Social Context-Aware Trust Network Discovery in Complex Contextual Social Networks

Guanfeng Liu, Yan Wang and Mehmet A. Orgun
Department of Computing
Macquarie University
Sydney, NSW 2109, Australia
{guanfeng.liu, yan.wang, mehmet.orgun}@mq.edu.au

Abstract

Trust is one of the most important factors for participants' decision-making in Online Social Networks (OSNs). The trust network from a source to a target without any prior interaction contains some important intermediate participants, the trust relations between the participants, and the social context, each of which has an important influence on trust evaluation. Thus, before performing any trust evaluation, the contextual trust network from a given source to a target needs to be extracted first, where constraints on the social context should also be considered to guarantee the quality of extracted networks. However, this problem has been proved to be NP-Complete. Towards solving this challenging problem, we first propose a complex contextual social network structure which considers social contextual impact factors. These factors have significant influences on both social interaction between participants and trust evaluation. Then, we propose a new concept called QoTN (Quality of Trust Network) and a social context-aware trust network discovery model. Finally, we propose a Social Context-Aware trust Network discovery algorithm (SCAN) by adopting the Monte Carlo method and our proposed optimization strategies. The experimental results illustrate that our proposed model and algorithm outperform the existing methods in both algorithm efficiency and the quality of the extracted trust network.

Introduction

Online Social Networks (OSNs) have been used as a means for a variety of activities. For example, according to a survey on 2600 hiring managers in 2009 by CareerBuilder (careerbuilder.com, a popular job hunting website), 45% of those managers used social networking sites to investigate potential employees. The ratio increased to 72% in January 2010. In such an activity, trust is one of the most important factors for participants' decision making, requiring approaches and mechanisms for evaluating the trustworthiness between participants who are unknown to each other.

In OSNs, each node represents a participant and each link between two nodes corresponds to a real-world or online interaction. For adjacent participants (i.e., those nodes with a direct link between them), the trust value between them could be explicitly given by one to another based on their direct interaction (e.g., T_{AB} and T_{BC}). As each participant usually interacts with many others, multiple trust paths may exist between nonadjacent participants (e.g., path $A \rightarrow B \rightarrow E \rightarrow G$ and $A \rightarrow B \rightarrow E \rightarrow F \rightarrow G$ in Fig. 1) from the source participant (e.g., A) to the target participant (e.g., G), which forms a *trust network*. As indicated in the disciplines of Social Psychology (Mansell and Collins 2005) and Computer Science (Golbeck and Hendler 2006; Liu et al. 2010), such a trust network can provide the basis for evaluating the trustworthiness of the target as it contains some important intermediate participants, the trust relations between those participants and social context. In the literature, there have been some existing trust evaluation approaches that evaluate the trust value between any two nonadjacent participants (Golbeck and Hendler 2006; Liu, Wang, and Orgun 2011a; 2011b). However, they all assume the trust network between the two participants have been identified. Namely, extracting such a contextual trust network between two nonadjacent participants is an essential step before performing any trust evaluation between them. The trust network discovery aims to identify a trust network between two nonadjacent participants with less nodes and/or less links but including the important intermediate participants, their trust relations, and the social context. With such a process, trust evaluation methods can be more efficient and effective.

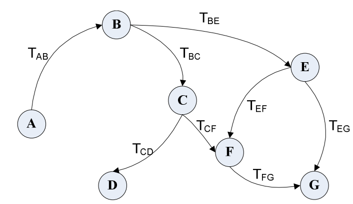

Figure 1: A social network

However, extracting such a trust network involves the NP-Complete problem of finding the longest simple path (a simple path is an acyclic path) in a graph (Baase and Gelder 2000). Alternatively, since the resource discovery problem in P2P networks (Adamic, Lukose, and Huberman 2003) has similar properties, the heuristic search strategies developed for P2P network resource discovery can be used. But these methods do not consider the social context in social networks, including social relationships, social positions, residential location and preferences of participants. As indicated in Social Psychology (Brass 2009; Lichtenstein and Slovic 2006), all the above information provides a social context and has significant influence on both social interaction and trust evaluation. In addition, a source may have different purposes to evaluate the trustworthiness of the target, e.g., looking for a potential employee, or a movie recommendation. Thus, to obtain a more trustworthy trust evaluation result, a source participant may specify some constraints of the values in social context as the trust evaluation criteria in trust network discovery. However, this feature has not been supported in the existing methods. Therefore, it is a significant and challenging problem to extract the contextual trust network satisfying the constraints specified by a source, which is expected to contain the most trustworthy trust evaluation result.

