IDENTIFYING HONEST RECOMMENDERS IN REPUTATION SYSTEMS

Farag Azzedin

King Fahd University of Petroleum and Minerals,
Information and Computer Science Department, Dhahran, 31261, Saudi Arabia.

ABSTRACT

Reputation systems aim to reduce the risk of loss due to untrustworthy peers. This loss is aggravated by dishonest recommenders trying to pollute the recommendation network. The objective of an honesty checking mechanism is to detect dishonest recommenders. Existing honesty checking mechanisms assume that contradicting recommendations are due to the dishonesty of the recommenders. However, such difference may be also due to the behavior change of the target peer. This paper shows the effect of such behavior change on the performance of existing honesty checking mechanisms. To the best of our knowledge, this is the first attempt at linking the behavior change to honesty checking.

Keywords: *Reputation Systems, P2P Systems, Honesty, Recommender-based Systems.*

1 INTRODUCTION

Being part in a peer-to-peer (P2P) system, a peer has the privilege of using pools of resources or services that would not be available to it otherwise. Unfortunately, the idea of having a virtual network framework is not attractive because of the risk associated with the notion of "sharing" resources or services [6, 8, 9, 12]. Because of the sensitivity and the vitality of data or information, such peers prefer to use their own "closed box" resources. This is not just costly and inefficient way to utilize resources, but also negates the advantages of P2P systems. As such, one of the fundamental challenges in the open and decentralized P2P environments is the ability to mitigate the risk of transacting with untrustworthy peers [6].

Many researchers have proposed reputation systems [2, 4, 6, 14] to assess the trustworthiness of a peer based on recommendations obtained from other peers. Recommenders play a vital role in the success of any reputation system because based on their recommendations, a derived reputation score will be computed and used to decide whether a transaction should take place. A false recommendation can result in committing a transaction with untrustworthy peers or avoiding a transaction with trustworthy peers. Recommenders with different motivations and malicious intentions can cause harm in such systems. Therefore, mechanisms to filter out undesirable recommenders are fundamental and are an integral part of the success of any online reputation-based community [3, 14]. Recommender filtering schemes are widely used in the literature [7, 21, 9, 15, 6] to minimize the effect of the undesirable recommenders in polluting the recommendation network

Hence, the objective of a reputation system is to have recommenders that positively contribute to the computed reputation score. A positive contribution helps narrowing the gap between the derived reputation score and the actual trustworthiness of the peer in question. To positively contribute, a recommender should be willing, active, and honest. Unwilling recommenders will not reply nor forward a recommendation request. Contacting unwilling recommenders will not contribute at all to the derived reputation score but it will result in inefficient bandwidth utilization. Incentive mechanisms are introduced to promote and encourage recommenders to participate in the recommendation network [13, 19]. Non-active recommenders will provide stale recommendation or a recommendation based on few trans-actions with target peer. Many researchers have investigated recommender activeness by considering the number of transactions, their values, and when these transactions were performed [17, 22].

Honesty checking has been also investigated by various researchers [7, 21, 9, 15, 6] but none has linked the behavior change of the peer in question to the honesty of recommenders. The primary goal of this paper is to shed light on the importance of this issue and its effect on recommenders' honesty. To the best of our knowledge, no existing literature tackles this issue, which we believe it is a vital dimension that should be considered when performing honesty checking.

This paper contributes to the honesty checking process by introducing another dimension that can affect the labeling of the dishonest recommenders. Existing honesty checking techniques [7, 21, 9, 15, 5] do not consider the behavior change a target peer as an important factor that might distort the image of a recommender. This issue is not tackled by existing literature and we believe that such an issue is a vital dimension that should be considered when identifying honest recommenders.

Throughout this paper, we refer to the peer that wants to assess a reputation as a source peer and the peer whose reputation is assessed as a target peer. The rest of the paper is organized as follows. Section 2 discusses the importance of honesty checking. Section 3 presents the performance metrics and the simulation setup to evaluate the existing mechanisms to filter out dishonest recommenders when the target peer changes its behavior. Also, results are presented and discussed in Section 3. Related works are discussed in Section 4. In Section 5, we conclude this paper and envision future directions to cope with the target peer's behavior change and its effect on honesty checking.

