


Trust evaluation based on evidence theory in online social networks

International Journal of Distributed
Sensor Networks
2018, Vol. 14(10)
© The Author(s) 2018
DOI: 10.1177/1550147718794629
journals.sagepub.com/home/dsn


Jian Wang¹, Kuoyuan Qiao² and Zhiyong Zhang²

Abstract

Trust is an important criterion for access control in the field of online social networks privacy preservation. In the present methods, the subjectivity and individualization of the trust is ignored and a fixed model is built for all the users. In fact, different users probably take different trust features into their considerations when making trust decisions. Besides, in the present schemes, only users' static features are mapped into trust values, without the risk of privacy leakage. In this article, the features that each user cares about when making trust decisions are mined by machine learning to be User-Will. The privacy leakage risk of the evaluated user is estimated through information flow predicting. Then the User-Will and the privacy leakage risk are all mapped into trust evidence to be combined by an improved evidence combination rule of the evidence theory. In the end, several typical methods and the proposed scheme are implemented to compare the performance on dataset Epinions. Our scheme is verified to be more advanced than the others by comparing the F-Score and the Mean Error of the trust evaluation results.

Keywords

Trust evaluation, evidence theory, online social networks, information flow prediction, decision making

Date received: 28 March 2018; accepted: 13 July 2018

Handling Editor: Shaojie Tang

Introduction

Online social networks (OSNs) are platforms or systems that people can interact with others by sharing or posting blogs online.¹ Social networking is very common, such as Facebook, Tweeter, Weibo and CyVOD.² These platforms provide a free space for everyone to unleash their mind and thoughts. However, it makes information leakage possible.³ The spammers spread malicious links and annoying messages to OSN users without target, and privacy information is unsafe for the cheating actions⁴ and blackmails.⁵ To prevent the malicious activities, many schemes such as Access Control⁶ and digital rights protection⁷⁻⁹ are proposed. In these schemes, trust degree is usually viewed as the main criterion for security policies to make the privacy management more feasible and effective.

As it is important to privacy preservation in OSNs, trust evaluation has become a research focus in recent years.¹⁰⁻¹³ Researchers try to find the relationship

between user features and trust decision. It is no doubt that trust decision is not only affected by objective features of each user but also affected by the subjective options of the user. For example, some people think the one who has a lot of fans in the OSNs is trustworthy, while others would rather choose the people who have higher credit or reputation. So, just a single model without individualization is insufficient to evaluate trust degree between users in OSNs.

Besides, most present schemes evaluate trust degree based on present state of each user. If user u has been judged to be trusted, the message transferred to him will

¹Zhengzhou University, Zhengzhou, China

²Henan University of Science and Technology, Luoyang, China

Corresponding author:

Jian Wang, Zhengzhou University, No. 100 Science Avenue, Zhengzhou 450001, Henan, China.

Email: iejwang@zzu.edu.cn



be deemed safe. Because message propagation takes time and the state of the user is not constant, privacy leaks may occur during message propagation due to changes in the state of the user. Therefore, trust evaluation should take the information flow risk into consideration to avoid privacy leakage in OSNs. The last but not the least, people trust has ambiguity; however, most present methods ignore it and give an absolute probability degree for Trusted or Untrusted. It is unreasonable unless there is no uncertainty in trust decision made by users and the result of trust evaluation is totally correct.

Aiming at the above problems, we provide an improved scheme to evaluate the trust between the users in OSNs. In our scheme, user features and information flow prediction result are mapped into trust evidence, and then combined based on evidence theory to obtain trust evaluation result.

The article is organized as follows. The “Related works” and “Preliminary” sections introduce the related works and the preliminary separately. In the “Trust evaluation based on the combination of evidence” section, the proposed scheme is illustrated in detail, including its design idea and practical implement approach. The performance of the scheme is mainly evaluated in the “Experiment and analysis” section. Finally, in the “Conclusion” section, we make some concluding remarks.

Related works

In most trust evaluation schemes, the inputs are user features such as similarity, intimacy, and reputation.¹⁴ Zhao and Pan¹³ proposed a trust evaluation method based on classifying user features. Brown and Feng¹⁵ proposed a trust degree calculation scheme based on user influence and a K-shell algorithm. Similar to Brown’s work, Silva et al.¹⁶ proposed another method based on user influence and information diffusion. Based on user credit and reputation, Tsolmon and Lees¹⁷ proposed a trust calculation scheme using the features, such as follower number, tweet number, and reputation, as the input of their algorithm. Relatively, Mármol and Pérez¹⁸ proposed a trust calculation method based on user behavior, user rating, and personal rewards.

Usually, users in real OSNs usually aggregate into communication groups to make their interactions facilitated. And the small world theory¹⁹ which divides users into groups is used in some present trust computing methods to improve the accuracy and efficacy of the result. Zhang and Wang²⁰ proposed a scheme based on community group, feedback, and trust decay. According to Yang et al.,²¹ users in a topic circle tend to trust the leaders in this circle and be likely to transmit the message sent by other users in the same circle,

and trust degree can be evaluated by computing the degree of influence between users.

Moreover, trust decisions made by users may affect the leaders and the experts in the OSNs.²¹ Tsolmon and Lee¹⁷ provided a leader-find method named hyperlink-induced topic search (HITS) based on computing the context transmitting in OSNs, ignoring the attributes such as follower number or favor count. In the work of the document,²² the problem of selecting top-k expert users in social group based on their knowledge about a given topic is addressed. Chiregi and Navimipour¹² proposed a trust evaluation scheme based on the leader and expert in the OSNs.

