SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2013, Vol. 56 108102:1–108102:14
doi: 10.1007/s11432-013-5005-4

SelfTrust: leveraging self-assessment for trust inference in Internetware

YAO Yuan[1,2], XU Feng[1,2]*, REN YongLi[3], TONG HangHang[4] & LÜ Jian[1,2]

[1]*State Key Laboratory for Novel Software Technology, Nanjing* 210023*, China;*
[2]*Department of Computer Science and Technology, Nanjing University* 210093*, China;*
[3]*School of Information Technology, Deakin University, Australia;*
[4]*City College, City University of New York, USA*

Received February 19, 2013; accepted May 17, 2013

Abstract Internetware is envisioned as a new software paradigm for software development in platforms such as the Internet. The reliability of the developed software becomes a key challenge due to the open, dynamic and uncertain nature of such environment. To make the development more reliable, it is necessary to evaluate the trustworthiness of the resource providers or potential working partners. To this end, we propose a novel trust inference approach to evaluating the trustworthiness of potential partners to guide the software development in Internetware. The main insight of our approach is to employ the self-assessment information in order to improve the trust inference accuracy. Especially, we first extend the balance theory and the status theory from social science to incorporate self-assessment, and then propose a machine learning framework to extract several features from the extended theories and infer trustworthiness scores based on these features. Experimental results on a real software developer network show that the self-assessment information truly helps to improve the accuracy of trust inference, and the proposed SelfTrust model is more accurate than other state-of-the-art methods.

Keywords trust inference, self-assessment, balance theory, status theory, Internetware

Citation Yao Y, Xu F, Ren Y L, et al. SelfTrust: leveraging self-assessment for trust inference in Internetware. Sci China Inf Sci, 2013, 56: 108102(14), doi: 10.1007/s11432-013-5005-4

1 Introduction

The explosive development of the WWW technology makes the Internet the biggest interactive on-line community, where Internetware software and systems are developed and deployed based on the interactions between users who are geographically away from each other. The Internet platform is known to be open, dynamic and uncertain, which allows the existence of selfish or even malicious service providers; therefore, finding the trustworthy partners in such environment becomes a key challenge to the reliability of the systems [1–3]. To make the development more reliable, it is necessary to evaluate the trustworthiness of the resource providers or potential working partners. In literature, many trust inference approaches have been proposed to help participants to evaluate the trustworthiness of the service providers [4,5]. Similar trust inference is also proposed in many other Web-based applications including e-commerce [6], peer-to-peer networks [7], and mobile ad hoc networks [8].

*Corresponding author (email: xf@nju.edu.cn)

Over the years, the existing approaches for trust inference are mainly based on the *transitivity* property [9] of trust, which basically means that if *Alice* trusts *Bob* and *Bob* trusts *Carol*, *Alice* would also trust *Carol* to a certain extent. Typically, these methods would take the existing trust ratings as input, and output a numerical score to indicate how much a trustor should trust a trustee. However, the input trust ratings as well as the output trustworthiness scores could imply different meanings for different users. This is also supported by social science where researchers have found that trust varies subjectively as a result of individual predispositions (i.e., one's inclination to take risks, degree of tolerance of potential disappointment, etc.) [10]. Therefore, to improve the accuracy of trust inference, it is necessary to incorporate individual personality into the process and tailor the inference result for each user.

In this paper, we take into consideration the self-assessment information to capture individual differences. Specifically, our approach is derived from the well-known social balance theory [11,12] and status theory [13]. Both of these two theories are proposed to model the attitude between human beings, and thus they are applicable in modeling trust which is a subcategory of attitude. To incorporate self-assessment, however, we need to make some extensions on the theories. For the balance theory, we first redefine the link signs by comparing the original trust rating to user's self-assessment. Then, we further employ the low-rank structure [14] of signed social networks for self-assessment leveraged trust inference. For the status theory, we extend the theory to the numerical setting, and take self-assessment into account by defining the status gap. The rational behind these extensions is that while trust ratings in trust networks signify how a user thinks of others, self-assessment indicates how the user thinks of himself/herself. Therefore, the difference between these thoughts could better reflect the attitude in balance theory and status theory. Finally, we derive our model by combining balance theory and status theory under an unified framework where several features are extracted from the extended theories.

The main contributions of this paper are as follows:

1) We extend the social balance theory and status theory to incorporate self-assessment in the context of trust inference. Our method could be applied in Internetware as well as many other Web-based applications.

2) Experiment results on a real software developer network show that the self-assessment truly helps to improve the accuracy of trust inference, and the proposed approach is more accurate and more effective than other state-of-the-art methods.

The rest of the paper is organized as follows. Section 2 presents the problem definition. Section 3 describes the proposed self-assessment leveraged SelfTrust model, which extends weak balance theory and status theory, and then combines them. Section 4 presents the experimental results. Section 5 covers related work, and Section 6 concludes the paper.

2 Problem definition

In this section, we define the trust inference problem which incorporates the self-assessment information. Specifically, we first review the objective trust inference problem where the goal is to infer a unique trustworthiness score for each trustee. Then, we present the basic subjective trust inference problem whose goal is to infer a pairwise trustworthiness score for each trustor-trustee pair. Finally, we define the problem in this research, which further incorporates the self-assessment into the basic subjective trust inference problem.

2.1 Notations

We first list the notations we use throughout the paper. We use $\mathcal{R}$ to denote the historical ratings, where $\mathcal{R}(i,j)$ is the trust rating from user i to user j. In this work, we adopt the definition where trust is defined as the subjective probability by which an individual expects that another individual will perform well on a given action. As a result, $\mathcal{R}(i,j)$ could range from 0 to 1. 0 means no trust and 1 means full trust. When discussing the extensions on the balance theory, we will transform the unsigned trust rating to its signed version, and also transform the directed trust network into its undirected version. We use

$\tilde{\mathcal{R}}(i,j)$ to denote the trust ratings after the first transformation, and we use $\overline{\mathcal{R}}(i,j)$ to denote the trust ratings after the second transformation. As for the self-assessment, we denote it with $\mathcal{S}$, where $\mathcal{S}(i)$ is the self-assessment of user i. We assume that the goal of trust inference is to infer the unseen trustworthiness score from the user u to another user v, where u is the trustor and v is the unknown trustee to u.