2 HONESTY CHECKING

The objective of a reputation system is to have recommenders that positively contribute to the computed reputation score. A positive contribution helps in narrowing the gap between the derived reputation score and the actual trustworthiness of the target peer. To positively contribute, a recommender should be willing, active, and honest. Unwilling recommenders will not reply nor forward a recommendation request. Contacting unwilling recommenders will not contribute at all to the derived reputation score but it will result in inefficient bandwidth utilization. Incentive mechanisms are introduced to promote and encourage recommenders to participate in the recommendation network [13, 19]. Non-active recommenders will provide stale recommendation or a recommendation based on few transactions with the target peer. Many researchers have investigated recommender activeness by considering the number of transactions, their values, and when these transactions were performed with the target peer [17, 22]. In this paper, we focus on the honesty aspect of the recommenders.

Reputation systems aim to reduce the risk of loss due to untrustworthy peers. This loss is aggravated by dishonest recommenders trying to pollute the recommendation network. The objective of an honesty checking mechanism is to detect dishonest recommenders. Existing honesty checking mechanisms assume that contradicting recommendations are due to the dishonesty of the recommenders. However, such difference may be also due to the behavior change of the target peer. This paper shows the effect of such behavior change on the performance of existing honesty checking mechanisms. To the best of our knowledge, this is the first attempt at linking the behavior change to honesty checking.

Relying on recommenders to estimate the reputation of the target peer, the source peer might be misinformed and form the wrong perception about the target peer. This is due to dishonest recommenders that try to pollute the environment by intentionally giving bogus reputation reports. Ideally, dishonest recommenders should be prevented from contributing to the target peer reputation. Measuring honesty is a dif cult task because in order for a source peer to determine the honesty of a recommender, the source peer needs to know what the recommender beliefs in. Since, this is impossible in P2P systems, most of the existing honesty checking algorithms [15, 9, 21, 7] use the consistency of the recommendation with an expected value in measuring honesty. In these honesty checking schemes, a recommender is marked as dishonest if it provides a recommendation that contradicts the expected value. In other words, existing honesty checking algorithms assume that if a recommendation is not in line with the expected value, then the recommender is the one to blame. On the contrary, the recommender might be honest and is reporting a behavior change of the target peer, while the expected value is reporting the old behavior of the target peer. Therefore, if there is an oscillating target peer that changes its behavior from trustworthy to untrustworthy, or vice versa, then such oscillating target peer will affect the recommenders' honesty.

A behavior change may be due to the target peer's attempt to gain pro t by building up its reputation and using it to perform untrustworthy transactions without being noticed. On the other hand, a behavior change can happen if a trustworthy peer is compromised by an untrustworthy peer and used to launch attacks.

3 PERFORMANCE EVALUATION

3.1 Simulation Setup and Performance Metrics

The objective of our simulation experiments is to measure the effect of the target peer's behavior change on the performance of existing honesty checking mechanisms. The simulation is a discrete event simulation. We use 1024 peers in constructing the overlay network and the number of transactions is 1024. We assume that all recommenders are honest, and when the nth transaction is attempted, the recommendations from the previous (n 1) transactions are