Furthermore, trust degree is also affected by the flow of information.²³ If the probability that a requestor shares privacy message to a malicious user is high, the server will consider denying the transmission to avoid a privacy leakage. Ranjbar and Maheswaran²⁴ proposed a method for computing information flow probability in OSN. Jiang et al.²⁵ proposed a scheme for generating the trust graph of an OSN. Later, with the trust graph, he proposed a novel trust evaluation method based on information leakage.¹¹ However, some researchers believe that predicting information flow in the future precisely is non-deterministic polynomial (N/P) hard, and the real value can never be approached.²⁶ In this article, Monte-Carlo simulation (MCS)²⁷ is used to predict the information flow probability between each two users to make the result infinitely close to the real value.

Preliminary

Evidence theory

Evidence theory is an effective tool to make decision from information with ambiguity.^{28,29} It is widely used in decision making,^{30–32} target recognizing,^{33,34} and OSNs analyzing.^{35–37} In evidence theory, the set of all the possible decisions is called discern-frame θ . The element in θ is named as focal-element. A brief example is shown to illustrate the difference between single-element focal-element and multi-element focal-element in evidence theory. Assuming three suspects Jim, Tom, and Kate are involved in a murder case, three officers hold different views based on the existing trace just as Table 1 shows.

In the last column, Tom, Jim, and Kate are single-element focal-elements, and the other ones are multi-element focal-element. In the first three columns of Table 1, $m(f); f \in F : \{Jim, Kate, Tom\}$ stands for the probability that the real murder is f . $m(f)$ is also called the basic probability assignment (BPA) of f .

All the BPAs that came from a same officer will constitute the body of evidence (BOE): $m : \{m(Tom), m(Jim), m(Kate), m(Tom, Jim), m(Tom, Kate), m(Jim, Kate), m(\theta)\}$. And the decision can be made by the

Table 1. View of each officer.

officer1	officer2	officer3	Focal-element
m(Tom) = 0.1	m(Tom) = 0.15	m(Tom) = 0.35	Tom
m(Jim) = 0.2	m(Jim) = 0.05	m(Jim) = 0.05	Jim
m(Kate) = 0.05	m(Kate) = 0.05	m(Kate) = 0.1	Kate
m(Tom, Jim) = 0.2	m(Tom, Jim) = 0.1	m(Tom, Jim) = 0.2	Tom or Jim
m(Tom, Kate) = 0.3	m(Tom, Kate) = 0.1	m(Tom, Kate) = 0.1	Tom or Kate
m(Jim, Kate) = 0.1	m(Jim, Kate) = 0.1	m(Jim, Kate) = 0.1	Jim or Kate
m(θ) = 0.05	m(θ) = 0.45	m(θ) = 0.1	Tom or Jim or Kate

combination of evidence based on combination rule. Dempster²⁸ proposed the first evidence combination rule. It realizes the combination of evidence but invalid when evidence is conflicting.^{38,39} And combination rule should make the correct decision whatever conflict evidence is contained.²⁷ Similarity Calculation^{40,41} and Ambiguity Measure^{42,43} of evidence are two main strategies for determining evidence weights, which is to reduce the impact of evidence conflicting. However, in our previous work,³⁸ it is found that similarity of evidence may collide and a combination rule is proposed to cope with the collision of similarity.

Sometimes, another evidence weight set $\psi : \{\psi_1, \psi_2, \dots, \psi_n\}$ may be appointed before the combination to indicate the importance of each evidence. Based on the conditions that ψ is appointed or not, we define the operator \oplus as the combination when evidence weight is not given and \oplus_ψ when ψ is provided.

Classify based on user features

In our scheme, classification is a necessary step to generate trust evidence, and the method we used is classification based on SVM (Support Vector Machine)³⁹ in RBF (Radial Basis Function)⁴⁴ type. Items in the training set are mapped into high dimension vectors, and the classification is realized by finding the optimize panel h that can tell all the vectors apart.

Denoting the normal vector of h as \vec{n} , to any point β , the distance to h is $dis(\beta) = \left| \overrightarrow{\alpha\beta} \right| \cdot \cos(\vartheta) = \overrightarrow{\alpha\beta} \cdot \vec{n}$, where ϑ is the angle between vector $\overrightarrow{\alpha\beta}$ and \vec{n} . In this article, the distance is obtained directly based on the tool provided by Chang and Lin.⁴⁵ And trust evidence is generated based on the concept that the greater the distance, the more uncertainty of the evidence. The process of generating trust evidence can be found in the “Trust evidence based on risk of information flow leakage” section.

MCS

Just as mentioned in the “Related works” section, the probabilities of message propagation among users are

calculated to measure the risk of privacy leakage. However, the number of paths between two users is tremendous, and the flow path may not be the shortest path. That is to say, information flow prediction based on computing flow probability on every edge is infeasible.²⁶ Even though it is hard to get the probability of information flow between two users in every detail, an approximate value can be obtained based on MCS. MCS²⁷ is a kind of method that rely on repeated random sampling to obtain the calculation result. The essential idea is to find undiscovered laws through a lot of experiments, and it is very useful when the solving process of traditional approach methods is too difficult or too complex. By modifying the network topology and detecting the existence of the path for many times, the probability of information flow is obtained based on the concept of MCS.

Trust evaluation based on the combination of evidence

There are four parts in our schemes. In the first part, the individualization of trust evaluation is computed. We name this kind of individualization as User-Will, which is constituted by a set of weights and a set of value scopes to generate trust evidence. In the second part, features of user being evaluated are converted into a set of trust evidence based on User-Will. In the third part, another piece of trust evidence is obtained based on predicting the risk of information flow in future. In

Table 2. Summary of symbols.