In this paper, we first propose a new complex contextual social network structure which contains social contextual impact factors, including trust, social intimacy degree, role impact factor, preference similarity and residential location distance. Then, we propose a new concept QoTN (Quality of Trust Network), taking these impact factors as the attributes to illustrate the capability of a contextual trust network to guarantee a certain level of trustworthiness in trust evaluation. After that we propose a social context-aware trust network discovery model. Since trust network discovery with QoTN constraints involves finding the longest simple path (a simple path is an acyclic path) in a graph, which has been proved to be an NP-Complete problem (Baase and Gelder 2000), we propose a novel Social Context-Aware trust Network discovery algorithm, called SCAN, by adopting the Monte Carlo method and our proposed optimization strategies. Experiments conducted on a real social network dataset, Enron emails[1], demonstrate the superior performance of the SCAN algorithm.

Related Work

Social Network Analysis

In 1960's, Milgram (1967) validated the *small-world*[2] characteristic in social networks. In recent years, sociologists and computer scientists have started to investigate the characteristics of popular OSNs, including MySpace (myspace.com) and Flickr (flickr.com), and validated the *small-world* and *power-law* characteristics of online social networks using data mining techniques.

Trust in Social Networks

Trust is a critical factor in the decision-making of participants in OSNs (Kuter and Golbeck 2007). Golback *et al*. (Golbeck and Hendler 2006) proposed a trust inference mechanism in social networks based on an averaging strategy. In addition, Liu et al. (2010) have proposed a heuristic algorithm to identify the most trustworthy social trust path between two nonadjacent participants. They further proposed a heuristic algorithm for the selection of top K $(K \geq 2)$ social trust paths (Liu, Wang, and Orgun 2011a) and a novel trust transitivity model in social networks (Liu, Wang, and Orgun 2011b). All of these trust models assume that the trust network from the source to the target have been extracted. Therefore, trust network discovery is a necessary step because it provides a foundation to apply the above promising trust evaluation models.

Network Discovery

To the best of our knowledge, in the literature, there are no approximation algorithms proposed for the NP-Complete trust network discovery problem in OSNs. Since the resource discovery problem in P2P networks has the similar properties as the trust network discovery problem, some search strategies developed for the resource discovery can be applied in trust network discovery. These strategies can be classified into three categories.

Flooding-Based Search (FBS) The flooding-based mechanism searches the network from the source by using the Breadth First Search (BFS) strategy, which was applied into Gnutella (rfc-gnutella.sourceforge.net). Since this search strategy consumes huge computation time, the Time To Live Breadth First Search (TTL-BFS) method (Filali and Huet 2010) was proposed. In TTL-BFS, the Time To Live (TTL) is set to an integer and its value is decreased by 1 or Vr $(0 < Vr < 1)$ after each layer of BFS. During the process, if the target is found, the search terminates. Otherwise, TTL-BFS repeats BFS until $TTL = 0$ or the target is found.

Random Walk Search (RWS) RWS (Gkantsidis, Mihail, and Saberi 2004) firstly searches all the neighboring nodes of the source. If the target is discovered, the search terminates. Otherwise, the method randomly selects one of the currently node's neighbors as the expansion node for the next step of the search.

High Degree Search (HDS) The HDS method (Adamic, Lukose, and Huberman 2003) firstly calculates the outdegree of each of the neighboring nodes of a source and selects the one with the maximal outdegree. If the selected node is the target, then the search terminates. Otherwise, HDS repeats the ourdegree calculation and node selection.

Summary The above search strategies have good performance in P2P networks which do not contain social contextual information. In addition, they do not support the trust evaluation criteria specification. Thus, existing methods cannot be expected to extract a trust network to deliver a trustworthy trust evaluation satisfying the constraints in social contexts.

Complex Contextual Social Networks

In this section, we propose a new complex contextual social network structure, containing social contextual impact factors, and hence reflecting the social networks in the real world better.

[1] http://www.cs.cmu.edu/ enron/

[2] The average path length between any two nodes is about 6.6 hops in a social network (Milgram 1967).