available. We have one target peer that would behave in an oscillating manner, switching from trustworthy to untrustworthy and vice versa. The starting behavior is trustworthy.
We simulate four honesty checking mechanisms, namely, filtering for Bayesian reputation system (BF) [21], outlier filtering (OF) [7], feedback consistency (FC) [9], and credibility factor (CF) [15]. If a recommendation is considered dishonest, the recommender is marked as dishonest and its recommendations are no longer accepted. The thresholds are set to 2.0, 0.01, and 0.5 for t in OF, q in BF, and T in FC respectively as used in [7, 21, 9].
For the performance metrics, we observe the number of recommenders misdetected as dishonest due to the oscillating behavior of the target peer. We experiment with behavior oscillation every 1 and 5 transactions. We calculate the reputation throughout the simulation to see if the transaction attempts are committed.
A recommendation contains the number of trustworthy and untrustworthy transactions committed with the target peer. In computing the derived reputation score, the basic Bayesian reputation system as described in [10] is used. This is done for all of the honesty checking mechanisms. We also measure the global reputation (GR) if all recommendations are used (no honesty checking is involved). The reputation value ranges from 0 to 1 and a peer is considered to be trustworthy if its reputation is 0.5. A transaction attempt is only committed if the target peer is predicted to be trustworthy.

3.2 Results and Discussion

Figure 1 and Figure 2 illustrate the number of misdetected honest peers and the global reputation if the target peer changes its behavior in every transaction. We can see that BF marks a high number (62.659%) of the recommenders as dishonest at the end of the simulation. The reputation of the target peer in BF is consistently high (>= 0.9) most of the time, indicating that peers reporting the untrustworthy behavior of the target peer are marked as dishonest. The marking starts when the target peer has had numerous transactions. In other words, the recommenders are punished for the target peer's behavior change if the target peer has performed many transactions.
Similar results are also shown by FC, although at a lower degree. At the end of the simulation, 32:454% of the peers are marked as dishonest. When the target peer performs an untrustworthy transaction, the consistency of the source peer decreases, and if it is the first transaction for the source peer, the source peer's feedback consistency is 0, thus, marked as dishonest. The reputation of the target peer in FC is lower than in BF; however, the target peer is still always considered trustworthy if the threshold of trustworthiness is less than or equal 0:6.
In OF, none of the recommenders are marked as dishonest because positive and negative recommendations have the same frequency when the behavior change occurs. It keeps the normalized value to be less than the threshold. It means recommendations from all peers are used throughout the simulation. This is shown by the reputation values which are equal to GR. This also means that all transaction attempts with the target peer are committed.
Using CF, only one recommender is marked as dishonest. The marked recommender is the one re-porting the first transaction of the target peer because in the second transaction attempt, the target peer is untrustworthy. Hence, the positive recommendation from the first recommender is no longer used and only the negative recommendation from the second recommender is available. This is reflected in the reputation of the target peer which is consistently low showing that the target peer is isolated in the rest of the simulation.

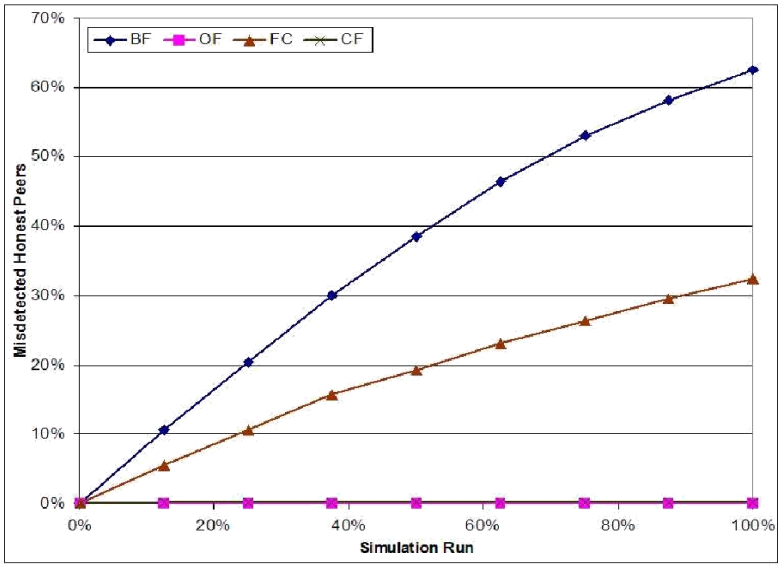

Figure 1. Misdetected honest peer (N =1).

