

Mining Implicit Correlations between Users with the Same Role for Trust-Aware Recommendation

Haifeng Liu, Zhuo Yang*, Jun Zhang, Xiaomei Bai, Wei Wang, Feng Xia

School of Software, Dalian University of Technology
Dalian 116620, China

[e-mail: zhuoyang@outlook.com]

*Corresponding author: Zhuo Yang

Received July 30, 2015; accepted September 23, 2015; published December 31, 2015

Abstract

Trust as one of important social relations has attracted much attention from researchers in the field of social network-based recommender systems. In trust network-based recommender systems, there exist normally two roles for users, truster and trustee. Most of trust-based methods generally utilize explicit links between truster and trustee to find similar neighbors for recommendation. However, there possibly exist implicit correlations between users, especially for users with the same role (truster or trustee). In this paper, we propose a novel Collaborative Filtering method called CF-TC, which exploits Trust Context to discover implicit correlation between users with the same role for recommendation. In this method, each user is first represented by the same-role users who are co-occurring with the user. Then, similarities between users with the same role are measured based on obtained user representation. Finally, two variants of our method are proposed to fuse these computed similarities into traditional collaborative filtering for rating prediction. Using two publicly available real-world Epinions and Ciao datasets, we conduct comprehensive experiments to compare the performance of our proposed method with some existing benchmark methods. The results show that CF-TC outperforms other baseline methods in terms of RMSE, MAE, and recall.

Keywords: Trust, Collaborative Filtering, Similar Trusters, Similar Trustees, Co-occurrence

1. Introduction

With the development of information technology, Recommender Systems (RSs) as an indispensable type of information filtering technique have attracted much attention in the past decades. Collaborative filtering (CF), a typical recommendation technique, which automatically predicts the interests of an active user by collecting rating information from other similar users or items, has been successfully applied in lots of domains such as e-commerce [1], e-learning [2], music [3], [4], movies [5], [6], mobile tourism [7], [8], etc.

Recently, with social networks becoming increasingly popular, more and more researchers have realized that social networks play an important role in recommender systems because they provide rich information on various social relations between users such as friendship to improve traditional recommendation methods. Trust, which is one of these important social relations, has been proved its strong connections to users' interests and its effectiveness in recommendation quality improvement in many studies [9], [10]. The rationale underlying trust-aware recommendation method is that, the taste of a user (truster) is similar to that of his trusted users (trustees).

Most of these studies focus on mining and utilizing explicit links between trusters and trustees to find similar neighbors for recommendation. However, implicit correlations between users with the same role (truster or trustee) are ignored, which are actually useful for generating accurate predictions. These implicit correlations between same-role users may be built based on explicit link between role-different users. Yao et al. [11] proposed to incorporate similar trusters and trustees into a traditional matrix factorization-based recommendation method. They considered two users with the same role (truster or trustee) to be related if they trusted (or are trusted by) the same trustees (or trusters), as shown in Fig. 1. In Fig. 1(a), two trusters u_1 and u_2 (with the same role) are similar to some extent because they trust a trustee v_2 simultaneously. In Fig. 1(b), two trustees v_1 and v_2 (with the same role) are similar to some extent because they are trusted by a truster u_1 simultaneously. Based on this theory, trust relation matrix was exploited to compute corresponding similarities between users with the same role. Additionally, implicit correlation between users with the same role may be built based on implicit co-occurrence between same-role users, as shown in Fig. 2. In Fig. 2(a), trusters u_1 , u_2 , and u_3 (with the same role) are co-occurring (because they trust a trustee v_1 simultaneously), trusters u_2 , u_3 , and u_4 are co-occurring (because they trust a trustee v_2 simultaneously). Thus, u_1 and u_4 are similar to some extent because they are described through the same co-occurrences, here trusters u_2 and u_3 . Likewise, Fig. 2(b) shows another example on inferring similarity between trustees v_1 and v_4 .

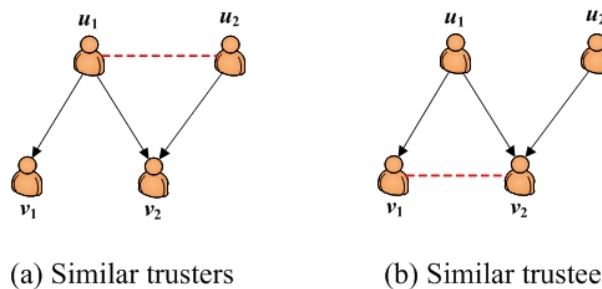


Fig. 1. Common trusters or trustees relations: (a) u_1 and u_2 have a common trustee v_2 ; (b) v_1 and v_2 have a common truster u_1