Social Contextual Impact Factors

1. **Trust:** Trust is the belief of one participant in another, based on their interactions, in the extent to which the future action to be performed by the latter will lead to an expected outcome (Liu et al. 2010). As pointed out in (Wang and Varadharajan 2007; Mansell and Collins 2005), the trust value between two people can be different in different domains. In our model, let $T_{AB}^{D_i} \in [0,1]$ denote the trust value that A assigns to B in domain i. If $T_{AB}^{D_i} = 0$, it indicates that A completely distrusts B in domain i, while $T_{AB}^{D_i} = 1$ indicates A completely believes B's future action can lead to the expected outcome in that domain.
2. **Social Intimacy Degree:** As illustrated in Social Psychology (Brass 2009), a participant can trust and have more social interactions with the participants with whom he/she has more intimate social relationships than those with whom he/she has less intimate social relationships. Let $SI_{AB} \in [0,1]$ denote the *Social Intimacy Degree* between A and B in online social networks. $SI_{AB} = 0$ indicates that A and B have no social relationship while $SI_{AB} = 1$ indicates they have the most intimate social relationship.
3. **Role Impact Factors:** As illustrated in Social Psychology (Adler 2001; Dalton 1959), in a certain domain of interest (e.g., hiring employees or product sales), an expert's recommendation is more credible than that from a beginner. Let $\rho_A^{D_i} \in [0,1]$ denote the value of the *Role Impact Factor*, illustrating the impact of participant A's social role in domain i. $\rho_A^{D_i} = 1$ indicates that A is a domain expert in domain i while $\rho_A^{D_i} = 0$ indicates that A has no knowledge in that domain.
4. **Preference Similarity:** As illustrated in Social Psychology (Zajonc 2011), a participant can trust and have more social interactions with another participant, with whom he/she shares more preferences (e.g., both of them like playing badminton) than those, with whom he/she shares fewer preferences. Let $PS_{AB}^{D_i} \in [0,1]$ denote the value of the *Preference Similarity* between A and B in domain i. When $PS_{AB}^{D_i} = 0$, A and B have no similar preference in the domain. When $PS_{AB}^{D_i} = 1$, they have the same preference in that domain.
5. **Residential Location Distance:** As illustrated in Social Psychology (Gimpel et al. 2008), a participant can trust more and have more social interactions with another whose residential location is close to that of the participant (e.g., in the same neighborhood) than those whose residential location is far away. Let $RLD_{AB} \in [0,1]$ denote the *Residential Location Distance* between A and B. When $RLD_{AB} = 1$, the residential location of A and B are the same. When $RLD_{AB} = 0$, it indicates that the residential location between them has the largest distance.

Although it is difficult to build up comprehensive social relationships, recommendation roles, preference similarity and residential location distances in all domains, it is feasible to build them up in particular applications (Liu, Wang, and Orgun 2011b). For example, in the work by Mccallum et al. (2007), through mining the subjects and contents of emails in the Enron Corporation[1], the social relationship between each email sender and receiver can be discovered and their roles can be predicted. Then the corresponding social intimacy degree and role impact factor values can be estimated based on probabilistic models. In addition, at Facebook, the preference similarity and the residential location distance between two participants can be mined from their profiles (Mislove et al. 2007). Detailed mining methods are out of the scope of this paper.

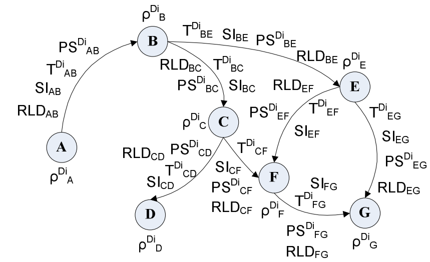

Figure 2: A complex contextual social network

A Complex Contextual Social Network Structure

Based on the above social contextual impact factors, we propose a new complex contextual social network structure as depicted in Fig. 2.

Social Context-Aware Trust Network Discovery Model

Social Context-Aware Social Interaction

As indicated in Social Psychology (Gimpel et al. 2008; Zajonc 2011; Brass 2009), some of the social contextual impact factors (i.e., *SI*, *PS*, *T* and *RLD*) have influence on social interactions. For example, based on the statistics of 1000 publications from 18 countries in ISI Web of Knowledge (apps.webofknowledge.com) (Wren et al. 2007), the first author and the last author had the same address in 54% papers, indicating that *RLD* of authors impacts on the social interactions in research. In addition, based on the statistics on Flickr(flickr.com)-an online photo sharing social network (Mislove et al. 2007), any two participants in photo sharing usually have similar preferences, indicating that *PS* impacts the social interactions.

Social Context-Aware Social Interaction Probability

In the real world, many social phenomenons approximately follow the *normal (Gaussian) distribution* (Bittinger 2002). For example, the IQ, memory, income and reading skills of people in a general population (Abell, Braselton, and Rafter 2012). The probability density function of the normal distribution is as Eq.(1) (see the function image in Fig. 3).

$$y = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

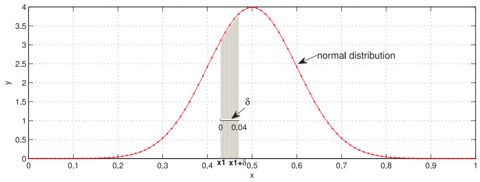

Figure 3: Normal distribution

where parameters μ and σ are the mean and standard deviation respectively.