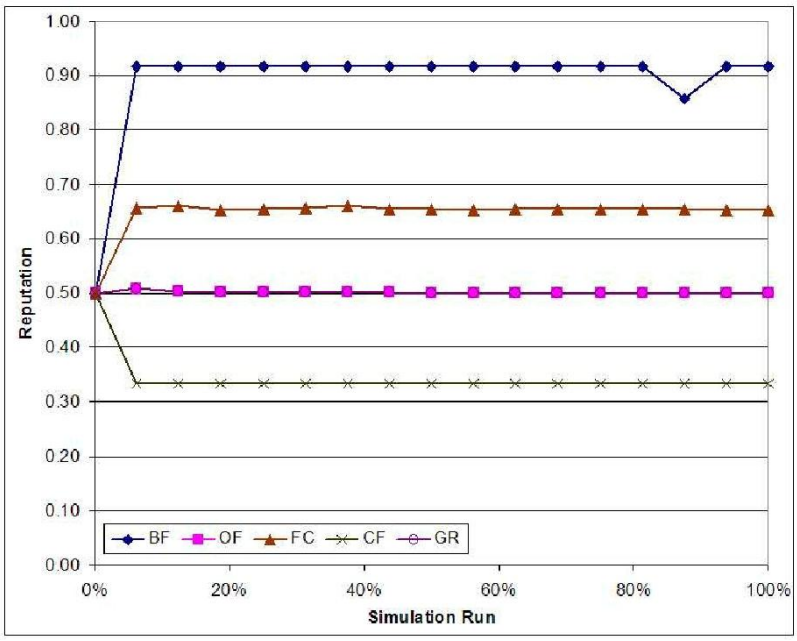

Figure 2. Global reputation of the target peer (N = 1).

Figure 3 and Figure 4 illustrate the number of misdetected honest peers and the global reputation if the target peer changes its behavior in every 5 transaction. Similar to the previous results, BF and FC mark 63.148% and 32.845% of the recommenders as dishonest, respectively, at the end of the simulation. All transaction attempts with the target peer are committed.

OF, also fails by misdetecting 39.198% of the recommenders. When the behavior change occurs, the target peer has invested 5 positive recommendations, so the normalized value for the negative recommendation is -2.041 which is outside the threshold. Hence, the recommender reporting the untrustworthy behavior is marked as dishonest. As the simulation continues, target peer gets more positive recommendations, and more peers reporting the untrustworthy

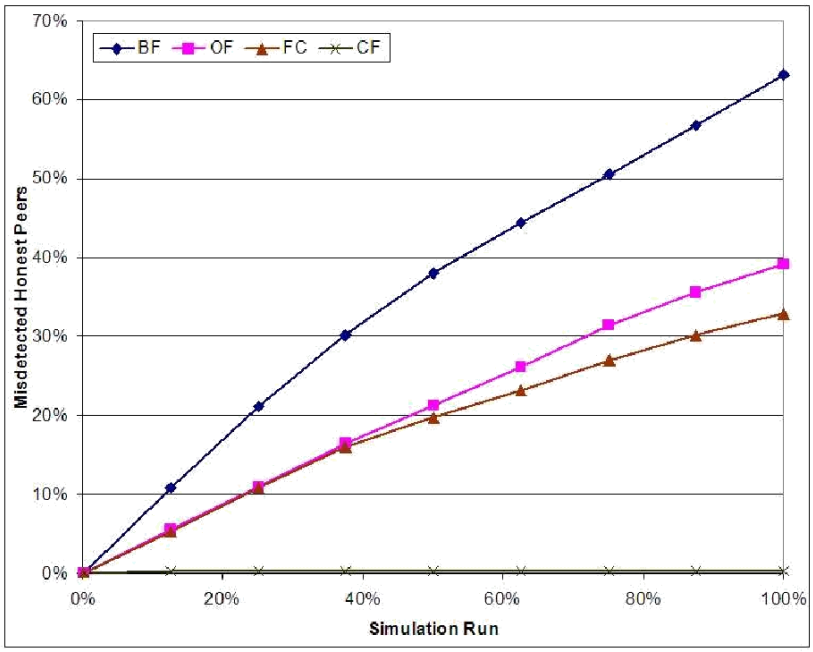

Figure 3. Misdetected honest peers (N = 5).

behavior are filtered out. The reputation of target peer in OF is also very high. CF, again, isolates the target peer while marking the recommenders reporting the earliest behavior of the target peer.