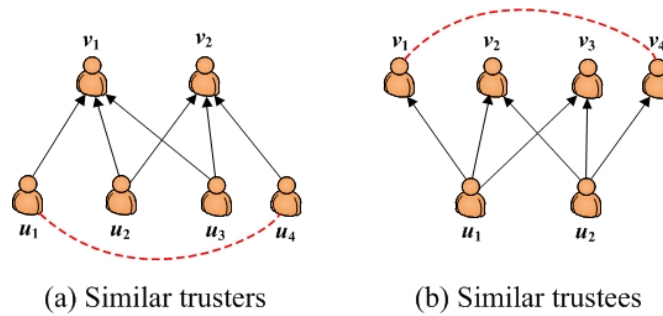


Fig. 2. Co-occurrence relations: (a) u_1 and u_4 are co-occurring with the same trusters u_2 and u_3 ; (b) v_1 and v_4 are co-occurring with the same trustees v_2 and v_3 .

In this paper, we focus on utilizing implicit correlations between users with the same role stated in the abovementioned scenario for recommendation and present a novel Collaborative Filtering recommendation method based on Trust Context (CF-TC). We first employ an association mining technique to determine whether two users with the same role are significantly co-occurring based on trust context. Then, each user is represented by these significantly co-occurred users. Finally, user similarities are computed based on user vectors and are fused into two traditional (memory-based and matrix factorization-based) collaborative filtering methods for rating prediction. We have conducted relevant experiments on two real-world datasets Epinions and Ciao, and the results demonstrate that our proposed method outperforms other baseline methods in terms of RMSE, MAE, and recall. This proves the effectiveness of proposed method.

The major contributions of this paper are summarized as follows:

(i) We propose to mine the implicit correlations between users with the same role based users' trust context.

(ii) We propose two methods to exploit implicit correlations between users with the same role for improving recommendation performance.

(iii) We conduct relevant experiments on two publicly available datasets, and verify the effectiveness of the proposed recommendation method using relevant metrics.

The rest of this paper is organized as follows. Section 2 summarizes the related work on trust-aware recommendation. Section 3 proposes our method and introduces the detail of its two variants. Section 4 presents related experiments and discusses the results. Section 5 concludes this paper.

2. Related Work

The challenges of existing recommender systems mainly include cold start [12], [13], data sparsity [14], [15], and attacks [16], [17]. Trust has been regarded as one kind of commonly-used auxiliary information to help design more efficient recommender systems by solving these challenges to some extent. In this section, we review some related work on trust-aware recommendation methods. In a trustware recommender systems, trust information can be used in one of the following approaches along with traditional recommender systems [18]: (i) Trust-aware memory-based CF approachers, which use memory-based CF techniques as their basic methods, and (ii) Trust-aware model-based CF approachers, which use model-based CF techniques as their basic methods.

Trust-aware memory-based CF approaches incorporate trust information to depress recommendations from distrusted users and boost recommendations from trusted users.

Golbeck [19] proposed a method called TidalTrust which performs a modified breadth first search in the trust network to compute trust values based on the following two observations: (i) shorter propagation paths produce more accurate trust estimates, and (ii) paths with higher trust values create better results. Massa and Avesani [9] proposed a new trust metric called MoleTrust which is similar to TidalTrust. Additionally, MoleTrust needs a predefined trust threshold to determine which users to consider in the rating aggregation process. Jamali and Ester [20] proposed a recommendation model called TrustWalker which combines trust-based and item-based recommendation. TrustWalker queries a user's direct and indirect friends' ratings for the target item as well as similar items by performing a random walk in online social networks. These recommendation approaches use trust information to either filter distrusted users or weigh the recommendation results made by all users.

Matrix factorization technique is a widely-used recommendation method in model-based CF. Ma et al. [21] considered that users and their trust networks should have similar ratings on items, and proposed to predict a missing rating for a given user as a linear combination of ratings from the user and his/her trust network. Tang et al. [22], [23] proposed to perform a co-factorization procedure in the user-item matrix and the user-user trust relation matrix by sharing the same user preference latent factor based on the assumption that a user shares the same user preference vector in the rating space (rating information) and the trust relation space. Menon et al. [22], [24] reconstructed the trust matrix T to perform trust relation prediction. Jamali et al. [25], [26] proposed methods to force a user's preference to be closer to that of users in the user's trust network. Specially, SocialMF [25] proposed to force the preferences of a user to be closer to the average preference of the user's trust network. Yang et al. [27] proposed a method called CircleCon which uses SocilMF [25] as a base method and focuses on inferring category-specific social trust circles from available rating information combined with social network information where social trust relations across all categories are mixed together. Trust-aware model-based MF approaches assume that users' preferences are similar to or influenced by their trusted users.