In our model, we assume the probability distribution of a social interaction between any two participants with the social contextual impact factors (denoted as $P(A \to B|X)$, X is the social impact factor value) also follows the normal distribution.

$$P(A \to B|X) = \int_X^{X+\delta} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} d(X) \quad (2)$$

Then, based on mathematical theory (Bittinger 2002), $P(A \to B|X)$ can be calculated by Eq. (2) (i.e., the Probability Density Function of normal distribution). In this equation, δ is the length of each small interval (the horizontal axis between 0 and 1 is divided into several small intervals (Bittinger 2002)). If X is in one of the intervals (e.g., in the interval $[x_1, x_1+\delta]$ in Fig. 3), $P(A \to B|X)$ is the integration of Eq. (1) with a lower limit X and a upper limit $X+\delta$ (in the case shown in Fig. 3, where $X = x_1$). Namely, $P(A \to B|X)$ is equal to the corresponding area of the trapezoid with curved edges in an interval $[X, X+\delta]$ (e.g., the shadowed area in Fig. 3). In addition, the parameters μ and σ in Eq. (2) can be computed by applying social statistics methods (Bittinger 2002), which is out of the scope of this paper. Finally, typically the social interaction with each factor is regarded as an independent event, based on probability theory (Bittinger 2002), the aggregated social interaction probability between A and B (denoted as $AP(A \to B)$) can be calculated by Eq. (3).

$$AP(A \to B) = \prod P(A \to B|X) \quad (3)$$

In our model, the aggregated social interaction probability will be considered in the node selection of trust network discovery, where the larger the probability of a node to have a social interaction with the target, the more likely for the node to be selected.

Quality of Trust Network (QoTN)

In addition to the influence of social context on social interactions, our model also considers different trust evaluation criteria of a source. We first propose a new concept, *Quality of Trust Network* as below.

Definition 1 Quality of Trust Network *(QoTN) is the ability of a contextual trust network to guarantee a certain level of trust in trust evaluation, taking T, SI, ρ, PS, RLD as attributes.*

In our model, a source participant can specify multiple constraints for QoTN attributes (i.e., T, SI, ρ, PS and RLD) for intermediate nodes and their links in a trust network, as the requirements of trust network discovery in different domains. Let $QoTN^{(\eta)}_{v_s,v_t}$ ($\eta \in \{T, SI, \rho, PS, RLD\}$) denote the QoTN constraints of η in the trust network from v_s to v_t (throughout this paper, v_s denotes the source and v_t denotes the target in a social network). For example, to *hire employees*, v_s, a hiring manager specifies the QoTN constraints as $\{QoTN^T_{v_s,v_t} > 0.3,\ QoTN^{SI}_{v_s,v_t} > 0.3,\ QoTN^{PS}_{v_s,v_t} > 0.3,\ QoTN^{RLD}_{v_s,v_t} > 0.3,\ QoTN^{\rho}_{v_s,v_t} > 0.8\}$, if he/she believes the role impact factor of each of the intermediate participants is more important in the domain of *recruitment*.

Trust Network Utility

In our model, we define the utility (denoted as $\mathcal{U}$) as the measurement of the trustworthiness of an extracted trust network. The utility function takes the QoTN attributes T, SI, ρ, PS and RLD as the arguments in Eq. (4)

$$\mathcal{U}(v_s, v_t) = \sum_{i=1}^{N} T_i + \sum_{i=1}^{N} SI_i + \sum_{i=1}^{M} \rho_i + \sum_{i=1}^{N} PS_i + \sum_{i=1}^{N} RLD_i \quad (4)$$

M is the number of the intermediate nodes and N is the number of the corresponding links in the trust network,

The goal of trust network discovery is to extract the optimal trust network from source v_s to target v_t that satisfies multiple QoTN constraints and yields the highest utility.

Social Context-Aware Trust Network Discovery Algorithm

To solve the NP-Complete trust network discovery problem with QoTN constraints, we propose a Social Context-Aware trust Network discovery (SCAN) algorithm, by adopting the Monte Carlo method and two optimization strategies.

Monte Carlo Method

Monte Carlo method (Gentle, Hardle, and Mori 2004) is a computational algorithm which relies on repeated random sampling to compute results. It is one of the techniques with good efficiency for solving NP-complete problems (Gentle, Hardle, and Mori 2004). In the literature, based on the Monte Carlo method, a number of algorithms have been proposed for solving NP-Complete multiple constrainted social trust path selection and composite service selection problems (Liu, Wang, and Orgun 2011b; Li, Wang, and Lim 2009).