The results show that checking the honesty by observing the pattern of the recommendations may lead to the failure of predicting the current behavior. If the target peer changes its behavior, it is dif cult for the source peer to predict the result of the current transaction attempt based on the past transactions. CF succeeds because it uses only the latest behavior, but it would punish the recommenders reporting the old behavior. If the target peer performs numerous trustworthy transactions before becoming untrustworthy, CF would mark many peers as dishonest.

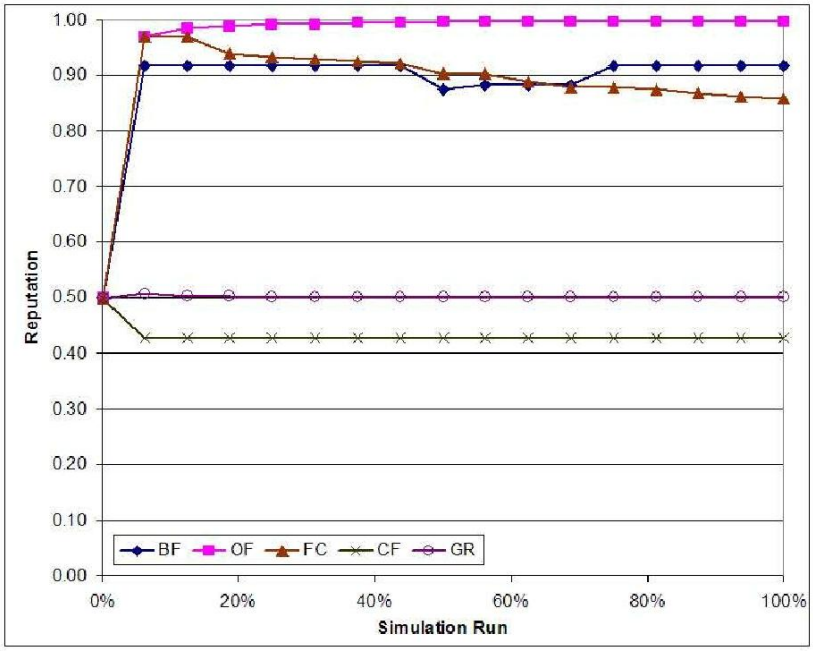

Figure 4. Global reputation of the target peer (N = 5).

4 RELATED WORK

Some reputation systems do not specifically tackle the honesty issue by correlating honesty and trust-worthiness [11, 13, 20], i.e., they assume that trustworthy peers provides honest recommendation and untrustworthy peers provide dishonest recommendations. Such reputation systems are vulnerable to badmouthing attacks. A trustworthy peer might provide dishonest recommendation to isolate other competing trustworthy peers and consequently to increase his profit. Other reputation systems [1, 16, 23] assume that the majority of the peers are honest and therefore cancel the effect of dishonest ones on the recommendation network.

The reputation system in [6] uses reply consistency to predict honesty. Consistent peers are assumed to be honest and vice versa. Each peer has a set of trusted allies through whom consistency check is performed. The checking is done by asking one or more of the trusted allies to send recommendation request for the target peer to the recommender. The source peer would compare the recommendation it gets directly with the one received by the trusted allies. Assuming the requests come in relatively short time, the recommender should give answers with no or little value difference. Therefore, if the difference is more than certain threshold, the recommender is being inconsistent. The recommender would be replaced from the source peer's recommender list and marked as dishonest so that it would not be included again in the list. However, this method can not detect dishonest peers that provide consistent replies.

