Peer-to-Peer Netw. Appl.
DOI 10.1007/s12083-011-0116-4

Group formation with neighbor similarity trust in P2P E-commerce

Felix Musau · Guojun Wang · Muhammad Bashir Abdullahi

Received: 30 April 2011 / Accepted: 17 October 2011

Abstract The simplicity with which products and prices are compared in e-commerce has introduced an attractive option for many online merchants. The completion of online business transactions with personal information provisioning has always been an act that beckons hesitation. Most online traders are highly conscious of various threats and attacks such as credit card fraud, identity theft, spoofing, hacking, phishing, and other abuses, leading to low trust in online business transactions. Online transactions take place among Peer-to-Peer (P2P) systems at the edge of the Internet. Peer communities are often established dynamically with peers that are unrelated and unknown to each other. In our proposed mechanism, peers form groups in order to ensure trust and security. Each group is established based on interest among peers. In this paper, we show how peers form groups, and select group leaders. A peer can belong to more than one group. Comparing with some existing work, our work reveals that peers can have common neighbors which have a similarity based on their interest. Our simulation results show that the model can deal with the malicious attacks efficiently by comparison with existing models.

Keywords P2P · Group · Similarity · E-commerce · Interest · Neighbor · Trust value

F. Musau · G. Wang (✉) · M. B. Abdullahi
School of Information Science and Engineering,
Central South University,
Changsha, Hunan Province 410083, People's Republic of China
e-mail: csgjwang@mail.csu.edu.cn

F. Musau
e-mail: musaunf@gmail.com

M. B. Abdullahi
e-mail: el.bashir02@gmail.com

1 Introduction

Compared with traditional networks, P2P networks are vulnerable to various attacks due to their characteristics. P2P systems are targeted for information sharing, file storage, searching, and indexing, often using overlay networks. P2P e-commerce expands the scope of P2P systems by forming groups based on interest, referred to as electronic communities. There are many examples of electronic communities, e.g., Usenet groups, Yahoo Groups, Google Groups, chat rooms and so on. Applications like IP telephony, video/audio conferencing, online gaming, and file sharing are all increasingly getting organized as groups of peers. Others may exist as social groups such as political movements, professional organizations, and religious denominations. Information sharing, not only within a community, but also among communities, is the major driving force behind trust in P2P networks. Such trust establishments typically involve peers that do not know each other. Cooperation among group members is a fundamental requirement due to anonymity, peer independence, high dynamics, and network conditions for effective security mechanism. There is a need for interactions of peer entities to be secure for each entity, and the environment in which the entity operates. The characteristics of peers in P2P systems pose a challenge in e-commerce where peers transacting business may not have any relation with each other. These results in malicious nodes, whitewashers, and free riders to exist in a system, making it very difficult for many business partners to embrace e-commerce. Many e-commerce websites have been developed or are emerging, such as eBay, Taobao, Yahoo, and Amazon. There is also the issue of peers pretending to gain advantage in business more than the others.

There is a lot of risk in trust transactions without knowing the reputation of those with whom the peers are

Published online: 10 November 2011

interacting with. At the same time there is effect of Sybil and collusion attacks among the peers which is out of scope of this paper. The risks have revived work on trust and reputation mechanisms in P2P e-commerce. The Cambridge dictionary defines trust as "the belief or confidence in the honesty, goodness or skill of a person, organization or thing". This is a measure of how a peer is willing to transact with another peer under the existing circumstances. Hosmer et al. [1] define trust as the expectation that the other parties will behave in accordance with commitments, negotiate honestly, and not take advantage, even when opportunity arises. This definition is more appropriate due to its applicability to virtual transaction based e-commerce.

In our work, any mention of reputation is a way of establishing trust; it is a risk management technique. According to ref. [2], reputation systems were built initially for the sole purpose of promoting cooperation in e-commerce sites like eBay, but have evolved to provide policies in P2P systems. Reputation has been realized to have a lot of drawbacks, i.e., 1) Handling of trust measures; 2) Handling dishonest feedbacks; 3) Handling malicious attacks on peers and taking care of free riders. From a practical point of view, reputation management schemes have already been used in many successful commercial systems. For example, eBay's reputation system discourages fraud behaviors because buyers will usually only bid on high-priced items from sellers whose reputation is high from numerous successful transactions.

In P2P e-commerce, some peers may have malicious intent in any transaction they make. Depending on the influence of the malicious peer, it can adversely affect the interaction of the victim peer with others in the system. A group of malicious peers actively try to subvert the system, they may spread negative feedback of good peers to isolate a particular peer, and the other peers cannot decide whom to trust. Colluding peers can also send positive accounts for each other to increase their trust values in the system [3]. Our approach takes group formation as an initial way in which some of the constraints in P2P e-commerce can be addressed. In addition to the groups maliciousness can be mitigated with duplicate data stored by multiple peers in the group. The scheme guarantees the scalability of peers in P2P e-commerce which remains a challenge.

A lot of drawbacks of the real P2P systems have disclosed that the performance cannot meet the expectation of users and system designers [4–6]. The major reason is lack of an effective cooperation mechanism inherently in P2P e-commerce systems, as not all participants are encouraged to take part in the systems actively and friendly. Our work leverages the directed and undirected graph analogy-based approach, and considers the common neighbor similarity in peer groups. Notion of similarity appear to play a fundamental role in human learning, and psychologists have done extensive research to model human similarity judgment. Tversky et al. [7] contrast model and ratio model represent an important class of similarity functions.

In summary, there are two facts in previous research, i.e., 1) Many e-commerce systems rely on individual peers for doing transactions, which is very risky; 2) Peers always interact with and anonymous peers in the dynamic environment [8].

Our main contribution is summarized as follows:

1. We present group formation of peers based on the interest to transact in an e-commerce environment. The groups increase supervision of peers hence improves security on transactions.
2. Peers which have common neighbors form a similarity group among the neighbors, which contributes to minimize maliciousness.
3. We present an easier way to search for products, as each group broadcasts the kind of goods or services it deals with. Other related products can be advertised among peers in a group.
4. The paper presents a new view of P2P systems at the edge of the internet.

The paper advances the area of P2P networks as it addresses the issue of maliciousness, which characterize the behavior of compromised peers due to their open and anonymous nature of the environment in which they operate. The idea presented serves to thwart the ability of compromised peers that collude to disrupt the P2P e-commerce transactions. Our paper also advances integrity, confidentiality, availability, and access control policies in decentralized P2P applications. We consider a way to establish a secure routing structure over which messages and data can reliably be exchanged in the presence of malicious peers. The P2P network structure harnesses the computing power of capable peers, and impels efficient ontology distribution across peers.

The remainder of this paper is organized as follows: In the next section, we describe the related work. Section 3 describes some preliminaries. Section 4 deals with group formation. Section 5 deals with overview of the proposed scheme. Section 6 deals with performance evaluation. Finally, Section 7 concludes this paper and discusses our future work.

2 Related work

Group-based approach in e-commerce has been studied for some time, both in centralized and distributed trust models. A well-known group-based distributed trust model is the Eigen group trust model in grouped P2P communities [2], which proposed an effective trust system built on top of a

P2P group infrastructure. The model is based on delegation system that not only manages trust within communities, but also between different communities. It had its own weakness, i.e., first it aimed at reducing the issue of overloading the network, which it never achieved satisfactorily. Second, it never took issue of user dynamics in which peers change their ways of operation and become malicious in course of transactions. In their approach how to effectively compute or, evaluate similarity degree between peers in a group is not investigated.

Kamvar et al. [9] proposed a distributed trust model based on global reputation from local reputation, called EigenTrust. In the model a peer collects the local trust values of all peers to calculate the global trust value of a given peer. EigenTrust relies on good choice of some pretrusted peers, which are supposed to be trusted by all peers. The main problem of EigenTrust lies in the following aspects as pointed out in refs [10–12]: 1) The precondition of iteration convergence is unreasonable; 2) It does not provide any punishment mechanism for bad behavior; 3) EigenTrust does not take into account user dynamics, and effect of credibility; and 4) EigenTrust does not bring security into consideration.

Tang et al. [13] proposed a grouping-based mechanism driven by reputation in P2P e-commerce (GDRep), in which peers are controlled by a central peer located in each group. The main problem of the mechanism is that: 1) There is no definite method in which a central peer is selected; 2) The method does not show how a peer can be punished after being dishonest in a transaction; 3) There is no clear method on how peers communicate to each other, and how the data is stored in a group.

In [9], newcomers are trusted based on trust probability which can be adjusted according to the actions of the newcomers. The work gives definite criteria in which the new peers are taken care of. The GRBTrust model [14] assumes that one peer belongs to only one group, which ensures enough security as members can monitor activities of others. In case of a peer belonging to more than one group, the peer can be looked as joining a different group with a different identity. The method has a disadvantage as it restricts a peer to belong to only one group, which is often not the case in practice.

Abdul-Rahman and Hailes [15] claim that given some predefined domain and context, e.g., communities of people reading books, people create ties of friendship and trust primarily with people resembling their own profile of interest. They fail to give direction on how the friendship and interest can be converted to trust if they were based on recommenders.

Chung-Wei et al. [16] proposed trust between a trusting and trusted party must have a basis in some direct relationship (and with respect to a relevant purpose). The relationship in question could be based on or arise from a commercial or social transaction, or through mere participation in common groups, or through an assessment of certain attributes that apply to each party. The work failed to bring the issue of interests in common groups. They propose that in real life, individuals and businesses give referrals and rely enormously on referrals to determine with whom to interact.