SCAN

In SCAN, initially, the source participant v_s is regarded as the current expansion node, and SCAN searches all the neighboring nodes of v_s (denoted as v_s*.neighboring_node*) to investigate whether the current node and its corresponding links satisfy the QoTN constraints. If all QoTN constraints can be satisfied, the neighboring node is called *a feasible node* (denoted as v_f). Given that the larger the outdegree of a node, the more likely for the node to have a connection with others (Adamic, Lukose, and Huberman 2003), SCAN calculates the *selection probability* of all the feasible nodes (denoted as $SCP(v_f \to v_t)$) based on $AP(v_f \to v_t)$ and the outdegree of v_f (denoted as $deg^+(v_f)$) by Eq. (5).

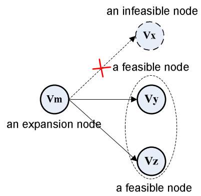

Figure 4: Unsatisfied nodes

$$SCP(v_f \to v_t) = AP(v_f \to v_t) \cdot \frac{deg^+(v_f)}{MAX(deg^+)} \tag{5}$$

where $MAX(deg^+)$ is the maximal value of the outdegree of the nodes in a social network.

After that, based on their selection probabilities, SCAN selects one of the feasible nodes as the next expansion node (denoted as v_{exp}), where the higher the selection probability of a node, the more likely for the node to be selected. During the process, a cycle in a path is avoided by the strategy in (Pearl 1984), as it leads to inefficiency and ineffectiveness of the network discovery (Mansell and Collins 2005). Finally, SCAN repeats the above search process at each v_{exp} until it finds v_t or reaches the threshold of search hops (denoted as λ_h, on average $\lambda_h \leq 7$ due to the *small-world* phenomenon of social networks[2]). During the search process, in addition to the basic Monte Carlo method, we adopt the following three optimization strategies to improve the efficiency of our algorithm.

Optimization Strategy 1: Avoiding Repeated Feasibility Investigations in Simulations. In each search step, the Monte carlo method investigates the feasibility of all the neighboring nodes of the current v_{exp} (e.g., investigating v_x, v_y and v_z in Fig. 4). In multiple simulations of the Monte Carlo method, a node may be selected as a v_{exp} more than once. In such a situation, the feasibility investigation needs to be performed repeatedly, leading to low efficiency. To address this issue, in SCAN, if a neighboring node is infeasible, (e.g., v_x in Fig. 4), the corresponding link from the current v_{exp} to the neighboring node (e.g., $v_m \to v_x$) will be removed. Then, upon reaching the same v_{exp} (e.g., v_m) in the subsequent simulations, SCAN does not investigate its neighboring nodes repeatedly as all of them are feasible.

Optimization Strategy 2: Avoiding Repeated Probability Calculations in Simulations. In multiple simulations of the Monte Carlo method, if the same v_{exp} is selected more than once (e.g., v_m in Fig. 4 is selected more than once), its feasible neighboring nodes' selection probabilities will have to be calculated repeatedly, leading to low efficiency. To address this issue, at each feasible node, SCAN records its selection probability (denoted as $v.selection$), thereby avoiding repeated calculations in the subsequent simulations.

Given a group of QoTN constraints, and a pair of v_s and v_t in a complex contextual social network, the process of SCAN includes the following steps.

Initialization: At each node v_i, set $v_i.probability\_status = 0$, which indicates the social interaction probabilities of v_i's neighboring nodes (denoted as $v_i.neighboring\_nodes$) have not been calculated. In addition, all the nodes are marked as unvisited ($v_i.visit = 0$), and set $v_{exp} = v_s$.

Step 1: Based on $v_{exp}.probability\_status$, SCAN performs the following search strategies.

If $v_{exp}.probability\_status \neq 1$, SCAN investigates the feasibility of v_j, $v_j \in \{v_{exp}.neighboring\_nodes\}$) as follows.

(a) **Case 1:** If $v_j = v_t$, then SCAN terminates the current search, and starts a new simulation from *Initialization*.

(b) **Case 2:** If $v_j \neq v_t$ and v_j is a feasible node, then SCAN calculates $SCP(v_j \to v_t)$, and sets $v_j.selection = SCP(v_j \to v_t)$ and $v_{exp}.probability\_status = 1$.

(c) **Case 3:** If $v_j \neq v_t$ and v_j is an infeasible node, based on the number of the infeasible neighboring nodes of v_{exp} (denoted as $v_{exp}.infeasible\_number$), SCAN performs the following search strategies.

(c-1-1): If $v_{exp}.infeasible\_number = v_{exp}.neighboring\_nodes$ and $v_{exp} = v_s$, then SCAN terminates, failing to return a trust network that satisfies QoTN constraints.

