Similarity Formation of Groups and Key Management in Dynamic Peer to Peer E-commerce

Felix Musau[1],
[1,2]School of Information Science and Engineering,
Central South University,
Changsha, China
[1]School of Engineering and Technology,
Kenyatta University,
Nairobi,Kenya
musaunf@gmail.com

[2]Muhammad Bashir Abdullahi
[1,2]School of Information Science and Engineering,
Central South University,
Changsha,China
el.bashir02@gmail.com

Abstract— **The ease of shopping and comparing products and prices online has made it an attractive option for many shoppers. When it comes to time to enter personal information or complete a business transaction on the Internet, most people hesitate, and understandably so. They've read about credit card fraud, identity theft, spoofing, hacking, phishing and other abuses. These business transactions take place among peer to peer systems at the edge of the internet. In peer-to-peer environments, a peer needs to interact with unknown entities. In the method proposed, the peers form groups in order to ensure security. Each group is based on interest among the peers and has its own way of operation. The paper gives a clear outline of how peers form groups, select group leaders. Peers can belong to more than one group to a maximum of n groups hence exempt a similarity behavior. This paper also addresses the distribution and management of keys among the peers.**

Keywords-group; e-commerce; keys; Interest; key management

I. INTRODUCTION

Peer-to-peer systems are targeted for information sharing, file storage, searching and indexing often using an overlay network. This paper expands the scope of peer to peer to include groups. A good cooperation among the group members is a fundamental requirement due to anonymity, peer independence, high dynamics and network conditions as a supplement to effective security mechanism. In the research security encompasses both the need for cryptographic security and the need for trust in e-commerce environment. The open, anonymous, uncertainty and dynamism of peers in P2P poses a challenge in e-commerce. Many e-commerce websites have been developed or are emerging, such as eBay [1] and Amazon [2]. The present trust models have not been able to fully address the issue of peers lying and having a conspiracy among them. A lot of drawbacks of the real P2P systems have been disclosed that performance of the systems can't reach or even be proximal to the expectation of users and system designers [3]. The work leverages the directed and undirected graph analogy-based approach and considers the similarity in peer groups. Our contributions are threefold:

1) Group formation of peers in a business transaction based on interest hence similarity model.
2) Key management in trust and security in group environment of dynamic peer to peer e-commerce.
3) Every peer forms a connection with another peer using the optimal path. Inlink weight and outlink weight can be applied to the transactions among the peers.

The paper is organized as follows: In the next section we describe the Related Issues: Section III describes the Preliminaries, section IV deals with group and group formation, Section V deals with Similarity Model in Dynamic P2P E-commerce, section VI key management, section VII deals with Analysis and simulation results to the performance of the new reputation system. Finally, we conclude this paper and shed some light on future works.

II. RELATED ISSUES

Group based approach has been studied for some time, the most acknowledged system is the grouping based mechanism (GDRep) [4]. The core of the GDRep is that the peers are grouped together and each group has a central peer. The main problem of this is that; 1) No definite method in which the central peer is selected; 2) The method doesn't show how the peers can be punished after being dishonest in a transaction; 3) There is no clear method on how the peers communicate to each other and how the data is stored in the group; 4) The method fails to bring security of the small transactions and also big transactions. The most famed type of distributed trust model is the global trust model named Eigen trust [5]. It states that if one wants to learn the global reputation of another peer he has first to study the variation information. GRBTrust[6] model divides the peers in the network into groups according to a single purpose.

III. PRELIMINARIES

Let $G=(V,E)$ consist of a finite, non empty set of vertices V which represent the peers available and a set of edges E, The edges are ordered pairs $(v, w)$of vertices representing a directed graph. v is called the trustor and w is called the trustee of the edge (v,w) . If $(v.w)$ is an edge in E, then we say vertex w is adjacent to vertex v, we can say edge (v, w) is from v to w, the number of vertices adjacent to v is called degree of v. v,w describes a symmetric neighbor relation among nodes $v,w \in V$. We use the notation D for the maximum distance along G from any node in a group or outside the group. The neighborhood of node v in the P2P network is $N(v)=\{w/(v,w)\in E\}$. Each node v maintains a set of identifiers of its neighbors in $N(v)$. Messages can be send from a node v to node w provided v knows the identifier of w. This paper assumes there is an overlay on top of an existing physical network (e.g. internet). Each edge in E, for example, from node i to node j, has two trust factors, namely, trust value $t(i,j)$ and risk level $r(i,j)$,both of which take values from a real interval (0, 1]. On each particular edge there are weights which are combined with each trust value. Trust value is to specify the trust estimation of node i to node j. Confidence value is to describe the accuracy of the evaluated trust value. Reputation-based frameworks,

