BPTrust: A Novel Trust Inference Probabilistic Model based on Balance Theory for Peer-to-Peer Networks

Zhenhua Tan[1], Liangyu Zhang[1], Guangming Yang[1]
[1]*Software College, Northeastern University,*
Wenhua Rd. 3-11, Heping Street, CN-110819 Shenyang, China, phone: +862483683990
tanzh@mail.neu.edu.cn

***Abstract*—In this paper, we propose a novel trust inference model named BPTrust (Balance and Probability based Trust model) based on balance theory and probability theory. Trust relations network, trust inference network and trust inference deep level are defined firstly before modelling. Based on balance theory, two inference rules are proposed, and trusted evidence chains generation algorithm is designed. In terms of a single trusted evidence chain, mathematics models are proposed to infer the trust value of sink node based on Markov chain theory while Bayesian theory is used to design inferring model during multi paths condition. Simulations proved the rightness and effectiveness.**

***Index Terms*—Distributed computing, information security, peer to peer computing.**

I. INTRODUCTION

Security and privacy issues have become critically important with the fast expansion of peer-to-peer networks owing to theirs openness and loose coupling. Trust models are being studied to ensure effective interactions among nodes and have got increasing attention recently. A well-defined trust model can provide meaningful decision support and help customer to reduce possible risk during an Internet transactions. In a social network, generally speaking, a person evaluates others' trust degree according to their direct communication history or according to evaluations from third parties. Thus, researchers modelling direct trust and indirect trust to evaluate trust degree in a peer-to-peer network. Direct trust could be acquired directly through the evidence history of transactions between nodes locally and formed a direct trust network [1], [2]. But indirect trust depends on recommendations by other nodes in most trust models and it is meaningless when no recommendations happened or no common nodes existed between trustor and trustee. Trust inference is a quiet new method for a trustor to compute a trustee's indirect trust only according to the direct local trust relationships [3], [8]. A trust chain from trustor to trustee would be found first from the direct trust network, and there exists two kinds of trust inference model to compute the indirect trust. The first one is based on multiplication which multiply all of the nodes' direct trust degree as the trustee's final trust degree [1], [3]; and the second one selects the max/min trust value or average trust value from the trust chain[4], [5]. Both are not rational. The multiplication model will lead the final result to be very small; even if the result is within the range of 0 and 1, it is not consistent with objective facts. And average model take the trust nodes as peer nodes and reduces the contribution from the nodes which have higher trust value in trusted computing. For example, in Fig. 1, do you think node 1 should trust node 11 or not? The above schema couldn't give a clear answer.

In this paper, we propose a novel trust inference model named BPTrust (Balance and Probability based Trust model) based on social psychological theory--Balance theory [9], and probability theory. BPTrust could analyze trust network and infer trusted evidence chain from it. Mathematics definitions are designed to infer the trust degree of sink node for source node even in completely strange condition.

II. TRUST RELATIONS NETWORK

To compute the trust inference, a trust relations network, which is constructed by direct trust model, should be provided in advanced. In BPTrust, a trust relations network is a directed graph $G_{Trn} =< V, E, W >$, where V is the set of IP nodes and $E = \{< i, j > | i \to j\}$ is the set of the directed relations between nodes; let $W = \{\omega_{i,j} \mid \omega_{i,j} \in [0,1] \wedge (i, j \in V) \wedge (< i, j > \in E)\}$ where each $\omega_{i,j}$ represents direct trust degree of node j from the perspective of node i. In this paper, integer number *i, j* represents node identification. Fig. 1 shows a simple trust network (a local trust relations network for a given node i $G_{LTrn}(i)$ is a sub-set of G_{Trn}, where $G_{LTrn}(i) =< V', E', W', i >$. V', E', W' in $G_{LTrn}(i)$ are subsets of V, E, W in G_{Trn} separately. $G_{LTrn}(i)$ represents a directed graph that start with node i).

Manuscript received March 19, 2012; accepted May 13, 2012.
This work is supported by the Doctor Program Foundation of Education Ministry of China (No. 20110042120027), China Postdoctoral Science Foundation under Grant No. 2012M511166, the Fundamental Research Funds for the Central Universities of China (No. N110417006, N110204003).