(c-1-2): If $v_{exp}.infeasible\_number = v_{exp}.neighboring\_nodes$ and $v_{exp} \neq v_s$, then SCAN terminates the current search and starts a new simulation from *initialization*.

(c-1-3): If $v_{exp}.infeasible\_number \neq v_{exp}.neighboring\_nodes$, go to *Step 2*.

Step 2. If $v_{sel}.visit = 0$, calculate the probability of v_j to be a v_{exp} by the following Eq.(6).

$$p(v_j) = \frac{v_j.selection}{\sum v_k.selection} \; v_k \in \{v_{exp}.neighboring\_node\} \tag{6}$$

Step 3. Select one of the feasible neighboring nodes (denoted as v_{sel}) based on their probabilities obtained by Eq. (6). Then based on *Optimization Strategy 1*, set $v_{sel}.visit = 1$, to avoid cycles in the subsequent search steps of the current simulation.

Step 4. Set $v_{exp} = v_{sel}$, and continue the search from *Step 1* until the number of searching hops reaches λ_h (on average $\lambda_h \leq 7(Milgram 1967)$).

The time complexity of SCAN is $O(mld)$, where m is the number of simulations; l is the average length of the social trust paths from v_s to v_t; d is the maximal outdegree of the nodes in social networks. In social networks, on average, $l < 7$ (Milgram 1967). Thus the time complexity of SCAN is $O(md)$, which is better than FBS (Flooding Based Search) with the time complexity of $O(d^{TTL})$ (TTL is Time To Live, introduced in *Section 2.3*), and the same as those of both RWS (Random Walk Search) and HDS (High Degree Search).

Experimental Evaluation

Experiment Setup

In order to evaluate the performance of our proposed algorithm on trust network discovery, we need a dataset which contains social network structures. The *Enron* email dataset[1] has been proved to possess the small-world and power-law characteristics of social networks and thus it has been widely

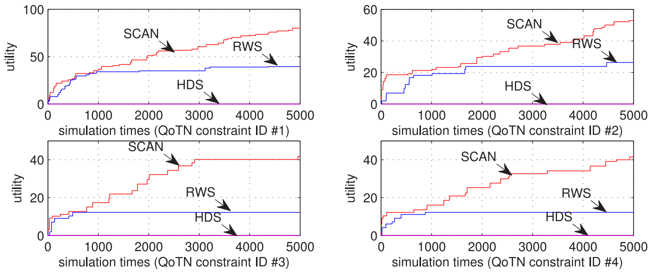

Figure 5: The utilities of extracted trust networks with 4 hops

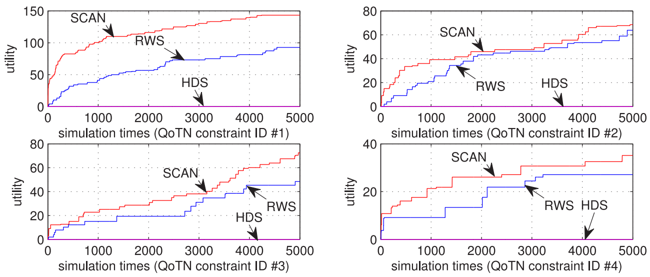

Figure 6: The utilities of extracted trust networks with 6 hops

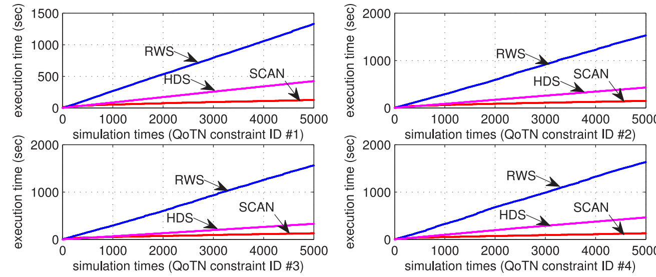

Figure 7: The execution time with 4 hops

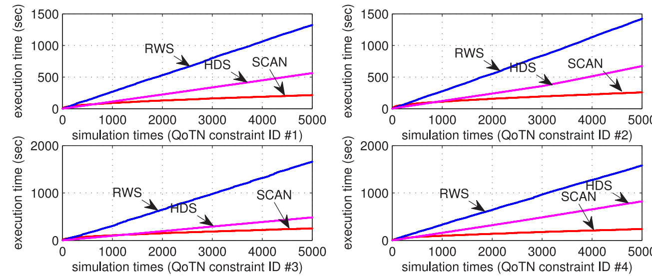

Figure 8: The execution time with 6 hops

used in the studies of social networks (Mccallum, Wang, and Corrada-Emmanuel 2007; S. Yoo and Moon 2009; Liu, Wang, and Orgun 2011b; 2011a). Thus, to validate our proposed algorithm, we select the *Enron* email dataset[1] with 87,474 nodes (participants) and 30,0511 links (formed by sending and receiving emails) as the dataset for our experiments. Secondly, we randomly select a source and a target from the dataset, and compare our SCAN with other methods in all the three categories, i.e., FBS, RWS and HDS. Thirdly, we set four groups of QoTN constraints as listed in Table 1 and set the social interaction probability to approximately follow the normal distribution with $\mu = 0.5$ and $\sigma = 0.1$. Finally, the social contextual impact factor values are generated by using function $normrnd(\mu, \sigma)$ in *Matlab*.