The methods above focus on utilizing direct or indirect trust relations to improve recommendation performance of traditional methods. However, users in trust network are associated with two different roles (truster and trustee). There may be implicit correlations between users with the same role. Therefore, our work proposes to mine these implicit correlations and then apply them in collaborative filtering for improving recommendation performance. Yao et al. [11] also take users' roles into consideration. In this paper, we take their work as a baseline to verify the effectiveness of our method.

3. Design of CF-TC

3.1 Overview

Our proposed CF-TC method mainly includes two components, as shown in Fig. 3: (i) mining implicit correlations between users with the same role and (ii) applying the mined implicit correlations for rating prediction. At the first step, for each user, we build his user representation using those who have the same role with the user. Then, we obtain the weight of each implicit correlation by computing the cosine similarity between any two users with the same role. At the second step, based on the computed weights, we present two variants of CF-TC, i.e., Memory-based CF-TC (CF-TC-Me) and Matrix Factorization-based CF-TC (CF-TC-MF). CF-TCMe and CF-TC-MF exploit the memory-based collaborative filtering and matrix-factorization-based collaborative filtering to predict ratings.

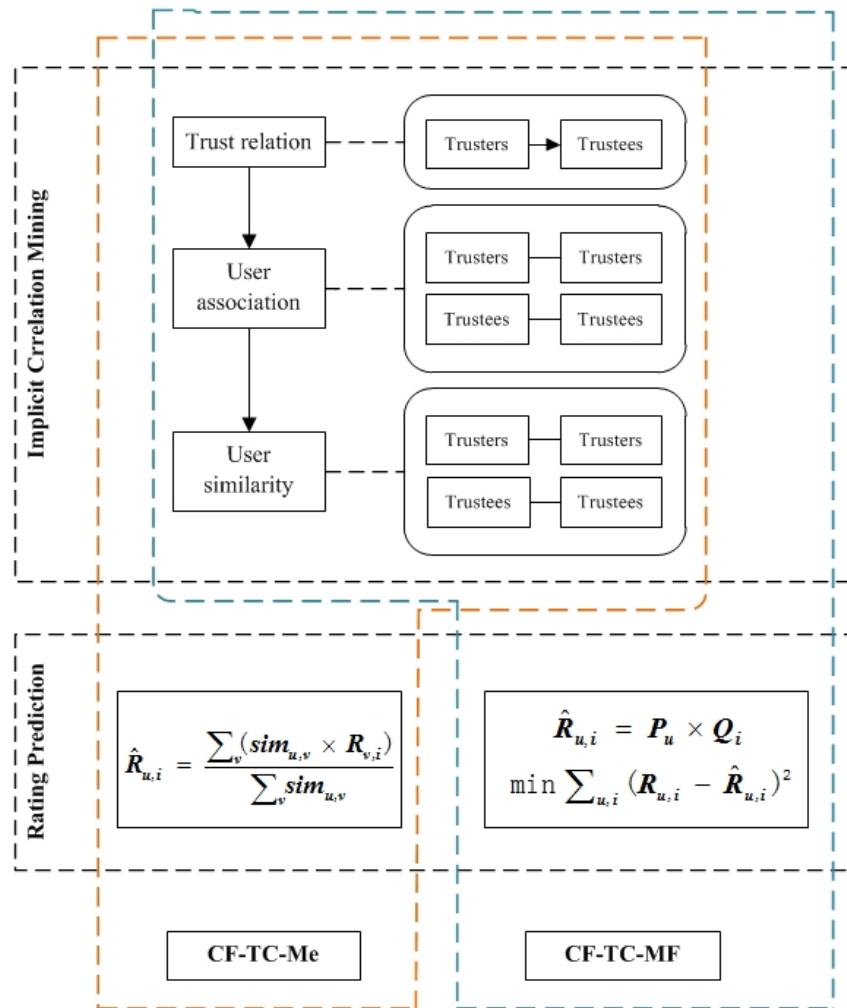


Fig. 3. Our proposed method.

