A Trust-Enhanced Topology Adaptation Protocol for Unstructured P2P Overlays

Changyong Niu
Shanghai Jiaotong University, Shanghai, China
cyniu@sjtu.edu.cn

Jian Wang
Shanghai Jiaotong University, Shanghai, China
jwang@sjtu.edu.cn

Ruimin Shen
Shanghai Jiaotong University, Shanghai, China
rmshen@sjtu.edu.cn

Abstract

In this paper, a novel probabilistic computational reputation based trust model is firstly introduced, which has the property of separating reputations between providing services and giving feedbacks. And then a topology adaptation protocol for unstructured P2P overlays is proposed based on this trust model. The basic principle of this adaptation protocol is that peers with higher service and feedback trust values have higher probabilities of being accepted as neighbors in the P2P overlay. We show via simulations that the proposed trust model has better accuracy of reputations generated. And the topology adaptation protocol can enhance trusts even further, resulting in fewer malicious transactions when highly dynamic peer behavior patterns exist in the P2P network.

1. Introduction

P2P systems are open communities and are often established dynamically with peers that are unrelated and unknown to each other. Peers have to manage the risk involved in making transactions without prior experience and knowledge about each other's trustworthiness. Reputation systems have been proposed to boost trust and enhance collaboration in such open communities. The basic problem related to reputation-based trust model in P2P network is that information about transactions performed between peers is dispersed throughout the network so that every peer can only build an approximation of the global situation in the network. And the situation is further complicated by the fact that peers storing and processing trust related data cannot be considered unconditionally trustworthy and their eventual malicious behaviour must be taken into account [3, 8]. Recently, there are solutions that aim to predict (typically using probability and AI techniques) the trustworthiness of a peer in the dynamic p2p overlay. In particular, the Beta Reputation System (BRS) [1] is a centralized probabilistic trust model based on the beta distribution. However, BRS does not show how it is able to cope with misleading information. The approach of [2] also builds on the BRS, but this method requires sufficient ratings to allow successful identification of unfair ratings. Another probabilistic estimation method leveraging Maximum Likelihood Estimation (MLE) [6] is quite a lightweight technique, and typically with good estimation results under certain circumstance. To give a rough estimate of performance, this method allows for limited or rare feedback on history transactions of the target peer. Although it is essentially mathematical sound, it uses posterior information (the fraction of liars in networks) as prior information in its reputation estimate. Even if various techniques exist in predicting such posterior information, the prediction accuracy loss due to dynamics in P2P networks has significant influence in its performance in reputation estimate.

In this paper, we present a different way to model reputation in P2P networks. Specifically, we distinguish between reputations between providing service and giving feedback, and estimate each reputation using probabilistic techniques. Through decoupling service and feedback reputation and operating only on limited number of feedbacks, rather than aggregating all the ratings in the p2p network, our model makes possible an efficient implementation of reputation generation. We then explore the use of Bayesian filtering techniques, namely, a Kalman filter, to estimate target peer's future reputation based only on its last performance state.

The rest of this paper is structured as follows: section two describes the reputation model, which formalizes the meaning of reputation representation

0-7695-3007-9/07 $25.00 © 2007 IEEE
DOI 10.1109/SKG.2007.140

and explains how it is used in reputation generation; section three elaborate the reputation measurement and estimation model; based on the new trust model proposed, a topology adaptation algorithm is introduced in section four, which leverages the benefits of separation between service and feedback reputations; evaluations and comparison are presented in section five, and section six concludes the work.

2. Reputation Model

A reputation is an expectation about a peer's behaviour based on information about or observations of its past behaviour [4]. Thus the objective of constructing reputation of a peer is to estimate its future behaviour to be of similar quality as its past behaviour. Following this definition we model reputation of a peer as the probability to behave loyally in providing services and giving feedbacks, denoted as R_s and R_f. Thus an R_s approaching 0 means a peer always being dishonest in providing services and an R_s approaching 1 means almost completely honest.