A method to filter out dishonest feedbacks for Bayesian reputation systems is presented in [21]. Bayesian reputation systems use the beta distribution in predicting a peer's reputation using the number of trustworthy and untrustworthy transactions as the distribution parameters. The parameters of the distribution are:

$$\alpha = N_T + 1$$

And $$\beta = N_U + 1$$

where N_T and N_U are the number of trustworthy and untrustworthy transactions with the target peer reported by the recommender respectively. Honesty checking is performed by identifying feedbacks whose expected probability is less than the *probability density function* (PDF) at a certain quantile, q, and those exceeding the PDF at (1 - q) quantile in the beta distribution of the aggregated recommendations. The recommenders providing those outliers are considered to be dishonest. The checking is performed iteratively until the all recommendations are within the PDF at [q, 1 - q] quantile.
In [7], an honesty checking mechanism is proposed, inspired by the concept in [21]. Recommendations are categorized in positive and negative recommendations, and values of 1 and 0 are assigned to them, respectively. If there are contradicting recommendations, these values are normalized so that their mean and standard deviation are 0 and 1, respectively. A threshold, t, is used to remove recommendations whose associated normalized values are $\leq -t$ or $\geq t$..

Transactions are evaluated by both the source peer and the target peer in [9]. A source peer's feedback is considered consistent if it agrees with the target peer's self-evaluation. Assuming most of the peers are trustworthy and honest, most inconsistencies would be in the case that the source peer reports a trustworthy transaction as untrustworthy (badmouthing) or an untrustworthy transaction as trustworthy (collusion to boost the target peer's reputation). Thus, a source peer is suspected to be providing false feedbacks if the proportion of inconsistent feedbacks exceeds certain threshold, T. Feedbacks from inconsistent peers are no longer accepted.

The reputation system in [15] uses a measurement called credibility factor. The credibility factor increases if a recommender provides a recommendation that matches the actual result of the transaction. From the credibility, discredibility factor can be derived. A recommender whose discredibility factor is higher than its credibility factor will be filtered out.

The honesty checking in [18] also compares the recommendations to the actual result of the transaction, if the transaction is committed. A feedback rating of 1 is given to recommenders whose recommendations match the transaction results whereas recommenders that provide contradicting recommendations get feedback rating of 0. The feedback rating is used to determine the acceptance of recommendations from the recommenders in next reputation assessment.

5 CONCLUSIONS AND FUTURE WORK

A peer has the motivation to change its behavior in order to pursue its own interest. Even worse, the behavior change may be due to a security breach to a trustworthy peer. Such behavior change, as shown in the experiments carried out in this paper, can have tremendous effect on misdetecting honest recommenders.

Recommenders reporting the new behavior are misdetected as dishonest in mechanisms that assume the honest recommendations are the ones that follow the mainstream opinion such as in the Bayesian filtering and outlier filtering. While systems that compare the recommendations to the current behavior misdetect recommenders reporting the old behavior as in the credibility factor. The change of behavior also affects the detected honesty in the feedback consistency. Misdetecting honest recommenders would reduce the availability of honest recommendations and may result in an incorrect derived reputation score, which would lead committing a transaction with an untrustworthy peer or avoiding a transaction with a trustworthy peer. As a future work, we are currently working on an honesty checking mechanism that considers the behavior change of the target peer in measuring the honesty of the recommenders.

REFERENCES

[1] K. Aberer and Z. Despotovic. Managing trust in a peer-2-peer information system. In *10th Int'l Conf. Information and Knowledge Management (CIKM'01)*, pages 310–317, Nov. 2001.

[2] K. Aberer and Z. Despotovic. Possibilities for managing trust in P2P networks. EPFL Technical Report, IC/2004/84, Lausanne, 2004.

[3] E. Anceaume and A. Ravoaja. Incentive-based robust reputation mechanism for p2p services. Principles of Distributed Systems, Lecture Notes in Computer Science, 4305:305–319, 2006.