Symbol	Description
p_i	Share probability of user u_i
$\rho_{i,j}$	Share probability between user u_i and user u_j
t	Share decision between user u_i and user u_j
$\sigma(u_i, u_j)$	Information flow probability from u_i to u_j
m_α	Trust evidence of attribute α
$m(\beta)$	Basic probability assignment of focal-element β
s_α	Scope value of attribute
ξ_i	User-Will of user u_i
F	The collection of user features

the last part, trust decision is made based on the combination of evidence and User-Will. Table 2 is the summary of symbols used in our scheme:

Determination of User-Will

Just as mentioned in the ‘‘Related works’’ section, different users may care about different features when making trust decision; we denote this kind of trust-making individualization as User-Will. It contains two parts, weight set and scope set. Weight set is the importance degree of each user features and scope set is a set of value range to generate trust evidence. The method for generating value range and trust evidence will be introduced later, but the way to determine weights will be introduced in the following part.

To compute the weight set in the User-Will of u_i , the items $us : \{(u_1, t_1), (u_2, t_2), \dots\}$ are extracted from the training set where all the trust decisions (t) are made by u_i , first. Then, define F as the collection of user features which u_i may cares about when making trust decisions, such as the number of fans, the number of blogs he or she posts in the last few days, and so on. To each user features $f \in F$, a piece of trust evidence is generated to represent the trust degree based on f . Assuming there are four elements in $F : \{f_1, f_2, f_3, f_4\}$, to each item u in us , four pieces of evidence m_1, m_2, m_3, m_4 will be generated based on the value of f_1, f_2, f_3, f_4 . m_{rst} is defined as the combination result of m_1, m_2, m_3, m_4 under the weight set w_1, w_2, w_3, w_4 . There are two cases that m_{rst} stands for the correct decision:

1. $m_{rst}(Trust) \geq m_{rst}(Distrust)$ and u_i trust u in the training set.
2. $m_{rst}(Trust) < m_{rst}(Distrust)$ and u is not trusted by u_i in the training set.

Define the function $g(w_1, w_2, w_3, w_4) = cor(w_1, w_2, w_3, w_4) / num(us)$ where $cor(w_1, w_2, w_3, w_4)$ represents the correct number of decision based on weight set w_1, w_2, w_3, w_4 . Then, the weights part of User-Will can be determined by finding out the parameters w'_1, w'_2, w'_3, w'_4 that make function g maximum.

To generate trust evidence, divide us into $vals_t$ and $vals_{dt}$, where all the users in $vals_t$ are trusted by u_i and the ones in $vals_{dt}$ are not. Taking feature $f \in F$, for instance, scopes sco_t, sco_{dt} can be determined as $sco_t = (\min(vals_t(f)), \max(vals_t(f)))$, $sco_{dt} = (\min(vals_{dt}(f)), \max(vals_{dt}(f)))$, where $vals_t(f)$ and $vals_{dt}(f)$ are the collections of f values of users in $vals_t$ and $vals_{dt}$, respectively. When $soc_t \cap sco_{dt} \neq \emptyset$, ambiguity scope sco_{am} will be $sco_{am} = soc_t \cap sco_{dt}$ or $sco_{am} = s_i - s_j$, where s_i and s_j satisfy the equation $s_i + sco_t = s_j + sco_{dt}$. Denoting v as a special scope $sco_v = [v, v]$, and the BPA of m_v will be obtained based on the

Algorithm 1: Generating User-Will of user u_i

Input: subset of train set which is marked as set_i , user u_i
Output: User-Will ξ_i of u_i

- 1: Define two maps $mt_s \leftarrow \emptyset, mdt_s \leftarrow \emptyset$; Define an ordered set $os \leftarrow \emptyset$;
- 2: Tell set_i apart into α_i and β_i where α_i is user trusted by u_i and β_i is not trusted by u_i
- 3: **for** $\forall f \in F$ **do**
- 4: extract max_f and min_f of f from α_i
- 5: extract max'_f and min'_f of f from β_i
- 6: set $sco_t = (\min_f, \max_f), sco_{dt} = (\min'_f, \max'_f)$
- 7: generate scope sco_{am} based on sco_t and sco_{dt}
- 8: set $\xi_i(sco_f) = \langle sco_t, sco_{dt}, sco_{am} \rangle$
- 9: Generate random evidence weights $\varpi : \varpi_1, \varpi_2, \dots, \varpi_f$ and set $g(\varpi_1, \varpi_2, \dots, \varpi_f) = 0$
- 10: **for** each item $\forall it \in set_i$: **do**
- 11: set $ms = \emptyset$
- 12: extract $sco_t, sco_{dt}, sco_{am}$ from $\xi_i(sco_f)$
- 13: **for** $\forall f \in F$ **do**
- 14: set the value of feature f belongs to it as $val(f)$
- 15: set the scope similarity between $[val(f), val(f)]$ and $sco_t, sco_{dt}, sco_{am}$ as $sim_t, sim_{dt}, sim_{am}$
- 16: set $sum = sim_t + sim_{dt} + sim_{am}$
- 17: set $m(t) = \frac{sim_t}{sum}, m(dt) = \frac{sim_{dt}}{sum}, m(am) = \frac{sim_{am}}{sum}$
- 18: add m to ms
- 19: $m_{\oplus} = m_1 \oplus_{\varpi} m_2 \oplus_{\varpi} \dots \oplus_{\varpi} m_f$
- 20: Set $g = g + \kappa$, κ is a binary value based on the rule bellows:
- 21: **if** $m_{\oplus}(t) > m_{\oplus}(t)$ and it is trusted by u_i **then**
- 22: $\kappa = 1$
- 23: **if** $m_{\oplus}(t) < m_{\oplus}(t)$ and it is distrusted by u_i **then**
- 24: $\kappa = 1$
- 25: $\varpi' = Argmax_{\varpi} g$
- 26: $\forall f \in F$, set $\xi_i(f) = \varpi'_f$ when $\varpi'_f \neq 0$
- 27: Return ξ_i as the User-Will belongs to user u_i

similarity among $sco_t, sco_{dt}, sco_{am}, sco_v$. Algorithm 1 is the detailed steps in finding User-Will of u_i .