Each of SCAN, FBS, RWS and HDS is implemented using Matlab R2008a running on an Lenovo ThinkPad SL500 laptop. The results are plotted in Fig. 5 to Fig. 8, where the execution time and the utilities of the extracted trust network for each of the algorithms are averaged based on 10 independent runs. In each run, we perform up to 5000 simulations.

Results and Analysis

During the execution of FBS with TTL=2, MATLAB runs out of memory. This could be caused by the large outdegrees of the nodes in the two search hops. Therefore, we compare the extracted trust networks' utilities delivered by SCAN, RWS and HDS and their execution time.

Table 1: The settings of QoTN constraints

ID	$QoTN(T)$	$QoTN(SI)$	$QoTN(PS)$	$QoTN(RLD)$	$QoTN(\rho)$
#1	0.1	0.1	0.1	0.1	0.1
#2	0.15	0.15	0.15	0.15	0.15
#3	0.2	0.2	0.2	0.2	0.2
#4	0.25	0.25	0.25	0.25	0.25

Fig. 5 to Fig. 6 plot the extracted trust networks' utilities with different QoTN constraints with 4 and 6 search hops (the figures for 5 and 7 hops are similar to those of 4 and 6). From them we could see that firstly, with the same simulation times, our proposed SCAN can deliver much higher network utilities than all other methods in all cases. In addition, since HDS cannot find v_t after 5000 times search from v_s in all cases, the extracted trust network's utility delivered by HDS is always *zero*. Thus, we compare the average utilities based on different QoTN constraints delivered by SCAN and RWS in Table 2. From them we could see that, on average, our SCAN can deliver extra 72.69%, 76.52%, 84.43%, 85.25% and 71.12% more utilities respectively than RWS with 1000, 2000, 3000, 4000 and 5000 simulations. This is because that SCAN takes into account the influence of social context on social interactions, where the larger the probability that a node has a social interaction with v_t, the more likely for the node to be selected. This method increases the probability of finding a trust path from v_s to v_t at each search. In addition, SCAN considers the QoTN constrains, and can avoid searching infeasible nodes, improving the effectiveness of each search.

Fig. 7 to Fig. 8 plot the execution time of SCAN, RWS and HDS with different QoTN constraints with 4 and 6 search hops (the figures for 5 and 7 hops are similar to those of 4 and 6). From the results, we could see that the execution time of SCAN is less than that of RWS in all cases. In addition, when the simulation times are less than 1000, the execution time of SCAN is similar to HDS. But with the increase of simulation times, HDS consumes much more execution time than SCAN. The average execution time of of HDS, RWS and SCAN in each of 4 to 7 search hops is listed in Table 3. Based on the statistics of all executions, on average, the execution time of SCAN is only 13.85% and

Table 2: The comparison of the utility

Simulations	Difference of path utility				
	4 hops	5 hops	6 hops	7 hops	total
1000	19.24%more	45.02%more	112.37%more	114.15%more	72.69%more
2000	62.40%more	61.68%more	63.29%more	118.71%more	76.52%more
3000	103.84%more	88.18%more	41.08%more	104.62%more	84.43%more
4000	114.80%more	101.68%more	40.13%more	84.42%more	85.25%more
5000	139.55%more	92.51%more	37.55%more	64.78%more	71.12%more

39.27% of that of RWS and HDS respectively. This is because SCAN can avoid repeated feasibility investigation (*by Strategy 1*) and repeated selection probability calculation (*by Strategy 2*).

