ETAF: An Extended Trust Antecedents Framework for Trust Prediction

Guibing Guo, Jie Zhang, Daniel Thalmann
School of Computer Engineering
Nanyang Technological University, Singapore
{gguo1, zhangj, danielthalmann}@ntu.edu.sg

Neil Yorke-Smith
American University of Beirut, Lebanon
and University of Cambridge, UK
nysmith@aub.edu.lb

***Abstract*—Trust is one source of information that has been widely adopted to personalize online services for users, such as in product recommendations. However, trust information is usually very sparse or unavailable for most online systems. To narrow this gap, we propose a principled approach that predicts implicit trust from users' interactions, by extending a well-known trust antecedents framework. Specifically, we consider both local and global trustworthiness of target users, and form a personalized trust metric by further taking into account the active user's propensity to trust. Experimental results on two real-world datasets show that our approach works better than contemporary counterparts in terms of trust ranking performance when direct user interactions are limited.**

***Keywords*—Trust prediction; user interactions; trust antecedents framework; user ratings;**

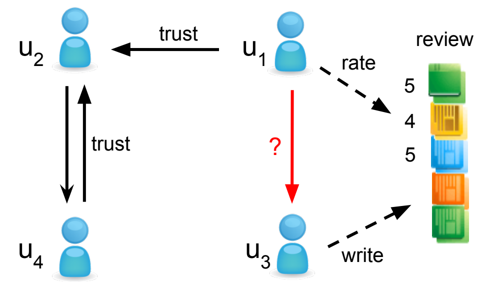


Fig. 1. A social rating network where users specify others as trustworthy and write (or rate) a number of reviews. Solid lines indicate the social trust relationships while dashed lines mean user behaviors (rating and writing).

I. Introduction

Trust helps users in online systems steer away from malicious users and firms, but connect and do business with reliable users and firms. It has been widely adopted in many personalized e-commerce applications. For example, trust-aware recommender systems can help resolve the issues (e.g, cold-start and data sparsity) of traditional recommender systems [1]; trust is an important factor to prevent from malicious users/attacks in security systems [2]; trust is also used to select trustworthy composite online service [3], etc. Two types of trust information have been studied in the literature. The situations where trust information can be directly specified by users themselves produce the *explicit trust* information. For example, users in the review website Epinions.com can add other users into their 'trust list' and form a 'Web of Trust'; Ciao.co.uk adopts a similar concept called 'Circle of Trust' in which users are treated as trustworthy. Explicit trust information, however, is usually unavailable or at best very sparse in most online systems. Hence, efforts have been made to predict *implicit trust* in terms of trust labels ('trust' or 'non-trust') or numerical trust values using computational methods.

Trust inference is usually based on social connections (explicit trust) or user interactions [4], [5], [6]. For example, as illustrated in Figure 1, users in a social rating network can specify others as trustworthy, e.g., user u_1 trusts user u_2 who mutually trusts user u_4. The explicit trust may be due to real-world relationships such as friends, colleagues, etc. Using the transitivity of trust, it can be inferred that user u_1 may trust user u_4 to some extent. However, it is lack of connections to infer the relationship between user u_1 and user u_3, and even worse if there are no social connections at all in the system. Fortunately, users' interactions can be used to infer implicit trust. For example, user u_3 wrote a number of reviews about many products or items some of which were positively rated by user u_1. From the viewpoint of user u_1, user u_3 may be an expert in writing reviews and her opinions are reliably consistent and thus trustworthy. Hence, more trust information can be inferred from user interactions, which is beneficial for trust-based applications such as recommendations [1]. The inference could be even harder when there are no direct interactions between two users. In such cases, we may have to rely on other users' interactions with the target user in order to get a hint about the user's trustworthiness. For example, a user who writes many high quality product reviews and honestly rates others' reviews tends to be trustworthy, though no direct interactions exist between the user and an active user.

This paper proposes a principled method called *ETAF* to predict trust values from users' interactions, by extending a well-known trust antecedents framework (TAF). To be more realistic, this paper assumes that there is no explicit trust in the systems. Computationally, trust in TAF is regarded as the result of a few general trust antecedents (see Figure 2), including trustor's propensity to trust (*trust propensity*), and trustee's trustworthiness determined by trust factors, namely *ability*, *benevolence*, and *integrity*. In addition to such computation of *local trustworthiness*, our ETAF model also takes into account the *global trustworthiness* of trustees based on all the users' interactions. Further accounting for the trust propensity, we obtain personalized trust values. Experimental results on two real-world datasets show that our approach performs better than counterpart methods for implicit trust prediction in producing a list of trustworthy users when there is a lack of direct user interactions.

In summary, our main contributions are in three-fold: (1) we propose an extended trust antecedents framework for trust value prediction; (2) we propose a set of formulations to implement the proposed framework; and (3) a series of experiments on two real-world datasets (i.e., CiaoDVDs and Epinions) are conducted to demonstrate the effectiveness of our approach in comparison with contemporary counterparts.

II. Related Work

Many trust-based applications (e.g., recommendations, web service composition) can be beneficial if more trust information is available. The methods of trust prediction can help infer implicit trust relationships for this purpose. Two types of approaches can be broadly classified to predict implicit trust from users' connections and interactions. The first type is to infer trust values from existing trust connections (i.e., trust network). The most common method is to propagate trust along the chains of trust. For example, Guha et al. [7] propose a trust propagation model with the concepts of co-citation, transpose trust and trust coupling. Golbeck [8], Massa and Avesani [4] propose the TidalTrust and MoleTrust approaches respectively to compute trust values by aggregating trust from multiple trust chains between two users. More recent works [5], [6] adopt the techniques of matrix factorization to factorize the trust relationship to trustor- and trustee-specific latent feature vectors.[1] Then, implicit trust can be computed as the inner product of vectors of a trustor and a trustee. The common drawback of these approaches lies in that they rely on explicit trust that is not available in most systems. In addition, sparse trust information further limits their performance. Our work is based on the assumption that no explicit trust information is available for trust prediction. Hence, trust propagation or factorization is not suitable to predict implicit trust in this case.

