A group-based trust propagation method

Fatime Rahman Easa, Abbas Ghaemi Bafghi, Hassan Shakeri
Department of Computer
Ferdowsi University of Mashhad
Mashhad-Iran
raheleh373@gmail.com, ghaemib@um.ac.ir, shakeri@mshdiau.ac.ir

***Abstract*--Trust is a concept taken from social sciences and is considered as a soft security approach that is effective in reducing risk. In this paper, for estimating the trust between unknown nodes, a group-based trust propagation method has been proposed. Most of the conventional trust propagation methods are not applicable for trust evaluation of today's large trust graphs. Our trust propagation method is scalable because of using grouping method. For better trust estimation inside group the confidence of trust value have been considered. We also consider two factors Intermediate Group Confidence (IGC) and Group Confidence (GC) for confidence of trust between two groups. To evaluate this method a real large data set of advogato.org is used. Evaluation of accuracy is based on correlation and mean absolute error(MAE). Comparing the proposed method with the Iterative Multiplication method (IMS) results suggests that the correlation and absolute mean error have been improved. In addition, due to the use of group-based method the speed of the proposed method has been improved.**

Keywords-trust management; trust propagation; group-based; trust evaluation; correlation.

I. INTRODUCTION

The concept of trust that is taken from social sciences is defined as "the degree of subjective belief about the behaviors of a particular entity[1].".

The main properties of trust is dynamicity, subjectivity, asymmetricity, context dependency and transitivity[2].

The first one who define "Trust Management" is Blaze et al. [3] who expressed that "Trust management provides a unified approach for specifying and interpreting security policies, credentials, and relationships.".

Solhaug et al. [4] defined trust management as "a special case of risk management with a particular emphasis on authentication of entities under uncertainty and decision making on cooperation with unknown entities.".

However, the application of trust management has been extended from authentication to various decision making mechanisms in communications and networking, such as access control, intrusion detection, key management. Trust management, including trust establishment, trust up-date, and trust revocation[2].

Trust is useful only in an environment characterized by uncertainty and where the participants need to depend on each other to achieve their goals[5].

Trust models can be classified into individual based models which the decision making process can be fulfilled according to the direct and indirect interactions among entities and system based models which set some rules, protocol and/or mechanisms such as incentive mechanisms that force the user to be trustworthy. By using the composition of these models the systems or networks can be more trustable [6].There are many ways to calculate the trust in both individual and system based models (some of which are described in [2, 5, 7]), but the problem is how to propagate this rating over the graph of trust and the problem become immense as the size of graph increase.

Nowadays, rapidly growing communications lead to trust graphs grow in size. Therefore, the probability of interaction between two nodes continuously degrades. On the other side, this increases the probability of malicious attacks like collusion attack, Sybil attack and On-Off attack.

Trust rate propagation in a large scaled graph, and trust value estimation between a pair of nodes are quite challenging. With limited information and communication about nodes, how is it possible to communicate to reduce risk, cost, and resource consumption? However, proposed methods in trust propagation have been able to challenge the possibilities in this problem effectively.

Some methods and approaches have been suggested by different researchers are not scalable. In this paper to improve the scalability of propagation method we apply a group management method. As Ren et al.[8] described "traditional group management schemes are mainly classified in three categories: 1) centralized management, in which all nodes must obey the management from a central authority; 2) hierarchical management, in which a network is divided into different layers based on some predefined rules, and some nodes are elected as leaders to manage their own layer; 3) clustering management, in which all nodes are clustered into different groups and each group has its own group head to control the whole group."In this paper the clustering management is used.

So, we suggest to apply multilevel grouping algorithm such that first divide the large trust graph to some smaller graphs(clusters), then propagate trust among nodes. We use the fast algorithm of Clauset, Newman and Moore (CNM)[9] with complexity near to linear complexity for this purpose.

For better estimation of trust we consider the confidence of trust value in both trust inside group and trust between

groups. The confidence of trust between groups is obtained from intermediate group confidence which is obtained from the links between groups and group confidence of the objective group.

The rest of the paper is organized as follows: second section describes related works. In third section, we introduce our proposed method. In forth section, we report the results of experiments. Finally we conclude the paper and present possible future research in fifth section

II. RELATED WORKS

A variety of research works have been done in the area of trust propagation, Hongjun et al[10]used a simple graph model to represent connections among nodes. In this case, for trust estimation between nodes find different paths between them. So, the final trust is calculated from the sum of multiplication of trust in every path divided by the direct connection of source node. In [11-13] average of weighted sum of trust scaling is used for trust aggregation. Zhao et al. [14] used most trusted path instead of shortest path. The weakness of this method is that the rating of the longest path is extremely punished and by default far nodes determined to be distrusted nodes and result in false positive while trust group may be formed.