3.2 Mining Implicit Correlations

3.2.1 Building User Representation

In this section, we introduce how to build user representation based on the trust context. In trust relations, there are two roles of users: trusters who expressed trust to others and trustees who were trusted by others. We consider that two trusters are co-occurring if they expressed trust to the same trustees, and two trustees are co-occurring if they were trusted by the same trusters. The following examples are taken to illustrate how to obtain the co-occurrence relations between trusters as well as trustees, and their user representations. **Table 1** shows the trust relations between trusters u_1, u_2, u_3, u_4 and trustees v_1, v_2, v_3, v_4, v_5 . Here, for simplicity, we consider two trusters to be significantly occurring if they expressed trust to at least one trustee simultaneously. In addition, a binary value of 1 or 0 is used for stating whether two trusters are co-occurring or not. **Table 2** shows the user (truster) vectors describing the four trusters calculated in this manner. As demonstrated in the table, trusters u_2 and u_4 are described by the same user vectors, therefore they are highly related. Similarly, we consider two trustees to be significantly co-occurring if they were trusted by at least one truster simultaneously. In this way, associations between trustees are computed and the user (trustee) vectors describing

the five trustees are shown in **Table 3**, where trustees v_3 and v_5 are described by almost the same user vectors, therefore they are also highly related. Actually, significant co-occurrences of users with the same role (truster or trustee) depend on the trust context. Therefore, we calculate a significance score based on contingency tables commonly used in statistical theory.

Table 1. Exemplary trust relation matrix

	v_1	v_2	v_3	v_4	v_5
u_1	1	1	1	1	-
u_2	-	1	-	-	1
u_3	-	-	-	1	1
u_4	1	-	1	1	-

Table 2. User vectors for trusters

	u_1	u_2	u_3	u_4
u_1	0	1	1	1
u_2	1	0	1	0
u_3	1	1	0	1
u_4	1	0	1	0

Table 3. User vectors for trustees

	v_1	v_2	v_3	v_4	v_5
v_1	0	1	1	1	0
v_2	1	0	1	1	1
v_3	1	1	0	1	0
v_4	1	1	1	0	1
v_5	0	1	0	1	0

Table 4. Contingency table

	v	$\sim v$
u	N_{11}	N_{12}
$\sim u$	N_{21}	N_{22}

Table 4 shows the contingency table for users u and v with the same role. When they are trusters, they expressed trusts to N_{11} trustees simultaneously, truster u expressed trusts to N_{12} trustees who were not trusted by truster v , truster v expressed trusts to N_{21} trustees who were not trusted by truster u , and they did not express trust to N_{22} trustees. On the other hand, when users u and v are trustees, they were expressed trusts by N_{11} trusters simultaneously, trustee u was trusted by N_{12} trusters who did not express trusts to trustee v , trustee v was trusted by N_{21} trusters who did not express trusts to trustee u , and they were not expressed N_{22} trusters. Based on the obtained contingency table, the χ^2 test is then used to measure the association degree between any two users, as shown in Equation 1. χ^2 test sums the squared z-scores for each cell in the contingency table and is correlated to the expected frequencies. The cumulative distribution function of χ^2 as shown in Equation 2, is used to compute association degree *deg* ranging from 0 to 1. A larger value of χ^2 will generate a larger association degree. The larger association degree indicates that the co-occurrence between two users is more significant. We use a threshold ts to determine whether the co-occurrence is significant or not. For two users with the same role, if their computed association degree is larger than ts , they are regarded to be significantly co-occurred, and vice versa.

$$\chi^2 = \frac{(|N_{11} \cdot N_{22} - N_{12} \cdot N_{21}| - \frac{N}{2})^2}{R_1 \cdot R_2 \cdot C_1 \cdot C_2} \quad (1)$$

where $R_1 = N_{11} + N_{12}$, $R_2 = N_{21} + N_{22}$, $C_1 = N_{11} + N_{21}$, $C_2 = N_{12} + N_{22}$, and $N = C_1 + C_2 = R_1 + R_2$.

$$deg = F(x/\nu) = \int_0^x \frac{t^{\nu/2-1} \cdot e^{-t/2}}{2^{\nu/2} \cdot \Gamma(\nu/2)} dt \quad (2)$$

where ν is the freedom degree and equals 1 for the contingency table here, $\Gamma(\cdot)$ is the Gamma function, x is the value of χ^2 computed using Equation 1.

3.2.2 Computing User Similarity

We calculate the similarity between any two users u and v with the same role using the cosine similarity which measures the angle between their user vectors, as shown in Equation 3. V_u and V_v denote their user vectors. The value of cosine similarity is ranging from 0 to 1.

$$sim_{u,v} = \frac{V_u \cdot V_v}{|V_u| |V_v|} \quad (3)$$

3.3 Memory-based CF-TC

To predict a target user's rating on a target item, memory-based collaborative filtering generally employs two classes of methods: (i) user-centric method, computes the similarities between users and then utilizes ratings of similar users on the target item to predict unknown rating; (ii) item-centric method, computes the similarities between items and then utilizes ratings of the target user on similar items to predict unknown rating. Here, we use the first method. We compute the predicted rating $\hat{R}_{u,i}$ of a user u on an item i by averaging the scores of neighbor users of u while each score is weighted by the corresponding similarity $sim_{u,v}$, as shown in Equation 4.

$$\hat{R}_{u,i} = \frac{\sum_v (sim_{u,v} \times R_{v,i})}{\sum_v sim_{u,v}} \quad (4)$$

where v is one of neighbor users of u .