The second type of approaches is built upon users' interactions. Liu et al. [9] propose a trust classification method to predict trust labels based on the evidence derived from the actions of individual users and from the interactions of pairwised users. They show that interaction factors have greater impact on trust prediction than user factors. Adopting the features identified in [9], Ma et al. [10] attempt to address the sparseness of explicit trust by deriving implicit trust from users' interactions. They find that the trust relationships involving more active trustors are easier to predict than those involving less active trustors. Matsuo and Yamamoto [11] investigate the bidirectional effects between trust and ratings. A support vector machine (SVM) classifier is trained according to the features related with two users' profiles, product ratings and other trust relations. They report that the product brand has an important effect on the bidirectional effect. Nguyen et al. [12] study and predict the degree to which a user will trust back if another user initializes a trust on him/her, i.e., the reciprocal trust. The prediction is obtained by extracting features from four trust related behaviors from which a SVM classifier is trained. In conclusion, most of these works rely on hand-crafted features to train a certain classifier so as to predict trust labels. In contrast, we are more interested in predicting trust values rather than trust labels, and underpinning our model on well-studied trust factors [13].

[1] Trustors are the users who trust other users, and trustees are those who are trusted by other users.

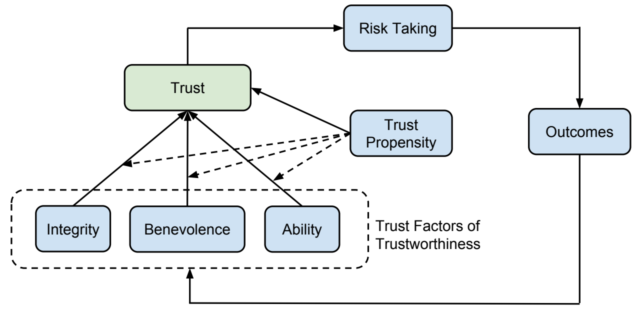


Fig. 2. Trust Antecedents Framework (TAF) including a number of trust factors, namely integrity, benevolence, ability and trust propensity.

The most relevant research to our work is as follows. Guo et al. [14] give an empirical study on trust prediction in recommender systems. They show that the existing trust metrics derived from user-item ratings, a kind of indirect user interactions, cannot provide satisfying trust ranking performance. Nguyen et al. [15] predict trust from users' rating data of the reviews written by other users, i.e., direct user interactions. They introduce a computational formulation of the trust antecedents framework proposed in management science [13]. They show that their method works better than the trust propagation approach, i.e., MoleTrust [4]. However, they only consider the trustee's local trustworthiness, and formalize the integrity factor based on trust networks which are not available in many systems. Kim and Phalak [16] propose to incorporate the global trustworthiness of trustees to address the sparseness of user interactions. Both trustees' expertise and trustors' preference in a specific category are considered. However, other trust factors such as benevolence and integrity are not considered in the computation of trustworthiness. On the contrary, we underpin our approach in a well-known trust antecedent framework [13] and extend it to incorporate the global trustworthiness of trustees. In addition, our implementations are under the assumption with no explicit trust. Hence, our approach holds the potential to be applied to any systems (for trust prediction) where direct user interactions are enabled.

III. Trust Prediction Frameworks

We first briefly introduce the original trust antecedents framework, and then propose an extended framework by involving the global trustworthiness of trustees.

A. Trust Antecedents Framework

The trust antecedents framework (TAF), illustrated in Figure 2, was proposed by Mayer et al. [13], who defined trust as the "willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party". Specifically, trust is a decision of the trustor regarding the perceived trustworthiness of a trustee. Four trust factors are generic and important during the trust evaluation, namely *ability*, *benevolence* and *integrity* of the trustee, and the *trust propensity* of the trustor. The decision to trust means that the trustor is willing to take any possible risk caused by the trustee or environment, no matter whether s/he has the ability to monitor or control the trustee or environment. Nevertheless,

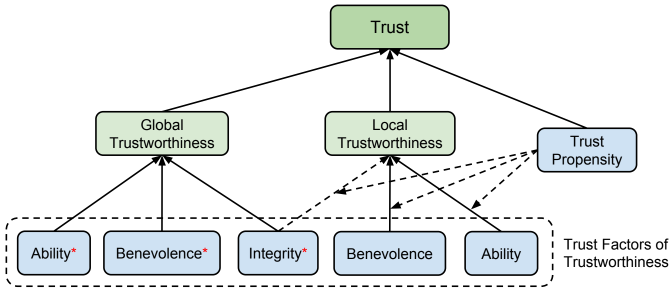


Fig. 3. Extended Trust Antecedents Framework (ETAF). Asterisks denote the global formalization of a trust factor; other factors are the local formalization of trust factors. Dashed lines indicate weak influence of a trust factor on the targets; solid lines strong influence.

the outcomes of the decision will return back to influence the future perceptions of the trustee's trustworthiness. In practice, researchers often ignore the loop influence of the outcomes but focus on the formalization of trust such as [15]. The reason is that the potential risk is negligible in the case of social networks where no monetary issues are related.

The four trust factors can be briefly explained as follows. Ability refers to the expertise or competence of the trustee to perform a specific and expected action. Benevolence refers to the extent to which the trustee is believed and will do good to the trustor, aside from an egocentric profit motive. Integrity refers to the consistence of the trustee to adhere to a set of good moral norms. Finally, trust propensity refers to the degree to which the trustor tends to trust a user. The first three *trustee factors* determines the trustworthiness of a trustee. Together with the last *trustor factor*, a personalized trust value can be derived from users' interactions.

B. Extended Trust Antecedents Framework

The TAF model formulates trustee factors according to the direct interactions between two users, and hence forms only the *local trustworthiness* of the trustee. However, such direct interactions could be sparse or unavailable between two specific users. In this case, the trustor cannot make a proper decision whether to trust a potential trustee or not. Therefore, it is necessary and helpful to further consider measures of *global trustworthiness* derived from other users' interactions with the target trustee. Gómez et al. [17] contend that four information sources can be used to derive trust values, namely *direct experience*, *advertisement*, *recommendation* and *global trust*. Although advertisement and recommendation are not applicable in our case, direct experience and global trust should be considered. In addition, studies in the literature such as [16] have already made use of the global trustworthiness in their computational models, but only the ability factor is considered. In this work, we extend the TAF model (denoted by ETAF) by incorporating the global trustworthiness, derived from all the interactions of users with target trustees. The ETAF model is illustrated in Figure 3.

