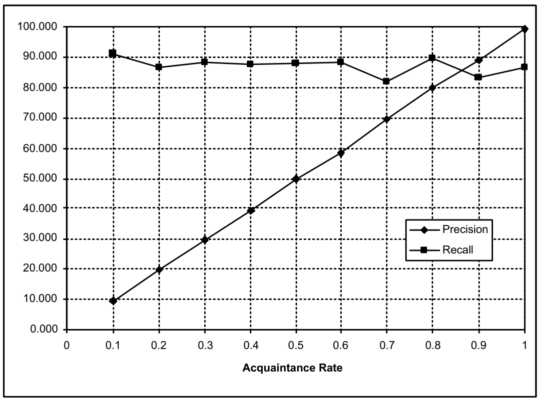

Fig. 2. Precision and recall for FL with 128 peers.

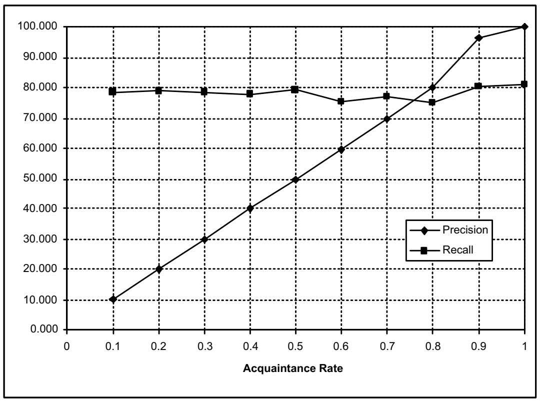

Fig. 3. Precision and recall for FL with 1024 peers.

Simulation results for P-Grid are shown in table II. The storage peer can be contacted in logarithmic hop count just as mentioned in [1]. This can improve the scalability of the reputation assessment process. One feature of P-Grid that we do not simulate is the use of replicas. It will improve the availability and also to decrease the risk of data tampering by the storage peer. However, using replicas will also increase the complexity as the replicas need to synchronize. P-Grid seems to be a promising approach. However, it is important to note that we do not include churn rate in the simulation. Also, the distribution of reputation information should be considered. It is possible that a peer that does few transactions has to provide more storage than others because the peer is the storage peer

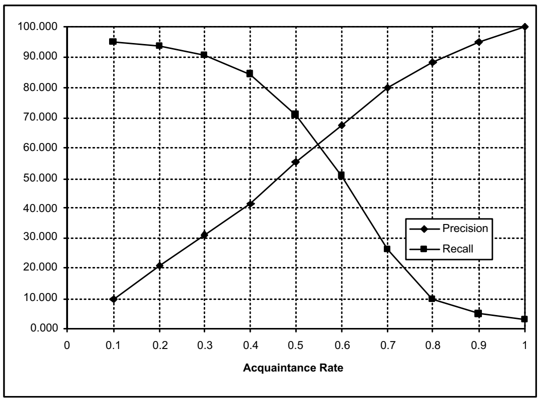

Fig. 4. Precision and recall for RT with 128 peers.

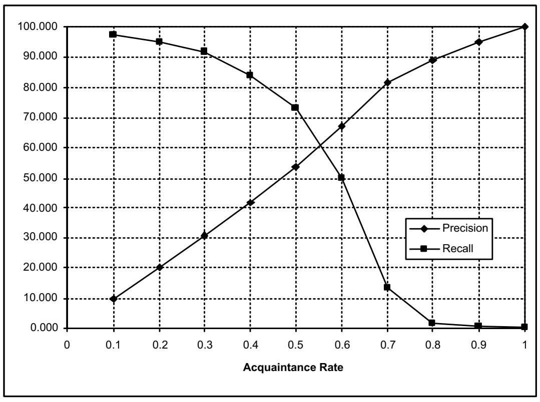

Fig. 5. Precision and recall for RT with 1024 peers.

for an active peer. This "unfairness" may impact the incentive to use the reputation system.

TABLE II
SIMULATION RESULTS FOR P-GRID.