In Algorithm 1, the training set is separated into two parts in lines 1 and 2. Trust scope, distrust scope, and ambiguity scope are generated in lines 3 to 8. A random weight set and function g are set in line 9. In lines 10 to 24, trust evidence is generated and combined. Similarity between scopes is defined as $sim_s(a, b) = 1 / (1 + \varepsilon(a, b))$; and $\varepsilon(a, b) = \int_{-0.5}^{0.5} \{ [0.5 \cdot (a_{min} + a_{max}) + x(a_{max} - a_{min})] - (1/2) \cdot [0.5 \cdot (b_{min} + b_{max}) + x(b_{max} - b_{min})] \}^2 dx$. In lines 25 to 27, the weights that make function g optimized are denoted as ρ' , and the User-Will of user u_i is denoted as ξ_i . Based on the process in Algorithm 1, the attributes that the user is not interested in when making trust decision are filtered out, and the importance degree of each user features is determined.

Trust evidence based on User-Will

We assume that F is constituted by activity degree, Bi-Jacard degree, group degree, and reputation degree.

Activity degree is defined as $acti = (N_{commit}^n + N_{tweet}^n)/n$. n is the threshold that how long the period is and $N_{commit}^n, N_{tweet}^n$ stands for the number of comments and the number of tweets posted in the last n months. Besides, people trend to accept the recommendation from the people who is close to them. Bi-Jacard number is defined as $|\tau_i \cap \tau_j|/|\tau_i \cup \tau_j|$, where τ_i is the set of followers who belong to user u_i . And group degree that belongs to u_i and u_j is defined as $gp_{i,j} = (|tags_i \cap tags_j|/|tags_i \cup tags_j|) \times (|groups_i \cap groups_j|/|groups_i \cup groups_j|)$; $tags_i$ and $groups_i$ are the interest tags and communion groups that belong to user u_i . And the last element is the reputation degree which is defined as $rep_{i,j} = (avg(R_j) + avg(R_i^i))/2$. R_j is the collection of scores that other users rank u_j , and R_i^i is the collection of scores ranked by u_i . The value ranges of activity degree, Bi-Jacard number, and group degree are all $[0,1]$; however, the range of reputation degree is $[0,\mu]$, where μ is the best rank of the current social network platform. Taking Epinions, for instance, the best rank is 5 and the range of reputation degree for it is $[0,5]$.

To any feature $f \in F$, extract scope set $sco_t, sco_{dt}, sco_{am}$ from $\xi(sco_f)$, and trust evidence $m_f = (m_f(t) = sim_t/sum, m_f(dt) = sim_{dt}/sum, m_f(am) = sim_{am}/sum)$. sum is the summation of $sim_t + sim_{dt} + sim_{am}$, and sim is the similarity between two scopes.

Trust evidence based on risk of information flow leakage

Besides, with user features mentioned, trust degree may be affected by the flow of information in the future. Privacy information may be available to the blacklist of the resource owner after being transmitted many times. And we take the scenario that if the probability that information flow between u_j to any one of the blacklist of u_i is high, the trust degree that u_i to u_j should be reduced. However, the flow path of message may be detoured or paralleled in most cases just as Figure 1.

In Figure 1, circles in dotted line and circles in solid line are the users in OSNs. The former one represents

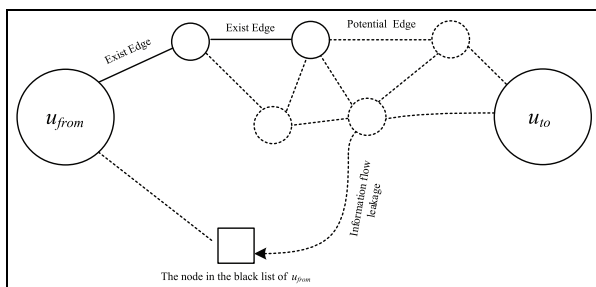


Figure 1. Information flow threat.

the remote users and the later one stands for the adjacent ones. User u_{to} may forward the message from u_{from} to the blacklist of u_{from} , which leads to a privacy leakage. Taking user u_{from} , for instance, the blacklist of u_{from} is $b : \{b_1, b_2, \dots, b_n\}$. When another u_{to} is requesting from u_{from} , the request will be denied if the probability of information flow between u_{to} to any one of b is too high. As the path of information flow is tremendous, the path may be parallel or detoured. The prediction of information flow is hard to be realized through predicting the flow on each edge but can be accomplished by looping MCS many times.

In the process of information flow, the flow between any two points is not unique. Even though u_i and u_j are linked directly, the flow path may through a third-party node for relationship between u_i and u_j is not strong enough. If one edge is removed, the flow of information can proceed on another path. However, if most edges are removed, the flow of information may not continue as there is no path between them. And information flow probability can be obtained based on computing the count of edges removed before the disappearance of the last path between u_i and u_j . Let σ' be the collection of information flow probabilities between u_j and each item of the blacklist of u_i . And vector $\langle \max(\sigma'), avg(\sigma'), t \rangle$ is one item in the training set for SVM classification (t is the tag showing whether u_j is trusted by u_i). The trust evidence based on information flow can be obtained based on classification result and classification distance which is shown in Algorithm 2.

Trust evidence combination

Taking trust evaluation between u_i to u_j , for instance, User-Will ξ_i of u_i is determined based on Algorithm 1. To each feature $f \in F$, trust evidence m_f can be determined based on ξ_i and the value of feature f belongs to u_j . Besides, trust evidence m_{flow} based on predicting information flow can be obtained based on Algorithm 2. ψ is the weight part of ξ_i and m_{will} is the combination result of all the evidence obtained by ξ_i : $m_{will} = \sum_{f \in F} \oplus_{\psi} m_f$. Based on m_{will} and m_{flow} , trust evaluation result $m_{RST} = m_{will} \oplus m_{flow}$.