Generally, the ETAF model is also based on the four generic trust factors. We separate the formalization of each trust factor into two possible perspectives: local and global. The local formalization means that the trust factors are implemented using the direct interactions between two users, whereas the global factors (denoted by asterisks) indicates that the trust factors are modelled based on the interactions among all the other users with the target trustor or trustee. In particular, the global versions of three trustee factors (ability, benevolence and integrity) form the global trustworthiness, while the local versions of ability and benevolence as well as the global integrity generate the local trustworthiness. Note that the integrity refers to the behaviors of the trustee towards all the users rather than a specific trustor by definition, hence it only has the global formalization which has influence on both the local and global trustworthiness. Both kinds of trustworthiness contribute to the overall trustworthiness of the trustee, and with the trustor's trust propensity, a personalized trust can be computed. The next section gives the formalization.

IV. Formalization of the ETAF Model

This section provides a formalization of the proposed ETAF model based on users' interactions. Recall that asterisks denote the global formalization of a trust factor. To facilitate the discussion, we introduce a number of notations. Let U, I, R denote all the users, reviews and the ratings on users' reviews, respectively. We preserve the symbols $u, v, p \in U$ for users, $i, j \in I$ for reviews. Let $r_{u,i,v} \in R$ be a rating given by user u on review i which is written by user v, and $R_{u,v}, I_{u,v}$ be the set of ratings and reviews that are given and rated by user u and written by user v, respectively. We assume that all the ratings are normalized to the range (0, 1] by dividing the maximum rating value. The task of trust prediction is to predict the trust value $t_{u,v}$ that user u will give to user v.

A. Ability*

We consider two types of abilities of a trustee in terms of the roles that the trustee v plays, namely the ability as a review writer (writer ability) ab_v^w and the ability as a review rater (rater ability) ab_v^r. The abilities are based on the quality of reviews that they write or rate, where the review quality can be measured as the average of the ratings given by users. In addition, Nguyen et al. [15] further point out that the ratings are also influenced by the users' *local leniency*, i.e., users with high leniency tend to give high ratings whereas users with low leniency tend to give low ratings. Hence, we obtain the following formula to compute the quality of a review:

$$q_i = w(|U_{i,p}|) \cdot \frac{\sum_{v \in U_{i,p}} ab_v^r \cdot r_{v,i,p}(1 - \beta l_{v,p})}{\sum_{v \in U_{i,p}} ab_v^r}, \quad (1)$$

where $U_{i,p}$ is the set of users who rated review i written by user p, $l_{v,p}$ is the local leniency of user v towards the reviews of user p, and $\beta \in [0, 1]$ controls the maximum adjustment on rating $r_{v,i,p}$, and $\beta = 0.5$ is set as suggested in [15]. Function $w(n)$ accounts for the effect of number n of ratings received by review i, given by: $w(n) = n/(n + 1)$.

Then, the rater ability of user v is updated by:

$$ab_v^r = w(|I_{v,\cdot}|)\Big(1 - \frac{\sum_{i \in I_{v,\cdot}} |r_{v,i,a(i)} - q_i|}{|I_v|}\Big), \quad (2)$$

where $I_{v,\cdot}$ is the set of reviews rated by user v and written by other users, $a(i)$ refers to the author of review i. The closer the given ratings to the review quality, the higher the rater ability is. Similarly, the local leniency is computed by the extent to

Algorithm 1: Global Computation of Trust Factors

Input : Users U, Reviews I, Ratings R
Output: Users' rater ability ab_v^r, review quality q_i and local leniency $l_{v,p}$

1 randomly initialize rater ability ab_v^r and local leniency $l_{v,p}$ with small values in $(0, 1)$;
2 **while** *not converged* **do**
3   **foreach** $i \in I$ **do**
4   compute review quality q_i by Equation 1;
5   **foreach** $v \in U$ **do**
6   update rater ability ab_v^r by Equation 2;
7   **foreach** $v \in U$ **do**
8   **foreach** $p \in U \backslash \{v\}$ **do**
9   update local leniency $l_{v,p}$ by Equation 3;
10 return ab_v^r, q_i, $l_{v,p}$ for all users, reviews, and user pairs;

which the ratings are deviated from the review quality. It is computed by:

$$l_{v,p} = \frac{1}{|I_{v,p}|} \sum\nolimits_{i \in I_{v,p}} \frac{r_{v,i,p} - q_i}{r_{v,i,p}}. \tag{3}$$

Negative leniencies mean that user u tends to give lower ratings than the review quality, whereas positive values indicate that higher ratings are possibly given to the reviews, and 0 means no deviation to the review quality.

The algorithm to derive global computation of trust factors is presented in Algorithm 1. Specifically, all the users, reviews and ratings are taken as input to the algorithm. First, we randomly initialize the rater ability and local leniency to the values in (0, 1) for any users and user pairs, respectively (line 1). Then, we compute the review quality for all the reviews using Equation 1 (lines 3-4), and update the rater ability by Equation 2 (lines 5-6) and local leniency by Equation 3 (lines 7-8) subsequently. The procedure is continued until the **while** loop is converged (line 2). The converging condition is satisfied if the squared error of updated review quality, rater ability, local leniency with respect to the previous ones is lower than a very small value, i.e., 10^{-5} in our case. In fact, our experiments show that the loops can be fast converged within 15 iterations. Since there are usually far more reviews than users in the systems, the most time-consuming part is to compute the quality of reviews (lines 3-4) with the time complexity $O(nm)$, where n is the number of reviews and m is the average number of ratings received by each review. Hence, the overall time complexity will be less than $O(3knm) \approx O(nm)$, where $k \approx 15$ is the number of echoes, and 3 means three parts of embedded **foreach** loops.