2.2 Objective trust inference

Some trust inference algorithms aim at computing an objective trustworthiness score for each node in the network. Such objective metrics are suitable in the environments where each node has a fixed and objective trustworthiness score. For example, in peer-to-peer networks, such objective trustworthiness score indicates the probability by which the node would provide non-polluted files [7,15].

Typically, the input of trust inference is the historical feedback of previous interactions, and the output could be a continuous trust value with higher value indicating more trustworthiness. Based on such input and output, we define objective trust inference as the method to evaluate a trustee's trustworthiness score, where the score indicates to what extent *all* the trustors can rely on the trustee on performing a given action. The objective trust inference problem could be defined as:

Definition 1. Objective Trust Inference. Given: historical ratings $\mathcal{R}$, and a trustee v; Find: the estimated objective trustworthiness score of v for all trustors.

2.3 Subjective trust inference

With the development of the social network based applications, researchers begin to realize the importance of the subjectivity property of trust, and many subjective trust inference algorithms have been proposed including [16–22]. In contrast to objective trust inference, subjective trust inference recognizes that different trustors can form different opinions on the same trustee, and thus aims to provide pairwise trust inference result for each trustor-trustee pair [23].

Compared to the goal of a unique trustworthiness score for each trustee in objective trust inference, the goal of subjective trust inference is to evaluate a trustee's trustworthiness score for a given trustor, where the score indicates to what extent the *given* trustor can rely on the trustee on performing a given action. More formally, the basic subjective trust inference problem can be defined as follows:

Definition 2. Basic Subjective Trust Inference. Given: historical ratings $\mathcal{R}$, a trustor u, and a trustee v; Find: the estimated subjective trustworthiness score of v for trustor u.

In this work, we also put our focus on the subjective trust inference problem, and aim to incorporate the self-assessment information to further improve the inference accuracy. By considering self-assessment, the inference results will be adjusted based on individual user's personality. Formally, the trust inference problem in this paper can be defined as follows:

Definition 3. Subjective Trust Inference with Self-assessment. Given: historical ratings $\mathcal{R}$, users' self-assessment $\mathcal{S}$, a trustor u, and a trustee v; Find: the estimated subjective trustworthiness score of v for trustor u.

3 The SelfTrust model

In this section, we propose our approach to solve the problem defined in Definition 3. Specifically, we first discuss how to incorporate self-assessment into the balance theory and status theory, respectively. Then, we combine the extended balance and status theories together in a generic framework, SelfTrust.

3.1 Extending balance theory

In this study, we consider two outputs from this theory: one is from the theory output in local scale (see Figure 1), and the other one is from its structure character in global scale.

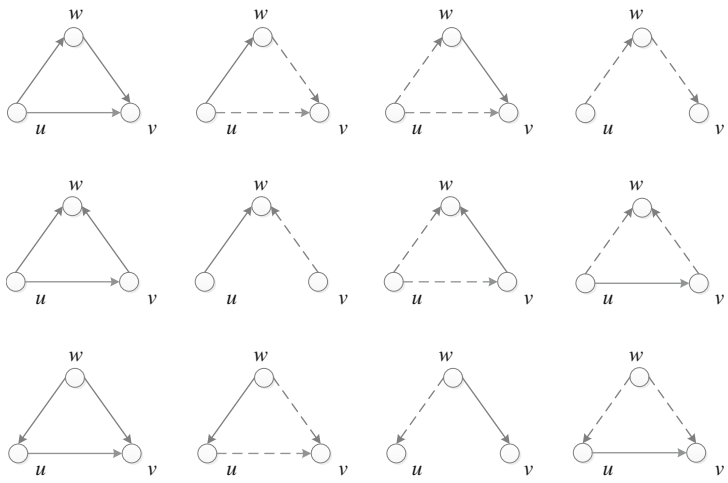

Figure 1 Illustration for weak balance theory in directed signed networks. Solid line indicates positive relationship and dashed line indicates negative relationship. The link between trustor u and trustee v is the theory output. The three triads in the figure do not have theory output because they fall into the fourth principle.

3.1.1 *Local scale of theory output*

Balance theory or the structural balance theory is proposed to model triadic social relationships in signed networks based on sentiments [11,24]. By categorizing the sentiments into two class, positive (friend) and negative (enemy), the balance theory can be interpreted by four principles: "the friend of my friend is my friend", "the enemy of my friend is my enemy", "the friend of my enemy is my enemy", and "the enemy of my enemy is my friend". These principles require that the number of positive signs should be odd in each triad. An extension of balance theory is proposed [12] by deleting the forth principle, and thus we have the definition of weak-balanced signed network:

Definition 4 (Weak-balanced signed network [12])**.** A signed network is weak balanced if and only if there is no triad that contains two positive links and one negative link in the network.

It has been found that this extended theory (weak balance theory) works better in trust-based social networks [25], and thus we will focus on this theory instead of the original balance theory.

We first interpret weak balance theory in directed signed networks, as shown in Figure 1. Unlike the undirected version in [24], we interpret weak balance in directed version because trust is asymmetric in nature, i.e., v trusts u cannot imply that u also trusts v. Figure 1 shows that all the triads that have output obey the above weak balance definition, i.e., no triad contains two positive links and one negative link. The triads that have output also follow the first three principles. For example, in the second row of the figure, the three triads that have output follow the rules "the friend of my friend is my friend", "the friend of my enemy is my enemy", and "the enemy of my friend is my enemy", respectively. The three triads in the figure do not have theory output because they fall into the fourth principle.

Next, we extend weak balance theory to deal with the unsigned trust networks with continuous trust ratings. Before that, we first give an illustrative example to explain the connection between self-assessment and balance theory for trust inference. Suppose a software developer *Alice* has worked with *Bob* in an Internetware environment. Based on the working experiences, *Alice* rates *Bob* as 0.5 which indicates the performance of *Bob* in the opinion of *Alice*. However, this rating alone does not necessarily convey the attitude from *Alice* to *Bob*. On the other hand, if we also know that *Alice* rates herself as 0.7, then we can know that the attitude from *Alice* to *Bob* is actually negative. This negative sentiment, instead of the rating from *Alice* to *Bob*, should be used as the input of the balance theory. Consequently, we need to generate a signed trust network where each sign indicates the positive/negative sentiment. We achieve so by redefining the link sign $\mathcal{R}(i,j)$ based on self-assessment as follows:

Definition 5 (Sign generation in trust networks)**.** $\tilde{\mathcal{R}}(i,j)$ is positive if $\mathcal{R}(i,j) - \mathcal{S}(i) \geqslant 0$; $\tilde{\mathcal{R}}(i,j)$ is negative if $\mathcal{R}(i,j) - \mathcal{S}(i) < 0$; where $\mathcal{R}(i,j)$ is the historical trust rating from user i to user j and $\mathcal{S}(i)$ is the self-assessment of user i.