[4] 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, 57(4):528–537, 2006.

[5] F. Azzedin and M. Maheswaran. A trust brokering system and its application to resource management. In International Parallel and Distributed Processing Symposium (IPDPS'04), Apr. 2004.

[6] F. Azzedin, M. Maheswaran, and A. Mitra. Trust brokering and its use for resource matchmaking in public-resource grids. Journal of Grid Computing, 4(3):247–263, 2006.

[7] F. Azzedin and A. Ridha. Honesty checking in reputation assessment for peer-to-peer systems. In 8th IEEE co-sponsored International Conference on Peer-to-Peer Computing (P2P'08), Submitted.

[8] B. Dragovic, S. Hand, T. L. Harris, E. Kotsovinos, and A. Twigg. Managing trust and reputation in the XenoServer Open Platform. In iTrust 2003, pages 59–74, 2003.

[9] Y. Jin, Z. Gu, and Z. Ban. Restraining false feedbacks in peer-to-peer reputation systems. Semantic Computing, 2007. ICSC 2007. International Conference on, pages 304–312, 17-19 Sept. 2007.

[10] A. Jøsang and R. Ismail. The beta reputation system. In Proceedings of the 15th Bled Electronic Commerce Conference, June 2002.

[11] S. Kamvar, M. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks. In 12th International World Wide Web Conference, pages 640–651, 2003.

[12] H. Nissenbaum. Will security enhance trust online, or supplant it? In R. Kramer and K. Cook, editors, Trust and Distrust Within Organizations: Emerging Perspectives, Enduring Questions, pages 155–188. Russell Sage Publications, 2004.

[13] T. Papaioannou and G. Stamoulis. An incentives' mechanism promoting truthful feedback in peer-to-peer systems. Cluster Computing and the Grid, 2005. CCGrid 2005. IEEE International Symposium on, 1:275–283 Vol. 1, 9-12 May 2005.

[14] P. Rodriguez, S. Tan, and C. Gkantsidis. On the feasibility of commercial, legal P2P content distribution. ACM SIGCOMM Computer Communication Review, 36(1):75–78, Jan. 2006.

[15] A. Selc¸uk, E. Uzun, and M. Pariente. A reputation-based trust management system for P2P networks. International Journal of Network Security, 6(3):235–245, May. 2008.

[16] S. Sen and N. Sajja. Robustness of reputation-based trust: Boolean case. In 1st Int'l Joint Conf. Autonomous Agents and Multi-Agent Systems (AAMAS-02), pages 288–293, July 2002.

[17] S. Song, K. Hwang, R. Zhou, and Y. Kwok. Trusted P2P transactions with fuzzy reputation aggregation. IEEE Internet Computing, 9(9):24–34, 2005.

[18] G. Swamynathan, B. Y. Zhao, K. C. Almeroth, and H. Zheng. Globally decoupled reputations for large distributed networks. Adv. MultiMedia, 2007(1):12–12, 2007.

[19] Y. Wang, Y. Hori, and K. Sakurai. Economic-inspired truthful reputation feedback mechanism in p2p networks. In FTDCS '07: Proceedings of the 11th IEEE International Workshop on Future Trends of Dis-tributed Computing Systems, pages 80–88, Washington, DC, USA, 2007. IEEE Computer Society.

[20] Y. Wang and J. Vassileva. Bayesian network-based trust model. Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on, pages 372–378, 13-17 Oct. 2003.

[21] A. Whitby, A. Jøsang, and J.Indulska. Filtering out unfair ratings in bayesian reputation systems. The Icfain Journal of Management Research, 4(2):48–64, Feb. 2005.

[22] L. Xiong and L. Liu. Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities. IEEE Trans. Knowledge & Data Engineering, 16(7):843–857, Jul. 2004.

[23] B. Yu and M. P. Singh. An evidential model for distributed reputation management. In 1st Int'l Joint Conf. Autonomous Agents and Multi-Agent Systems (AAMAS-02), pages 294–301, July 2002.