978-1-4244-5392-4/10/$26.00 ©2010 IEEE

does not include confidence value into opinion formation. In social networks individuals decide with whom they want to establish relationships. The groups are a reflection of shared interest. They manifest themselves in networks as groups of nodes that are densely connected to each other than by random chance. Group identification can be solved based on the maximization of quality function M, called modularity [7]. For any given partition P of a weighted graph into groups, the modularity is defined as

$$M(P) = \frac{1}{2L}\sum_{ij}\left[w_{ij} - \frac{s_i s_j}{2L}\right]\delta_{m_i m_j}$$

Where the sum is over all nodes, w_{ij} is the weight of edge (i,j), s_i is the sum of the weights of all of node i's edges. , m_i is the group to which node i belongs (in partition P) and δ_{ij} is the Kronecker symbol ($\delta_{ab}=1$ if a=b and $\delta_{ab}=0$).The modularity function shows how much edge weight falls within groups as opposed to between groups. If there were no communities in the network, the total connection strength s_i of each node would be evenly distributed among all the other nodes, so that the weights w_{ij} would be proportional to s_i and s_j . The heuristic approach proposed by Duch et al. [8] is used efficiently to explore the space of possible partitions, as the technique provides a compromise between accuracy and speed [18].

IV. GROUPS AND GROUP FORMATION

According to [9] a group is a set of peers or processes which are said to be members of the group, [10] also defines a group as a community that sets up for a certain purpose. A group can be mathematically expressed as a set. Each set has its own members; i.e. if S and T are sets then $S=T$ iff(*)$x \in S$, iff $x \in T$. The notation $\{a,b,c\}$, means the set whose members are a,b,c. The notation $\{x|p\}$ means the set of x such that p is true. The number of elements of a group S will be called the order of S and denoted by $0(S)$. A group S is a singleton, Iff $o(S)=1$,and an n-ton if $o(S)=n \in N$. A set S is a subset of a set T written as S⊂T, * iff(*)if x∈S, then x∈T, If S is a set and T is a set for each i∈S, then

$\cup\{T_i| \ i \in S\}=\{x| \in i \in S$ such that $x \in T_i\}$

$\cap\{\text{Ti}|\ \text{i}\in\text{S}\}=\{\text{x}|\text{if}\ \ \text{i}\in\text{S}$ then $x \in Ti\}$

Groups have no empty set as they are formed by members, each member is 1 or the n members. In our case a group arises from peers transacting business based on interest among themselves. A peer or process becomes a member by requesting to join the group. Each group is associated with a logical group id. Our research recommends that a group should have a membership service of each particular group which can be invoked by the group leader in case any information on membership is needed. If it doesn't get a group to join, it forms a new group based on its interest which we refer as attribute in this paper. It then authenticates the other members of the group. For security of the group and any transaction among its members the group leader initiates a key management procedure.

The Group Authentication algorithm

Group_Authentication_algorithm()

{

Initialize peers in the network;

New peer u broadcasts authentication query to neighboring peers expressing its interest;

IF peer u receives t or more authentication replies THEN

u generates system certificate and joins the group;

ELSE

Peer u broadcasts proxy finding query to neighboring peers; neighboring peers search authentication peer from peer u;

IF peer u gets t authentication replies THEN

u informs neighboring peer to stop finding;

u generates system certificate and join the group;

ELSE

Authentication fails;

}

If the peer is accepted it is issued with a key which is shared by the all the members, the key in our case is called session key, it also has its own private key which is not shared. This key can be combined with a key called pairwise key which is shared by the peer with the group leader called individual key. After the joining peer gets a key signed certificate, new peer v first validates the signature with a verifiable secret sharing scheme (VSS). If the signature is valid, it reserves the certificate, and waits for the next signed certificate .When new node v gets t key signed certificate, it merges them to recover the keys, and joins a group successfully.