The problem with the above schemes is that they do not take into account the neighbor similarity interest of a peer in their dynamic and open nature. In addition the schemes do not consider the trading patterns of peers and the malicious impact. Our work ensures that security is enhanced by peers policing each other. If a peer misbehaves its reputation is affected and with time may reach a threshold level in which the peer is expelled from the group. We propose credibility in our work in addition to normalization for peers to be able to monitor each other in the group, and report any malicious behaviors. In our proposed scheme each peer has a responsibility in the administration of a group in which it belongs. In the group there is a group leader, selected by voting. In ref. [9], it is assumed that a peer belongs to only one group. In contrast, we propose that a peer can belong to more than one group but with its first group as the base group, unless decided otherwise. How a peer makes a decision to change its base group is beyond the scope of this paper. Previous work does not address the idea of choosing the recommenders. Our work proposes that recommenders are initially selected from the neighbor peers who are well known to the concerned peer. Peers have an incentive as they can be able to identify potential business partners according to the trust levels.

3 Preliminaries

In this section, we give some definitions and explanations to form the basis of our scheme.

Definition 1 (Neighborhood graph) A graph G is a tuple $\langle V, E\rangle$, where V is a set of vertices and E is a set of edges. Specifically, $V = \{v_1, v_2, ..., v_x\}$ represents the peers available, and $E = \{e_1, e_2, ..., e_y\}$ represents the edges among the peers. An edge is an ordered pair (v, z) of vertices, where v is called a trustor, and z is called a trustee. If vertex z is adjacent to vertex v, there is an edge (v, z) in E from v to z. Notice that if there is an edge (v, z) in E, then there is also an edge (z, v) in E.

The neighborhood of a node v in a P2P network is $N(v) = \{z/(v,z) \in E\}$. Each node v maintains a set of identifiers of its neighbors in $N(v)$ in which each one is unique. Messages can be sent from a node v to a node z, provided that v knows the identifier of z. Any packet broadcasted by a node is received by all its neighbors. There is an overlay network on top of an existing physical

network for example the Internet. Each edge in E, for example, from node a to node b, has two trust factors, namely trust value $t(a, b)$, and risk level $r(a, b)$, both of which take values from a real interval (0, 1]. For every pair of nodes (v, z) there would be a path from v to z with a short distance. The cost of an edge from u to v is most often different as v to u where $(u, v)=E$.

Definition 2 (Nodes distribution) A graph representing a P2P network should have a low degree, for each node in the graph to ensure a low maintenance cost, easy update in case of arrivals or departures of nodes, and changes in their positions. The nodes are distributed in a 2-dimensional Euclidean space represented by a set of points $V \subset \mathbb{R}^2$, which can also be extended to higher dimensions [17]. Given any pair of nodes $u = (u_x, u_y)$, $v = (v_x, v_y) \in \mathbb{R}^2$,

$$\|uv\| = \sqrt{(u_x - v_x)^2 + (u_y - v_y)^2},$$

denotes the Euclidean distance between u and v, sequence of nodes $s = (a_1, a_2, ..., a_k)$, and any $\delta \geq 0$, $\|s\|^\delta = \sum_{i=1}^{k-1} \|a_i a_{i+1}\|^\delta$, denotes the δ-cost of s. The graph $G=(V, E)$, has a node sequence $s = (a_1, a_2, ..., a_k)$ called a path in G if $(a_i, a_{i+1}) \in E$, for all $1 \leq i < k$.

For a directed graph $G=(V, E)$, and two nodes $a, b \in V$, the δ distance $d_G^\delta(a, b)$ of a and b in G is the maximum δ -cost $\|p\|^\delta$ over all paths p from u to v in G. If $\delta \geq 0$, then $\delta_G^\delta(a, b)$ gives the topological (hop) distance of u and v in G, and if $\delta=1$, $d_G^\delta(a, b)$ gives the Euclidean distance of a and b in G. A trust value specifies the trust estimation that node i puts in node j. A similar concept can be seen in the real world, e.g., in Facebook, edges are friendships among people; in citation networks, nodes are papers, and edges are citations; in web graphs, nodes are webpages, and edges are hyperlinks. The simplest trust value cannot reflect the accuracy of formed trust, which shows subjectiveness.

Definition 3 (Peer modularity) In P2P e-commerce, individuals decide with whom they can establish business relationships based on trust. The number of groups and their sizes are unknown, which makes the problem of community identification qualitatively different, and much challenging than graph partitioning. Group identification could be solved based on the maximization of quality function Q, usually called modularity [18], For any given partition R of a weighted graph, the modularity is defined as:

$$Q(R) = \frac{1}{2N} \sum_{ij} ((2N)w_{ij} - s_i s_j)(2N)^{-1} \delta_{m_i m_j}$$

where the sum is for all peers, $w_{i\ j}$ is the weight of edge (i, j), s_i is the sum of the weights of peer i's edges. m_i is the group to which peer i belongs (in partition R) and δ_{ij} is the Kronecker symbol ($\delta_{ij}=1$, if $i=j$ and $\delta_{ij}=0$) [19]. We note that modularity is calculated considering the whole network even if some of the peers are removed due to misbehaving, or seceding.

Definition 4 (Vertex similarity) Blondel et al. [20] proposed a vertex similarity measurement between graphs. Given two directed graphs G_A with n_A vertices, and G_B with n_B vertices, a similarity matrix M is an $n_A \times n_B$ matrix where s_{ij} is the similarity score between node i in G_A, and node j in G_B. M can be calculated by a convergent iterative process:

$$M_{k+1} = \frac{BM_kA^T + B^TM_kA}{||BM_kA^T + B^TM_kA||_F} \tag{1}$$

where A and B are the adjacency matrices of G_A, and G_B. Here M has all entries equal to 1, and $\|.\|_F$ is the square root of the sum of the squares in all entries. The denominator normalizes M_{k+1} to [0, 1]. The limit of this convergent process is M. The convergence can be determined by:

$$\|M_{k+1} - S_k\|_F < \in \tag{2}$$

where $\in$ denotes error tolerance. Each vertex B_i in G_B, is associated with two similarity scores, say m_{i1} (for A_1) and m_{i2} (for A_2), each of which corresponds to the similarity between one vertex in G_A and B_i. Both scores are initialized to one, and then updated according to the mutually reinforced relation:

$$\begin{cases} m_{i1} = \sum_{j:(i,j)\in E_B} m_{j2} \\ m_{i2} = \sum_{j:(j,i)\in E_B} m_{j1} \end{cases} \tag{3}$$

The vertex is similar to A_1 if it connects many vertices that are similar to A_2, hence is also similar to A_2, if it is connected by many vertices that are similar to A_1. The update process is iterated as the scores m_{i1} and m_{i2} mutually reinforce each other. The work shows the update process converges to a state, which corresponds to the similarity scores between A_1, A_2, and B. The idea is based on calculating the vertex similarity between the trust network, and a structure graph. The trust-based recommendation problem can be transformed into a graph similarity problem, and the scores can be viewed as a measurement of how many good connection.

4 Group formation

Keidar et al. [21] proposed a group as a set of peers, or processes. Ji et al. [22] defined a group as a community that is set up for a certain purpose. When a group is established,

it should announce the purpose of establishment, and declare that it carries out the transactions related to its purpose. A group can mathematically be expressed as a set. A set supports operations, i.e., membership, union, intersection, subset, power set, cartesian product, and complementation. So in a group $x \in (A \cup B) \leftrightarrow x \in A$ or $x \in B$ and $x \in (A \cap B) \leftrightarrow x \in A$ and $x \in B$, as clearly shown by Fig. 1:

A group has no empty set, i.e., it can only exist if there is at least one member, and at most n members. A group is dynamic in nature, as it involves peers transacting business with varying interest changes [13], where peers with similar interests form interest community. A peer may join a group by sending an application message. Receiving a positive acknowledgement message is an indication a peer is qualified to apply for membership. The group leader gives the new peer temporary Id, then the peer continues to negotiate for admission. The peer may then send a join message, and receive a second acknowledgement. The second acknowledgement message include a credential, and advertisement to assure the peer full membership of the group. In case it does not get any to join, it forms a new group. Each group introduces a group charter that specifies the rules each member has to follow. Peers join groups with different motives in P2P e-commerce, i.e., to share services, or content, transact with other peers, publish a peer group advertisement that allow potential peer members to discover peers with same interests. For security of a group, the group leader initiates a key management.