3.4 Matrix Factorization-based CF-TC

Matrix Factorization (MF) is a traditional model-based collaborative filtering method. Matrix Factorization-based CF performs a low-rank matrix factorization on the user-item rating matrix based on the assumption that a few latent patterns influence user rating behaviors. Let $P_u \in R^K$ and $Q_i \in R^K$ be the user preference vector for user u and item feature vector for i respectively, where K is the number of latent vectors. The objective function of MF method is

$$L_1(P, Q) = \frac{1}{2} \sum_{u \in U, i \in I} W_{u,i} (R_{u,i} - P_u Q_i)^2 + \left(\frac{1}{2} \lambda (\|P\|_F^2 + \|Q\|_F^2) \right) \quad (5)$$

where $W_{u,i}$ is an indicator function that is equal to 1 if user u expressed rating on item i and equal to 0 otherwise, $\|\cdot\|_F^2$ denotes the Frobenius norm, λ is the regularization coefficient. By performing gradient descent in P_u and Q_i , the minimum of the objective function above can be found and then P and Q can be obtained. Next, the prediction rating $\hat{R}_{u,i}$ can be computed as follows:

$$\hat{R}_{u,i} = \bar{r} + P_u Q_i \quad (6)$$

where \bar{r} is a (global) offset value.

The theory of trust-based recommendation stipulates that the tastes of two users existent in trust relations have to be similar. The similarity between users means they have similar interests to some extent. Thus, we minimize the following two objective functions to integrate the information on implicit correlations between trusters and trustees, respectively.

$$L_2(P) = \frac{1}{2} \sum_{u \in U_1} \left\| P_u - \sum_{v \in U_1} sim_{u,v} P_v \right\|^2 \quad (7)$$

$$L_3(P) = \frac{1}{2} \sum_{u \in U_2} \left\| P_u - \sum_{v \in U_2} sim_{u,v} P_v \right\|^2 \quad (8)$$

where U_1 and U_2 denote the set of trusters and the set of trustees, $sim_{u,v}$ is the similarity between users u and v computed in previous section.

The unified object function for CF-TC-MF is defined as

$$L(P, Q) = L_1(P, Q) + \alpha L_2(P) + \beta L_3(P) \quad (9)$$

where α and β are non-negative parameters that are used to trade off the three objective functions. The minimum of the objective function can be found by performing gradient descent in P_u and Q_i :

$$\begin{aligned} \frac{\partial L}{\partial P_u} &= \sum_{u \in U, i \in I} W_{u,i} (P_u Q_i - R_{u,i}) Q_i + \lambda P_u \\ &\quad + \bar{W}_{u,U_1} \times \alpha (P_u - \sum_{v \in U_1} sim_{u,v} P_v) \\ &\quad + \bar{W}_{u,U_2} \times \beta (P_u - \sum_{v \in U_2} sim_{u,v} P_v) \end{aligned} \quad (10)$$

$$\frac{\partial L}{\partial Q_i} = \sum_{u \in U, i \in I} W_{u,i} (P_u Q_i - R_{u,i}) P_u + \lambda Q_i \quad (11)$$

$$P_u = P_u - \theta \times \frac{\partial L}{\partial P_u} \quad (12)$$

$$Q_i = Q_i - \theta \times \frac{\partial L}{\partial Q_i} \quad (13)$$

where $W_{u,i}$ is the indicator function that is equal to 1 if u expressed a rate score on i , and equal to 0 otherwise. Similarly, \bar{W}_{u,U_1} (or \bar{W}_{u,U_2}) is equal to 1 if u is a truster (or a trustee) and equal to 0 otherwise. θ is the learning rate for iteration. The initial values of P_u and Q_i are sampled from the normal distribution with zero mean. In each iteration, P_u and Q_i are updated based on the latent variables from the previous iteration. The learned P_u and Q_i can be used to predict ratings for user-item pairs.