Based on the above sign generation definition, we can generate a signed network from the trust network. Then, we can apply the weak balance theory, which is illustrated in Figure 1, to the generated signed network. We will use the theory output of the generated signed network as features in our model.

3.1.2 *Global scale of network structure*

As we mentioned above, we use another output of balance theory by employing the low-rank structure of balanced social networks. To simplify discussion, we first transform the directed signed network into its undirected version by the following rules:

Definition 6 (Direction degeneration in trust networks)**.** $\overline{\mathcal{R}}(i,j)$ (or equivalently, $\overline{\mathcal{R}}(j,i)$) is positive if $\tilde{\mathcal{R}}(i,j)$ is positive and $\tilde{\mathcal{R}}(j,i)$ is not negative, or $\tilde{\mathcal{R}}(j,i)$ is positive and $\tilde{\mathcal{R}}(i,j)$ is not negative; $\overline{\mathcal{R}}(i,j)$ (or equivalently, $\overline{\mathcal{R}}(j,i)$) is negative in all other cases.

Please note that the second rule includes the case where $\tilde{\mathcal{R}}(i,j)$ is negative and $\tilde{\mathcal{R}}(j,i)$ is not positive, or analogously $\tilde{\mathcal{R}}(j,i)$ is negative and $\tilde{\mathcal{R}}(i,j)$ is not positive. It also includes the case where $\tilde{\mathcal{R}}(i,j)$ and $\tilde{\mathcal{R}}(j,i)$ are of opposite signs though this is in fact rare in real datasets. The following theorems are all based on such transformed undirected signed networks unless otherwise specified.

For weak balance theory, the constraints on local triads could imply the following global structure:

Theorem 1 (Global "weak balance" structure [12])**.** A complete signed network is weakly balanced if and only if all links are positive, or the nodes can be divided into several groups such that within-group links are positive and between-group links are negative.

In other words, weak balance theory can induce a graph with several clusters where links within clusters are all positive and links between clusters are all negative. Clustering is also a common phenomenon in social networks [26]: people with similar interests tend to trust each other and form clusters. As a result, weak balance theory actually reflects the underlying network structure in global scale by regularizing balance constraints on local triads. We apply such connection and connect weak balance theory with low-rank approximation. Firstly, we define weak-balanced trust network as:

Definition 7 (Weak-balanced trust network)**.** A trust network is weak balanced if and only if there is no triad containing two positive links and one negative link in the resulting network which is derived from the sign generation in Definition 5 and the direction degeneration in Definition 6.

Then, we have the following theorem:

Theorem 2 (Low-rank structure of trust networks)**.** By applying the sign generation in Definition 5 and the direction degeneration in Definition 6 on a complete weak-balanced trust network, the adjacency matrix of the resulting network has rank k for $k > 2$ where k is the number of clusters in the complete weak-balanced trust network.

Proof. Theorem 3 in [14] has already shown that the adjacency matrix of a complete weakly-balanced signed network has rank k for $k > 2$. Therefore, based on Definition 7, we only need to show that after applying the sign generation and direction degeneration on a complete weak-balanced trust network, the resulting network is a complete signed network. First, by Definition 5, after sign generation on a complete directed graph with continuous values on the links, we will have a complete directed graph with signs on the links; second, the transformation from directed graph to undirected graph contains all the possible combinations of bidirectional links between two users. Thus, by applying the sign generation and the direction degeneration, the resulting network is a complete signed network, which completes the proof.

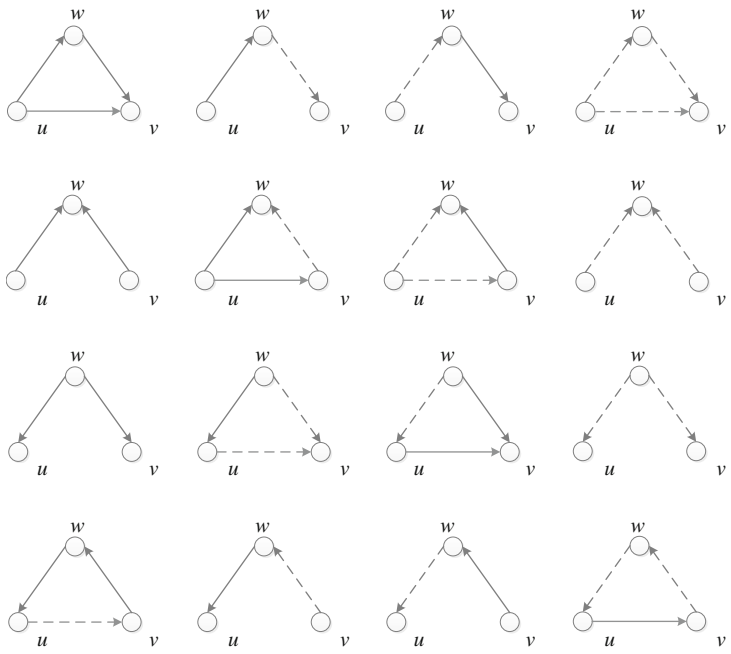

Figure 2 Illustration for status theory in directed signed networks. Solid line indicates positive relationship and dashed line indicates negative relationship. The link between trustor u and trustee v is the theory output. The triads in the figure do not have theory output because the transitivity property cannot apply.