Experiment and analysis

Experiment

In experiment part, we implement methods AveR-MaxT, AveR-WaveT, MaxR-MaxT, MaxR-WaveT, SWTrust*,²⁵ and GFTrust¹¹ which are proposed recently, and the experiment is based on the dataset Epinions, which is available on http://www.trustlet.org/extended_epinions.html. There are three files in the dataset, and the first file is the record of trust decision between each two users. Column MY_ID stands for

Algorithm 2: Generating trust evidence based on information flow prediction

Input: topology of the OSNs, users u_i, u_j

Output: Trust evidence m_{flow} based on predicting information flow

- 1: Extract the sub-graph of OSNs as $SG_{u_i} = (v', e')$ where all nodes in SG_{u_i} can reach u_i within 3 hops.
- 2: extract $B : b_1, b_2, \dots, b_k$ from v' as the blacklist of u_i and the rest as $W : w_1, w_2, \dots, w_s$
- 3: $s_{svm} \leftarrow \emptyset; O \leftarrow \emptyset$
- 4: **for** each user $k \in v'$ **do**
- 5: set $\sigma' = \emptyset$
- 6: **for** each element $b \in B$ **do**:
- 7: set $n_b = 0$
- 8: Select a random edge e from e' , mark the two nodes on e as δ and γ .
- 9: Denote the inflow and outflow of γ as γ_{in} and γ_{out}
- 10: Denote $p_\gamma = \gamma_{out}/\gamma_{in}$ if $\gamma_{in} \geq \gamma_{out}$ or $p_\gamma = 1$
- 11: Generating a random t_k ranging from 0 to 1
- 12: Drop the edge from SG_{u_i} if $t_k > \rho_{\delta, \gamma}$
- 13: Set $n_b = n_b + 1$ if there is a path exists between u_j and b
- 14: Repeat steps 7–12 for n_r times
- 15: Set the information flow rate between u_j to b as $\sigma(u_j, b) = n_b/n_r$
- 16: Add $\sigma(u_j, b)$ into σ'
- 17: Add $s = (tr_k, \max(\sigma'), \text{avg}(\sigma'))$ into s_{svm} , where $tr_k = 1$ if u_i trust k , or $tr_k = 0$
- 18: Classify s_{svm} by RBF svm type, mark the accuracy and hyperplane as acc and h
- 19: Compute the collection of information flow probability σ'' that u_j to each item in B
- 20: Denote $dis(u_j)$ as the distance that point $p : \langle \max(\sigma''), \text{avg}(\sigma'') \rangle$ to h
- 21: Set $m_{flow}(Trust) = dis(u_j) \times acc$ if p is above the plane h
- 22: Set $m_{flow}(Distrust) = (1 - dis(u_j)) \times acc$ if p is below the plane h
- 23: Set $m_{flow}(Trust) = (1 - dis(u_j)) \times acc$ if p is below the plane h
- 24: Set $m_{flow}(Distrust) = dis(u_j) \times acc$ if p is above the plane h
- 25: Set $m_{flow}(Ambiguity) = 1 - acc$
- 26: Return m_{flow} as the trust evidence of u_i based on information flow predict

the ID of the user who made the decision and column OTHER_ID is the ID of the user being evaluated. And VALUE column is the trust decision, where 1 stands for trust and -1 stands for distrust. In the second file, there are three columns CONTENT_ID, AUTHOR_ID, and SUBJECT_ID in it. CONTENT_ID is the ID of the comment or blogs that send by AUTHOR_ID. SUBJECT_ID denotes the ID of items being commented by user AUTHOR_ID, and the ID of this comment is COMMENT_ID. In the third file, there are eight columns in it but only three of them, OBJECT_ID, MEMBER_ID, and RATING, are useful. OBJECT_ID is the comment or blogs ranked by the user with its ID MEMBER_ID; the score of ranking is RATING which is ranging from 1 to 6. We use the same method in Jiang et al.^{11,25} to extract the same subset of this dataset, which is used in this experiment.

Just as mentioned in the “Trust evaluation based on the combination of evidence” section, trust evaluation result is based on User-Will and trust evidence. User-Will of each user is extracted by Algorithm 1, and BOE of each trust evidence is shown in Table 3.

The first column “MY_ID” in Table 3 stands for the ID of the user who made the trust decision and the weight of each user features belongs to “MY_ID.” And the value in column “OTHER_ID” stands for the ID of the user being evaluated. Based on Algorithm 2, weight of each evidence is determined, which is shown in the third column. Evidence m_{acti} , m_{rep} , m_{bj} , and m_{gp} are converted from features activity degree, reputation degree, Bi-Jacard degree, and group degree. Besides, the threat of information flow in the future is mapped to evidence m_{gp} , which is also shown in the fourth column. By combining the trust evidence above, evaluation result is obtained and shown in the column “Combination result.” The last column in Table 3 stands for the correct decision.

By counting the correct decisions that each method made, the comparison of accuracy and the comparison of F-Score are shown in Figures 2 and 3.

From Figures 2 and 3, it is apparent that the accuracy and F-Score of our scheme are the largest. As the trust decisions made by users are Boolean, both two trust evaluation results may be correct even though they are not same. Taking trust degree between u_i and u_j equals 0.6, for instance, both 0.9 and 0.7 can make the correct decision but the difference between 0.9 and 0.6 is larger than the difference between 0.6 and 0.7. We use the technique in Richardson et al.⁴⁶ to transform the trust values to be continuous in $[0,1]$. And we compute “Mean Error” of each method, which is defined in Jiang et al.,¹¹ to illustrate this kind of difference, which is shown in Figure 4 and Table 4.