After obtaining the review quality, the writer ability of user v can be estimated by the average quality of reviews written by herself:

$$ab_v^w = w(|I_{\cdot,v}|) \frac{\sum_{i \in I_{\cdot,v}} q_i}{|I_{\cdot,v}|},$$

where $I_{\cdot,v}$ is the set of reviews rated by others and written by user v. The number of written reviews is also taken into account by $w(|I_{\cdot,v}|)$.

Different from Kim and Phalak [16] who only consider the writer ability, we compute the global ability by a linear combination of both writer and rater abilities, given by:

$$ab_v^* = \gamma \cdot ab_v^w + (1 - \gamma) \cdot ab_v^r, \tag{4}$$

where $\gamma \in [0, 1]$ represents the importance of writer ability. Although more complex and non-linear combinations could be possible, we use linear combination for the sake of simplicity and generality. Hereafter, we also adopt the linear combination to model other factors if applicable.

B. *Benevolence**

Benevolence is about the trustee's characteristics in doing good to the trustor. Nguyen et al. [15] associate it with user's leniency in giving ratings. We adopt the same strategy. First, the global leniency of user v is obtained by the average of local leniency of user v with other users, given by:

$$ln_v = \frac{1}{|U_{v,\cdot}|} \sum_{p \in U_{v,\cdot}} \frac{l_{v,p} - \min l}{\max l - \min l},$$

where $U_{v,\cdot} = \{p | p \in U, I_{v,p} \neq \emptyset\}$ is the set of users who have interactions with user v, and $\min l, \max l$ are the minimum and maximum local leniency between any two users, respectively. Second, the global benevolence is measured as the normalized global leniency by:

$$be_v^* = \frac{ln_v - \min ln}{\max ln - \min ln}, \tag{5}$$

where $\min ln$ and $\max ln$ are the minimum and maximum global leniency of all the users respectively.

C. *Integrity**

Integrity often refers to the consistency of users' behaviors to adhere to a set of social norms. Similarly, we identify two kinds of consistency in terms of users' behaviors: consistency as a rater (rater integrity) and consistency as a writer (writer integrity). For rater integrity, the consistency means the ratings given by the trustee are always close to those of majority users (which are regarded as social norms), i.e., the review quality. In this regard, we formalize the rater integrity as the rating similarity with majority users, which is computed by the Pearson correlation coefficient (PCC) [1]:

$$in_v^r = \frac{w(|I_{v,\cdot}|)}{2} \left(1 + \frac{\sum\limits_{i \in I_{v,\cdot}} (r_{v,i,a(i)} - \bar{r}_v)(q_i - \bar{q})}{\sqrt{\sum\limits_{i \in I_{v,\cdot}} (r_{v,i,a(i)} - \bar{r}_v)^2} \sqrt{\sum\limits_{i \in I_{v,\cdot}} (q_i - \bar{q})^2}} \right),$$

where $w(|I_{v,\cdot}|)$ works as a shrinkage to account for the number of reviews rated by user v, and $\bar{r}_v, \bar{q}$ are the average rating and review quality of user v and all users, respectively. Note that the value range of PCC is $[-1, 1]$; we normalize it to $[0, 1]$ by $\frac{1}{2}(1 + \text{PCC})$. Other more advanced similarity measures such as Bayesian similarity [18] can be used as well. For writer integrity, we regard the consistency as the high quality of reviews written by the user and the low deviation among all the quality of reviews. In other words, high consistency indicates high mean μ_v and low standard deviation σ_v of the distribution of review qualities. Therefore, we obtain the following formation:

$$in_v^w = w(|I_{v,\cdot}|) \cdot \mu_v \cdot (1 - \sigma_v).$$

Thus, the overall integrity is derived by a linear combination of both rater and writer integrities:

$$in_v^* = \eta \cdot in_v^w + (1 - \eta) \cdot in_v^r, \tag{6}$$

where η denotes the importance of writer integrity.

Recall the formalization of global ability in Equation 4, and we note that integrity and ability are overlapping to some extent. Similarly, the rater ability reflects the distances between a rater's ratings and the review quality, while the rater integrity indicates the similarity between a rater's ratings and the review quality. Differently, the writer ability only considers the mean rating whereas the writer integrity further requires low deviation among ratings.

D. Ability

Nguyen et al. [15] suggest that two features should be considered to formulate local ability, namely average rating received from user u, and the interaction intensity between users u and v, given by:

$$ab_v^u = \psi(|I_{u,v}|; \alpha, \mu) \cdot \frac{\sum_{i \in I_{u,v}} r_{u,i,v}}{|I_{u,v}|}, \tag{7}$$

where $\psi(x; \alpha, \mu)$ is a logistic function to account for the number of interactions:

$$\psi(x; \alpha, \mu) = \frac{1}{1 + e^{-\alpha(x-\mu)}},$$

where α, μ control the slope and midpoint of the sigmoid curve, respectively. As suggested in [15], we set $\alpha = 0.1$ and $\mu = 5$.

E. Benevolence

Similarly with the Benevolence*, the local benevolence is measured as the normalized local leniency of user v. It is defined by:

$$be_v^u = \frac{l_{u,v} - \min l}{\max l - \min l}. \tag{8}$$

F. Trust Propensity

Trust propensity refers to the extent to which a trustor tends to trust another user. Two possible ways to model trust propensity are indicated in [15]: (1) as the global leniency to give ratings; (2) as the function value of $\psi(x, \alpha, \mu)$, where x is the number of users trusting user u. Nguyen et al. [15] show that the first way works better than the second. Hence, the trust propensity is simply formulated as: $tp_u = ln_u$.