The above theorem shows that a complete weak-balanced network is of low rank. Consequently, it is rational to view trust inference problem as low-rank matrix completion problem, i.e., filling those unknown trust relationships so that the resulting network is still weak balanced. Specifically, we apply low-rank approximation to infer the unknown trust as follows:

$$\min_{\boldsymbol{P},\boldsymbol{Q}} \sum_{(i,j)\in\mathcal{R}} (\mathcal{R}(i,j) - \mathcal{S}(i) - \boldsymbol{P}(i,:)\boldsymbol{Q}(j,:)')^2 + \lambda(\|\boldsymbol{P}\|_F^2 + \|\boldsymbol{Q}\|_F^2), \tag{1}$$

where λ is a regularization parameter, $\|\cdot\|_F$ is the Frobenius norm of the matrices and both $\boldsymbol{P}$ and $\boldsymbol{Q}$ are of low rank k, which is the possible number of clusters in the network. We could view each row in $\boldsymbol{P}$ ($\boldsymbol{Q}$) as the relationships between the trustor (trustee) and the k clusters. As a result, such relationships can be aggregated to form the trustworthiness between two users. In the above formulation, we adopt square loss instead of the sign loss defined in [14] because of the continuous nature of trust network. In addition, the self-assessment information is considered by deleting $\mathcal{S}(i)$ before applying low-rank approximation to compute $\boldsymbol{P}$ and $\boldsymbol{Q}$. We will also use the linear combination of $\boldsymbol{P}$ and $\boldsymbol{Q}$ as features in our model.

3.2 Extending status theory

Compared to balance theory, status theory is relatively new and barely visited. Basically, a positive link from user i to user j in status theory indicates that i believes j has a higher status than i, and a negative link from i to j analogously implies that i believes j has a lower status [13]. As a result, a positive link from i to j is equal to a negative link from j to i. Moreover, the status is transitive: if i has a higher status than j, and j has a higher status than w, then i has a higher status than w. By using such transitivity property, we illustrate the theory output of status theory in Figure 2. As we can see in this figure, all the triads that have theory output follow the status transitivity. Take the second triad in the second row as an example. In this example, u thinks w has a higher status than u, and v thinks w has a lower status than v (or equivalently, v thinks v has a higher status than w). The theory therefore infers that v has a higher status than u which is a positive link from u to v. The triads in the figure do not have theory output because the transitivity property cannot apply.

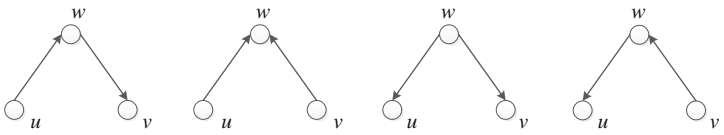

Figure 3 Illustration for our extended status theory for directed trust networks.

Different from the balance theory which is originally proposed for undirect signed networks, status theory is devised for directed signed networks. Although we can still use the signed network generated by Definition 5, we choose to take a further step by introducing the concept of status gap. Essentially, status gap is a quantification to measure how much user i thinks user j is higher or lower than i in status. Although directly using $\mathcal{R}(i,j)$ as the gap is one way to quantify this measurement, $\mathcal{R}(i,j)$ is usually a non-negative probability in trust networks resulting in only half of the triads available in trust networks. In addition, it would be more reasonable if we incorporate self-assessment into status gap, because trust rating indicates how the user thinks of others and self-assessment reflects how the user thinks of himself/herself. Let us revisit the example mentioned before, where a software developer *Alice* has worked with *Bob* in an Internetware environment. In this example, if we only know that *Alice* rates *Bob* as 0.5, we cannot decide whose status is higher in the opinion of *Alice*. On the other hand, if we also know that *Alice* rates herself as 0.7, then we can decide that *Alice* actually thinks that *Bob* has a lower status than herself. In this work, we define status gap as follows:

Definition 8 (Status gap)**.** Given user i and user j, the status gap from i to j is $g(i,j) = \mathcal{R}(i,j) - \mathcal{S}(i)$, where $\mathcal{R}(i,j)$ is the trust rating from i to j and $\mathcal{S}(i)$ is i's self-assessment.

Based on the above definition, the status gap from user i to user j is the difference between the rating that i gives to j and the rating that i gives to himself/herself. Given the trust rating $\mathcal{R}(i,j)$ and $\mathcal{S}(i)$, i believes its status is $g(i,j)$ higher than j if $g(i,j) > 0$, $-g(i,j)$ lower than j if $g(i,j) < 0$, and equal to j if $g(i,j) = 0$.

One advantage of status gap is that we can now infer a numerical theory output for each triad in Figure 2. The reason is that, instead of only the link sign, we now know the exact status gap on each link and the status gap is additive. To further illustrate our extended status theory, let us consider the four types of triads that are possibly available in trust network as shown in Figure 3. Take the first triad as an example. If $g(u,w) = 0.3$ and $g(w,v) = -0.1$, then the status gap between u and v is $g(u,v) = g(u,w) + g(w,v) = 0.2$. In addition, the four types of triads in the figure could derive all the 16 triads in Figure 2, depending on the sign of the status gap. For example, the first triad in Figure 3 will derive the four triads in the first row of Figure 2 under the conditions of $g(u,w) \geqslant 0$ and $g(w,v) \geqslant 0$, $g(u,w) \geqslant 0$ and $g(w,v) < 0$, $g(u,w) < 0$ and $g(w,v) \geqslant 0$, and $g(u,w) < 0$ and $g(w,v) < 0$, respectively. The output of our extended status theory is summarized in the following theorem:

Theorem 3 (Extended status theory output)**.** Given the status gap in Definition 8, the output $\mathcal{R}(u,v)$ from trustor u to trustee v for the four triads in Figure 3 is computed as $\mathcal{R}(u,w) + \mathcal{R}(w,v) - \mathcal{S}(w)$, $\mathcal{R}(u,w) - \mathcal{R}(v,w) + \mathcal{S}(v)$, $-\mathcal{R}(w,u) - \mathcal{R}(v,w) + \mathcal{S}(u) + \mathcal{S}(w) + \mathcal{S}(v)$, and $-\mathcal{R}(w,u) + \mathcal{R}(w,v) + \mathcal{S}(u)$, respectively.

Proof. Let us first take a look at the first triad in Figure 3. Based on the status gap definition, the status gap between u and w, w and v is $g(u,w) = \mathcal{R}(u,w) - \mathcal{S}(u)$ and $g(w,v) = \mathcal{R}(w,v) - \mathcal{S}(w)$, respectively. Then, based on the additive property of status gap, $g(u,v)$ can be computed as $g(u,v) = g(u,w) + g(w,v) = \mathcal{R}(u,v) - \mathcal{S}(u) = \mathcal{R}(u,w) - \mathcal{S}(u) + \mathcal{R}(w,v) - \mathcal{S}(w)$, and thus $\mathcal{R}(u,v) = \mathcal{R}(u,w) + \mathcal{R}(w,v) - \mathcal{S}(w)$. Similarly, for the the rest three triads, we can get the results by substituting $g(i,j)$ as $\mathcal{R}(i,j) - \mathcal{S}(i)$ into the following three formulae: $g(u,v) + g(v,w) = g(u,w)$, $g(w,u) + g(u,v) = g(w,v)$, and $g(u,v) = -g(w,u) - g(v,w)$, respectively.