Table 3: The comparison of execution time (5000 simulations)

Algorithms	The sum of the average execution time (sec)				
	4 hops	5 hops	6 hops	7 hops	total
SCAN	532.5590	772.2890	910.7940	1.1280e+003	3.3436e+003
RWS	6.0538e+003	5.8048e+003	5.9819e+003	6.3017e+003	2.4142e+004
HDS	1.6466e+003	2.0008e+003	2.3238e+003	2.5434e+003	8.5146e+003
SCAN/RWS	0.0880	0.1330	0.1523	0.1790	0.1385
SCAN/HDS	0.3234	0.3836	0.3919	0.4435	0.3927

Conclusions and Future Work

In this paper, we have proposed a complex contextual social network structure containing social contextual impact factors. Then, we have proposed a general concept QoTN (Quality of Trust Network), and a novel social context-aware trust network discovery model. Finally, we have proposed a new Social Context-Aware trust Network discovery algorithm (SCAN) by adopting the Monte Carlo method and our proposed optimization strategies. Based on the above experimental results and analysis, we conclude that our proposed SCAN outperforms all the existing methods significantly in both *execution time* and *the quality of the extracted trust networks*. Therefore, SCAN is an efficient and effective algorithm for the trust network discovery problem with QoTN constrains in complex contextual social networks.

In future work, we plan to incorporate our models and algorithms in a new generation of contextual social network based recommendation systems.

References

Abell, M.; Braselton, J.; and Rafter, J. 2012. *Statistics with Maple*. Elsevier Science and Technology.

Adamic, L. A.; Lukose, R. M.; and Huberman, B. A. 2003. Local search in unstructured networks. In *Handbook of Graphs and Networks*. Wiley.

Adler, P. S. 2001. Market, hierarchy, and trust: The knowledge economy and the future of capitalism. *Organization Science* 12(12):215–234.

Baase, S., and Gelder, A. 2000. *Computer Algorithms Introduction to Design and Analysis*. Addision Wesley.

Bittinger, M. 2002. *Basic Mathematics*. Addison Wesley.

Brass, D. J. 2009. *A Socal Network Prespective On Industral/organizational psychology*. Industrial/Organizational Handbook.

Dalton, M. 1959. *Men Who Manage*. New York: Wiley.

Filali, I., and Huet, F. 2010. Dynamic ttl-based search in unstructured peer-to-peer networks. In *CCGrid*, 438–447.

Gentle, J.; Hardle, W.; and Mori, Y. 2004. *Handbook of Computational Statistics*. Springer.

Gimpel, J.; Karnes, K.; Mctague, J.; and Pearson-Merkowitz, S. 2008. Distance-decay in the political geography of friends-and-neighbors voting. *Political Geography* 27:231–252.

Gkantsidis, C.; Mihail, M.; and Saberi, A. 2004. Random walks in peer-to-peer networks. In *IEEE INFOCOM*, 120–130.

Golbeck, J., and Hendler, J. 2006. Inferring trust relationships in web-based social networks. *ACM Transactions on Internet Technology* 6(4):497–529.

Kuter, U., and Golbeck, J. 2007. Sunny: A new algorithm for trust inference in social networks using probabilistic confidence model. In *AAAI'07*, 1377–1382.

Li, L.; Wang, Y.; and Lim, E. 2009. Trust-oriented composite services selection and discovery. In *ICSOC'09*, 50–67.

Lichtenstein, S., and Slovic, P. 2006. *The construction of preference*. Cambridge University Press.

Liu, G.; Wang, Y.; Orgun, M.; and Lim, E.-P. 2010. A heuristic algorithm for trust-oriented service provider selection in complex social networks. In *SCC*, 130–137.

Liu, G.; Wang, Y.; and Orgun, M. A. 2011a. Finding k optimal social trust paths for the selection of trustworthy service providers in complex social networks. In *ICWS'11*, 41–48.

Liu, G.; Wang, Y.; and Orgun, M. A. 2011b. Trust transitivity in complex social networks. In *AAAI'11*, 1222–1229.

Mansell, R., and Collins, B. 2005. *Trust and crime in information societies*. Edward Elgar Publishing.

Mccallum, A.; Wang, X.; and Corrada-Emmanuel, A. 2007. Topic and role discovery in social networks with experiments on Enron and academic email. *Journal of Artificial Intelligence Research* 30(1):249–272.

Milgram, S. 1967. The small world problem. *Psychology Today* 2(60).

Mislove, A.; Marcon, M.; Gummadi, K.; Druschel, P.; and Bhattacharjee, B. 2007. Measurement and analysis of online social networks. In *ACM IMC'07*, 29–42.

Pearl, J. 1984. *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. Addison-Wesley.

S. Yoo, Y. Yang, F. L., and Moon, I. 2009. Mining social networks for personalized email prioritization. In *KDD'09*, 967–976.

Wang, Y., and Varadharajan, V. 2007. Role-based recommendation and trust evaluation. In *IEEE EEE'07*, 278–295.

Wren, J.; Kozak, K.; Johnson, K.; Deakyne, S.; Schilling, L.; and Dellavalle, R. 2007. A survey of perceived contributions to papers based on byline position and number of authors. *EMBO Report* 8(11):988–991.

Zajonc, R. 2011. Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science* 10(6):224–228.