In the proposed trust model, we compute (or estimate) reputations of service and feedback by two normal distributions, $N(s;\mu_s,\sigma_s)$ and $N(f;\mu_f,\sigma_f)$. Consequently, a perfect estimate of R_s will be $N(s;\mu_s = R_s,\sigma_s \to 0)$. If a μ_s value is something like 0.8, it is interpreted as, statistically, the corresponding peer associated with μ_s will provide honest service with a probability of 80%, and σ_s will describe the degree of uncertainty in this estimation. Then using μ and σ, the system can easily characterize its belief of the estimation of each peer's reputation.

According to the definition above, our peer model involves three types of behaviour patterns in the network when providing services or giving feedbacks, namely, honest, dishonest and strategy. Honest service providers always are truthful in providing services, but may not be trustworthy in giving feedback, thus being a strategy or dishonest feedbacker.

2.1 Reputation Generation

As in [8, 9], trust (as well as reputation) itself can be experience-based and recommendation-based. In the former, trust is based on peers' individual experience and in the latter it is based on the information provided by others, consequently generating totally subjective trust value and vice versus. Also like [9], we try to combine these two kinds of trust, but in a totally different way. If a querying peer gets two items of feedback from A (denoted as $f_{sA,}$) and B (denoted as f_{sB}) when tries to estimate the reputation of another peer C and the system currently characterizes the feedback reputations of A and B by A~ N(μ_{fA}, σ_{fA}) and B~ N(μ_{fB}, σ_{fB}), the estimate on the service reputation of peer C ($\mu_{sC,}\sigma_{sC}$) can be generated from the following equations.

$$\mu_{sC} = [\sigma_{fB}^2/(\sigma_{fA}^2+\sigma_{fB}^2)]*f_{sA}*\mu_{fA}+[\sigma_{fA}^2/(\sigma_{fA}^2+\sigma_{fB}^2)]*f_{sB}*\mu_{fB} \quad (1)$$

$$1/\sigma_{sC}^2 = 1/\sigma_{fA}^2 + 1/\sigma_{fB}^2 \quad (2)$$

As illustrated, the form given in equation 1 makes good sense. If σ_{fA} were equal to σ_{fB}, which is to say the system thinks the two feedbacks are of equal uncertainty, the equation says the optimal estimation of the service reputation of peer C is simply the average of the two feedbacks, as would be expected. On the other hand, if σ_{fA} is larger than σ_{fB}, which is to say that the uncertainty involved in the feedback μ_{sA} is larger than the feedback μ_{sB}, then the equation dictates "weighting" μ_{sB} more heavily than μ_{sA}. Finally, σ_{sC} of the estimate is less than both that of A and B, denoting a decrease in the uncertainty involved in the estimate, which means a lower level of uncertainty in the estimate even a 'poor' feedback from quite unbelievable peer A is retrieved. The rationale behind this is that the system tends to believe that even poor feedbacker can occasionally provide an honest feedback.

3. Reputation Estimate

On the basis of feedback aggregation from (1) and (2), we employ a Kalman Filter [10, 11] as an 'observer' to estimate reputation in P2P settings. In terms of using Kalman Filter to observe the reputations of peers in P2P networks, it is assumed that each peer presented in the network has a pattern of behaviour. So we can map each behaviour patter to a point between [0, 1] with a value approaching 0 meaning always behaving dishonestly, and vice versa for 1. Most importantly, peers whose behaviour swing between dishonest and honest are modelled by mapping each of them to a real number between (0, 1) with that number meaning the probability of honest behaviour. Given this, the proposed trust model is able to model honest, dishonest and strategy behaviours in a P2P network. By using a Kalman Filter as parameter estimator for with no state transformation, the system can just maintain the last state of the reputation of each peer (the state function of Kalman filter) and combine the reputation generation model (the measurement

function of Kalman filter) as measurement to recursively produce more accurate reputation estimate. Using the reputation generation model and a Kalman Filter estimate, a limited number of feedback cycles on a target peer at each step can lead to enhanced reputation estimates. We use the same method as [6] to get feedback when needed. As our main concern in this paper is reputation estimate, we will not elaborate it here.