4. Experimental Results and Analysis

4.1 Datasets

Epinions (www.epinions.com) and Ciao (www.ciao.it) are well-known consumer opinion websites. From these two websites, two kinds of information can be extracted: (i) rating

information, users can make reviews to familiar items and further assign these items numeric ratings in the range of 1 (min) to 5 (max); (ii) trust relations, users express their trusts to other users and add these users to their trust networks if they find their reviews consistently interesting and helpful. We used the versions of Epinions and Ciao datasets [28] for our performance evaluation. In order to explore the impact of our proposed trust relation-based method (CF-TC) on recommendation quality, users without trust relations and their ratings were removed. The statistics of these preprocessed datasets are presented in Table 5.

Table 5. Data statistics

Features	Epinions	Ciao
Number of users	8518	2342
Number of items	153593	77524
Number of ratings	345022	146753
Sparsity of ratings	0.9997	0.9992
Number of trusters	8514	2190
Number of trustees	8510	2325
Number of trust links	299936	57544
Sparsity of trust links	0.9917	0.9790
Number of users	8518	2342

4.2 Experimental Setup

We randomly divided each of the two above-mentioned datasets into two parts: training set (80%) and test set (20%). The evaluation metrics in our experiments are the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE), as these are some of the most popular accuracy measures in the literature of recommender systems. RMSE and MAE are defined using Equations 14 and 15.

$$RMSE = \sqrt{\frac{\sum_{R_{u,i} \in R_{test}} (R_{u,i} - \hat{R}_{u,i})^2}{\sum |R_{test}|}} \quad (14)$$

$$MAE = \sqrt{\frac{\sum_{R_{u,i} \in R_{test}} |R_{u,i} - \hat{R}_{u,i}|}{\sum |R_{test}|}} \quad (15)$$

In addition, we employed top-N recall as another evaluation metric in our experiments. As to computing the top-N recall, for each user u , we rank the items i according to the predicted rating $\hat{R}_{u,i}$. An item is defined as relevant to a user in the test set if s/he finds it appealing or interesting (e.g., the assigned rating in the test data is above a certain threshold). For instance, in our experiments with Epinions data, the rating value ranges from 1 to 5, and we consider rating value of 5 as relevant (i.e. the user definitely liked these items), while other rating values and missing rating values are considered not relevant. Other choices led to similar results. Now the recall can be defined as the fraction of relevant items in the test set that are in the top-N of the ranking list, denoted by $Num(N, u)$, from among all relevant items, $Num(u)$. For each user u , the individual top-N recall is given by

$$Recall(u) = \frac{Num(N, u)}{Num(u)} \quad (16)$$

which can be aggregated over all users to obtain the average recall for the test set [29]. The average top-N recall is computed as follows:

$$Recall = \frac{\sum_u Num(N, u)}{\sum_u Num(u)} \quad (17)$$

In our experiments, we compared the recommendation results of the following two classes of methods in order to demonstrate the effectiveness of our proposed CF-TC method. The core of the first class of methods is memory-based collaborative filtering, including CF-Me (This is the traditional user-centric collaborative filtering method), CF-TM-Me (It computes the user similarities based on trust matrix for applying memory-based collaborative filtering method [11]), and CF-TC-Me (This is the first variant of CF-TC). The core of the second class of methods is matrix factorization, including CF-MF (This is the traditional matrix factorization-based recommendation method, which represents user and item with feature vector [30]), and CF-TM-MF (It fuses two kinds of correlations between users into traditional matrix factorization model for recommendation [11] and these correlations are computed based on trust matrix), and CF-TC-MF (This is the second variant of CF-TC).

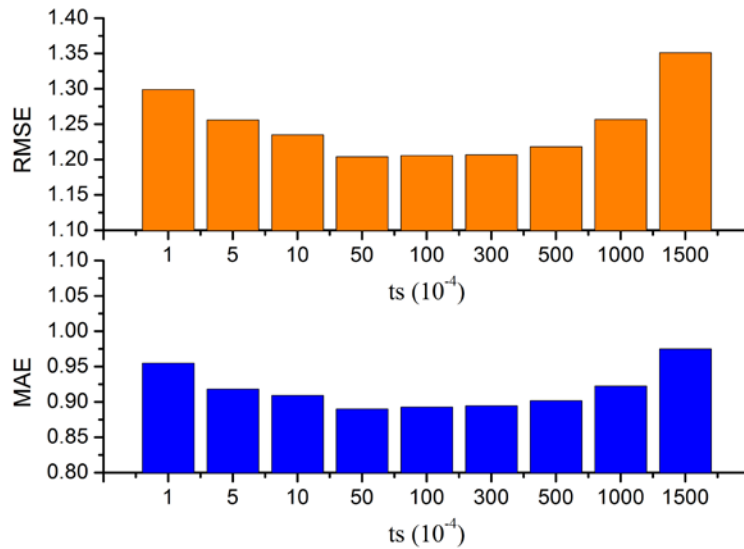


Fig. 4. Impact of ts on CF-TC-Me (truster) on Epinions.