In some works a bio inspired techniques have been applied. Gómez Mármol et al. [15]applied ant colony and fuzzy methods. The output of bio-inspired ant colony technique is optimal while fuzzy make the model inferable. Selvaraj et al. [16] used genetic algorithm for selecting trusted paths from service requester to service provider. The trust value of path is used as fitness function.

Group-based methods have been used in some works. Ejei et al[17] used the concept of metagraph[18] to represent trust relationship between person-person, person-group, group-person and group-group. Wen et al.[19] used grouping method CMW for detection of Sybil attack and collusion attack. He also used the similarity of node's behaviors for this purpose. Gummadi et al.[20] to reduce the storage space and communications in peers used grouping method. So, initially construct a matrix of local trust based on the previous interactions of peers. Then, the matrix is divided into groups. The trust value between groups has been obtained by simple averaging the local trust value between the members of groups. Each peer only save one value for trust between groups and the trust values of its group members.

III. OUR PROPOSED METHOD

We propose a group-based trust method for large graph of trust. We use CNM algorithm [9] which is an agglomerative algorithm with a very low complexity close to linear complexity compared to other general clustering algorithms to find group structure in a graph. Then, for trust estimation inside each group and between two groups we derived some equations based on global and local trust values. For better trust estimation we considered the global trust confidence inside group. For trust calculation between groups we consider two factors for confidence of trust. One of them is Inter Group Confidence (IGC) which is obtained from the edges between groups and group confidence (GC) of the target group.

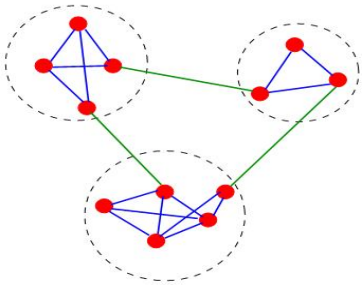

Figure1 : grouping: the connections in group is more than between groups

A. *Grouping*

For large graphs with immense number of nodes and edges, the conventional methods of trust estimation have high computational and time complexity. There are different algorithms to find group structure in networks. For more details refer to [21]. Some clustering methods need a lot of features and have high computational complexity. So, they are not applicable to very large graphs. The other clustering methods do not require many features but they have less accuracy. In this case, there is a tradeoff between accuracy and computational complexity. We use CNM method which has computational complexity near to linear complexity, $O(n \log^2 n)$, and also an acceptable accuracy. In CNM modularity is used for grouping. Modularity measures when the division is good one, in the sense that there are many edges within groups and only a few between them[9].

Let i and j denote two different groups. The Modularity Q is defined as follow:

$$Q = \sum_i (e_{ii} - a_i^2) \qquad (1)$$

$$e_{ii} = \frac{\text{edges in group i}}{\text{edges in the whole network}} \qquad (2)$$

$$a_i = \frac{\text{edges connecting to group i}}{\text{edges in the whole network}} \qquad (3)$$

High value of modularity corresponds to good group structure. In CNM algorithm the group structure is found when the highest value of modularity is gotten[19]. For more details refer to[6].

After grouping, nodes with close relationship reside in the same group. As shown in Fig. 1 the connections in group are more than between two groups.

B. *Trust Estimation Inside Group*

In our proposed method if there is an edge between nodes the trust value is equal to the weight of the edge. Otherwise, if there is no direct edge between nodes, it has to

consider experience of others. The group trust expresses opinions of other ones in group about one node. In the proposed method, the confidence of trust value is used to increase the accuracy of trust estimation. The inside group trust estimation is explained below.