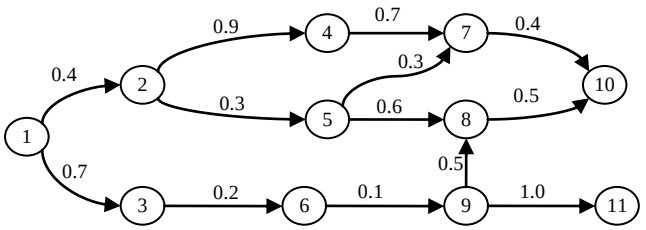

Fig. 1. Demo of trust network.

III. TRUST INFERENCE NETWORK

In a peer-to-peer network, each node (such as node i) fixes a trust threshold (such as τ_i) to make sure other nodes are trustworthy or not. In order to infer from trust relations network for node i, we should convert the $G_{LTrn}(i)$ into a trust inference network firstly. Thus, a local trust inference network for a given node i $G_{L-TIN}(i) =< V', E', Opers, i >$ is a directed graph to represent trust or distrust relations in $G_{LTrn}(i)$, where $Opers(i, j) = \begin{cases} +, when & \omega_{ij} \geq \tau_i \\ -, when & \omega_{ij} < \tau_i \end{cases}$, and "+" means trust while "-" means distrust. Fig. 2 is converted from Fig. 1 (assume all of the trust threshold is 0.5).

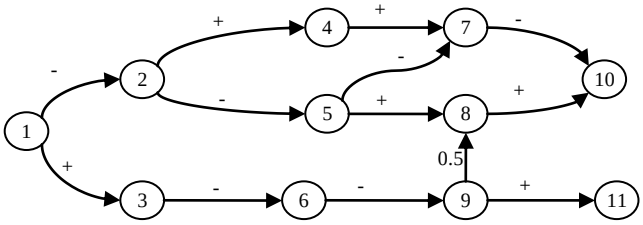

Fig. 2. Demo of a local trust inference network.

IV. TRUST INFERENCE RULES

Balance theory [9] which originated in social psychology in the mid-20th-century, constructs a triangle with two people and the event between them. Each vertex of the triangle has a positive or negative relationship with the other two vertices. To judge the current status of the triangle, we first pick up the signs of the three edges (positive be 1, negative be -1), then multiply the three signs. If the result is "1", the triangle is balanced. Otherwise, the triangle is unbalanced.

As shown in Fig. 3, triangles with three positive signs (T3) or two negative sign (T1) tend to be balance. On the contrary, triangles with two positive signs (T2) or three negative signs (T0) tend to be unbalance. And J. Leskovec found the universality of T3 and T1 in real trust relations [10].

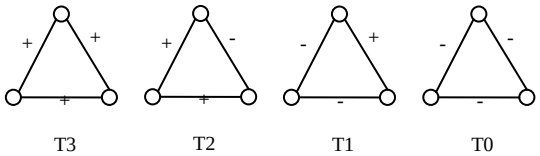

Fig. 3. Balance and unbalance relations.

If node A trusts node B and B trusts node C, we can indicate that A trusts C according to the principle of "The friend of my friend is my friend". This result is consistent with case T3 in Fig. 2. Another situation is that if A trusts B, but B distrusts C, then we can indicate that A distrusts C according to the principle of "The enemy of my friend is my enemy". This result is also consistent with case T1 in Fig. 3. Therefore, we design two important rules for trust inference:

$$Rules = \begin{cases} rule1: A(+)B \wedge B(+)C \Rightarrow A(+)C, \\ rule2: A(-)B \wedge B(-)C \Rightarrow A(+)C, \end{cases} \quad (1)$$

where $A(+)B$ means A trust B and $A(-)B$ means A distrust B.

On the basis of above definition, we could define how to infer from trust relations networks. Let $TI(i, j) =< G_{L-TIN}(i), Rules, j >$, which means node i can infer trust value of node j in the trust inference network $G_{L-TIN}(i)$, where i is the inference source node while j is the sink node.

V. TRUST INFERENCE LEVEL

Using $L(i)$ to express the inferring level (deep degree) of trust inference, where i would be the first level, and i's neighbours would be the second level, and so forth. According to small world theory, we assume $L(i) \leq 6$ to help the search efficiency. Fig. 4 shows the level demo.