Any peer joining a group is assigned a unique identifier n_j, where j=*0, 1, ..., (N - 1)*, and N is the number of peers in the group. A peer identifier that is computed as in Chord, by hashing the IP address of the node. A peer p is a member of a group G defined as:

$$a^n = \begin{cases} a^n = aa...a & n > 0 \ (n \text{ of } a) \\ e & n = 0 \\ a^{-1}a^{-1}...a^{-1} & n < 0 \ (|n| \text{ of } a^{-1}) \end{cases}$$

The order $|G|$ of a group G is its cardinality. A finite group whose order is a power of a prime p is called a p-group. In

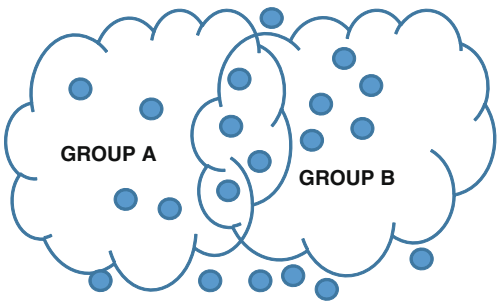

Fig. 1 Group showing set operations where peers exist in more than one group

case there is another group in which the element is to power m, the rule holds that:

$$a^m a^n = a^{m+n}, (a^m)^n = a^{mn},$$
$$a^m a^n = a^{m+n}, (a^m)^n = a^{mn}, \text{ all } m, n \in \mathbb{Z}.$$

From above, the set is: $\{n \in \mathbb{Z} | a^n = e\}$

A peer admitted to a group is issued with a key shared by all the members, called session key. Each peer has also a unique pairwise key which it shares with its neighbors hence can interact with each of them, and a private key used for digital signature which is not shared with others. Encryption is suitable for scenarios where an authorized peer outside the network needs to send a private query to a peer inside. After a peer joins a group, it gets a key signed certificate, in which a signature can be validated with a verifiable secret sharing scheme (VSS) [23]. The joining and leaving of a group can be modeled as a continuous time stochastic process. The process is characterized by a rate parameter, also known as intensity, such that the number of events in time interval $(t, t+\tau]$, follows a poison distribution with associated parameter $\lambda\tau$. The relation is given as:

$$P[(N(t+\tau) - N(t))] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}, \ k = 0, 1, ...,$$

where $N(t+\tau)-N(t)$ is the number of events in time interval $(t, t+\tau]$. The process is characterized by its rate parameter λ, which is the expected number of events.

In summary P2P e-commerce groups are not only a natural extension for arranging distributed systems, but also to enhance the capabilities of each member. Groups are useful in structuring the information storage space, discovering resources, and pruning the search space.

4.1 Group formation algorithm

The formation process of a group is to gather the M peers into N groups. The basis of group formation is to use the distance between group leaders in existing groups to estimate the distance between two member peers. We assume a peer will form or join a group which is near to it, if trade items are of interest. The optimization criterion for group formation is to minimize the n-Average Error as follows:

$$\text{Minimise } \frac{1}{M'} \sum_{\substack{i,j \in [1..n] \\ i<j}} \sum_{x \in G_i, y \in G_j} (|(x,y) - (S_i, S_j)|)(x,y)^{-1},$$

$$\text{where } M' = \sum_{\substack{i,j \in [1..n] \\ i<j}} |G_i||G_j|.$$

A peer should join a group by checking through the group leader to know the proximity. When a peer x joins a

network an algorithm is used to decide whether it needs to create a new group or join an existing group:

- Peer x joins a network, i.e., the internet, and sends a request with a specified Time-to-Live (TTL) value to nearby group leaders.
- If a leader recieves the request, it replies by providing its group ID.
- If the peer receives a reply within a latency, T_{nearby} and joins the corresponding group.
- Otherwise x creates a new group, and acts as a group leader automatically.

When the group leader leaves the group, a leader voted performs the above process. In the bootstrapping stage, to avoid all peers joining at the same time, each peer sets an exponentially distributed delay timer, and joins the system when its timer expires. In our scheme, we noted a newly joined peer tries to join a group whose leader has nearest distance to it. After a period of time, the peer can decide to join or not the other group nearer, the joining algorithm can be expressed as:

$$S' = \arg\min_{S_j}\{x, S_j\} < (x, S_i), j \in [1, ..., n].$$

The peer x in G_i does not know the exact distances to other group leaders. It only knows (S_i, S_j) by intergroup communications.

The algorithm 1, show group algorithm during the admission of new members. It requires synchronization to function in asynchronous environments, which generate large number of information. The group algorithm takes advantage of local broadcasts. It completes in R steps and produces valid groups. The upper bound on the range of random numbers R, is chosen so as to minimize the number of leadership conflicts. The leaders are at distance r apart. Group leaders should belong to only one group.

Algorithm1: Group formation algorithm

```
1:  Input: Integer R(Upper bound for random numbers)
    Output: Groups ()
2:  Boolean Group Leader, member=false
3:  t_i:=R
4:  r_i:=random[0,R)
5:  While(not member and not groupleader)
6:    t_i=t_i-1
7:    if(t_i-1)
8:      r_i:=r_i-1
9:      If(no CertificateInformation(TrustInformation))
10:       If(first (TrustInformation)="newmember")
11:         member=true
12:   else
13:       Groupleader:=true
14:       Broadcast("newmember")
15: While(t_i≥0)
16:       Listen for other leaders
17:       t_i=t_i-1
18: Return Groups
```

4.2 Certificate renewal and revocation

Certificate enrollment is introduced to allow peers automatically request certificates from the Certificate Authority (CA). Key and certificate have a time limit, after which a renewal is done. A peer renews its certificate before it expires, and also if a corresponding peer secret key is changed or compromised. A certificate renewal is carried out by specifying a renewal time, $T_{renew.}$ During the certificate renewal, a peer broadcasts its current valid certificate, and a future expiration time, $T <$ *(current time+* T_{renew}*)*, to its k one-hop neighbors. Certificate revocation mechanism takes place based on the assumption that all peers monitor the behavior of their one-hop neighbors, and maintain their own certificate revocation lists. The neighboring peers check the system public key, and the certificate revocation list to determine whether to accept, or deny a request from its neighbors.

A certificate revocation is carried out by two mechanisms as suggested in [24], i.e., implicitly, or explicitly. In the implicit mechanism, the certificates are revoked if the expiration time *(*T_{expire}*)* is lesser than the time of issue, plus the time of renewal *(*T_{renew}*)*. We note that existing certificate and key pairs are deleted immediately after key generation.

In summary, for a group to exist, there are three points considered according to [25]:

- *Membership*: Structural features that influence whether a given individual will join a particular group.
- *Growth*: *S*tructural features that influence whether a given group will grow significantly.
- *Change*: A given group generally exists a purpose at any point in time, in our datasets, for example, groups are focused on particular business.

A group A may change its focus of interest over time to become more like a group

5 Overview of the proposed scheme

In this section, we propose interest similarity trust model in P2P e-commerce, and common neighbor similarity trust algorithm.

5.1 Interest similarity trust model in dynamic P2P E-commerce

Peers organize themselves in groups, where there is intersection depending on common interests. A Peer in two groups has an architecture represented as a Venn structure in a geometrical plane. Let $C=\{C_1, C_2, \ldots, C_n\}$ denote a family of n simple closed curves in a plane. The curves finitely intersect, i.e., Let X_i, i=1, 2, …, n, be either the open bounded interior, or the open unbounded exterior

of the curve C_i. We say that C is a Venn structure if all of the $2n$ open regions $X_1 \cap X_2 \cap \ldots \cap X_n$ are nonempty, and connected. If the connection condition is dropped, the diagram is called an independent family. An i-region in a Venn diagram is a connected region that is interior to exactly i of the curves. Group diagrams can be seen as FISC [26], a FISC is a family of Finitely Intersecting Simple closed curves in the plane, with the property that the intersection of the interiors of all the curves is not empty.

Theorem 1.1 *In a FISC of n convex k-gons there are at most* $\binom{n}{2}2k$ *vertices. k-gon designates any convex polygon with at most k sides.*

Proof A pair of convex k-gons can intersect with each other at most $2k$ times; there are $\binom{n}{2}$ pairs.

Theorem 1.2 *In a simple n-Venn diagram of k-gons,*

$$k \geq \left\lceil (2^{n-1} - 1) / \binom{n}{2} \right\rceil.$$

Proof Euler's formula for plane graphs, is combined with the fact in a simple diagram where all vertices have degree four, which implies that the number of vertices in a simple Venn structure is $2n-2$. Combining this with Theorem 1.1, which gives an upper bound on the number of vertices, the inequality as Theorem 1.2 gives, for each n, on the minimum k required to form a simple n-Venn diagram of k-gons.

The idea of similarity show a peer belonging to different groups can still have different identities as per interest. Among the two groups, there should be one group which is referred as a base group, or a reference group. This is the group where the "arrow" originates, as shown in Fig. 2. In a given Peer group, the higher the similarity degree, the more trustworthy recommendations. A peer may have interest to transact with others in distant neighborhood.

We assume that each peer x keeps $k=k(x)$ pointers to other peers. The peers are denoted as $l = \{l_1, l_2, \ldots, l_k\}$, where $l_i \triangleq$ distance between x and i-th pointer. Without loss of generality, l is in strictly ascending order, i.e., $l_i < l_2 \ldots < l_k$. When a request destined for key y reaches peer x, x will forward it to the peer $x+l_i$, where $l_i \leq y-x \leq l_i+1$.