The group trust of node b is computed according to the following equation:

$$GT_b = \frac{\sum_{x \in M_b} LT_{x,b}}{|M_b|} \quad (4)$$

GT_b is the group trust of node b in its group which is saved in the group head. $LT_{x,b}$ is the local trust of each node x that has edge to node b. M_b is the set of nodes that have edges to node b. $|M_b|$ is the size of M_b. It should also be noted that the amount of confidence about others opinions must be considered. The confidence of trust value is affected by dispersion distribution of the opinion, number of people that contributes in trust estimation and the credibility. If the distribution of these opinions is Less dispersed, more people are involved, and contributed nodes credibility is considered; the collective experience of others will be more credible. In formula 5 dispersion effect is considered which is the standard deviation of group trust of node b.

$$D(b) = \sqrt{\frac{1}{M_b} \sum_{x \in M_b} (LT_{x,b} - GT_b)^2} \quad (5)$$

In formula 6 the number of participants has been considered in terms of:

$$F_b = 1 - \frac{1}{|M_b| + 1} \quad (6)$$

F_b is defined as the fans of node b that reflects the effect of the number of nodes that have edges to b.

$$S_b = \sum_{x \in M_b} \frac{GT_x}{|M_b|} \quad (7)$$

S_b is defined as the support of node b which reflects the effect of group trust value of nodes that have links to b.

Finally, the confidence of the global trust value of node b obtained by the weighted sum of three mentioned factors as follow:

$$C_b = \alpha_1(1 - D(b)) + \alpha_2 P_b + \alpha_3 S_b \quad (8)$$
$$\alpha_1 + \alpha_2 + \alpha_3 = 1$$

The coefficients reflect the importance of each factor in the confidence value.

So, If the nodes have an interaction with each other, the trust value of them is according to their interaction and equal to local trust value($LT_{a,b}$).If the nodes have no direct interaction with each other then the trust value is computed as follow:

$$ET_b = C_b . GT_b \quad (9)$$

Input:

- *a sparse matrix representing the trust relationships between every two peers.*
- *Nodes a (source) and b (target)*

Output

Estimated Trust from a to b

// Initialization: divide the graph using the CNM algorithm.

1. *Calculate the Global Trust for each node using eq.(4)and save it in its group head table*
2. *Determine the group of each node.*
3. *Trust Estimation:*

If a and b are in the same Group

//Trust estimation in group

a. Calculate deviation of local trusts using eq.(5)
b. Calculate fans of node b using eq.(6)
c. Calculate support value of b using eq.(7)
d. Calculate confidence of b using eq.(8)
e. Estimate the trust value using eq.(9)

Else

//the pairs are not in the same group

a. Calculate Mean of the Global Trust using eq.(10)
b. Calculate Support B using eq.(11)
c. Calculate group confidence using eq.(12)
d. Calculate IGC using eq.(13)
e. Calculate overall confidence using eq.(14)
f. Estimate the trust value using eq.(15)

Go to step2 and select another pair

Figure 2. A group-based trust propagation method

C. Trust Estimation Between Groups

As we mentioned before, the nodes that locate in group have more close relation with each other than that reside in another group. So, the interactions between groups are weak compared to inside group. To estimate the trust value between two nodes in two different groups, because of the limited edges between groups, the overall confidence must be less than inside group confidence. We use two factors for the confidence calculation between groups. One of them is Intermediate group confidence (IGC) which obtained from direct experience between two groups (direct edges between groups) and the other is the group confidence of the group which is obtained from the target group.

$$MeanGT_B = \frac{\sum_{x \in B} GT_{B,x}}{|B|} \quad (10)$$

$MeanGT_B$ is the average of the global trust value of group B's nodes that are located in group's head table. $GT_{i,x}$ is the group trust value of group i about node x that is located in group head table. B is the set of group B nodes that are located in group head table. |B| is the number of nodes that are located in group head table.

$$S_B = \frac{\sum_{x \in B} C_x GT_x}{\sum_{x \in B} C_x} \quad (11)$$

S_B reflects the importance of global trust value of group. C_x is the confidence of node x in group B.

$$GC_B = \beta MeanGT_B + \gamma S_B \quad (12)$$
$$\beta + \gamma = 1$$

GC_B is the group confidence of group B

$$IGC_{A,B} = 1 - \frac{\sum_{(a,b) \in OL_{A,B}} |LT_{a,b} - GT_b|}{|OL_{A,B}|} \quad (13)$$

$$OL_{A,B} = \{(a,b) | a \in A, b \in B, a \ \ give\ comments\ about\ b\}$$

$IGC_{A,B}$ is defined as intermediate group confidence which reflects the confidence between two groups. This reflects the effect of links between groups and the difference of local view between two groups and group view about the node.

$$C_{A,B} = \mu_1 GC_B + \mu_2 IGC_{A,B} \quad (14)$$
$$\mu_1 + \mu_2 = 1$$

$C_{A,B}$ is the confidence of group A to group B.$\mu_1$and μ_2 show the importance of each factors.