V. SIMILARITY MODEL IN DYNAMIC P2P E-COMMERCE

Peers are grouped together to form peer groups. Some peers belong to several groups. Among the peers in a group one peer acts as a group leader. As peers can belong to different groups, this structure forms a Venn diagram structure.

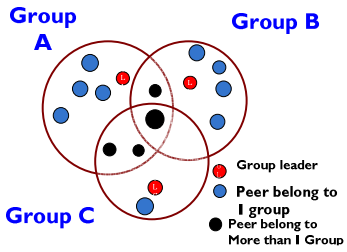


Figure 1: Peer Groups in represented as a Venn diagram

The idea of similarity show that a peer which belongs to different group should broadcast its trust to the group it joins. There should be one group which is referred as the base group. A group can be considered as the reference group. In our method we come up with two types of similarity group

- High intra-class similarity: this one is cohesive in a group
- Low inter-class similarity: distinctive between groups.

The figure below represents peer group e-commerce community such that the peer in a group will have interest similarity to another and different from, or unrelated to peers in other groups.

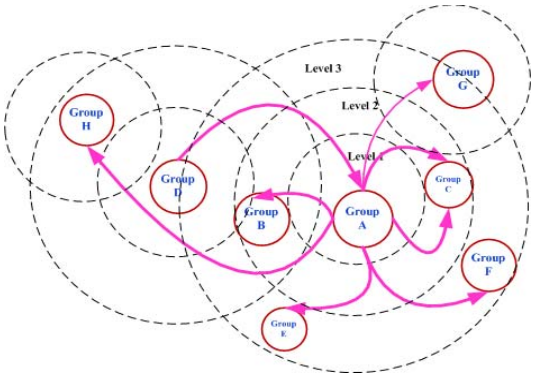


Figure2: Groups formed at different levels based on similarity

In the small world network model, the characteristic path length between nodes is small based on random network in which aggregation coefficient is high. This depends on the level. In our model of similarity the group leader has to be the peer which formed the group. In idea of similarity peers communicate in optimal path. In our model peers at long distances are perceived to have more opportunities to the group than short distances. Trade will be based on similarity of peers to trade on common items. Reputation function is different for interval scaled, Boolean, categorical, ratio, and vector variables. In P2P, peers in different levels help marketers discover distinct groups and potential in their customer bases. In case the group leader leaves, Peers exchange messages and then choose the group leader which takes the administrative roles. Reputation-based trust management is widely recognized as an effective way for an open system to identify, avoid malicious nodes, protect the system from possible misuses and abuses in a decentralized networked computing environment [11]. This method can be a solution to many buyers and sellers who are afraid of anonymous traders in e-commerce transactions. If a peer in London declares an interest in wombats, and a peer in china also declares the same interest, then the two peers become implicit, undiscovered community.

Computing Expected Length

Computing the length between two peers our method uses the greedy routing algorithm on Paths of polylogarithmic expected length between any pair of nodes. The number of nodes in a given distance (level) 2*r* from a given node is at most a constant times the number of nodes within distance *r* or level. The greedy routing algorithm computes paths of polylogarithmic expected length between any pair of nodes [12]. The greedy routing considered proposes that given two nodes, task is to determine which one is the closest to some target node in the graph before augmentation[13]. Dynamically maintaining a precise topology is difficult. The key to small world augmentation processes is that the random distribution of the new links fits the underlying graph. In our method we refer to the set of nodes in a group within distance r from some node u as the ball centre u and radius r, denoted by *Bu(r)* we can denote by *bu(r)* its cardinality. Multi-layers sample scheme is used which is in charge of the link computations of other nodes. Our augmentation process requires the addition of one long range contact per node.