The peer-pair neighborhood (e.g., distance) between nodes x and y is denoted as a function, (x, y). It has been shown that the distance satisfies the triangle inequality [27]. That is, for any three nodes x, y, z in the network, inequality $(x, y) \leq (x, z)+(z, y)$ holds. We can further derive that.

$$|(x,z) - (z,y)| \leq (x,y) \leq (x,z) + (z,y).$$

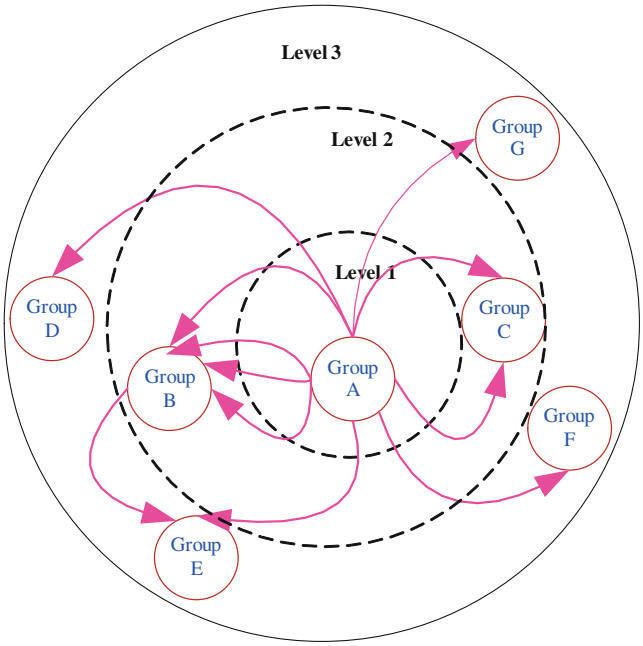

Fig. 2 Groups based on peer interest similarity

In our method, we consider two types of interest similarity groups.

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

In Fig. 4, we compare the similarity of two peers based on their common neighbors. Peers with similar neighbor interest, and service trust value are connected as neighbors. The connection adopts the small world network phenomenon with a characteristic path length. At a random network, the aggregation coefficient from a peer to another is high, but the path length is small, depending on the level as in Fig. 2. A Peer is edged to its neighbors as per similar interest. Peers in a group establishes trust relationships with their distant peers by small world network communication in an optimal path.

Trust, risk, and recommendations are propagated through paths which connect a peer to others on the same or different levels. Before a search, peers should know their prerequisites: 1) local document vector; 2) neighbor peer set; 3) a specified TTL. At each hop, the TTL is reduced by one. If the neighbors do not have feedback for a peer, they forward the path discovery message to their neighbors until the TTL parameter is zero or the peer with the feedback is found. Each time the data is not found, which is detected by a timeout, the source node increases the TTL. In our work, TTL is defined by the levels shown in Fig. 2 and further illustrated by Fig. 3. The TTL value is the maximum search depth. In the topology, groups are formed virtually to ensure security, and reduce risks.

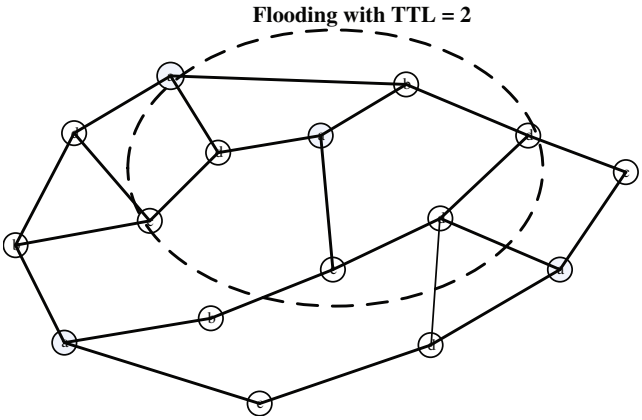

Fig. 3 Showing Flooding with TTL

The directed flooding as in Fig. 3 is being used to the neighbor peers which are of significance to the peer. Our goal is to let peers retrieve the most relevant results they are interested by using little bandwidth. The work in ref. [28] proposed a method of finding good peers in P2P network. Each peer *P* keeps monitoring successful responses. When a peer is found to frequently provide good results and is only two hops away from *P*, *P* then attempts to move closer to that peer by creating a new connection to it.

Study of small-world network started with Milgram's study on social networks which show that any two people in the world are connected via a chain of six acquaintances on average. In particular, Kleinberg considered a family of random graphs having long links with a *2Dn x n* grid with n^2 nodes. In the model each node is with a small set of local contacts and one long-range contact defined by a harmonic distribution. With greedy routing, the path-length between any pair of nodes is $O(log^2 N)$ hops w.h.p. Symphony adapts Kleinberg's construction to arrive at a randomized P2P routing network. It uses a ring (one dimension) and each node with multiple long links (instead of one). The average length of greedy routes has been shown to be $O(log^2 N/k)$, where *k* denotes the number of links per node. Other models of randomized P2P include: randomized-hypercube, randomized-Chord, and skip-graphs [29, 30].

Our model adapts Kleinberg's construction to perceive peers at long distances as having more opportunities to form a group than their counterparts at short distances. Similarity is expressed in terms of reputation function, which is different from scaled, Boolean, categorical, ratio, and vector variables. In P2P e-commerce systems, peers at different levels help marketers discover distinct groups, and potential in their customer bases. The tasks handled by a group leader in addition to performing normal operations include, linking the group to other groups [31], administration, locating a specific file index, and coordination of issuing group key to other peers. In our model of similarity, a group leader has to be a peer which formed a group. The administrator communicates with each of group, its main rights are:1) Make minimum threshold of group trust value, and should be less than its trust value; 2) Manage the peer in group, by allow or refuse other peers to join in, or evict the extreme malicious peer from group. If the administrator withdraws from this group, elect the peer that has relatively large trust value. Voting can be done to ensure fairness, and have a selected peer acceptable to majority.

The groups may be located at different levels which signify small world distances. The organization of peers in a group show exemplary levels of aggregation among the members. Groups and communities can implicitly be formed, i.e., 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. A peer may belong to many groups in a community [32], as it may have varied interest depending on the goods it transacts. We organize peers with similar interest, and service trust value vectors as neighbors in a network, and hence increase the resource location efficiency as well as resource response quality. Peers automatically change their primary and secondary groups over time, adapting to changes in the interests they discover.

5.2 Aggregation in group formation

P2P tasks, has a consideration that an aggregation service should be able to operate in spite of node, and network failures. In the small world network model, the characteristic path length between peers is small, almost random network, and the aggregation coefficient is quite high.

The P2P e-commerce aggregation problem considers a distributed system consisting of a set of *N* peers $\{x_1, x_2, \ldots, x_N\}$.

Aggregation is performed when a single peer initializes a query. The current value should be made available to the querying peer. Since peers may be added and removed from the system, and the value stored on a peer may change during the aggregate calculation, it is not possible to find a fully consistent value. In our work we ensure that the network based on interest group has a higher division of aggregation.

There are two common generic methods of routing and aggregating input data in peer-to-peer systems. The first involves building a tree structure over which data can travel from the sources to a root node, with partial aggregates calculated as branches joined along the way. A peer can edge two groups in which it acts as an intermediary as shown in Fig. 4. The second involves unstructured gossiping of information in which data are randomly routed through the system and the aggregate is calculated at each node. Our method adopts the second method due to the nature of peers which exist as decentralized entities, where each peer acts as both server and client [8].

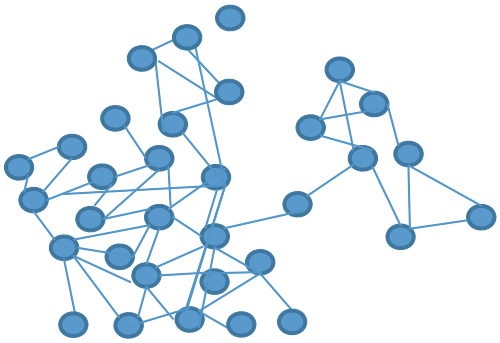

Fig. 4 Peer groups edged by a common peer

5.3 Interest-based neighborhood

If peer x is a neighbor of peer y, the two peers maintain intermediate variables to track each other's values. Peer $V(x; y|x)$ is the value of peer x as known by peer y stored in node x, similarly $V(y; x|x)$ denotes values of peer y as known by peer x stored in node x. In case peer x has n neighbors, $\{y_1, y_2, \ldots, y_n\}$, it stores intermediaries from each destination. Its own value, $V(x; x|x)$, the value of its neighbors as known to itself, $V(y_i; x|x)[y_i \in \{y_1, y_2, \ldots, y_n\}]$. And the value of itself as known to its neighbors, $V(x; y_i|x)[y_i \in \{y_1, y_2, \ldots, y_n\}]$.

This will be $2n+1$values for each destination. This is $O(N.\ d)$ storage complexity in a network with N destination nodes and average node degree d. The variables $V(y_i; x|x)$ and $V(x; y_i|x)$ are called intermediate variables.

A peer can be able to know another peer with the same interest by studying the kind of goods in which they transact. This improves performance as a peer can discover even related goods. Figure 7, illustrates interest similarity, a $peer_i$ is looking for similar goods A, B, and C from $peer_i$ to m. $Peer_j$, has no goods matching the ones needed by $peer_i$. $Peer_m$ has all the three goods; therefore, it shares the same interest with $peer_i$. The interest represents a group of peers trading on a certain type of goods, namely, $\{A, B, C\}$. The goal is to identify peers of the same interest and group them together, or consider them as neighbors of the $peer_i$.

In Fig. 5, a peer i need to have one or more logical neighbors. If it has four neighbors, j, k, l and m then we say it has four edges, $i \rightarrow j$, $i \rightarrow k$, $i \rightarrow l$ and $i \rightarrow m$. Our edge network is based on selected neighbors. Adding edges improves search speed and information access. They are like friendships and improve trust relationships among the peers

Following Liben-Nowell and Kleinberg [33], we define the attributes in Fig. 7 of given pair of peers as intersection of sets of similar products. The function is zero when two peers share no products, it creates a smooth distribution by interpolating between the normalized Adamic-Adar score and a preferential attachment model. Peers that share similar interest create shortcuts to one another. The shortcuts are used to locate products.