The estimated trust value between nodes from different groups:

$$ET_{A,b} = C_{A,B}.GT_b \quad (15)$$

The overall algorithm is represented in Fig. 2.

IV. Experiments and Results

The objective is to determine whether it is possible to make an acceptable estimation about the trust value of unknown entity based on the proposed trust propagation method. For this purpose we check the correlation between direct and indirect or propagated experience. We compare our method with IMS.

A. Dataset

A real trust graph from advogato.org[22] is used for the dataset of the experiments. Advogato is an online community site dedicated to free software development in which users can certify each other into 4 distinct certification levels: Observer, Apprentice, Journeyer, and Master. The dataset is a text file of graph that including about 71000 lines of data which contain about 14000 vertices (users) and about 51000 directed edges (links). Mapping these levels into the numbers in the range [0, 1] is left to the user. We considered the numbers 0, 0.33, 0.66 and 1 as the numerical equivalent of observer, Apprentice, Journeyer, and Master, respectively as used in [21]. We also considered 0 for the cases where a programmer has not stated any opinion on another programmer.

B. Experiment

To evaluate and compare the accuracy of the proposed method, we used the leave-one-out technique which is a common validation method in trust research works. So that, we calculate the estimated trust between every pairs of node that there is an edge between them and then compute the difference. For the measure of accuracy we calculate correlation as defined in[23] and mean of absolute error of direct trust and estimated trust.

The coefficients that we used in these functions are obtained by search in some test data. The value of coefficients has been selected as follow: the coefficients α_1, α_2 and α_3 take values 0.85, 0.1 and 0.05 respectively. The coefficients β, γ, μ_1 and μ_2 take values 0.5, 0.95 and 0.05 respectively.

In IMS method the estimated trust value is calculated by multiplication of every edge in the trust chain. If there are multiple paths between two nodes the trust value obtained through average, maximum or minimum trust value of different path.

So, first we extracted the direct edge weights (certificate level) from the graph then we calculated the estimated trust value according to the group-based trust calculation method and IMS method. Finally we obtain the MAE and correlation of each method to estimate the accuracy of the methods.

C. Results

The results are shown in table I. According to the results, mean of absolute error has been decreased in our method about 0.06 and the correlation increased about 0.06 comparing to IMS. The results indicate that the accuracy of group-based trust propagation has been improved compare to IMS method. A scatter plot of the direct trust values and the corresponding propagated trust values is given in Fig. 3 and Fig. 4 The results indicate strong positive linear correlation exists between direct trust and propagated trust in our proposed method.

The values of direct trust and propagated trust are obtained independent of each other in the experiment. Considering large amount of nodes in graph, the results are satisfiable.

TABLE I. THE RESULTS

Method	Correlation	MAE
IMS	0.61	0.25
Proposed method	0.67	0.19

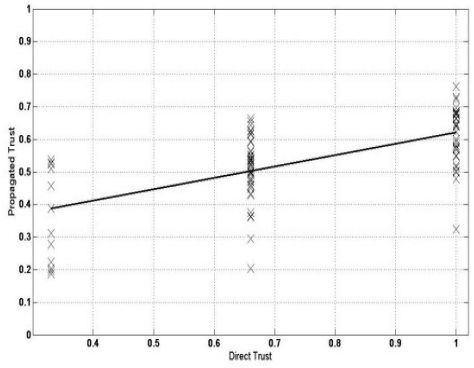

Figure 3. Correlation between Direct Trust and Propagated Trust in a group-based propagation method

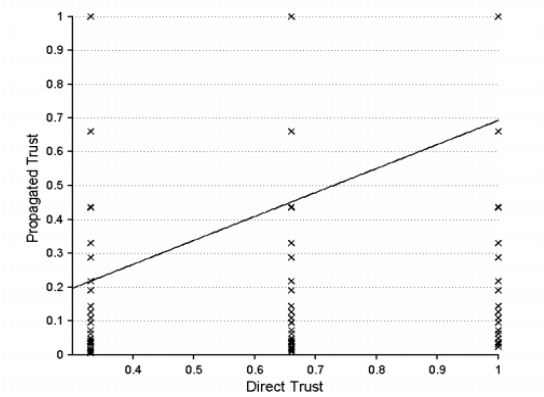

Figure 4. Correlation between Direct Trust and Propagated Trust in IMS

V. CONCLUSION AND FUTURE WORK

In this paper we proposed the group-based trust propagation method. By considering the confidence of trust value in both inside group and between groups our proposed method has acceptable accuracy and the results show that this method is improved compared to IMS method. In addition, using group-based method improved computational complexity and speed. In the future work, we aim to use optimization methods to compute the coefficients in the equations. Moreover, the structure of the groups is hierarchical and solely depends on the connections of nodes. Therefore, considering other features of trust in grouping method could increase the accuracy. This method assumed to be static, but interactive environments are almost dynamic environments and nodes may be added to the group or leave the group by which affects the estimated trust value.