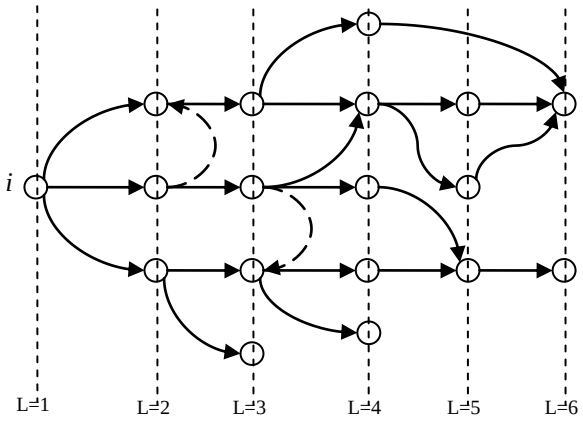

Fig. 4. The hierarchy model of trust relations network.

During the trust inference, each level should have different level factor for the inference computing. So level factor could be denoted by $f_L(x)$ and used exponential function to express decreasing degree

$$f_L(x) = 1 - e^{x - L(i) - 1}, \quad (2)$$

where $x \in [1, L(i)]$.

Obviously, $f_L(1) > f_L(2) > \ldots > f_L(L(i))$.

VI. DISCOVERY ALGORITHM FOR TRUSTED EVIDENCE CHAIN

A trusted evidence chain TEC(i, j) means a directed chain from node i to node j, which is caught from $G_{LTrn}(i)$ and also the basis to compute trust inference probability. Fig. 5 is a demo of a TEC according to Fig. 1 and 2.

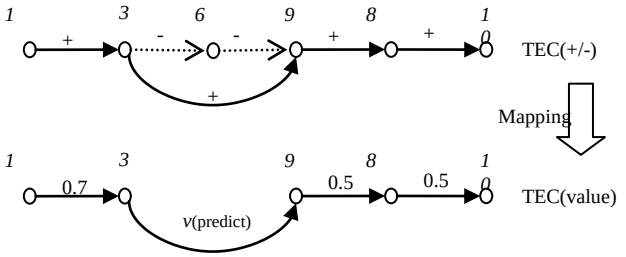

Fig. 5. Demo of TEC.

We design an algorithm below to find TEC from $G_{LTrn}(i)$.

Input: $G_{LTrn}(i)$ **, *i, j***

Output: TEC Array with TEC(i,j)

(1) while ($L(i)$ >0) {

(2) Converting $G_{LTrn}(i)$ into $G_{L-TIN}(i)$ according to each node trust threshold;

(3) $L(i)$ --;}

(4) Search path(i,j) from $G_{L-TIN}(i)$, build into TEC array;

(5) goto step (4) if exists more path;

(6) For each TEC(i,j) in TEC Array

(7) Computing TEC(i,j) with *Rules1,2* into *new TEC(i,j)*;

(8) return TEC array.

VII. TRUST INFERENCE COMPUTING

Let

$$p_d(i,j) = \begin{cases} \omega_{ij}, when\ i \neq j, \\ 1, when\ i = j, \end{cases} \tag{3}$$

which means the direct trust from i to j, and i is adjacent with j. In this paper, we assume that any node would completely trust itself.

Markov chain describes the fact that the current state of the node is just associated with the adjacent node and the trust evidence chain (TEC) is corresponded to this Markov property. In Markov model, the transfer probability of Markov chain in k-step, denoted by $p\{X(n+k)=j \mid X(n)=i\}$, which means that the condition probability in state i to j after k-step. Since Markov is stable, the k-step transfer can be solved through one step. $p[k] = p_{i,i+1} \times p_{i+1,i+2} \times \ldots \times p_{i+k-1,j}$, each $p_{x,y}$ is the transfer probability of adjacent node x, y. According to Markov model, we could calculate the TEC(i,j) by the following equation

$$p_{TEC}(i,j) = \prod_{k=1}^{L(i)} p(i+k-1, i+k). \tag{4}$$

However, the above equation will return a very small (even to zero) value to TEC(i,j) because of the too many multiplication operations, and it's not objective. Each node should have a confidence to its trust evaluation, and different transfer level should have difference level factor (just like described as (2)), while the equation 3 doesn't consider any of them.

Confidence is the confident level of the trust evaluation value. According to the degree of the confidence, we divide the trust level into three categories: *NT*(Not very Trust), *GT*(General Trust) and *VT*(Very Trust).