Algorithm Design

Similarity Algorithm (SimTrust)implicitly show how to compute trust metrics when applying the highest link cost of algorithms . The metrics have strong relationship with the trusted graphs from one peer to peer and from one group to another. Given a trusted Graph *G=(V,E)* and two peer nodes v and w, we find the trust value from v to w, denoted SimTrust, and then find the highest edge-disjoint from v to w,

Algorithm:Edge Preprocessing-based SimTrust Algorithm from the perspective of *Highest Trust Value)*

I*nput*: A trusted graph *G= (V, E), v,w* ∈ *V,* and the Trust value *t(i,j),* $\forall$ *(i,j)* ∈ E

Output:The normalized trust value T For each t*(i,j)*Do

{Let δ_i be the node degree of I *(inlink +outlink)*

Let δ_j be the node degree of j*(inlink + outlink)*

If $\delta_i \geq \delta_j$ Then $t'(i,j) = \frac{t(i,j)}{\delta_i}$

Else $t'(i,j) = \frac{t(i,j)}{\delta_j}$ }

Let SimTrust be the highest trust value of G

From v to w with *t'(i,j)*as the edge capacity

Return T=SimTrust

Dynamism

P2P network is dynamic if nodes can subscribe or unsubscribe from the network at will. Nodes that exhibit such behaviors are called faulty nodes. In case a node n detects a fault at node *v* ,*n* removes *v* from its neighbors *N(v)* and recovers from the unsubscribe of *v*.

VI. KEY MANAGEMENT

Chan, Perring, and song proposed a *q*-composite random key distribution scheme [14] for key management. Later Du, Deng, Han, and Varshney proposed a new key pre-distribution scheme [15] which can improve simultaneously the resilience of the network. This scheme has an excellent threshold property in which when the number of compromised peers is less than the threshold the probability that any peers than these compromised nodes can be affected is close to zero. The method proposed is the Mkeying method.

M Keying of peer to peer

Mkeying [16] supports the establishment of five keys for each peer i.e. pairwise key, individual key, session key, encryption key and message authentication code (MAC) key. *Pairwise key*: Each peer shares a pairwise key with each of its immediate neighbors. The pairwise key generate the encryption key and the MAC key. *Individual key*: Each peer has a unique key shared pairwise with the Group Leader. This key is used by the group leader to authenticate individual peers. *Session key*: This is a global key shared by all Peers in the network. The need for the session key is motivated by the key revocation scheme. We present an efficient key revocation scheme, KeyRev, for a Peer with group leader using a globally distributed session key. In case a peer leaves the group, all the keys are revoked and updated in all peers. Before peers are

grouped, the setup server randomly generates a bivariate *t*-degree polynomial

$$f(x, y) = \sum_{i,j=0}^{t} a_{ij} x^i y^i$$

over a finite field F_q where q is a prime number that is large enough to accommodate a cryptographic key such that it has the property of $f(x, y) = f(y, x)$. By exchanging the identifiers between a Group Leader and a peer, both the group leader and the peer can find the shared individual key.

Encryption key and message authentication code key

Encryption key and the MAC key can be used to ensure security of the communication in peers. In e-commerce peers have vital information which is constantly propagated across the nodes. The encryption key and the MAC key are derived from the pairwise key and the session key. We use encryption key K_{encr} and message authentication code key K_{mac} to ensure the security of a message. Let A and B be two entities in a Peer group, K_{AB} be the pairwise key shared by A and B, K_j be the current session key:

$$K_{encr} = F(MAC(K_{AB}, K_j), 1)$$
$$K_{mac} = F(MAC(K_{AB}, K_j), 2)$$

F is a pseudo-random function

MAC(K, R) computes the MAC of message. R with key K.. The complete message A sends to B is:

$$A \rightarrow B : \{M \mid T_s\}_{K_{encr}}, MAC(K_{mac}, \{M \mid T_s\}_{K_{encr}})$$

where M is the message and T_s is the timestamp when sending the message.

VII. ANALYSIS

Let max(x, y) be the maximum value of x and y, Considering the direct relationship, $T(G_1G_2)$ denotes how much group G_1 trusts group G_2. In a system with N groups, trust value is calculated as follows.