Based on the estimate, the querying peer may or may not make transaction with the target peer. If a transaction is made the system makes updates to the feedback reputation of the peers providing them, based on the outcome of the transaction. A simple logic is used in feedback reputation update. If peer *i* gave a consistent feedback with the outcome of the transaction made, a value is added to μ_i to make it approaching 1 and a value is minused to σ_i to decrease the uncertainty involved in its future feedback. An opposite strategy is used when an inconsistent feedback is got from peer *i*.

$$\mu_i = \mu_i + \alpha * (\sigma_i^2 / (\sigma_i^2 + \sigma_q^2)) * (1 - \mu_i) \quad (3)$$

$$\sigma_i = \sigma_i - \beta * (\sigma_i^2 / (\sigma_i^2 + \sigma_q^2)) * abs(0 - \mu_i) \quad (4)$$

For example, where there is consistent feedback, equation (3) and (4) are used to update feedback reputation. α And β are parameters for controlling the strength of the update and μ_i is updated approaching 1 and σ_i is updated approaching 0. The update is also weighed by the feedback reputation of the querying peer σ_q. If σ_q is large, which means that the former querying peer update is somewhat untrustworthy, (3) and (4) produce relatively small updates.

4. Topology adaptation algorithm

To further leveraging the benefits of separating the reputations between providing services and giving feedbacks, it is a straightforward optimal strategy for all the peers making transactions with the peers with highest reputation in providing services. On the other hand, it is optimal for the service providers only serving those with higher feedback reputations in order to truly reflect its service reputation in the network. Thus the topology adaptation protocol is two-folded based on connection sender and receiver with both the sender and receiver trying to optimize their own neighbour sets. For the simplicity of expression, we only consider single context of transactions. For multiple contexts of making transactions, the proposed trust model and topology adaptation algorithm can be easily extended using the view model proposed in [5, 7].

Primitives of the adaptation protocol are defined in table 1.

Table 1. Primitive and semantic

Primitives	Semantics
Disconnection_request(u,v)	Peer u disconnecting from peer v and sending disconnecting message to peer v; peer v updating its neighbour list by deleting peer u after receiving the request
Get_neighbor(u,Nu)	Peer u getting is neighbour list and saving it in Nu
Get_transaction(u,PRu)	Peer u getting the peer list from cash, with which it made transactions, and saving it in PRu
Add_neighbor(u,v)	Peer u adding v to is neighbour list

Table 2. Adaptation protocol algorithm

	Connect_Sender(u) //peer u adaptating its neighbour list and trying to connect to peers with higher service reputations
1	{
2	Get_neighbor (u, N_u)
3	Get_transaction (u, PR_u)
4	if (PR_u- N_u) <> Null
5	for {each j in (PR_u- N_u)}
6	T_{sj}=R_{sj} //getting j's service reputation
7	T_{sv}=min{R_{sv} v∈N_u} // getting the peer v in u's neighbour list with minimum service reputation
8	if sizeof(N_u) < sizeof(u.neighbortable) //if the neighbour table of u is not full
9	Add_neighbour(u, j);
10	elseif T_{sj} > T_{sv} and Connect_reciever(u, j)== True // if j's service reputation is larger than the one in u's neighbour list with minimum services reputation
11	Disconnect_request(u,v);
12	Add_neighbour(u, j);
13	endif //corresponding line 8
14	endfor //corresponding line 5

15	endif //corresponding line 4
16	}
	Connect_reciever(u, j) // peer j responses the connecting request from peer u
1	{
2	Get_neighour(j, N_j)
3	$T_{fu}=R_{fu}$; //getting u's feedback reputation
4	$T_{fv}=\min(R_{fv}, v \; v \in N_j\}$// getting the peer v in j's neighbour list with minimum service reputation
5	if sizeof (N_j) < sizeof(j.neighbortable)
6	Add_neighbor(j, u);
7	return True;
8	elseif $T_{fu} > T_{fv}$
9	Disconnection_request(j, v)
10	Add_neighbor(j, u);
11	return True;
12	else
13	return False;
14	endif
15	}