Based on the above theorem, our extended status theory could output a numerical value for each triad in the trust network. In the next subsection, we will use these numerical values as features in our model.

3.3 The SelfTrust learning framework

To infer unknown trustworthiness scores in trust networks, we extract a collection of features from the extended balance theory and status theory, and then apply logistic regression to combine these features.

Let us first consider the two outputs from the extended balance theory. For the first local output, we only know the sign of the output for each triad. In view of this, we could count the number of the nine triads with theory output in Figure 1, and use the number as features. The intuition behind these features is that each of the nine triads can provide evidence about the sign from u to v. However, in our continuous trust inference problem setting, such a sign is not sufficient. We further incorporate the self-assessment of u as an additional feature since our sign generation in Definition 5 depends on user's self-assessment. Consequently, the number of the nine triads for each link $\langle i, j\rangle$ in the network as well as the self-assessment of i can be encoded as a 10-dimensional vector $\boldsymbol{v}_b$, and this vector is used as the features in our model.

For the second output of extended balance theory, based on (1), we could compute the low-rank matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$. The estimated trustworthiness score from u to v could then be computed as

$$\mathcal{R}(u, v) = \boldsymbol{P}(u, :)\boldsymbol{Q}(v, :)' + \mathcal{S}(u). \tag{2}$$

Please note that the self-assessment information is added back to adjust the inference result. We denote the resulting score by s_b, and directly use it as the feature in our model. Overall, we have 11 features from our extended balance theory, i.e., the 10-dimensional vector $\boldsymbol{v}_b$ and the scalar s_b.

As for the extended status theory, we could compute the numerical output for each triad in Figure 3 as summarized in Theorem 3. Given link $\langle i, j\rangle$, we could then derive a set of numerical values for each of the triad in the figure. Therefore, we could use more meaningful features in our extended status theory, instead of simply using the number of triads as features in the extended balance theory. Here, each output for each triad can be seen as a piece of evidence that could be used for trust inference. Then, the amount of evidence, the disposition of evidence, and the variance of evidence could all be meaningful. For example, if we have more evidence, we would be more confident about the trust decision; if all the evidence shows that someone is trustworthy, we would probably trust this person; if some of the evidence shows that someone is trustworthy while other evidence shows the opposite, we might be hesitant on the decision. To capture these important intuitions about evidence, we consider the sum, average, maximum, and minimum characters for each set of numerical values that belong to a certain triad. The combination of four characters and four triads results in a 16-dimensional vector $\boldsymbol{v}_s$, and we use this vector as the features in our model.

In total, we extract 27 features from our extended balance theory and status theory, and we combine these features under logistic regression which learns a model in the following form:

$$\mathcal{R}(u, v) = \exp\left(\frac{1}{1 + \exp[w_0 + \Sigma_{i=1}^{10} w_i \boldsymbol{v}_b(i) + w_{11} s_b + \Sigma_{i=12}^{27} w_i \boldsymbol{v}_s(i-11)]}\right), \tag{3}$$

where w_i are the coefficients to be estimated, and $\boldsymbol{v}_b$, s_b, and $\boldsymbol{v}_s$ are the features for trustor-trustee pair $\langle u, v\rangle$. This trust inference algorithm is termed SelfTrust in this paper, as it incorporates the self-assessment information during the learning process.

4 Experiments

In this section, we evaluate the proposed SelfTrust method on a real dataset of software developer network. The goal of our experiments is to show that 1) self-assessment truly helps to improve the accuracy of trust inference, and 2) our proposed method is more accurate than other state-of-the-art methods.

Table 1 High level statistics of advogato dataset

Nodes	Edges	Avg. degree	Avg. clustering [26]	Avg. diameter [27]	Date
5 428	51 493	19.0	0.31	4.82	2011-06-23

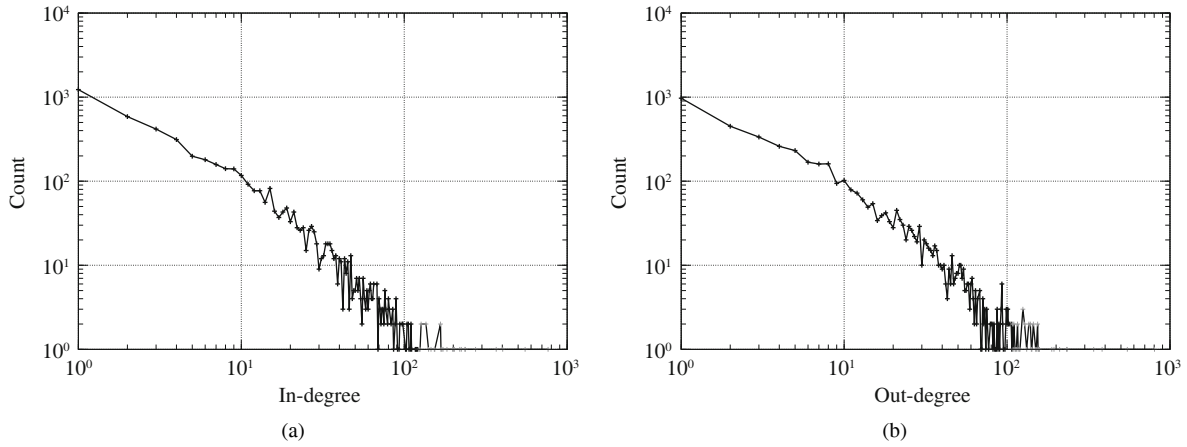

Figure 4 The degree distribution of advogato. (a) In-degree distribution of advogato; (b) Out-degree distribution of advogato.