G. Personalized Trust

Putting these factors together, the ETAF model considers both local and global trustworthiness of user v to form his/her overall trustworthiness. Including the trust propensity, we obtain the following personalized trust value:

$$t_{u,v} = \Big(\alpha \cdot lt_{v,u} + (1 - \alpha) \cdot gt_v\Big) \cdot tp_u, \tag{9}$$

where $lt_{v,u}, gt_v$ are the local and global trustworthiness of user v (from the perspective of user u), given by:

$$lt_{v,u} = ab_v^u \cdot be_v^u \cdot 0.5, \qquad gt_v = ab_v^* \cdot be_v^* \cdot in_v^*,$$

The constant 0.5 is used to balance the value range of local and global trustworthiness. It can be regarded as the average of 'local' integrity relative to the factor integrity*. Therefore, we can form a personalized trust for any two users u and v, no matter whether they have direct interactions or not. The derived trust value (9) could be small due to the multiplication operations, but what matters in this paper is the ranking of trust values (in descent order) rather than the values themselves.

V. Evaluation

The main objective of our evaluation is to investigate the effectiveness of the proposed ETAF model in comparison with other approaches for trust prediction.

A. Datasets

Two real-world datasets are utilized in our experiments, namely CiaoDVDs and Epinions. In both datasets, users can write textual reviews to products that they purchased or used in the past; and some other users can rate the helpfulness of these reviews in terms of rating scales. If one's reviews are consistently valuable to a specific user, the user may specify the review writer as trustworthy and add him/her in the trust list. These two datasets reflect typical social rating networks, and are often adopted by previous studies [1], [10], [5].

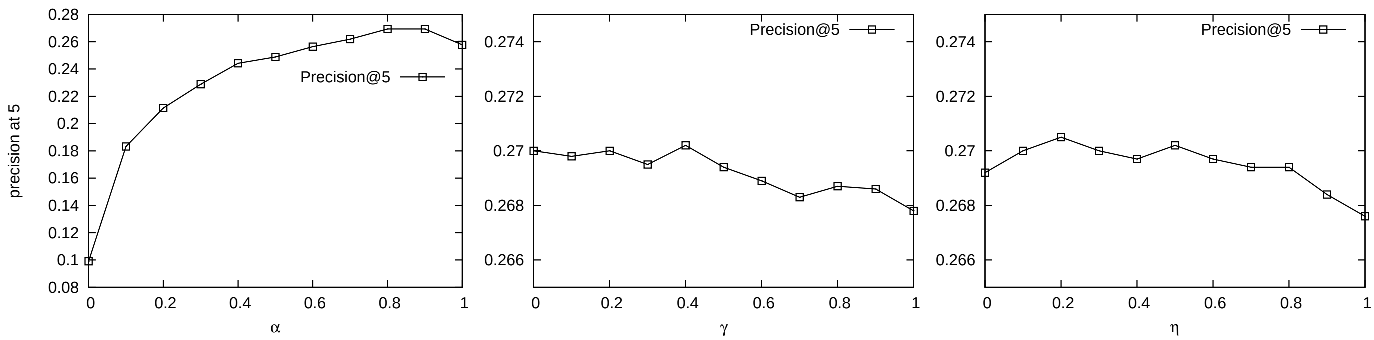
User ratings and reviews are used to predict users' trust while explicit trust is used as ground truth to evaluate the effectiveness of trust prediction methods. Although the data of Ciao (ciao.co.uk) has been used in the previous works [5], the published datasets[2] do not include users' review information. Hence, we crawled the CiaoDVDs dataset in December 2013 from the category of DVD in the ciao.co.uk where users can add others into their circle of trust. The review ratings take values from 0 ('off topic') to 5 ('exceptionally useful'). To deal with the value 0, we shift rating scales to 1-6 and then normalize them to (0,1] by dividing the maximum value 6. The second dataset is sampled from the Extended Epinions dataset[3] by randomly selecting 1,500 trustors and keeping all the review and trust ratings given by the trustors and trustees. The statistics of the datasets are illustrated in Table I.

TABLE I. Statistics of the case study datasets

Features	CiaoDVDs	Epinions
Writers	920	6,167
Reviews	20,469	429,093
Reviews/Writer	22.25	69.58
Raters	3,951	6,028
Reviews	20,455	230,891
Review Ratings	641,810	6,512,699
Ratings/Rater	162.44	1080.41
Trustors	1,438	1,500
Trustees	4,299	6,156
Trust Ratings	40,133	11,310
Density	0.65%	0.12%
Direct Interactions	5.65	7.67
Total Users	4,658	7,551

[2]http://www.public.asu.edu/~jtang20/datasetcode/truststudy.htm

[3]http://www.trustlet.org/wiki/Epinions_datasets



(a) $\gamma = 0.5, \eta = 0.5$ (b) $\alpha = 0.9, \eta = 0.5$ (c) $\alpha = 0.9, \gamma = 0.4$

Fig. 4. The effects of importance of local trustworthiness (α), writer ability γ, and writer integrity η on the ETAF's performance

B. Experimental Settings

Four approaches are compared in our experiments: (1) **EPT** denotes the method proposed by Kim and Phalak [16] where users' trusts are derived by linearly combining both local and global trust parameterized by a function of $N_{\min}$, where $N_{\min}$ is the minimum number of user interactions. The best performance derived by varying $N_{\min} \in [1, 10]$ is adopted. (2) **TAF** is proposed by Nguyen et al. [15] based on the TAF model, where only local trustworthiness is considered and the integrity is modelled based on the number of users who trust the trustee. Since we assume that no explicit trust is available, we ignore the integrity factor. The other suggested parameter values are adopted in our experiments. (3) **ETAF** is our approach by incorporating both local and global trustworthiness of trustees. The best performance is obtained by tuning parameters α, γ and η, corresponding to Equations 9, 4 and 6, respectively. (4) **ETAF*** is a variant of the ETAF model where only local trustworthiness is used by fixing parameter $\alpha = 1$ for Equation 9.

The trust information in the datasets are used as test data in three different views: (1) **All Users** refers to the view where all the trustors are used as testing users. (2) **Cold Start** involves the cold-start users who have rated or written no more than 5 reviews. (3) **Warm Start** consists of the warm-start users who have rated or written at least 20 reviews. Note that it does not mean two warm-start users must have many direct interactions.