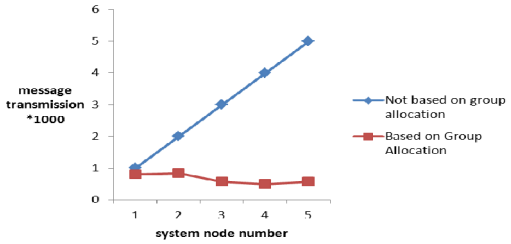
$$Tp_i(p_k) = C(G_1G_2) * \ln\left(\sum_{j=1}^{n} \frac{Tp_i(p_j) * Tp_j(p_k)}{n}\right) + 1$$

$$C(G_1G_2) = \begin{cases} \dfrac{\max(T)(G_1G_2), 0}{\sum_{i=s} \max(T(G_1G_2), 0)}, \sum_{j=s} \max(T(G_1G_2), 0) \neq 0 \\ 1, otherwise \end{cases}$$

The simulations will be done with a peers based on group allocation and others not based on group allocation. In collusive setting malicious peers act similarly to the non-collusive setting. The simulation is based on iteration. In every iteration, each peer randomly selects peers. If the selected peer is a malicious one, it contributes a malicious service with probability q%. Finally, the selected peer is given a rating value according to service provided. When the percentage of malicious peers is no more than 50% (i.e. $p\%<=0.5$) and $q\%=1$), the service selection success ratio is close to 1 after about 20 iterations. When p%=0.5 and q%=0.5, the success ratio drops a little, which still has a high success ratio (about 0.9).

Simulation Results

Simulations on peers are run. The graph shown in the figure shows the results of our simulation. They demonstrate nodes based on group allocation and also not based on group allocation. Peers will discover groups



depending on their interest.

Figure 3: Graph showing peers based on group allocation and not based on group allocation.

VIII. CONCLUSION

The paper concludes with a brief summary of contributions and presentation of future work, including a discussion of assumptions. The similarity of peers in groups is supposed to be the same if the peer belongs to more than one group. The issue of undiscovered community has not yet been addressed, where peers collude to trade together but doesn't disclose to other members of the group that they have another engagement.

ACKNOWLEDGMENT

This paper is supported by the National Natural Science Foundation of China for Major Research Plan under Grant No. 90718034, Fusion-Based Trust Model and Trust Propagation Mechanism for Trusted Network Computing, January 2008-December 2010.

REFERENCES

[1] eBay, http://www.ebay.com

[2] Amazon, http://www.amazon.com

[3] E. Friedman, P. Resnick. "The Social Cost of Cheap Pseudonyms" Journal of Economics and Management Strategy, 10 (2):173-199, 1998

[4]Liu Tang(2009),Grouping based mechanism driven by reputation in p2p e-commerce. WISA'09, pp.510-515.

[5]Kamvar S and Schlosser M. The eigen trust algorithm for reputation management in P2P networks. In: Proc. of the 12th Int'l Conf. on World Wide Web. New York: ACM Press, 2003: PP.640-651.

[6] Li Chen, Liang Ju, The Research of Trust Model Based on Group-Recommend in P2P Network,2008

[7] Newman, M. E. J., and Girvan, M Finding and evaluating Community structure in networks. Phys. Rev. E 69, 2 (2004).

[8] Duch, J., and Arenas, A, Community detection in complex networks using extremal optimization. Phys. Rev. E 72 (2005).

[9] I. Keidar, J.Sussman, K.Marzullo,and D.Dolev, A group membership service for WANs.ACM 2002.

[10] W Ji, S Yang, D Wei, W Lu (2007), GARM: A Group - Anonymity Reputation Model in Peer-to-Peer System, IEEE, p.1-8.

[11]Z Dehua1, Y Zhang, and Y Zhou(2003), Research of Security Architecture for P2P Network Based on Trust Management System.

[12] M.S. Khambatti, K.D. Ryu, P. Dasgupta(2002),Efficient Discovery of implicitly formed peer-to-peer communities, International Journal of Parallel and Distributed Systems and Networks, Vol. 5, No. 4.

[13] N Liu, J Li, L Hao, Y Wu, P Yi(2008), Group-based Trust Model in P2P System Based on Trusted Computing,CSSE,pp.797-801

[14] H. Chan, A. Perrig, and D. Song, "Random key pre-distribution schemes for sensor networks," in IEEE Symposium on Security and Privacy, Berkeley, California, May 11-14 2003, pp. 197–213.

[15] W. Du, J. Deng, Y. S. Han, and P. K. Varshney, "A pairwise key pre- distribution scheme for wireless sensor networks," in Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS), Washington, DC, USA, October 27-31 2003, pp. 42–51.

[16]Y Wang, B Ramamurthy, and Y Xue(2008), A Key Management Protocol for Wireless Sensor Networks with Multiple Base Stations,pp.1625-1629.