4.3 Results and Analysis

As the threshold ts is an important impact parameter for determining significant co-occurrences between users with the same role, we conducted relevant experiments to discuss its impact. Fig. 4 shows the results of our proposed CF-TC-Me method based on the truster-truster similarities in terms of RMSE and MAE on Epinions when ts is ranging from 0.0001 to 0.15. From Fig. 4 we can observe that, CF-TC-Me achieves smaller values of RMSE and MAE when ts is equal to an intermediate value (e.g., 0.005) than those when ts is equal to a smaller value (e.g., 0.0001) or a larger value (e.g., 0.15). This is due to the fact that, too large value of ts leads to filtering too many useful co-occurrence relations, and too small value of ts leads to keeping too many useless co-occurrence relations, thereby obtaining so inaccurate similarities.

Fig. 5 shows the results of our proposed CF-TC-Me method based the trustee-trustee similarities in terms of RMSE and MAE Epinions when ts is ranging from 0.0001 to 0.15. It can also be seen from **Fig. 5** that RMSE and MAE are smaller when ts equals an intermediate value (e.g., 0.005). In addition, in order to further discuss the impact of ts on CF-TC-Me method, we conducted other experiments on Ciao and obtained the similar results as shown in **Figs. 6** and **7**. and the results demonstrated the importance of an appropriate value of ts . Therefore, an appropriate value of ts is important for improving recommendation quality.

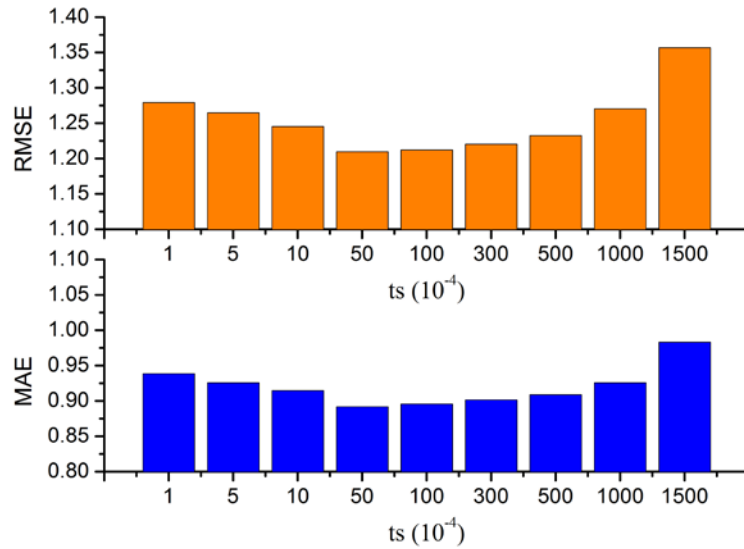


Fig. 5. Impact of ts on CF-TC-Me (trustee) on Epinions.

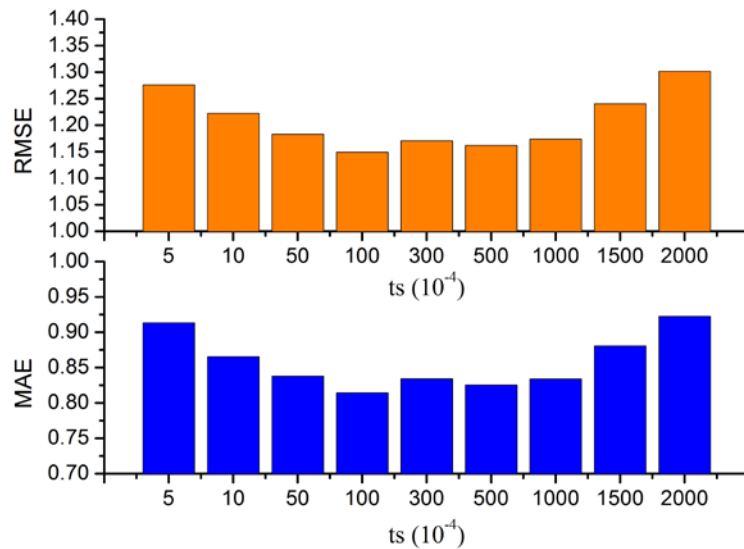


Fig. 6. Impact of ts on CF-TC-Me (truster) on Ciao.