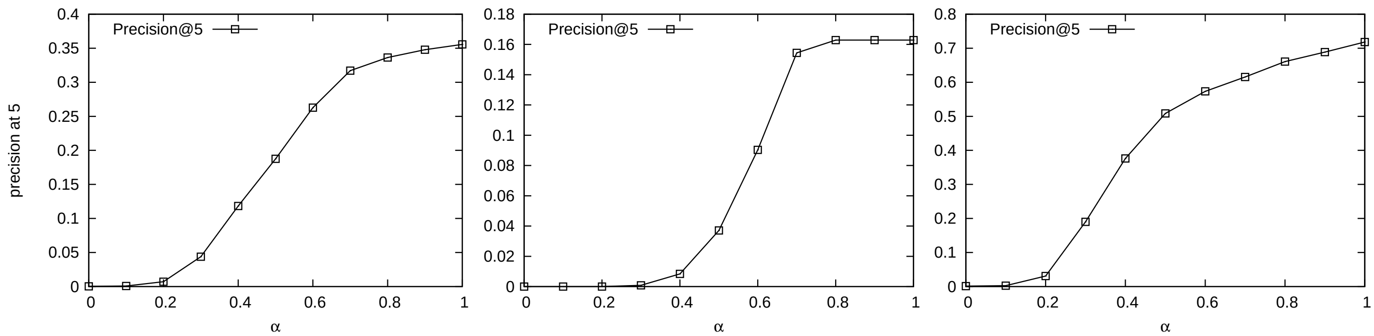
All the approaches are used to predict the trust values for each trustor in the testing views towards other users, and then generate a list with the top 20 most trustworthy users. Guo et al. [14] suggest that trust ranking performance and measures are useful to investigate the performance of trust metrics. Hence, a number of ranking-based measures are used to evaluate the performance, including precision and recall (cutoffs at 5 and 10, and denoted by Pre@5/10, Rec@5/10), mean average precision (MAP), normalized discounted cumulative gain (NDCG) and mean reciprocal rank (MRR) [19]. The higher these measures are, the better performance it achieves.

C. Case Study 1: CiaoDVDs

Sensitivity Analysis. To determine a proper set of values for the three parameters of our model, we first consider equally weighting the rater and writer abilities via setting $\gamma = 0.5$, and setting $\eta = 0.5$ to equally weigh the rater and writer integrities. After that, we tune the importance of local trustworthiness (i.e., α) from 0.0 to 1.0 stepping by 0.1 in the view of *All Users*. The results are illustrated in Figure 4 (a). Although we only show the performance in terms of Pre@5, other measures in all the testing views follow similar trends. The figure shows that the best performance is achieved when α is set around 0.8 or 0.9. It indicates that: (1) combining both local and global trustworthiness can obtain better gains (than when setting $\alpha = 1.0$), and (2) local trustworthiness should be weighed more than the global trustworthiness in order to achieve better performance of predicting implicit trust.

By setting $\alpha = 0.9$, we continue to adjust the importance of writer ability (i.e., γ) in the range [0, 1] with step 0.1. The results are depicted in Figure 4 (b) from which the best precision is obtained when setting γ around 0.4. That is, rater ability (in providing reliable ratings) is more effective than writer ability (in writing reliable reviews) for trust prediction. This observation is different from the previous works [15], [16] which focus only on the writer ability. Further, we consider the effect of parameter η for combining writer and rater integrities by fixing $\alpha = 0.9, \gamma = 0.4$, and illustrate the results in Figure 4 (c). It shows that the setting $\eta = 0.2$ works the best, indicating that: (1) the combination of rater and writer integrities are useful; (2) rater integrity is more important than writer integrity. Nevertheless, we also note that the varying ranges of performance when tuning the three parameters are different. Specifically, the parameter of local trustworthiness α has greater and significant effects on the overall performance than γ and η which only function on the global trustworthiness and slightly impact the whole performance. This conclusion also holds for the other testing views, namely Cold Start and Warm Start where $\alpha = 0.8$ or 0.9 achieves the best performance and tuning γ and η has small effect.

Performance Comparison. The ranking performance as well as the percentages of improvements relative to our approach ETAF in predicting the top-5/10 trustworthy users are presented in Table II. The best results for each method are adopted by exclusively grid searching the best parameter values, i.e., by tuning each parameters from 0 to 1 with step 0.1. The table shows that both of our approaches (i.e., ETAF* and ETAF) consistently outperform the others across three different testing views. Specifically, the EPT method exhibits the poorest performance since it only considers the ability factor for trust prediction and ignores the other trust factors. TAF is more



(a) All Users (b) Cold Start (c) Warm Start

Fig. 5. The effects of importance (parameter α) of local trustworthiness on the ETAF's performance when fixing $\gamma = 0.5, \eta = 0.5$

TABLE II. THE TRUST RANKING PERFORMANCE OF ALL THE METHODS IN CIAODVDS

View	Method	Pre@5	Pre@10	Rec@5	Rec@10	MAP	NDCG	MRR
All Users	EPT	0.1270	0.1080	0.0236	0.0358	0.0247	0.0662	0.2577
	Improve	113%	102%	219%	185%	221%	147%	92%
	TAF	0.2529	0.2067	0.0733	0.0992	0.0757	0.1550	0.4614
	Improve	6.96%	5.66%	2.73%	2.82%	4.89%	5.55%	7.11%
	ETAF*	0.2577	0.2107	0.0702	0.0950	0.0733	0.1547	0.4678
	Improve	4.97%	3.65%	7.26%	7.37%	8.32%	5.75%	5.64%
	ETAF	**0.2705**	**0.2184**	**0.0753**	**0.1020**	**0.0794**	**0.1636**	**0.4942**
Cold Start	EPT	0.0072	0.0042	0.0142	0.0179	0.0116	0.0167	0.0314
	Improve	949%	883%	686%	552%	762%	696%	
	TAF	0.0663	0.0332	0.1057	0.1057	0.0968	0.1208	0.2509
	Improve	13.88%	24.40%	5.58%	10.41%	3.31%	10.02%	
	ETAF*	0.0663	0.0332	0.1057	0.1057	0.0973	0.1206	0.2487
	Improve	13.88%	24.40%	5.58%	10.41%	2.77%	10.20%	
	ETAF	**0.0755**	**0.0413**	**0.1116**	**0.1167**	**0.1000**	**0.1329**	**0.2616**
Warm Start	EPT	0.2849	0.2411	0.0327	0.0504	0.0397	0.1159	0.4785
	Improve	55.77%	52.26%	109%	91.07%	103%	66.70%	40.15%
	TAF	0.3836	0.3281	0.0654	0.0922	0.0771	0.1852	0.6081
	Improve	15.69%	11.89%	4.34%	4.45%	4.67%	4.32%	10.28%
	ETAF*	0.4301	0.3651	0.0664	**0.0972**	**0.0810**	**0.1938**	0.6583
	Improve	3.19%	0.55%	2.86%	-0.93%	-0.37%	-0.31%	1.87%
	ETAF	**0.4438**	**0.3671**	**0.0683**	0.0963	0.0807	0.1932	**0.6706**