Compared with the group concept in sociology, the interest group we propose means a collection of user with similar interests in e-commerce applications. They don't live in a certain geographical area, so the interest group is indeed a virtual group. The basic elements of interest group are users and the interests of users. The users in the same group can share, communicate and cooperate. Users with similar interests will form an interest group and construct direct connections. The connections are also constructed between any two groups. A user may belong to several groups [34]. Sufficient and reliable trust relationships are constructed in the group without relying on any fixed networking infrastructure or centralized entities. Reliable trust relationships are constructed in the group without relying on any fixed networking infrastructure or centralized entities. In our approach similarity concept can also be expressed when a peer purchases a certain product because of purchasing another, i.e., when two items seem to have a similarity interest among the buyers.

In a group, the average distance between two peers P_i and P_j in the network graph G is: $d_{avg}(P_i, P_j)$. The distance defined above allows a given peer P_i to connect to the peers that has same interest within the smallest distance [35].

The transaction works in reference with time: $T = t_{i+1} - t_i$

Let the similarity between any two peer be: $A(P_i, P_j)$. Let Sp_i and Sp_j be the interest spaces for P_i and P_j, respectively. For example, for Fig. 5, the interest space of P_i is ABC and that of P_m is $ACDE$.

The similarity:

$$A(P_i, P_j) = \frac{|S_{P_i} \cap S_{P_j}|}{\min(|S_{P_i}|, |S_{P_j}|)}$$

Intersection between P_i and P_m is A and C. Thus, the similarity of P_i and P_j is computed as follows:

$$A(P_i, P_j) = \frac{2}{3}.$$

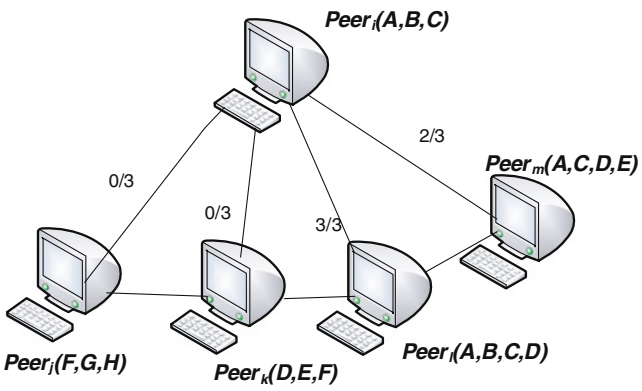

Fig. 5 Interest based neighborhood

When two peers P_i and P_j do not share any interests, the distance $d(P_i, P_j)$ will be infinite.

5.4 Common neighbor similarity trust

Similarity trust is derived from the similarity of the same set of neighbors based on interest in a pair of peers, i.e., p_i, and p_j. We use the Jaccard metric in which the similarity of p_i and p_j is defined as follows:

$$sim(p_i, p_j) = \frac{|p_i \cap p_j|}{|p_i \cup p_j|}.$$

where |.| indicates the cardinality of a set. Note that the denominator $|p_i \cup p_j|$ cannot be null. If $sim(p_i, p_j)$ is not smaller than the similarity threshold S, then the interests of p_i and p_j are similar. The similarity relationship is symmetric, i.e., $sim(p_i, p_j) = sim(p_j, p_i)$.

We can determine the dissimilarity between peers to see if it exists by:

$$sim_\delta(p_i, p_j) = 1 - sim(p_i, p_j) = \frac{|p_i \cup p_j| - |p_i \cap p_j|}{|p_i \cup p_j|}$$

N_i is the set of peer p_i's neighbors, and N_j is the set of peer p_j's neighbors. N_{ij} is the set of common neighbors of p_i and p_j assuming that the feedback is given by the peers which trade with that peer. The common set of neighbor peers that have interacted with both p_i and p_j, is denoted by $N_{ij}^{\bullet} = N_i \cap N_j$, which are in the same or different groups defined as $N_{ij} = N_i \cup N_j$. S_{ij} is the similarity between p_i's trust values and p_j's trust value, about the same set of neighbors. It can be defined by the feedback of p_i's and p_j's trust value about the same neighbors. If $L_i(H_u, j)$ (or $L_j(H_u, i)$) represents p_i's (p_j's) local feedback about p_j (p_i), this shows p_j's (p_i's) behavior in different transactions with the common neighbors, respectively. Considering the set of common neighbors of p_i, and p_j:

$$N_{ij}^{\bullet} = (H_1, H_2, H_3, ..., H_n).$$

Assume $\vec{Q}_i = < L(i, H_1), L(i, H_2), ..., L(i, H_n) >$ is the p_i's trust vector about common neighbors; $\vec{Q}_j = < L(j, H_1), L(j, H_2), ..., L(j, H_n) >$ is p_j's trust vector [14].

We note the importance of credibility, and the normalization of trust values in P2P e-commerce. Credibility, Cr, is used to represent the measurement of recommendation trust. A peer may award higher trust value to the friendly neighbors. To address the problem, aggregation of the trust value can be done to have discrete value between −1 and 1, in consideration to credibility of peers. In EigenTrust [9], the normalization will be between 0 and 1. We normalize using:

$$\lambda_{ij} = \frac{L_i(H_u, j)}{\sum_{u=1}^{n} L_i(H_u, j)}$$

The credibility can be obtained as follows:

$$Cr_{ij} = \sum\nolimits_u L(i, H_u) L_i(H_u, j) \lambda_{uj}$$

And Cr_{hk} is the matrix $[Cr_{hk}]$, and local trust value

Global trust view produced by recursive view of transitive trust,

$$T_{ij} = \frac{Cr_{ij}}{u}$$

Suppose S_{ij} is the similarity between p_i and p_j trust values, about the same set of neighbors, and defined as the cosine of the angle between $\vec{Q}_i$ and $\vec{Q}_j$, then S_{ij} is calculated as follows:

$$S_{ij} = \frac{\vec{\lambda}_i \vec{\lambda}_j}{\left|\vec{\lambda}_i\right| \times \left|\vec{\lambda}_j\right|}$$

$$= \begin{cases} \dfrac{\sum_{x, N_{ij}} \lambda_{ix} * \lambda_{jx}}{\sqrt{\sum_{x, N_{ij}} \lambda_{ix}^2 \sum_{x, N_{ij}} \lambda_{jx}^2}} & if \left\|\vec{Q}_i\right\| ! = 0 \; and \left\|\vec{Q}_j\right\| ! = 0 \\ 0 & if \left\|\vec{Q}_i\right\| = 1 \quad or \quad \left\|\vec{Q}_j\right\| = 0 \end{cases}$$

S_{ij} denotes the similarity of p_i and p_j, while $[S_{ij}]$ denotes the matrix of common neighbor similarity trust, and n denotes number of peers in P2P e-commerce environment. We use cosine similarity, widely known for information retrieval, among most popular approaches used for measuring profile proximity, because we are not interested on detecting negative correlation coefficient. Negative correlation occurs when users have completely diverging interests. Similarity will yield the value of −1 as meaning exactly opposite, *0* meaning independent and +1 meaning exactly the same, with in-between values indicating intermediate similarities or dissimilarities. In peers which do not have any known neighbors, recommendation is used, and can be computed by the following formula:

$$(RS)_{ij} = \sum_k S_{ik} S_{kj}$$

p_i can get indirect similarity with the help of p_h and p_k whose similarity can be calculated directly. It is sensible to weigh p_j's similarity by the similarity of p_h and p_k while taking S_{ih} and S_{ik} as the trustworthiness of p_h's and p_k's feedback about p_j as shown by Figs. 6 and 7 respectively. Indirect similarity can be computed as follows:

$$(RS)_{ij} = S_{ih} * S_{hj} + S_{ik} * S_{kj}.$$

A group formed is viewed at higher level in terms of ability to detect malicious nodes. We explicitly use similarity to characterize the reputation feedback, and credibility to identify malicious peers. Global credibility of a peer in a specific field is determined by recommendations, and transactions it makes

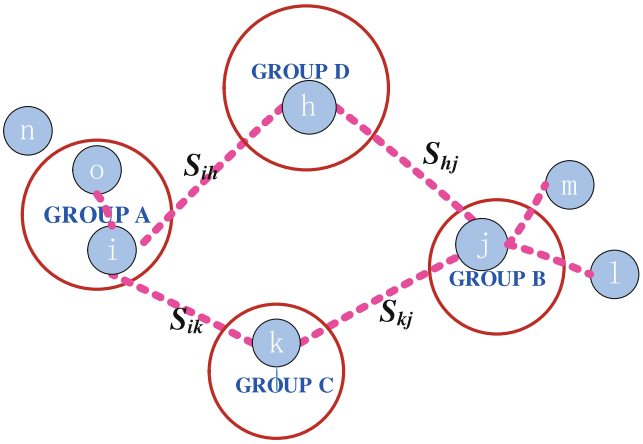

Fig. 6 Indirect common neighbor similarity trust between peers

with others. In peer group, the number of transactions between peers with exterior peers is less, as most peers would like to utilize the benefits of the group. With the neighbor similarity trust method, the malicious behaviors in P2P e-commerce are reduced. The paper compares relatively the limited number and fluctuation rates of peer interests.