According to Table 4, the F-Score of our method is the largest and the Mean Error of our scheme is the smallest, which means the trust evaluation is more precise and accurate.

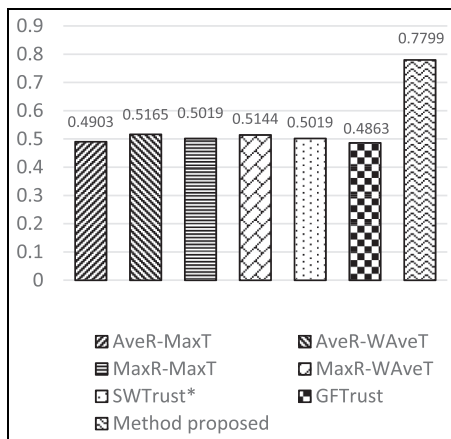
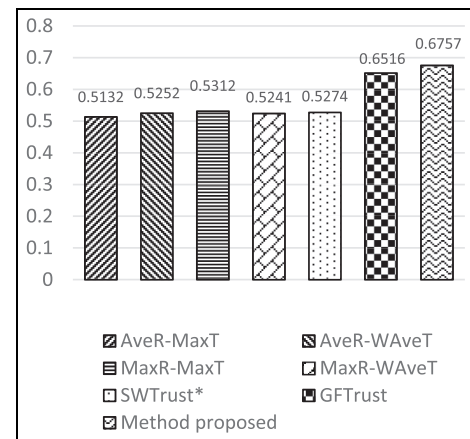
Analysis

In the above experiments, we compared the performance of methods proposed in Jiang et al.^{11,25} and our method from such aspects as Accuracy, Recall, F-Score, and Mean Error. It shows that our scheme has better performance in these aspects. For this result, we think the reason is mainly as bellows.

All the methods above, including ours, try to build the model of the mapping relationship between the trust decision and the user features, which is a classification process to divide the users into “trusted” and “untrusted” by their features, in essence. However, as we all know, trust is very subjective, that is, the same person usually has different trust values in the eyes of

Table 3. Trust evidence and combination result.

MY_ID	OTHER_ID	Trust evidence	Combination result
513794 $\xi_{acti} = 0.196$ $\xi_{rep} = 0.352$ $\xi_{bj} = 0.095$ $\xi_{gp} = 0.357$	433727	$m_{acti} = (5.7e - 4, 0.4998, 0.4996)$ $m_{rep} = (0.3347, 0.3326, 0.3326)$ $m_{bj} = (0.3169, 0.3423, 0.3408)$ $m_{gp} = (0.3371, 0.3314, 0.3315)$ $m_{flow} = (0.8988, 0.0321, 0.0689)$	$m_{RST} = (0.7141, 0.2765, 0.0094)$
	3045494660	$m_{acti} = (5.8e - 4, 0.4997, 0.4996)$ $m_{rep} = (0.3337, 0.3331, 0.3332)$ $m_{bj} = (0.3168, 0.3417, 0.3413)$ $m_{gp} = (0.3374, 0.3312, 0.3314)$ $m_{flow} = (0.2791, 0.6519, 0.0689)$	$m_{RST} = (0.2411, 0.7526, 0.0063)$
	19733	$m_{acti} = (5.9e - 4, 0.4997, 0.4996)$ $m_{rep} = (0.3298, 0.3351, 0.3351)$ $m_{bj} = (0.3428, 0.3271, 0.3301)$ $m_{gp} = (0.3256, 0.3372, 0.3372)$ $m_{flow} = (0.4534, 0.4777, 0.0689)$	$m_{RST} = (0.3912, 0.5972, 0.0116)$
201915 $\xi_{acti} = 0.379$ $\xi_{rep} = 0.336$ $\xi_{bj} = 0.159$ $\xi_{gp} = 0.126$	2774503300	$m_{acti} = (0.1062, 0.4468, 0.4468)$ $m_{rep} = (0.3433, 0.3149, 0.3416)$ $m_{bj} = (0.3252, 0.3495, 0.3252)$ $m_{gp} = (0.3365, 0.3317, 0.3317)$ $m_{flow} = (0.9157, 0.0105, 0.0736)$	$m_{RST} = (0.5498, 0.4351, 0.0151)$
	293737	$m_{acti} = (0.1062, 0.4468, 0.4468)$ $m_{rep} = (0.3435, 0.3154, 0.3411)$ $m_{bj} = (0.3246, 0.3506, 0.3246)$ $m_{gp} = (0.3363, 0.3318, 0.3318)$ $m_{flow} = (0.0529, 0.8733, 0.0736)$	$m_{RST} = (0.2411, 0.7526, 0.0063)$
	2776928132	$m_{acti} = (0.1062, 0.4468, 0.4468)$ $m_{rep} = (0.3434, 0.3149, 0.3417)$ $m_{bj} = (0.3251, 0.3498, 0.3251)$ $m_{gp} = (0.3364, 0.3317, 0.3317)$ $m_{flow} = (0.0418, 0.8844, 0.0736)$	$m_{RST} = (0.1176, 0.8698, 0.0124)$

**Figure 2.** Accuracy of each method.**Figure 3.** F-Score of each method.

different people. If the trust decisions in the training set are made by different users, it will be very difficult to build the classification model, because the trust criterion is different from person to person. In our scheme, we choose the sample data from one user's trust decisions to build the training set and obtain the

classification model for himself. Since the inclination of one person is much more obvious, the classification model is easier to be built correctly, and, at the same time, one model for one user embodies the individualization and subjectivity quite well. The experiments prove it as we expect.