TABLE III. THE TRUST RANKING PERFORMANCE OF ALL THE METHODS IN EPINIONS

View	Method	Pre@5	Pre@10	Rec@5	Rec@10	MAP	NDCG	MRR
All Users	EPT	0.1700	0.1164	0.2840	0.3124	0.2753	0.3267	0.4203
	Improve	109%	105%	94.37%	87.04%	105%	95.81%	97.69%
	TAF	0.3509	0.2361	0.5478	0.5824	0.5599	0.6361	0.8275
	Improve	1.34%	1.23%	0.77%	0.33%	0.75%	0.57%	0.41%
	ETAF	**0.3556**	**0.2390**	**0.5520**	**0.5843**	**0.5641**	**0.6397**	**0.8309**
Cold Start	EPT	0.0471	0.0235	0.1933	0.1933	0.1895	0.2011	0.2216
	Improve	246%	246%	245%	245%	243%	234%	243%
	TAF	0.1629	0.0814	0.6660	0.6660	**0.6529**	**0.6799**	**0.7651**
	Improve	0.00%	0.00%	0.00%	0.00%	-0.46%	-0.35%	-0.55%
	ETAF	**0.1629**	**0.0814**	**0.6660**	**0.6660**	0.6499	0.6775	0.7609
Warm Start	EPT	0.5775	0.4739	0.1633	0.2264	0.2293	0.3689	0.7946
	Improve	24.38%	28.99%	33.31%	36.35%	48.89%	34.05%	12.90%
	TAF	0.7085	0.6042	0.2108	0.3040	0.3307	0.4846	**0.8994**
	Improve	1.38%	1.18%	3.27%	1.55%	3.24%	2.04%	-0.26%
	ETAF	**0.7183**	**0.6113**	**0.2177**	**0.3087**	**0.3414**	**0.4945**	0.8971

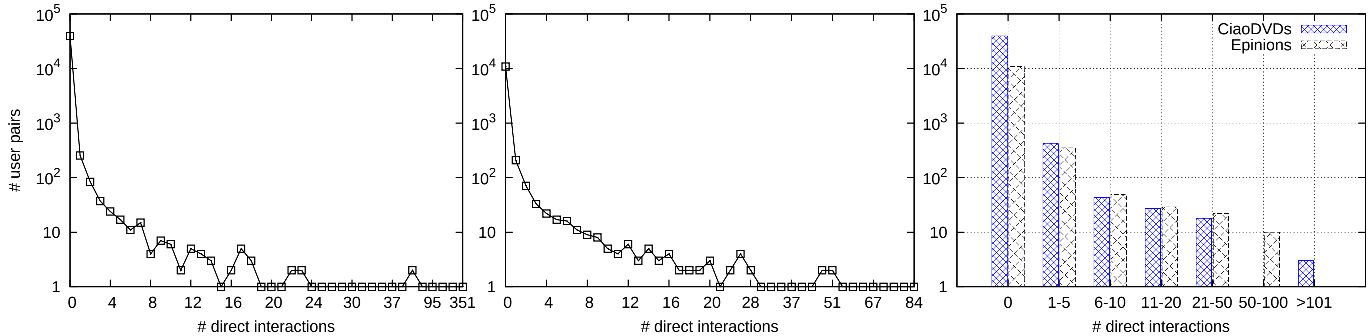
effective by considering more trust factors (i.e., benevolence and integrity) in predicting local trustworthiness, demonstrating their usefulness in trust prediction. Our approaches further improve the performance by combining both local and global trustworthiness. Taking the Pre@5 as an example, the maximum percentage of improvements (over TAF) is up to 6.96, 13.88 and 15.69, corresponding to the views of All Users, Cold Start and Warm Start, respectively. More intuitively, by comparing with ETAF* which only considers local trustworthiness, we note that ETAF achieves even better performance by considering a small amount of global trustworthiness since the best performance is obtained when setting $\alpha = 0.9$.

D. Case Study 2: Epinions

We proceed to study the performance in Epinions. Similarly as in CiaoDVDs, we first analyze the impact of local trustworthiness (α) on our approach in Epinions by fixing $\gamma = 0.5, \eta = 0.5$. The results across three testing views are illustrated in Figure 5. However, different from CiaoDVDs, the usefulness of global trustworthiness is negligible since: (1) the best performance is achieved when $\alpha = 1$; and (2) the precision produced by merely global trustworthiness is near to 0 (only 0.0004 in Warm Start) when $\alpha = 0$.

Another point of view is from the results presented in Table III. Note that ETAF* is not presented since it is equivalent with ETAF when $\alpha = 1$. Specifically, EPT works the worst among all the methods due to partial considerations of trust factors. TAF performs better than EPT by adopting more trust factors. Finally, our approach ETAF further improves TAF, especially in the case of Warm Start. This indicates that our approach outperforms TAF, even if only local trustworthiness is used. In other words, we propose a better formalization of local trustworthiness — although the ability is modelled the same as TAF, the values of local leniency (leading to benevolence, see Equation 8) is the real difference from the TAF method. Recall that local leniency is obtained by computing the global trustworthiness from Algorithm 1. We can draw a conclusion that even if global trustworthiness of trustees is less effective to predict trust, it is still capable to influence and improve the values of local trustworthiness. In this regard, considering global trustworthiness is useful and important for trust prediction.