REFERENCES

[1] K. S. Cook, *Trust in Society* vol. 2. New York: Russell SageFoundation Series on Trust 2003.

[2] C. Jin-Hee, A. Swami, and C. Ing-Ray, "A Survey on Trust Management for Mobile Ad Hoc Networks," *Communications Surveys & Tutorials, IEEE*, vol. 13, pp. 562-583, 2011.

[3] M. Blaze, J. Feigenbaum, and J. Lacy, "Decentralized Trust Management," in *Proc. IEEE Symposium on Security and Privacy*, ed, 1996, pp. 164 - 173.

[4] D. E. B. Solhaug, and K. Stolen, " , and V. , Austria, pp. 11-18, "Why Trust is not proportional to Risk?," in *Proc. 2nd Int'l Conf. on Availability, Reliability, and Security*, 2007.

[5] J. Esch, "A survey of trust and reputation management systems in wireless communications," *Proceedings of the IEEE*, vol. 98, pp. 1752-1754, 2010.

[6] J. Tweedale and P. Cutler, "Trust in multi-agent systems," vol. 4252 LNAI - II, ed, 2006, pp. 479-485.

[7] Y. W. a. M. P. Singh, "Formal trust model for multiagent systems," presented at the in Proc. 20th Int. Joint Conf. Artif. Intell 2007.

[8] Y. Ren and A. Boukerche, "A secure group management scheme for mobile ad hoc networks," 2010, pp. 429-432.

[9] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 70, pp. 066111/1-066111/6, 2004.

[10] D. Hongjun, J. Zhiping, and D. Xiaona, "An entropy-based trust modeling and evaluation for wireless sensor networks," 2008, pp. 27-34.

[11] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt, "An integrated trust and reputation model for open multi-agent systems," *Autonomous Agents and Multi-Agent Systems*, vol. 13, pp. 119-154, 2006.

[12] R. Baraglia, M. Mordacchini, P. Dazzi, and L. Ricci, "A P2P recommender system based on gossip overlays (PREGO)," 2010, pp. 83-90.

[13] A. Salehi-Abari and T. White, "Towards con-resistant trust models for distributed agent systems," 2009, pp. 272-277.

[14] H. Zhao and X. Li, "VectorTrust: trust vector aggregation scheme for trust management in peer-to-peer networks," *Journal of Supercomputing*, pp. 1-25, 2011.

[15] F. Gómez Mármol and G. Martínez Pérez, "Providing trust in wireless sensor networks using a bio-inspired technique," *Telecommunication Systems*, vol. 46, pp. 163-180, 2011.

[16] C. Selvaraj and S. Anand, "Peer profile based trust model for P2P systems using genetic algorithm," *Peer-to-Peer Networking and Applications*, vol. 5, pp. 92-103, 2012.

[17] M. Ejei, B. T. Ladani, and N. Movahedinia, "A Group Based Trust Model," in *Proceedings of the 18th Iranian Conference on Electrical Engineering (ICEE 2010)*, Isfahan University of Technology, Isfahan, Iran, May 2010.

[18] b. Amit and W. B. Robert, "Metagraphs and their Applications(Integrated Series in Information Systems)," *Springer-Verlage New York*, 2006.

[19] J. Wen, Y. Shoubao, and C. Bo, "A Group-based trust metric for P2P networks: Protection against sybil attack and collusion," 2008, pp. 90-93.

[20] A. Gummadi and J. P. Yoon, "Modeling group trust for peer-to-peer access control," 2004, pp. 971-978.

[21] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, pp. 75-174, 2010.

[22] R. Levien and A. Aiken, "Attack-resistant trust metrics for public key certication.," in *In Proceedings of the 7th USENIX Security Symposium*, ed. San Antonio, Texas,, January 26-29 1998.

[23] O. Hasan, L. Brunie, and J. M. Pierson, "Evaluation of the iterative multiplication strategy for trust propagation in pervasive environments," 2009, pp. 49-53.