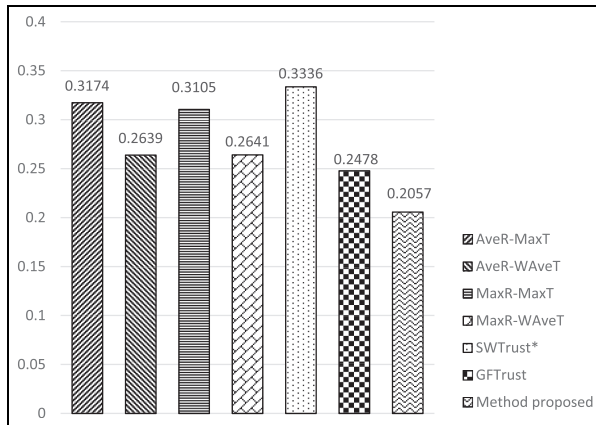


Figure 4. Mean Error of each method.

Table 4. Mean Error and F-Score of each method.

Method	Mean Error	Recall	Accuracy	F-Score
AveR-MaxT	0.3174	0.5385	0.4903	0.5132
AveR-WAveT	0.2639	0.5342	0.5165	0.5252
MaxR-MaxT	0.3105	0.5641	0.5019	0.5312
MaxR-WAveT	0.2641	0.5342	0.5144	0.5241
SWTrust*	0.3336	0.5556	0.5019	0.5274
GFTrust	0.2478	0.9872	0.4863	0.6516
Method proposed	0.2057	0.5961	0.7799	0.6757

The F-Score and Mean Error of our method are 0.675 and 0.205, respectively. And the max F-Score and min Mean Error of methods being compared are 0.651 and 0.247. The experiment result shows that the trust decisions obtained through the proposed scheme are more accurate and better agreed with the real decisions according to the data from Epinions.

Conclusion

In this article, a new trust evaluation scheme based on evidence theory is proposed.

Our study achieved a better performance by focusing the issues as follows:

1. Determine the importance degree of each user features that each user cares about in making the decision to realize the individualization of trust evaluation.
2. Quantifying the risk of privacy leakage by information flow prediction to make the trust evaluation more comprehensive.
3. Use trust evidence to indicate the probability of trust, probability of distrust, and probability of ambiguity at the same time.

Compared with the existing methods, our proposed method achieves the highest accuracy and minimal error in the dataset Epinions. However, the weight part of User-Will does not contain the weight of trust evidence based on information flow risk, and the weight determination of this evidence will be our future work.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work was sponsored by National Natural Science Foundation of China, Grant No. 61772174 and Grant No. 61370220; Plan for Scientific Innovation Talent of Henan Province, Grant No. 174200510011; Program for Innovative Research Team (in Science and Technology) in University of Henan Province, Grant No. 15IRTSTHN010; Program for Henan Province Science and Technology, Grant No. 142102210425; Natural Science Foundation of Henan Province, Grant No. 162300410094; and Project of the Cultivation Fund of Science and Technology Achievements of Henan University of Science and Technology, Grant No. 2015BZCG01.

References

1. Wang X, Zhang Z, Wen J, et al. A novel crowd evaluation method for security and trustworthiness of online social networks platforms based on signaling theory. *J Comput Sci* 2017; 26: 468–477.
2. Zhang Z, Sun R, Zhao C, et al. CyVOD: a novel trinity multimedia social network scheme. *Multimed Tools Appl* 2016; 76: 18513–18529.
3. Zhang Z, Sun R, Wang X, et al. A situational analytic method for user behavior pattern in multimedia social networks. *IEEE Trans Big Data*, 2017; 99: 1–1.
4. Wang AH. Don't follow me: spam detection in twitter. In: *Proceedings of the international conference on security and cryptography*, Athens, 26–28 July 2010, pp.142–151. Piscataway, NJ: IEEE.
5. Miller Z, Dickinson B, Deitrick W, et al. Twitter spammer detection using data stream clustering. *Inform Sciences* 2014; 260(1): 64–73.
6. Hu VC, Kuhn DR, Ferraiolo DF, et al. Attribute-based access control. *Computer* 2015; 48(2): 85–88.
7. Zhang Z, Wang Z and Niu D. A novel approach to rights sharing enabling digital rights management for mobile multimedia. *Multimed Tools Appl* 2015; 74(16): 6255–6271.
8. Lah U and Lewis J. The effect of expertise on the usability of a digital rights management sharing application. *IEEE Software*. Epub ahead of print January 2015. DOI: 10.1109/MS.2015.104.