Furthermore, we investigate the distributions of the amount of trustor-trustee user pairs with respect to the number of direct interactions between the two users. The results over the two datasets are illustrated in Figure 6. It shows that there are a lot more user pairs in CiaoDVDs without any direct interactions than in Epinions (39625 vs. 10849), and these user pairs occupy the majority of both datasets. This phenomenon also gives a possible explanation why the overall precision



(a) CiaoDVDs (b) Epinions (c) Histogram

Fig. 6. The distribution and histogram of the amount of trustor-trustee pairs with respect to the number of direct interactions.

and recall (presented in Tables II and III) are relatively small. The user pairs in Epinions have flatter distribution over the number of direct interactions (when greater than 5). In contrast, there are only 4 user pairs with direct interactions greater than 50 in CiaoDVDs. Statistically, the average of direct interactions in Epinions is around 7.67 (see Table I), which is much greater than that (i.e., 5.65) in CiaoDVDs. As a result, the local trustworthiness could be sufficient to predict trust (if predictable) when a number of direct interactions are available, and thus further limits the importance of global trustworthiness. This is not conflicting with our motivation where we posit that global trustworthiness is needed if there is a lack of direct user interactions. Nevertheless, we have analyzed that the computation of global trustworthiness can help derive more accurate values of local leniency by which the local trustworthiness is improved.

VI. Conclusion and Future Work

This paper proposed an extended trust antecedents framework for trust value predictions. Specifically, both local and global trustworthiness of trustees were taken into account, computed from three trust factors, namely ability, benevolence and integrity from the perspectives of local and global views, respectively. Together with the trust propensity of the trustor, a personalized trust metric was derived. The resulting ETAF model can alleviate the situation where users' direct interactions are sparse or zero. Experimental results on two real-world datasets indicate that our approach outperformed contemporary counterparts in terms of trust ranking performance. Global trustworthiness has both direct impact on the ranking performance and indirect influence on the local trustworthiness. For future work, we intend to take into consideration more features of user interactions, such as interaction duration and frequency, to better formulate the trust factors.

Acknowledgment

Guibing Guo thanks the Institute for Media Innovation for a PhD grant at Nanyang Technological University, Singapore. NYS thanks the Operations group at the Judge Business School and the fellowship at St Edmund's College, Cambridge.

References

[1] G. Guo, J. Zhang, and D. Thalmann, "A simple but effective method to incorporate trusted neighbors in recommender systems," in *Proc. of UMAP*, 2012, pp. 114–125.

[2] P. Lamsal, "Understanding trust and security," *Department of Computer Science, University of Helsinki, Finland*, 2001.

[3] A. Josang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decision Support Systems*, vol. 43, no. 2, pp. 618–644, 2007.

[4] P. Massa and P. Avesani, "Trust-aware recommender systems," in *Proc. of RecSys*, 2007, pp. 17–24.

[5] J. Tang, H. Gao, and H. Liu, "mTrust: Discerning multi-faceted trust in a connected world," in *Proc. of WSDM*, 2012, pp. 93–102.

[6] Y. Yao, H. Tong, X. Yan, F. Xu, and J. Lu, "MATRI: a multi-aspect and transitive trust inference model," in *Proc. of WWW*, 2013.

[7] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust," in *Proc. WWW'04*. ACM, 2004, pp. 403–412.

[8] J. Golbeck, "Computing and applying trust in web-based social networks," Ph.D. dissertation, 2005.

[9] H. Liu, E. Lim, H. Lauw, M. Le, A. Sun, J. Srivastava, and Y. Kim, "Predicting trusts among users of online communities: an epinions case study," in *Proc. of EC*, 2008, pp. 310–319.

[10] N. Ma, E.-P. Lim, V.-A. Nguyen, A. Sun, and H. Liu, "Trust relationship prediction using online product review data," in *Proceedings of the 1st ACM International Workshop on Complex Networks Meet Information & Knowledge Management*, 2009, pp. 47–54.

[11] Y. Matsuo and H. Yamamoto, "Community gravity: measuring bidirectional effects by trust and rating on online social networks," in *Proc. of WWW*, 2009, pp. 751–760.

[12] V.-A. Nguyen, E.-P. Lim, H.-H. Tan, J. Jiang, and A. Sun, "Do you trust to get trust? a study of trust reciprocity behaviors and reciprocal trust prediction," in *Proc. of SDM*, 2010, pp. 72–83.

[13] R. Mayer, J. Davis, and F. Schoorman, "An integrative model of organizational trust," *Academy of management review*, vol. 20, no. 3, pp. 709–734, 1995.

[14] G. Guo, J. Zhang, D. Thalmann, A. Basu, and N. Yorke-Smith, "From ratings to trust: an empirical study of implicit trust in recommender systems," in *Proceedings of the 29th Annual ACM Symposium on Applied Computing (SAC)*, 2014.

[15] V.-A. Nguyen, E.-P. Lim, J. Jiang, and A. Sun, "To trust or not to trust? predicting online trusts using trust antecedent framework," in *Proc. of ICDM*, 2009, pp. 896–901.

[16] Y. Kim and R. Phalak, "A trust prediction framework in rating-based experience sharing social networks without a web of trust," *Information Sciences*, vol. 191, no. 0, pp. 128–145, 2012.

[17] M. Gómez, J. Carbó, and C. B. Earle, "Honesty and trust revisited: the advantages of being neutral about others cognitive models," *Autonomous Agents and Multi-Agent Systems*, vol. 15, no. 3, pp. 313–335, 2007.

[18] G. Guo, J. Zhang, and N. Yorke-Smith, "A novel bayesian similarity measure for recommender systems," in *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI)*, 2013.

[19] G. Shani and A. Gunawardana, "Evaluating recommendation systems," *Recommender Systems Handbook*, pp. 257–297, 2011.