5. Simulations

5.1 Comparison with Previous Probabilistic Trust Model

In this section, we compare the proposed reputation model to MLE (Maximum Likelihood Estimation) presented in [6, 12], the most similar system we have identified. As the MLE simulation is based on mean absolute error of the estimate (based on different percentage of dishonest peers and number of feedback reports), the comparison is based on both identical and similar simulation settings. The detailed simulation settings and results are listed in Table 1. It shows that with identical simulation settings, our model works at the same performance level when the fraction of dishonest peers in the network is relatively low (typically less than 30%), but the new model outperforms MLE by more than 200% when that fraction goes up. We have also conducted the comparison using different simulation settings. It shows that although the feedback decreased according to the number of peers in the network (setting II), it always worked a little bit better than with the setting (I). It shows that in a medium-sized P2P network, as long as each peer makes transactions with approximately 5% of other peers in the network and provides feedbacks based on these transactions (honestly or dishonestly), our model will provide enhanced accuracy of the target's reputation estimate. Also, the fraction of interaction among peers in a P2P network goes down as the scale of the network goes up.

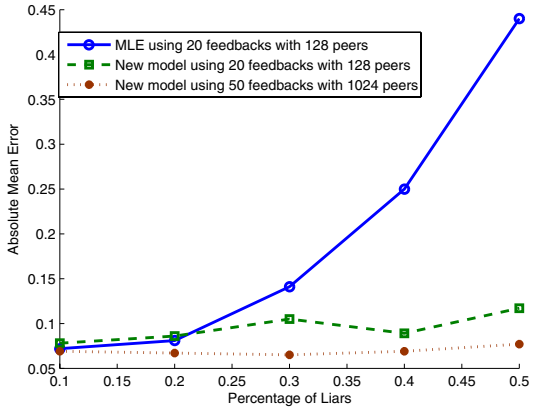

Figure 1. Comparison with MLE

5.2 Effectiveness of adaptation protocol

To further examine the effectiveness of the topology adaptation algorithm in preventing malicious behaviours, we implement simulations in a file sharing p2p network. The network simulation proceeds in cycle and we assume that ever peer in the network makes one transaction in each query cycle. For simplicity, we assume only one content category exists in the network. And the file distribution and query responses as well as file popularity are all defined to follow a Zipf distribution. And the initial topology graphs are generated from the GT-ITM Topology Generator and the number of connections of a peer is kept unchanged when performing topology adaptation. We set the number of peers in the network to be 500 and simulate 100,000 transactions. In Figure 1 it is showed that the evaluation of the number of malicious transactions with and without topology adaptation with only static peer models. That is honest peers are always provide honest services and feedbacks, and dishonest peers always provide malicious services and feedback. We vary the percentage of malicious peers from 10% to 70%. As showed in Figure 1, both approaches, with or without topology adaptation, performance well in limiting the number of malicious transactions, with the latter being generally more effective.

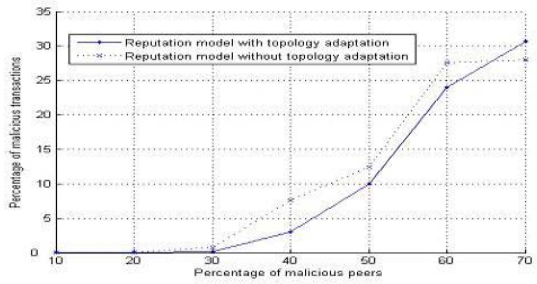

Figure 2. Malicious transactions with and without topology adaptation protocol (50,000 transactions)

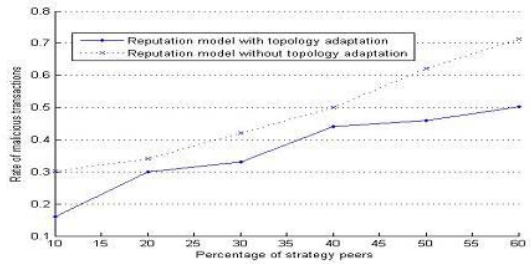

Figure 3. Malicious transactions with and without topology adaptation protocol 35%malicious nodes, with percentage of strategy nodes varies