4.1 Dataset

The data we experiment with is the advogato dataset, which is a snapshot on June 23, 2011 containing 5 428 nodes and 51 493 edges[1)]. Advogato is an online community and social networking site dedicated to free software development, where users could certify and rate each other. This dataset is applied in this study for two reasons. The first reason is that, in addition to rating others, many users in advogato network rate themselves, and these self-ratings provide the self-assessment information. The other one is that, advogato is a trust-based social network and it contains multilevel trust assertions. Specifically, there are four levels of trust assertions in the network, i.e., 'Observer', 'Apprentice', 'Journeyer', and 'Master'. These assertions can be mapped into real numbers in [0,1]. In our experiments, we map 'Observer', 'Apprentice', 'Journeyer', and 'Master' to 0.1, 0.4, 0.7, and 0.9, respectively. For those users who do not provide self-ratings, we set the them at 0.5 by default.

The statistics and the degree distribution of the dataset are shown in Table 1 and Figure 4, respectively. It is obvious that, the advogato graph is a typical small-world social network because it exhibits the properties of power-law degree distribution, high clustering coefficient, and low diameter [28].

4.2 Experimental setup

To evaluate the performance of the proposed method, we compare SelfTrust with several state-of-the-art trust inference algorithms, including HCD [14], MoleTrust [17], and GKRT [16]. Moreover, while trust inference algorithms are usually unsupervised, researchers have found that supervised method is usually better than unsupervised one in link prediction accuracy [29]. Thus, we also compare SelfTrust with the supervised method LHK [24]. Specifically, the comparisons are conducted on three groups of features (i.e., weak balance, status, and all triads) in LHK, and we denote them by LHK-1 (for weak balance), LHK-2 (for status), and LHK-3 (for all triads), respectively.

For evaluation metrics, we conduct 5-fold cross validation on the dataset, and apply both the root mean squared error (RMSE) and the mean absolute error (MAE) between the estimated and the true trustworthiness scores: $\text{RMSE} = \sqrt{\sum_{i=1}^{n}(\hat{x_i} - x_i)^2/n}$, $\text{MAE} = \sum_{i=1}^{n}|\hat{x_i} - x_i|/n$, where n is the number of trust ratings, x_i is the true trustworthiness scores, and $\hat{x_i}$ is the estimated trustworthiness scores by SelfTrust.

1) http://www.trustlet.org/wiki/Advogato_dataset.

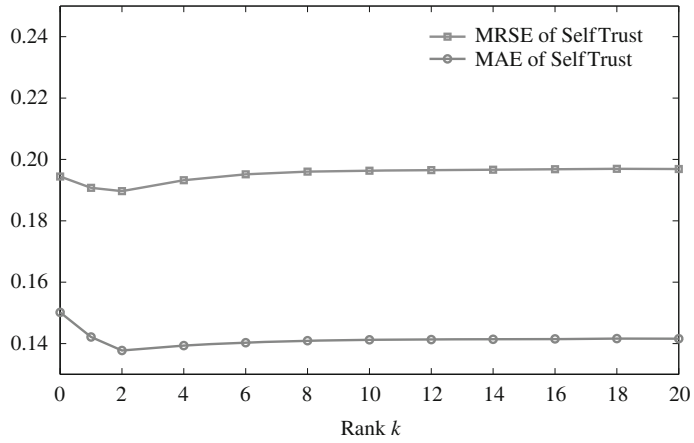

Figure 5 Prediction error of SelfTrust with different rank k.

Table 2 The effect of self-assessment on the three components of SelfTrust. Self-assessment truly improves the inference accuracy in all three components, where component 1 is local scale of theory output from extended balance theory, component 2 is global scale of network structure from extended balance theory, and component 3 is theory output from extended status theory

	RMSE	MAE
Component 1: without self-assessment	0.2340	0.1807
Component 1: with self-assessment	0.2190	0.1690
Component 2: without self-assessment	0.2546	0.1790
Component 2: with self-assessment	0.2091	0.1514
Component 3: without self-assessment	0.2067	0.1594
Component 3: with self-assessment	0.2036	0.1589

4.3 Experimental results

To show the performance of the proposed method, we first examine the impact of the parameter k in SelfTrust (the rank k in (1)). This k indicates the possible number of clusters in the underlying social network of the dataset. The impact of k is shown in Figure 5. Obviously, both RMSE and MAE stay stable when $k > 10$, and the best result comes from $k = 2$. This indicates that there are possibly two big clusters in the advogato network, and we set $k = 2$ in our experiments unless otherwise specified.

Then, we show the effect of self-assessment by presenting the RMSE and MAE results on the three components in SelfTrust, i.e., the two outputs from extended weak balance theory and the one output from extended status theory. For comparison, we delete the self-assessment in the extended theories, and the results are shown in Table 2. It is observed that, self-assessment truly improves the inference accuracy in all three components. Specifically, in the second component from extended balance theory, incorporating self-assessment achieves 17.9% and 15.4% improvement in RMSE and MAE, respectively.

Next, we compare the effectiveness of SelfTrust with the unsupervised methods HCD, MoleTrust, and GKRT, as shown in Table 3. It is observed from the table that SelfTrust significantly outperforms all the compared methods in terms of both RMSE and MAE. MoleTrust and GKRT perform poorly as they infer trust solely based on the paths from trustors to trustees. SelfTrust also outperforms HCD by 40.5% in RMSE and 40.4% in MAE. The reason is that while HCD makes use of the global structure of social networks, we also consider two other groups of features from extended balance theory and status theory.

Finally, we compare SelfTrust with the supervised method LHK-1, LHK-2, and LHK-3. We first show the result in Table 4. As we can see, SelfTrust performs better than all compared methods, and it achieves up to 19.6% improvement in prediction error compared to the three variants from LHK. Furthermore, we evaluate how the prediction error changes as more information is available. As shown in Figure 6, we plot the prediction results against the embeddedness threshold, where embeddedness is the number of common

Table 3 Comparison of SelfTrust with unsupervised methods HCD, MoleTrust, and GKRT. SelfTrust performs best

	RMSE	MAE	RMSE improvement of SelfTrust	MAE improvement of SelfTrust
SelfTrust	0.1897	0.1377		
HCD	0.3188	0.2309	40.5%	40.4%
MoleTrust	0.4093	0.2758	53.7%	50.1%
GKRT	0.6343	0.5909	70.1%	76.7%

Table 4 Comparison of SelfTrust with supervised methods LHK-1, LHK-2, and LHK-3. SelfTrust performs best