5.5 Expected length

Computing the length between two vertices u and v is a set of edges connecting u and v (possibly via intermediate vertices), and having a minimum sum of edge weights. Let a distance $d(u, v)$ be the weight of the shortest u-v path, and a distance between unconnected vertices defined be infinite. The diameter of a graph $dia(G) := \max_{u,v \in V} d(u, v)$, is the length of the longest path between any two vertices. The diameter influences the time of information distribution in the whole network. To get a better view on the whole network, it is important to study the average distance $\overline{d} := \sum_{u,v \in V} d(u, v)$, i.e., the average length of the shortest path between two vertices of G. However, the measure jumps to infinity as soon as the graph becomes disconnected. It is more interesting to look at the average connected distance $\overline{d}$ regarding the paths between connected vertices.

The greedy routing algorithm also computes paths of poly-logarithmic expected length between any pair of peers. The combination of low diameter and ability for each node to discover short paths, without the global knowledge of other connections, shows the small world phenomenon. The greedy routing considered proposes that given two peers, the task is to determine which is the closest to some target node in the graph before augmentation [36, 37].

The key to the small world augmentation process is the random distribution of the new edge to fit the underlying graph so as to create shortcuts at all distance scales. The edge distribution depends on the graph structure properties. The computation roughly requires the exact knowledge of all graph connections. We refer to the set of peers in a group within distance r from some peer as the peer center, u and radius r, denoted by $pu(r)$, as the cardinality. A long range peer becomes a contact neighbor after the augmentation.

The small world augmentation process is inspired from the random link distribution. Random link distributions are caused by external human factors, e.g., a business company employee who uses her computer to occasionally connect to her company, while simultaneously participating as a member of the P2P e-commerce group. The multi-layers sample scheme is used which is in charge of the link computations of other peers.

5.6 Algorithm design

The common neighbor similarity algorithm implicitly shows how to compute trust metrics when applying the highest edge cost. The metrics have strong relationship with the trusted graphs among the peers as well as the groups. There is a strong relationship between the sizes of the trusted graphs, and the highest number of edge-disjoint paths of a trusted graph. Edge disjoint paths problem is NP-Complete and is closely related to the multi-commodity flow problem. A graph is called k-edge-connected if $\gamma \geq k$, i.e., if there exist at least k edge-disjoint paths between. Similarly, it is called k-vertex-connected if $\gamma \geq k$, i.e., between every pair of unconnected vertices if there exist at least k vertex-disjoint paths. If the paths are only required to be edge-disjoint, they can be constructed in polynomial time, using standard maximum flow algorithms. Given a trusted graph $G=(V, E)$, and two peer nodes v and w, we find the trust value from v to w, and

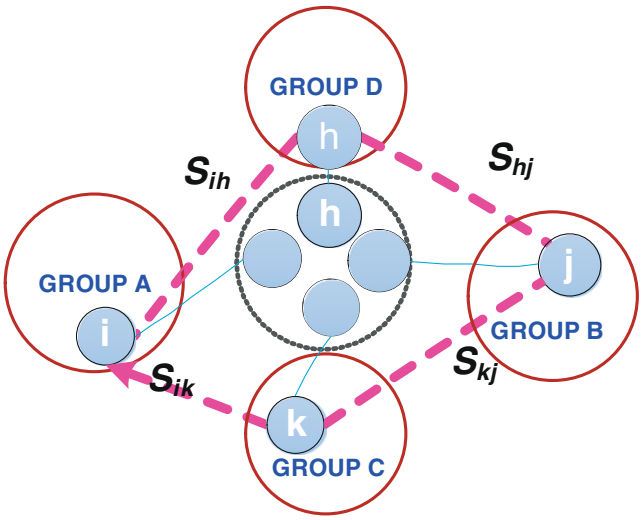

Fig. 7 Common neighbor similarity trust between peers

then the highest edge disjoint. The proposed neighbor similarity algorithm is as follows:

Algorithm 2: Common Neighbor Similarity Trust

Input: Graph $G= (V, E)$, $v, w \in V$, and the Trust value $t (i, j), \forall (i, j) \in E$
Output: Trust value $t (i,j)$
For $i = 1$ to n
  For $j = 1$ to n
    δ_i = (in-edgeSimilarity+out-edgeSimilarity)
    δ_j = (in-edgeSimilarity+ out-edgeSimilarity)
    If $\delta_i \geq \delta_j$ **then**
      $t^{\bullet}(i,j) = t(i,j)\delta_i^{-1}$
    else
      $t^{\bullet}(i,j) = t(i,j)\delta_j^{-1}$
    end if
  end
end

Out-edgeSimilarity is the percentage of outgoing connections from a peer to another peer. *In-edgeSimilarity* is the percentage of incoming connections to a peer that thrive directly from other peers.

5.7 Peer churn

P2P can tolerate structural dynamism due to node joining or leaving. A peer may unsubscribe from a group without broadcasting some information to the others. In real life, behavior of subscribing and unsubscribing may occur when a peer computer is switched off, in which others cannot monitor. Peers that exhibit such behaviors are called faulty peers. The decentralization and dynamism of P2P systems introduce distributed decision making problem, which can be addressed by group formations. Considering the dynamism of P2P e-commerce, the peers should update their neighbors regularly. Peers will change easily their interests due to availability of new products and business opportunities. The dynamics of P2P system could impair the network structure and reduce system performance in terms of, e.g. routing efficiency and fault resilience [12, 38]. The metrics of routing efficiency include for instance, query path length, routing table size and message/time complexity [39]. The resilience of a P2P e-commerce system lies in its ability to maintain its performance in the presence of dynamics.

We consider a protocol that allows a peer to join or leave the network, and then properly recover the network connectivity in view of the change. Chord uses stabilization protocol [40] to regulate the dynamically formed topology. Some peers can exhibit a dynamic personality that is, switching between a honest and dishonest behavior. Behavior changes can be based on the type or value of the transaction or the party involved at the other end. Reputation milkers, or oscillating peers, are one type of peer personality that builds a good reputation and then takes advantage of it to do harm. In a P2P network, nodes may join/leave the network at any time. When a peer p leaves the network, its DHT table will be taken over by the closest neighbor P'. In order to deal with abrupt departure, P' should cache the information kept at P.

5.8 Trust relationships

An algorithm is used to build the relationship between Interest-based groups and peers. The peer looks for those peers of same interest. When joining the network, each peer declares its interest. Since a peer accumulates no trust value at this time, node interest bias similarity (the number of same interest domains between the two nodes) to choose neighbor nodes. If there are N nodes in the network, the definition of network average group coefficient:

$$C = \frac{1}{N} \sum_{i=1\ldots N} \frac{2p_i}{q_i(q_i - 1)}$$

where q_i denotes the number of neighbors of peer H_i, and p_i represents the number of logical connections between the q_i neighbors. Larger network average group coefficient value means better group capability of peers that have similar interest, which can help decrease resource location time. In every trust relationship, trust value has to be used to determine if a transaction is to take place or not. The trust value may be propagated through a transaction pipe (i.e., path). The trust value of a target for a single path, $V_{path}(T)$, from peer S to peer T through peers P_i, $(I, 2, \ldots, n)$ may be calculated as follows:

$$V_{path}(T) = \frac{1}{4n}\left(\sum_{i=1}^{n} V(P_i)\right) \times V(T)$$

$V(P_i)$ denotes the trust value of the peer, P_i, who provides the information. $V(T)$ is the trust value of the target peer, T.

For multiple paths, the final trust value may be the average of all propagated trust values. Assuming there are two paths from peer A to peer D, the first path is through peer B and C. The second one is through B, E and F. C trusts D with a value 3, B trusts C as a recommender with a value 2, and A trusts B as recommender with a value of 3.

$$V_1(D) = \frac{V(B) + V(C)}{8} \times V(D)$$

Using the same method, the trust value of the second path $V_2(D)$to the *nth* path can be calculated and then the average computed.

A pair of neighbor peers may have direct or indirect relationship with each other. The edge values are a measure of how much p_A trusts p_B in e-commerce transactions. Let

max(x, y) be the maximum value of x and y, considering the direct relationship, $T(G_A G_B)$ denotes how much group G_A trusts group G_B. In a system with Q groups, trust is calculated:

$$T(p_i, p_k) = R(G_A, G_B) * \ln\left(\sum_{j=1}^{n} (T(p_i, p_j) * T(p_j, p_k)\right) n^{-1} + 1$$

6 Performance evaluation

Our network model consists of peers interacting and making business transactions. For comparison purpose, we used the Eigen Group Trust model which is able to work in intra group and inter group Trust. The simulation tool used for our experiments is C++. Eigen Group Trust is considered a good model to compare with our model because it aggregates trust information from peer by having them perform a distributed calculation. Our method is superior as it considers the similarity of peers' feedback information to compute the recommendations. Other models which exist include [2, 9], and [13] which we have summarized towards the end of this section.

Different parameters were used in the simulation as shown on Table 1. Freely evolving P2P networks have been shown to exhibit power-law network characteristics [16], hence we organize peers into a power-law network. Upon joining the network, peers connect to a peer i with probability:

$$\frac{d_i}{\sum_{j \in N} d_j},$$

where N is the set of peers currently in the network, and d_i is the peer degree of peer i. We create peer groups and assign peers at random to the groups based on peer's affinity towards a particular category of interest, or a peer's malicious intent. In our work we vary the number of malicious peers that will exist in various groups. Malicious peers are more likely not to reward good services they have received. In the overview, groups have an advantage in business transactions as they are mechanisms to assure confidence in transactions, and peers can identify related products. We provide multiple simulation cycles to simulate multiple queries being generated.