Assume $\alpha \in [0,1]$ is the cut-off point of *NT* and *GT* while $\beta \in [0,1]$ is the cut-off point of *GT* and *VT*. The range of the direct trust values in TEC(i,j) is:

$$R_k = P_{(\max)} - P_{(\min)}, \tag{5}$$

where $P_{(\max)}$ is the maximum direct trust value among TEC(i, j), and $P_{(\min)}$ is the minimum. Then:

$$\alpha = P_{(\min)} + R_k / 3, \tag{6}$$

$$\beta = P_{(\max)} - R_k / 3. \tag{7}$$

So, $NT \in (0, \alpha]$, $GT \in (\alpha, \beta]$ and $VT \in (\beta, 1]$.

Using $C(i,j)$ to express the confidence of direct trust evaluation from node i to node j. And, define $C(i,j)$ as:

$$C(i,j) = \begin{cases} \dfrac{1}{1+(\alpha - p_d(i,j))} \times \alpha, if\ p_d(i,j) \in NT, \\ \dfrac{1}{1+(\beta - p_d(i,j))} \times \beta, if\ p_d(i,j) \in GT, \\ \dfrac{1}{1+(P_{(\max)} - p_d(i,j))} \times P_{(\max)}, if\ p_d(i,j) \in VT. \end{cases} \tag{8}$$

Therefore, we adjust the $p_{TEC}(i,j)$ as

$$p_{TEC}(i,j) = \frac{\sum_{k=1}^{L(i)} \left(p_d(i+k-1, i+k) \times f_L(k) \times C(i+k-1, i+k) \right)}{\sum_{k=1}^{L(i)} \left(p_d(i+k-1, i+k) \right)}. \tag{9}$$

The above equation can be used to compute a single trust evidence chain (or a single path). Bayesian network could be applied while multi-path (such as $path_1$,$path_2$,...,$path_m$) exists between node i and j. Assume there are m paths existed between i and j, according to total probability

$$\sum_{x=1}^{m} Path_x = 1. \tag{10}$$

To compute each path's weight, let

$$Path_x = \frac{p_{TEC}(x)}{\sum_{y=1}^{m} p_{TEC}(y)}, \tag{11}$$

where $p_{TEC}(x)$ means the trust inference value for the x^{th} path according to (6). Finnaly, we can obtain the trust inference (node i to node j) method to computing multi-path TEC as

$$p(i,j) = \sum_{x=1}^{m} \left(Path_x \times p_{TEC}(x) \right). \tag{12}$$

VIII. SIMULATIONS AND ANALYSIS

In order to verify the rightness of BPTrust, we design a simulation platform by dataset Epinion, which provides the transaction nodes and trust relationship with Trust and Distrust status. To simulate BPTrust, we design a random function to generate trust probability ranged between [0, 1]. More than 1,500,000 trust relations are simulated for 20,000 nodes.

Firstly, we search trusted evidence chains in the trust relations network. We simulate the capability of searching TEC path of BPTrust with other algorithm. Table I shows the result.

TABLE I. SIMULATIONS FOR GENERATING TEC.

TEC Rules	Trust relations scale	TEC Numbers
All trust	312	12
Distrust as filter	312	2
BPTrust with rule1	312	17
BPTrust with rule2	312	19
BPTrust with rule1 and rule 2	312	22

As we can see from the above table, the BPTrust with rule1 and rule 2 will get most trust evidence chains while distrust as filter get the least amount. It proves that the BPTrust with balance theory will get more TEC information for trust inference.

Then, we compare the BPTrust with Average policy and Multiplication Policy in computing the trust inference. Table II shows the simulated data of $G_{LTrn}(1)$.

TABLE II. DATA OF TRUST RELATIONS NETWORK $G_{LTrn}(1)$.

Form Node i	To node j	P_d(i,j)
1	2	0.8
2	16	0.8
1	3	0.7
3	4	0.3
4	5	0.4
5	6	0.6
6	16	0.7
1	7	0.8
7	8	0.7
8	9	0.5
9	16	0.5
1	10	0.4
10	11	0.7
11	12	0.6
12	13	0.5
13	16	0.7
1	14	0.7
14	15	0.8
15	16	0.7

We discovered 5 TEC(1,16) in the above $G_{LTrn}(1)$. After computing the confidence, level factor and etc, we conclude the result in Table III.