	RMSE	MAE	RMSE improvement of SelfTrust	MAE improvement of SelfTrust
SelfTrust	0.1897	0.1377		
LHK-1	0.2175	0.1698	12.8%	18.9%
LHK-2	0.2201	0.1712	13.8%	19.6%
LHK-3	0.2062	0.1590	8.0%	13.4%

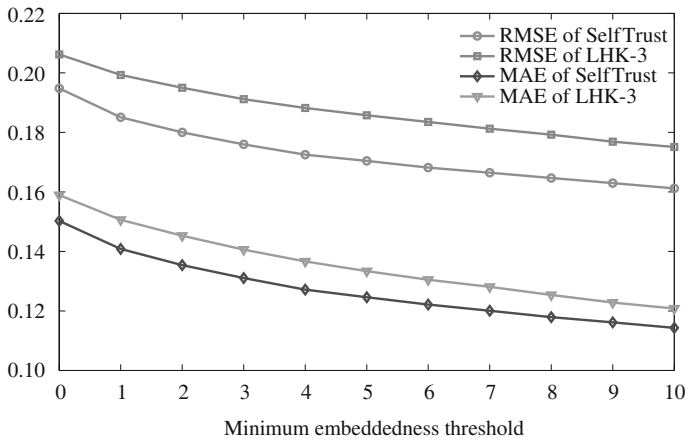

Figure 6 Prediction error of trust inference algorithms with different levels of embeddedness. The x-axis represents the embeddedness threshold. We can see that SelfTrust consistently gives us lower error than LHK-3 for all thresholds on both RMSE and MAE.

neighbors and thus it indicates the amount of available information for prediction. For clarity, we only compare the result of SelfTrust with LHK-3 which is the best among the three compared methods. As we can observe from the figure, SelfTrust consistently gives us lower error than LHK-3 on both RMSE and MAE regardless of the embeddedness. Moreover, it is also observed that SelfTrust performs better as more information is available.

5 Related work

In this section, we briefly review related trust inference methods from literature.

For the trust inference models that employ the transitivity property, we first divide them into two categories: *recommender-based approaches* and *flow-based approaches*.

Recommender-based approaches focus on measuring the recommendation credibility of available recommenders. For example, PeerTrust system for peer-to-peer e-commerce environment proposes two measures for the recommendation credibility, namely, reputation-based measure and similarity-based measure [30]. In mobile ad-hoc networks, Buchegger et al. [31] use a deviation test to eliminate unreliable recommendations from the final computation of trust. In multi-agent systems, Sabater et al. determine the recommendation credibility by applying fuzzy logic [32] on the social relationships between the recommenders and the trustee [33]. Travos system in multi-agent system evaluates the recommen-

dation credibility by tracking the past recommendation behavior of recommenders [34]. The trust value is then computed based on the recommendation content and the credibility of these recommendations. Other recommender-based approaches include [6,8,35].

Later, with the popularity and availability of underlying social networking structures, researchers begin to explore *flow-based approaches* which compute trust values by finding a set of connected paths or a connected component from the trustor to the trustee. Trust then flows from the trustor to the trustee over these paths or components. For example, in peer-to-peer networks, EigenTrust assumes that trust is recursively transitive through the network [7]. The EigenTrust algorithm computes global trust values by calculating the left principal eigenvector of the matrix representing the local trust values. Similar to EigenTrust, PowerTrust also uses a flow-based trust calculation approach, with special consideration of the power-law distribution of feedback [36]. In multi-agent system, Wang et al. [19] define two operators to deal with trust propagation along a path, and trust aggregation among different paths. To deal with large graphs, MoleTrust first destroys the cycles in the graph to ensure that every node is visited only once, and then infers the trustworthiness score in a flow manner [17]. To avoid the loss of information, FlowTrust uses the network flow theory on a connected component instead of finding connected paths [37]. Other flow-based approaches include [16,18,20,38,39].

In addition to the transitivity property which has been widely used by the existing approaches, there are other properties that could be useful and important for trust inference. In this work, we put our focus on the self-assessment of users. Particularly, we incorporate such self-assessment information into two social science theories: the weak balance theory and the status theory. In our work, the balance theory and status theory are based on the triangle structure, making our approach fall into the category of recommender-based approaches. Actually, recommender-based approach can be viewed as the special case of flow-based approach if we fix the trust transitive distance in two steps. Generalizing the theories to longer distance has been considered by several researchers [25,40], while we leave it as our future work.

In machine learning and data mining domains, there are several pieces of work that focus on edge sign prediction. For example, Nguyen et al. [41] derive several features from social science theory to infer trust. However, their method needs the user-item ratings and thus becomes infeasible when such information is not available. In contrast, our method is solely based on the user-user trust ratings, and therefore it has a broader applicability. Leskovec et al. [24] formulate the link sign prediction problem and connect the problem to the theories of balance and status. In our work, we extend the theories to incorporate self-assessment information so that the personalization property of trust is captured. Hsieh et al. [14] find the connection between structural balance and low-rank approximation. In our approach, we also make use of such connection to derive one output from balance theory.

6 Conclusions

In this paper, we have proposed a trust inference approach to evaluating the trustworthiness of potential partners to guide the software development in Internetware. In the context of trust inference, the personality of each individual user has significant effect on the trust information to other users. To this end, we investigate the personalization property of trust by leveraging self-assessment, and based on this, we propose a trust inference model SelfTrust to infer trust between potential partners. Specifically, we first extend the balance theory and the status theory to incorporate the self-assessment information, and then propose a machine learning-based framework to infer the trustworthiness scores. Experiment results on a real software developer network show that the self-assessment information truly helps to improve the accuracy of trust inference, and the proposed SelfTrust is more accurate than other state-of-the-art methods.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant Nos. 91318301, 60736015, 61073030), National High-tech R&D Program of China (863 Program) (Grant No. 2012AA011205), and National

Basic Research Program of China (973 Program) (Grant No. 2009CB320702).