Table 1 Simulation parameters

Number of peers	100
Percentage of malicious peers	0%–80%
Number of interactions	<5000
max number of categories of interest	40
Number of simulation runs	100

6.1 Effect of grouping

In our set of experiments we show the effectiveness of grouping in a P2P e-commerce environment. If the number of peers that have interest similarity p is N, and of which M pairs are neighbors, we define the connection ratio of interest similarity p as, $IR(p) = \frac{M}{N}$. Specific peers have set interest similarity p, the higher of $IR(p)$, the better of its group effect. Managing a group in P2P systems improves scalability of the network, and increases the time taken to undertake a business transaction. The effect of grouping holds in the presence of malicious peers. The groups help to weed out the malicious peers and isolate maliciousness to specific groups rather than allow it to spread out throughout the network. We ran an experiment consisting of 100 peers involved in 100 simulation runs resulting in a total of 1000 interactions as shown in Table 1. Each peer interacted with a random number of peers defined by a uniform distribution. Our P2P e-commerce community has a total of 40 different categories of interest. The transaction interactions between peers with similar interest can be defined as successful or unsuccessful, expressed as positive or negative respectively.

We run three set of experiments: a) Group allocation and without group allocation as shown on Fig. 8; b) Transactions with 0% to 80% malicious peers as shown on Fig. 9, the graph show comparison of Eigen Group Trust and Neighbor Similarity Trust; c) Peers in groups compared with ordinary peers. Every peer in the group will have a particular interest and will interact only with other members of similar interest. Grouping provides peers with similar interest vectors opportunity to get together in a group.

6.2 Peers based on group allocation

The graph in Fig. 8 shows the results of our simulation. It demonstrate comparison of peers based on group allocation and not. It further illustrates the message transmission among

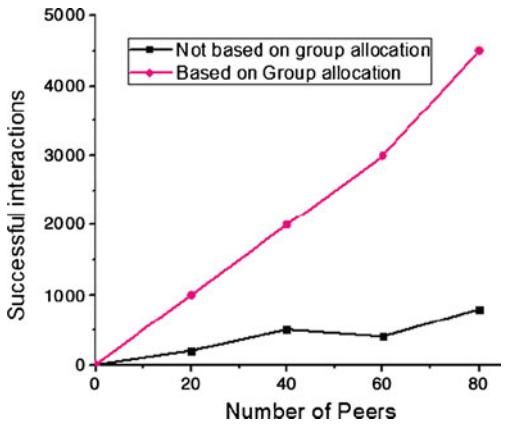

Fig. 8 Comparison of peers based on group allocation and not based on group allocation

the peers in a group. Peers discover groups depending on their interest. Each group has a number of peers which depend on the charter governing it. A group has to be updated as peers attain membership, as trust is a dynamic evolutionary process, which requires assessment of the current trust value after the transaction trust values and history stack. Cosine function is used to simulate the peers. As the number of groups increase, the percentage of successful interactions increase.

As the number of groups increase, the percentage of Monitoring in P2P e-commerce increases.

6.3 Improvement over Eigen Group trust

Evaluation performance of the proposed groups with neighbor similarity interest is compared to the Eigen Group Trust model proposed by Ravichandran and Yoon [2]. For a typical P2P e-commerce, peers can request for services and respond to requests. The services can be provided to conclude an interactional process. The network consists of good peers and malicious peers. Malicious peers are more likely not to reward good services they receive. At times a peer fails to provide the requested service or information when acting maliciously during a transaction. We consider several combinations with different peers in the network. In short, we add a number of uncertain peers to the network such that malicious peers make up between 0% and 80% of all peers in the network. We chose the range to illustrate the variation depending on different percentage number of malicious peers as the ones used by Eigen Group Trust. Our proposed model is based on groups with neighbor similarity trust and compared to Eigen Group trust. The interaction with different percentage of malicious peers is shown in Fig. 9:

From the figures it can be noted that as the number of groups in the e-commerce environment increases, the percentage of successful interactions increase. Less percentage of transactions are bad transactions in our proposed trust model compared to the Eigen Group Trust.

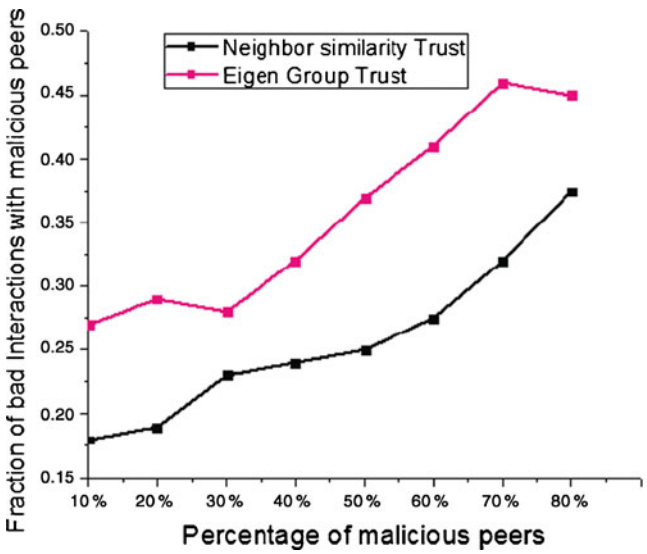

Fig. 9 Transactions between peers in peer groups and malicious peers

6.4 Neighbor similarity and ordinary groups

The simulation is based on iteration, in each iteration, peers randomly select others. If a selected peer is a malicious one, it contributes a malicious service. Finally, it is given a rating value according to whether the provided service is satisfactory or not. Our work addresses situations based on varying group interest. The proposed method enhances reduction of maliciousness among the group members; it also makes it possible to establish trust relationship between peers based on common interest. Groups can organize peers that have similar interest together much better, and thus it helps increase peer resource query hit rate, and decrease the resource location time. At the construction of peers service trust, malicious peers are repelled to network edge, and thus the average malicious path length keeps on increasing. Both interest and trust takes effectiveness in groups to get honest partners. Peers that transact with each other more times, and have similar interests become neighbors with each other.

Figure 10 shows the results of the simulation, the graph demonstrates comparisons of maliciousness in groups based on neighbor similarity, and ordinary groups.

In P2P e-commerce uncertain peers also exist which are more likely to reward the service received than malicious peers. Compared with the peers only based on local reputation, the peers in group are more capable to avoid from interacting with malicious peers. Our experiments show the proposed model based on the neighbor similarity trust is superior to the Eigen Group Trust model, and other models.

6.5 Comparison with other models

We compare the efficiency of our proposed model with the existing models in [2, 9, 13]. The existing schemes in [2, 9]

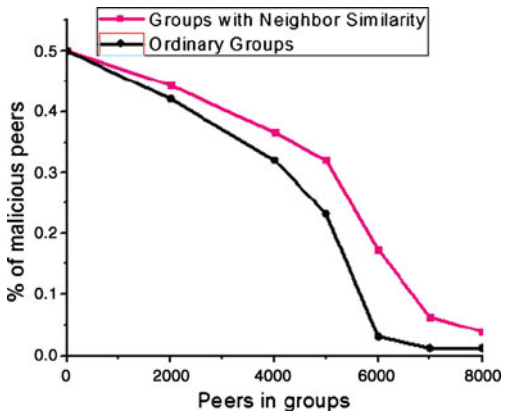

Fig. 10 Comparison of maliciousness in neighbor similarity groups and ordinary groups

are not dynamic. In our scheme peers are dynamic, as they can leave, and join a group anytime. Groups can change interests, as well as their members. In [9] we have pre-trusted peers, while in our scheme peers join groups with a certain set threshold trust value. A peer is proved by the kind of transaction it does. Our scheme and [2] peers form groups depending on the interest, while in ref. [13] a group is formed by peers having same tag. Our model is the only which goes further to form neighbor similarity groups depending on interest. In ref. [2], and our model, delegation is used whereby peers are given some tasks to perform on behalf of others. Credibility in security is a very important aspect for consideration. In the models considered all of them took care of credibility of the peers except ref. [9]. All the models have several issues in common, they all encounter malicious peers during their transactions which acts as a draw back in P2P e-commerce systems. In all models we considered for comparison, whitewashers and free riders exist. Peers face also Sybil, and collusion attacks. In all from the comparisons we can clearly proof that our method is more superior to others. Our method is able to reduce malicious peers to a large extend, as a group has a lot of control to its members.

6.6 Security analysis

In order to ensure the security of the P2P e-commerce transactions, we employ two mechanisms: *key revocation* and *group key refreshing*. When a malicious peer is detected, the group leader broadcasts a revocation message including the identity of the malicious peer to all the group members and the group key is refreshed to prevent possible attacks. When a new peer join the "similarity net" the group key must be refreshed to prevent *backward secrecy*. The *forward secrecy* is ensured as it is computationally infeasible for any set of revoked peers before a particular transaction period T_i to compute a session key K_i used for that period with non-negligible probability, given its session key K_{i-1}. Our model is also better as it effectively compute, or evaluate neighbor similarity degree between peers in a group which is not addressed by the other models.