9. Lah U and Lewis JR. How expertise affects a digital-rights-management-sharing application's usability. *IEEE Software* 2016; 33(3): 76–82.
10. Liu L and Jia H. Trust evaluation via large-scale complex service-oriented online social networks. *IEEE T Syst Man Cy S* 2015; 45(11): 1402–1412.
11. Jiang W, Wu J, Li F, et al. Trust evaluation in online social networks using generalized network flow. *IEEE T Comput* 2016; 65(3): 952–963.
12. Chiregi M and Navimipour NJ. A new method for trust and reputation evaluation in the cloud environments using the recommendations of opinion leaders' entities and removing the effect of troll entities. *Comput Hum Behav* 2016; 60: 280–292.
13. Zhao K and Pan L. A machine learning based trust evaluation framework for online social networks. In: *Proceedings of the IEEE 13th international conference on trust, security and privacy in computing and communications*, Beijing, China, 24–26 September 2014, pp.69–74. Piscataway, NJ: IEEE.
14. Zhang Z and Gupta BB. Social media security and trustworthiness: overview and new direction. *Future Gener Comp Sy* 2016; 86: 914–925.
15. Brown P and Feng J. Measuring user influence on twitter using modified k-shell decomposition. In: *The social mobile web, papers from the 2011 ICWSM workshop*, Barcelona, Catalonia, Spain, July 2010. DBLP.
16. Silva A, Meira W and Zaki M. ProfileRank: finding relevant content and influential users based on information diffusion. In: *Proceedings of the 7th workshop on social network mining & analysis*, Chicago, IL, 11–14 August 2013, pp.1–9. New York: ACM Press.
17. Tzolmon B and Lee KS. A graph-based reliable user classification. *Lect Notes Electr En* 2014; 285: 61–68.
18. Mármol FG and Pérez GM. Towards pre-standardization of trust and reputation models for distributed and heterogeneous systems. *Comp Stand Inter* 2010; 32(4): 185–196.
19. Watts DJ. Networks, dynamics, and the small-world phenomenon. *Am J Sociol* 1999; 105(2): 493–527.
20. Zhang Z and Wang K. A trust model for multimedia social networks. *Soc Netw Anal Min* 2013; 3(4): 969–979.
21. Yang L, Qiao Y, Liu Z, et al. Identifying opinion leader nodes in online social networks with a new closeness evaluation algorithm. *Soft Comp* 2018; 22: 453–464.
22. Bozzon A, Brambilla M, Ceri S, et al. Choosing the right crowd: expert finding in social networks. In: *Proceedings of the 16th international conference on extending database technology*, Genoa, 18–22 March 2013, pp.637–648. New York: ACM Press.
23. Wang G and Wu J. FlowTrust: trust inference with network flows. *Front Comput Sci Chi* 2011; 5(2): 181–194.
24. Ranjbar A and Maheswaran M. Using community structure to control information sharing in online social networks. *Comput Commun* 2014; 41(5): 11–21.
25. Jiang W, Wang G and Wu J. Generating trusted graphs for trust evaluation in online social networks. *Future Gener Comp Sy* 2014; 31(1): 48–58.
26. Garey MR and Johnson DS. *Computers and intractability: a guide to the theory of NP-completeness*. W. H. Freeman, 1986.
27. Peters GW and Sisson S. Bayesian inference, Monte Carlo sampling and operational risk. *J Oper Risk* 2017; 1(3): 27–50.
28. Dempster AP. Upper and lower probabilities induced by a multivalued mapping. In: Yager RR and Liu L (eds) *Classic works of the Dempster-Shafer theory of belief functions*. Berlin; Heidelberg: Springer, 2008, pp.57–72.
29. Lin G, Liang J and Qian Y. An information fusion approach by combining multigranulation rough sets and evidence theory. *Inform Sciences* 2015; 314: 184–199.
30. Wang J, Hu Y, Xiao F, et al. A novel method to use fuzzy soft sets in decision making based on ambiguity measure and Dempster-Shafer theory of evidence: an application in medical diagnosis. *Artif Intell Med* 2016; 69: 1–11.
31. Li Z, Wen G and Xie N. An approach to fuzzy soft sets in decision making based on grey relational analysis and Dempster-Shafer theory of evidence: an application in medical diagnosis. *Artif Intell Med* 2015; 64(3): 161–171.
32. Huang S, Su X, Hu Y, et al. A new decision-making method by incomplete preferences based on evidence distance. *Knowl-Based Syst* 2014; 56(3): 264–272.
33. Dong G and Kuang G. Target recognition via information aggregation through Dempster-Shafer's evidence theory. *IEEE Geosci Remote S* 2017; 12(6): 1247–1251.
34. Wang P, Shang CX and Han ZZ. Target recognition fusion based on improved evidence theory. *Electr Optics Contr* 2015; (9): 46–49.
35. Mekouar S, Ibrahim K and Bouyakhf EH. Inferring trust relationships in the social network: evidence theory approach. In: *Proceedings of the international wireless communications and mobile computing conference*, Nicosia, Cyprus, 4–8 August 2014, pp.470–475. IEEE.
36. Jiang J, Xiang J, Zhou H, et al. Trust calculation model based on social network and evidence theory. In: *Proceedings of the international joint conference on service sciences*, Taipei, Taiwan, 25–27 May 2011, pp.173–177. Piscataway, NJ: IEEE.
37. Zhao QY, Zuo WL, Tian ZS, et al. A method for assessment of trust relationship strength based on the improved d-s evidence theory. *Chin J Comput* 2014; 37: 873–884.
38. Wang J, Qiao K, Zhang Z, et al. A new conflict management method in Dempster-Shafer theory. *Int J Distrib Sens N*. Epub ahead of print 9 March 2017. DOI: 10.1177/1550147717696506.
39. Chang CC and Lin CJ. LIBSVM: a library for support vector machines. *ACM T Intel Syst Tec* 2001; 2(3): 27.
40. Wang J, Xiao F, Deng X, et al. Weighted evidence combination based on distance of evidence and entropy function. *Int J Distrib Sens N*. Epub ahead of print 25 July. DOI: 10.1177/155014773218784.
41. Deng Y, Zhu ZF and Liu Q. Combining belief functions based on distance of evidence. *Decis Support Syst* 2004; 38(3): 489–493.
42. Han DQ, Yong D, Han CZ, et al. Weighted evidence combination based on distance of evidence and uncertainty measure. *J Infrared Millim W* 2011; 30(5): 396–400.
43. Jousselme AL, Grenier D and Boss L. A new distance between two bodies of evidence. *Inform Fusion* 2001; 2(2): 91–101.

44. Govindarajan M. Recognition of handwritten numerals using RBF-SVM hybrid model. *Int Arab J Inf Techn* 2016; 13: 344–350.
45. Chang CC and Lin CJ. LIBSVM: a library for support vector machines. *ACM T Intel Syst Tec* 2011; 2: 27.
46. Richardson M, Agrawal R and Domingos P. Trust management for the semantic web. In: *Proceedings of the international semantic web conference*, Sanibel, FL, 20–23 October 2003, pp.351–368. New York: ACM Press.