References

1 Lü J, Ma X, Tao X, et al. Research and progress on Internetware (in Chinese). Sci China Ser E-Inf Sci, 2006, 36: 1037–1080
2 Li L, Wang Y, Lim E P. Trust-oriented composite service selection and discovery. In: Proceedings of the 7th International Joint Conference on Service-Oriented Computing, Stockholm, 2009. 50–67
3 Hang C W, Singh M P. Trustworthy service selection and composition. ACM Trans Auton Adapt Syst, 2011, 6: 5
4 Wang Y, Vassileva J. A review on trust and reputation for web service selection. In: Proceedings of the 27th International Conference on Distributed Computing Systems Workshops, Toronto, 2007. 25–25
5 Jøsang A, Ismail R, Boyd C. A survey of trust and reputation systems for online service provision. Decis Support Syst, 2007, 43: 618–644
6 Jøsang A, Ismail R. The Beta reputation system. In: Proceedings of the 15th Bled Electronic Commerce Conference, Bled, 2002. 41–55
7 Kamvar S D, Schlosser M T, Garcia-Molina H. The Eigentrust algorithm for reputation management in p2p networks. In: Proceedings of the 12th International Conference on World Wide Web, Budapest, 2003. 640–651
8 Li F, Wu J. Uncertainty modeling and reduction in MANETs. IEEE Trans Mob Comput, 2010, 9: 1035–1048
9 Golbeck J. Computing and applying trust in web-based social networks. Dissertation for the Doctoral Degree. University of Maryland, 2005
10 Gambetta D. Can we trust trust. In: Gambetta D, ed. Trust: Making and Breaking Cooperative Relations. University of Oxford Press, 2000. 213–237
11 Cartwright D, Harary F. Structural balance: a generalization of heider's theory. Psychol Rev, 1956, 63: 277–293
12 Davis J A. Clustering and structural balance in graphs. Hum Relat, 1967, 20: 181–187
13 Leskovec J, Huttenlocher D, Kleinberg J. Signed networks in social media. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems, Atlanta, 2010. 1361–1370
14 Hsieh C J, Chiang K Y, Dhillon I S. Low rank modeling of signed networks. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, 2012. 507–515
15 Ayday E, Fekri F. Iterative trust and reputation management using belief propagation. IEEE Trans Dependable Secur Comput, 2012, 9: 375–386
16 Guha R, Kumar R, Raghavan P, et al. Propagation of trust and distrust. In: Proceedings of the 13th International Conference on World Wide Web, New York, 2004. 403–412
17 Massa P, Avesani P. Controversial users demand local trust metrics: an experimental study on epinions.com community. In: Proceedings of the 20th National Conference on Artificial Intelligence, Pittsburgh, 2005. 121–126
18 Ziegler C N, Lausen G. Propagation models for trust and distrust in social networks. Inf Syst Front, 2005, 7: 337–358
19 Wang Y, Singh M P. Trust representation and aggregation in a distributed agent system. In: Proceedings of the 21st National Conference on Artificial Intelligence, Boston, 2006. 1425–1430
20 Liu G, Wang Y, Orgun M. Trust inference in complex trust-oriented social networks. In: Proceedings of the International Conference on Computational Science and Engineering, Vancouver, 2009. 996–1001
21 Hang C W, Wang Y, Singh M P. Operators for propagating trust and their evaluation in social networks. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems, Budapest, 2009. 1025–1032
22 Wang G, Wu J. Multi-dimensional evidence-based trust management with multi-trusted paths. Future Gener Comput Syst, 2011, 27: 529–538
23 Yao Y, Tong H, Xu F, et al. Subgraph extraction for trust inference in social networks. In: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Istanbul, 2012. 163–170
24 Leskovec J, Huttenlocher D, Kleinberg J. Predicting positive and negative links in online social networks. In: Proceedings of the 19th International Conference on World Wide Web, Raleigh, 2010. 641–650
25 Yao Y, Xu F, Yang Y, et al. PatTrust: a pattern-based evaluation approach for trust and distrust in Internetware. In: Proceedings of the 3rd Asia-Pacific Symposium on Internetware, Nanjing, 2011
26 Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks. Nature, 1998, 393: 440–442
27 Leskovec J, Kleinberg J, Faloutsos C. Graphs over time: densification laws, shrinking diameters and possible explanations. In: Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, 2005. 177–187
28 Yao Y, Zhou J, Han L, et al. Comparing linkage graph and activity graph of online social networks. In: Proceedings of the 3rd International Conference on Social Informatics, Singapore, 2011. 84–97
29 Lichtenwalter R N, Lussier J T, Chawla N V. New perspectives and methods in link prediction. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington DC, 2010. 243–252

30 Xiong L, Liu L. Peertrust: supporting reputation-based trust for peer-to-peer electronic communities. IEEE Trans Knowl Data Eng, 2004, 16: 843–857
31 Buchegger S, Le Boudec J Y. A Robust Reputation System for Mobile Ad-Hoc Networks. Technical Report. KTH Royal Institute of Technology, Theoretical Computer Science Group, 2004
32 Zadeh L A. Fuzzy logic and approximate reasoning. Synthese, 1975, 30: 407–428
33 Sabater J, Sierra C. Reputation and social network analysis in multi-agent systems. In: Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems, Bologna, 2002. 475–482
34 Patel J, Teacy W T L, Jennings N R, et al. A probabilistic trust model for handling inaccurate reputation sources. In: Proceedings of the 3rd International Conference of Trust Management, Paris, 2005. 193–209
35 Teacy W T L, Patel J, Jennings N R, et al. Travos: trust and reputation in the context of inaccurate information sources. Auton Agents Multi-Agent Syst, 2006, 12: 183–198
36 Zhou R, Hwang K. Powertrust: a robust and scalable reputation system for trusted peer-to-peer computing. IEEE Trans Parall Distrib Syst, 2007, 18: 460–473
37 Wang G, Wu J. Flowtrust: trust inference with network flows. Front Comput Sci China, 2011, 5: 181–194
38 Mui L, Mohtashemi M, Halberstadt A. A computational model of trust and reputation. In: Proceedings of the 35th Annual Hawaii International Conference on System Sciences, Big Island, 2002. 2431–2439
39 Liu G, Wang Y, Orgun M A. Optimal social trust path selection in complex social networks. In: Proceedings of the 24th AAAI Conference on Artificial Intelligence, Atlanta, 2010. 1391–1398
40 Chiang K Y, Natarajan N, Tewari A, et al. Exploiting longer cycles for link prediction in signed networks. In: Proceedings of the 20th ACM International Conference on Information and knowledge Management, Glasgow, 2011. 1157–1162
41 Nguyen V A, Lim E P, Jiang J, et al. To trust or not to trust? predicting online trusts using trust antecedent framework. In: Proceedings of the 9th International Conference on Data Mining, Miami, 2009. 896–901