7 Conclusion and future work

Trust has been studied for many years, particularly in P2P networks, social networks, and mobile ad-hoc networks. Trust in P2P e-commerce is relatively new. Our proposed group formation offers a solution to reduce malicious peers, which exist in the P2P infrastructure. This creates confidence among business partners. The proposed trust model in P2P e-commerce, employs similarity interest trust, based on neighbors which is more operative and secure than other trust models, to maintain trustworthiness. The method reduces malicious behaviors by comparing relatively the limited number, and fluctuation rates of peer's interests. The issue of undiscovered group has not yet been addressed. More on common neighbor similarity should be investigated by use of social communities. Game theory should be incorporated to secure group P2P e-commerce transactions, as the success of one entity is based upon the choices of others

Acknowledgements This work is supported by National Natural Science Foundation of China under grant numbers 61073037 and 61103035, and Hunan Provincial Science and Technology Program under grant numbers 2010GK2003 and 2010GK3005.

References

1. Hosmer LT (1995) Trust: the connecting link between organizational theory and philosophical ethics, Academy of management Journal. Acad Manag Rev 20(2):379–403
2. Ravichandran A, Yoon J (2006) Trust management with delegation in grouped peer-to-peer communities, SACMAT, ACM, pp 71–80
3. Chen P-T, Laih C-S (2011) A challenge-based trust establishment protocol for peer-to-peer networks, Journal of security and communication networks. 4(1):71–78
4. Friedman E, Resnick P (2001) The social cost of cheap pseudonyms. JEMS 10(2):173–199
5. Saroiu S, Gummadi KP, Gribble SD (2001) A measurement study of peer-to-peer file sharing systems. Technical report, Univ. of Washington, Seattle, pp 1–15
6. Adar E, Huberman B A, (2000) Free riding on gnutella, Journal of First Monday, 5(10):1–14
7. Tversky A (1977) Features of similarity. Psychol Rev 84(2):327–352
8. Zhang Y, Zheng H, Liu Y, Li K, Qu W (2011) A GroupTrust model based on service similarity evaluation in P2P networks. Int J Intell Syst 26:47–62
9. Kamvar SD, Schlosser MT, Garcia-Molina H (2003) The EigenTrust algorithm for reputation management in P2P networks, in: Proc. of ACM WWW, pp 640–651
10. Sheng PD, Chuang L, Dong LW (2008) A distributed trust mechanism directly evaluating reputation of nodes. Journal of Software 19(4):946–955
11. Xu JS, Zhong LJ (2007) A reputation-based trust mechanism for P2P e-commerce systems. Journal of Software 18(10):2551–2563
12. Wen D, Huai-Min W, Yan J, Peng Z (2004) A recommendation-based peer-to-peer trust model. Journal of Software 15(04):571–583
13. Tang L (2009) Grouping based mechanism driven by reputation in P2P e-commerce. Academy publisher, WISA:ISBN 978-952-5726-00-8 (Print), pp 510–515
14. Sun L, Jiao L, Wang Y, Cheng S, Wang W (2005) An adaptive group-based reputation system in P2P networks. WINE, LNCS 3828, Springer-Verlag Berlin Heidelberg, pp 651–659
15. Abdul-Rahman A, Hailes S (2000) Supporting trust in virtual communities. In: Proc. of Hawaii international conference on system sciences, IEEE, Maui, HI, USA, pp 1–9
16. Hang CW, Wang Y, Singh MP (2009) Operators for propagating trust and their evaluation in social networks. In: Proceedings of the 8th international conference on autonomous agents and multi-agent systems (AAMAS), Columbia, SC, IFAAMAS, pp 1025–1032
17. Goel A, Sharma A (2002) Performance analysis of mobile Ad-hoc network using AODV protocol. International Journal of Computer Science and Security (IJCSS), pp 333–343

18. Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Journal of Phys, Rev E69, The APS, pp 1–15
19. Musau F, Abdullahi MB (2010) Similarity formation of groups and key management in dynamic peer to peer e-commerce, In: Proc computational intelligence and software engineering(Cise), IEEE, ISBN:978-1-4244-5391-7, pp 1–4
20. Blondel VD, Gajardo A, Heymans M, Senellart P, Dooren P (2004) A measure of similarity between graph vertices: applications to synonym extraction and web searching. SIAM Review, 46(4):647–666
21. Keidar I, Sussman J, Marzullo K, Dolev D (2000) A Group membership service for WANs, ACM Transactions on Computer Systems, (2002), 20(3):1–48
22. Ji W, Yang S, Wei D, Lu W (2007) GARM: A Group - anonymity reputation model in peer-to-peer system. In: Proc of GCC, IEEE, ISBN: 0-7695-2871-6, pp 481–488
23. Dehkordi MH, Mashhadi S (2008) Verifiable secret sharing schemes based on non-homogeneous linear recursions and elliptic curves. Elsevier, (9):1776–1784
24. Capkun S, Buttyan L, Hubaux JP (2003) Self-organized public-key management for mobile Ad hoc networks. IEEE Transactions on Mobile Computing, (2):52–64
25. Backstrom L, Huttenlocher D, Kleinberg J (2006) Group formation in large social networks: membership, growth, and evolution. ACM press 2006, Philadelphia, Pennsylvania, USA, pp 1–11
26. Bultena B, Grunbaum B, Ruskey F (1999) Convex drawings of intersecting families of simple closed curves, in:Proc. CCCG, pp 18–21
27. Francis P, Jamin S, Jin C, Jin Y, Raz D, Shavitt Y, Zhang L (2001) IDMaps: a global internet host distance estimation service. IEEE/ACM Transactions on Networking, 9(5):525–540
28. Ramanathan MK, Kalogeraki V, Pruyne J (2002) Finding good peers in peer-to-peer network. In:Proc. parallel and distributed processing symposium (IPDPS), ISBN: 0-7695-1573-8, pp 24–31
29. Zhang H, Goel A, Govindan R (2003) Incrementally improving lookup latency in distributed hash table systems. In ACM SIGMETRICS, pp 114–125
30. Harvey N J A, Dunagan J, Jones M B, Saroiu S, Theimer M, and Wolman A (2002) SkipNet: A Scalable Overlay Network with Practical Locality Properties, Technical report MSR-TR-2002-92, Microsoft Research, pp 1–14
31. Lu JC, Lin H, Huang K Constructing peer-to-peer networks using common interest grouping. In: Proc. on Consumer Electronics and Signal Processing (WCEsp 2005), pp 1–6
32. Chen L, Liang J (2008) The research of trust model based on group-recommend in P2P network. In: Proc on Computer Science and Software Engineering IEEE, ISBN: 978-0-7695-3336-0, pp 845–848
33. Liben-Nowell D, Kleinberg J (2003) The link prediction problem for social networks. In: Proc. CIKM, ACM Press, pp 1–4
34. Mo Hai, Yan Tu (2010) A P2P e-commerce model based on interest community, international conference on management of e-commerce and e-government. IEEE, pp 362–365
35. Daishi K, Kaoutar E, Kazuo K, Keiji Y, Pietro M (2010) A scalable interest-oriented peer-to-peer pub/sub network, peer-to-peer Netw. Appl.-science and business media, springer science and business media, LLC, pp 165–177
36. Schlosser MT, Condie TE, Kamvar SD (2003) Simulating a file-sharing P2P network. Technical report, Stanford University, USA, pp 1–11
37. Liu N, Li J, Hao L, Wu Y, Yi P (2008) Group-based trust model in P2P system based on trusted computing. CSSE, IEEE, ISBN: 978-0-7695-3336-0, pp 797–801
38. Xuan D, Chellappan S, Wang X (2004) Resilience of structured P2P systems: analysis and enhancement. In handbook on theoretical and algorithmic aspects of sensor, Ad Hoc wireless and peer-to-peer networks. CRC Press LLC, ISBN: 0849328322
39. Xu J, Kumar A, Yu X (2004) On the fundamental tradeoffs between routing table size and network diameter in peer-to-peer networks. IEEE Journal on Selected Areas in Communications 22(1):151–163
40. Stoica I, Morris R, Karger D, Kaashoek MF, Balarikishnan H (2001) Chord: a scalable peer-to-peer lookup service for internet applications. ACM SIGCOMM, San Deigo, USA, pp 149–160

Felix Musau received BED (Honors) in Computer Science/ Mathematics from Kenyatta University, Nairobi-Kenya, a Postgraduate Diploma in Computer Science and Technology, and M. Eng in Computer Science and Technology from Central South University, China. He is currently a Ph.D. candidate at the Trusted Computing Institute, Central South University, China. His current research interests include network and information security, trust management in peer to peer e-commerce, and Internet of things.

Guojun Wang received B.Sc. in Geophysics, M.Sc. in Computer Science, and Ph.D. in Computer Science, from Central South University, China. He is now Chair and Professor of Department of Computer Science at Central South University. He is also Director of Trusted Computing Institute at Central South University. He has been an Adjunct Professor at Temple University, USA; a Visiting Scholar at Florida Atlantic University, USA; a Visiting Researcher at the University of Aizu, Japan; and a Research Fellow at the Hong Kong Polytechnic University. His research interests include network and information security, Internet of things, and cloud computing. He is a senior member of CCF, and a member of IEEE, ACM and IEICE.

Muhammad Bashir Abdullahi received B.Tech (Honors) in Mathematics/Computer Science from Federal University of Technology, Minna-Nigeria and M.Sc. in Computer Science from Abubakar Tafawa Balewa University, Bauchi-Nigeria. He is currently a Ph.D. candidate at the Trusted Computing Institute, Central South University, P. R. China. His current research interests include trust, security and privacy issues in wireless sensor and ad hoc networks, Internet of things and network and information security.