When introducing strategy behaviour in the other experiment as illustrated in figure 2, the malicious peers may try to provide honest service or feedback in some cases but dishonest service or feedbacks in others, consequently introducing highly dynamics in peer behaviour patterns in the network. We use a network with 35% malicious peers who will provide both bad service and bad feedback and vary the percentage of strategy peers from 10% to 60%, with the remaining nodes being totally honest in services and feedbacks. Figure 2 demonstrates that with topology adaptation, the network is enhanced in preventing malicious behaviours. And we note that the relatively high percentage of malicious transaction is due to the high number (35%) of initial malicious nodes in the P2P network setup.

5. Conclusion

This study presents a new approach to modelling and estimating reputation in P2P networks. Although various reputation systems have been proposed or built, the proposed represents a novel approach to modelling reputation and is capable of modelling random behaviours in a network. By distinguishing between service and feedback reputation, our method models reputation with clearer semantics. The simulation results strongly suggest that our model provides good performance in a variety of situations and outperforms similar work. Based on the new reputation model, a topology adaptation protocol is proposed, and we show via simulation that this protocol can enhance trust behaviours when highly dynamic peer behaviours patterns exist in the network.

In this study, the behaviour probability associated with each peer is assumed to be independent from each other. However, in practice, peers may be in collusion with each other to boost their reputation estimate or bad mouthing other peers in a network. Modelling and estimating reputation in such settings and further identifying these collusion peers in the overlay topology are the most import part of our future research

References

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

[2] A. Whitby, A. Jøsang, J. Indulska, "Filtering out unfair ratings in Bayesian reputation systems", In *Proceedings of the 7th International Workshop on Trust in Agent Societies*, 2004 ..

[3]L. Xiong and L. Liu, "Peertrust: Supporting reputation-based trust in peer-to-peer communities", *IEEE Transactions on Knowledge and Data Engineering (TKDE), Special Issue on Peer-to-Peer Based Data Management,* 2004

[4]S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "Eigentrust: Reputation management in p2p networks", In *Proceedings of the World Wide Web Conference*, Budapest, Hungary, 2003.

[5] H. Zhuge, X. Chen, and X.P. Sun, "Trust-based probabilistic search with the view model of peer-to-peer networks", *Concurrency and Computation: Practice and Experience*, 2006

[6]Z. Despotovic, and K. Aberer, "A Probabilistic Approach to Predict Peers' Performance in P2P Networks", *Eighth International Workshop on Cooperative Information Agents*, 'Best Paper' at CIA 2004

[7] H. Zhuge, and X. Li, "Peer-to-Peer in Metric Space and Semantic Space", *IEEE Transaction on Knowledge and Data Engineering,* Vol 19, 2007

[8]G. Swamynathan, B. Y. Zhao and K. Almeroth: "Decoupling Service and Feedback Trust in a Peer-to-Peer Reputation System", *Proceedings of the 1st International Workshop on Applications and Economics of Peer-to-Peer Systems (AEPP)*, 2005.

[9]N. Griffiths, "Enhancing Peer-to-Peer Collaboration Using Trust", in *The International Journal of Expert systems with Applications*, Elsevier, 2006

[10] M., S. Grewal, and A. Andrews, *Kalman Filtering Theory and Practice Using MATLAB* (Second Ed.). New York, NY USA, John Wiley & Sons, Inc. 2001

[11]G. Welch and G. Bishop, "*An Introduction to the Kalman Filter*" Presented in ACM SIGGRAPH 2001

[12] Z. Despotovic, K. Aberer, "P2P reputation management: Probabilistic estimation vs. social networks", *Journal of Computer Networks, Special issue on Management in Peer-to-Peer Systems: Trust, Reputation and Security*, Elsevier, 2006