# of peers	Run 1	Run 2	Run 3	Run 4	Run 5	Avg.	$\frac{log_2(N)}{2}$
128	3.49	3.49	3.45	3.55	3.40	3.47	3.5
256	4.10	4.07	3.99	4.03	4.03	4.03	4.0
512	4.56	4.54	4.49	4.49	4.49	4.52	4.5
1024	5.01	4.97	5.02	5.03	5.01	5.01	5.0

V. Conclusions and Future Work

The methods in recommender selection and recommendation retrieval contribute to the effectiveness and efficiency of a reputation based trust management system. In this paper we analyze three different methods to compare their performances. The flooding method is simple and can reach most part of the network resulting in high and stable coverage. However, the precision depends on the acquaintance rate. In larger networks, the increase of traffic is not desirable from the scalability point-of-view . On the other hand, in RT as acquaintance rate increases, the precision increases while the recall decreases. If all or most recommenders are active i.e. having recent transactions with the target peers, few recommendations (i.e. low recall) may be sufficient. As the network getting larger, acquaintance rate is more likely to decrease which in turn will generate deeper recommendation trees and higher recall but it also means higher traffic. In fact, RT may use more bandwidth (i.e. contacts more peers) compared to FL for low AR and high diameter networks. P-Grid seems to offer a very good solution to retrieve recommendations. Further research should be focused on the cost of maintaining the structure in networks with high churn rate.

VI. Acknowledgment

The authors wish to thank King Fahd University of Petroleum and Minerals and its Information Technology Center for the facilities utilized to perform the present work and for their support.

References

[1] K. A. et al., "P-Grid: a self-organizing structured p2p system," *ACM SIGMOD Record*, Vol. 32, 2003.
[2] A. Abdul-Rahman, *A Framework for Decentralised Trust Reasoning*, PhD thesis, Dept. of Computing, University College London, July 2004.
[3] R. Arienghieri, E. Damiani, S. Vimercati, S. Paraboschi, and P. Samarati, "Fuzzy techniques for trust and reputation management in anonymous peer-to-peer systems," *JASIST*, Vol. 57, No. 4, 2006, pp. 528–537.
[4] F. Azzedin and M. Maheswaran, "A trust brokering system and its application to resource management in public-resource grids," *2004 International Parallel and Distributed Processing Symposium (IPDPS 2004)*, Apr. 2004.
[5] F. Azzedin, M. Maheswaran, and A. Mitra, "Trust brokering and its use for resource matchmaking in public-resource grids," *Journal of Grid Computing*, Vol. 4, No. 3, 2006, pp. 247–263.
[6] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley, Boston, MA, 1999.
[7] M. A. M. Gupta and M. Ahamad, "Trade-offs between reliability and overheads in peer-to-peer reputation tracking," *Computer Networks*, Vol. 50, No. 4, 2006, pp. 501–522.
[8] Q. Gu, K. Bai, H. Wang, P. Liu, and C. Chu, "Modeling of pollution in p2p file sharing systems," (3rd IEEE Consumer Communications and Networking Conference), Jan.
[9] D. Hughes, G. Coulson, and J. Walkerdine, "Free riding on gnutella revisited: The bell tolls," *IEEE Distributed Systems Online*, Vol. 6, No. 6, 2006.
[10] A. Jsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," (Decision Support Systems), 2006.
[11] S. Kamvar, M. Schlosser, and H. Garcia-Molina, "The eigentrust algorithm for reputation management in p2p networks," (12th International World Wide Web Conference), May 2003.
[12] J. Liang, R. Kumar, Y. Xi, and K. Ross, "Pollution in p2p file sharing systems," (24th Annual Joint Conference of the IEEE Computer and Communications Societies), July 2005.
[13] P. Rodriguez, S. Tan, and C. Gkantsidis, "On the feasibility of commercial, legal p2p content distribution," *ACM SIGCOMM Computer Communication Review*, Vol. 36, No. 1, Jan. 2006, pp. 75–78.
[14] S. Song, K. Hwang, R. Zhou, and Y. Kwok, "Trusted p2p transactions with fuzzy reputation aggregation," *IEEE Internet Computing*, Vol. 9, No. 9, 2006, pp. 24–34.
[15] iTrust: Working group on trust management in dynamic open systems, "itrust," `http://www.itrust.uoc.gr`, June 2002.