TABLE III. DIFFERENT TRUST INFERENCE COMPUTING MODEL.

Method	Path 1	Path 2	Path 3	Path 4	Path 5	Trust Inference
BPTrust	0.79	0.54	0.69	0.53	0.74	0.674
Average Method	0.8	0.54	0.625	0.58	0.733	0.656
Multiply With Sqrt	0.8	0.51	0.61	0.57	0.73	0.636

As we can see, the BPTrust could normally infer a rational value for node 1 while the multiplication needs a SQRT function. In fact, during the processes of our experiments, the average method and multiplication method didn't work at all when very small distrust value appeared. The result is variable when the network has different levels. From the data, the algorithm reflects the trust inference objectively. Trust inference response to the fact that the value of trust inference is between the minimum trust probability and the maximum trust probability.

IX. CONCLUSIONS

In this paper, we propose a new trust inference model BPTrust to compute the indirect trust value in a strange condition. Trust relations network, trust inference network and trust inference deep level are defined firstly before modelling. Based on balance theory, two inference rules are proposed to generate trusted evidence chains. Markov and Bayesian probability theory are used to infer sink node's trust value and mathematics models are designed in BPTrust. Simulations proved the rightness.

However, it will be a long time to study the trust inference model for a distributed system. There are many problems waiting to be improved and solved. In our future work, more inference rules and experiments will be studied as well as trust/distrust group discovery algorithms.

REFERENCES

[1] S. Kamvar, M. Schlosser, H. Garcia Molina, "The eigentrust algorithm for reputation management in p2p networks", in *Proc. of the 12th International World Wide Web Confidence (WWW 03),* ACM press, 2003, pp. 640–651. [Online]. Available: http://dx.doi.org/10.1145/775152.775242

[2] Z. H. Tan, X. W. Wang, W. Cheng, G. R. Chang, Z. L. Zhu, "A Distributed Trust Model for Peer-to-Peer Networks Based on Multi-Dimension-History Vector", *Chinese Journal of Computers*, Science Press of China, vol. 33, no. 9, pp. 1725–1735, 2010. [Online]. Available: http://dx.doi.org/10.3724/SP.J.1016.2010.01725

[3] F. Walter, S. Battistion, F. Schweitzer, "A model of a trust-based recommendation system on a social network", *Autonomous Agents and Multi-Agent Systems*, Springer press, vol. 16, no. 1, pp. 57–74, 2007. [Online]. Available: http://dx.doi.org/10.1007/s10458-007-9021-x.

[4] W. Hines, D. Montgomery, D. Goldsman, *Probability and Statistics in Engineering.* John Wiley, 2003, p. 655.

[5] J. Bi, J. Wu, W. Zhang, "A trust and reputation based anti-spim method", in *Proc. of the IEEE 27th International Conference on Computer Communication (IN-FOCOM'08),* IEEE Press, 2008, pp. 2458–2493. [Online]. Available: http://dx.doi.org/10.1109/INFOCOM.2008.319

[6] U. Kuter, J. Golbeck, "Using probabilistic confidence models for trust inference in Webbased social networks", *ACM Transactions on Internet Technology*, ACM press, vol. 10, no. 2, pp. 1–23. [Online]. Available: http://dx.doi.org/10.1145/1754393.1754397

[7] L H. Vu, K. Aberer, "Effective usage of computational trust models in rational environments", *ACM Transactions on Autonomous and Adaptive Systems*, ACM press, vol. 6, no. 4, pp. 1–25, 2011. [Online]. Available: http://dx.doi.org/10.1109/WIIAT.2008.172

[8] V. Patricia, C. Chris, D. C. Martine, M. Ankur, "Trust-and-Distrust Based Recommendations for Controversial Reviews", *IEEE Intelligent Systems*, IEEE Computer Society, vol. 26, no. 1, P. 48–54, 2011. [Online]. Available: http://dx.doi.org/10.1109/MIS.2011.22

[9] F. Heider, *The Psychology of Interpersonal Relations.* Wiley, 1982, p. 336.

[10] J. Leskovec, D. Huttenlocher, J. Kleinberg, "Signed Networks in Social Media", in *Proc. of the 28th ACM Conference on Human Factors in Computing Systems (CHI).* ACM press, 2010, p. 1361–1370. [Online]. Available: http://dx.doi.org/10.1145/1753326.1753532