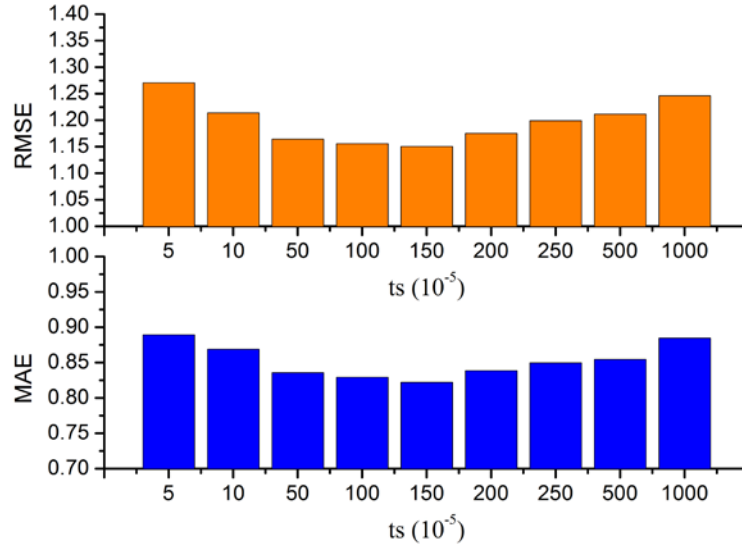


Fig. 7. Impact of ts on CF-TC-Me (trustee) on Ciao.

We compared the results of the above-mentioned first three methods (CF-Me, CF-TM-Me, and CF-TC-Me) on Epinions and Ciao. Table 6 summarizes their performance comparisons when taking each truster as recommendation target and setting ts to 0.005 for Epinions, 0.01 for Ciao. From Table 6, it can be observed that, CF-TC-Me and CF-TM-Me achieve smaller values of RMSE and MAE than CF-Me. This proves once again that trust relations can improve recommendation quality of traditional CF method, as demonstrated in other related work [9], [10]. Another observation is that CF-TC-Me performs better than CF-TM-Me in terms of RMSE and MAE. This indicates that, trust context-based similarity yields more accurate user relations being discovered and exploited when compared to the trust matrix-based recommendation method. Additionally, Table 7 summarizes their performance comparisons when taking each trustee as recommendation target and setting ts to 0.005 for Epinions, 0.0015 for Ciao. The results also demonstrate that CF-TC-Me outperforms CF-TM-Me and CF-TMMe outperforms CF-Me. This also verifies the effectiveness of our proposed method.

Table 6. Performance comparison for trusters

Datasets	CF-Me	CF-TM-Me	CF-TC-Me	Metrics
Epinions	1.4157	1.2579	1.2040	RMSE
	1.0697	0.9219	0.8901	MAE
Ciao	1.2924	1.2005	1.1490	RMSE
	0.9236	0.8545	0.8143	MAE

Table 7. Performance comparison for trustees

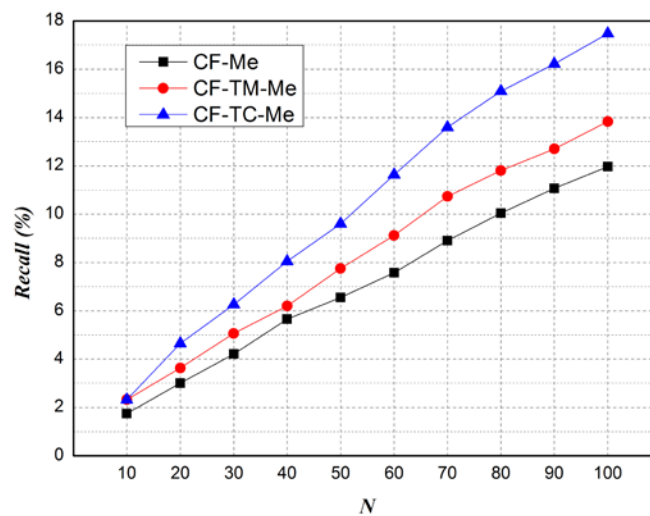
Datasets	CF-Me	CF-TM-Me	CF-TC-Me	Metrics
Epinions	1.4154	1.3032	1.2096	RMSE
	1.0695	0.9509	0.8920	MAE
Ciao	1.2806	1.1988	1.1508	RMSE
	0.9158	0.8435	0.8221	MAE

Table 8. RMSE of different CF-TC-Mes

Datasets	CF-TC-Me (truster)	CF-TC-Me (trustee)	CF-TC-Me (hybrid)
Epinions	1.1910	1.1985	1.1220
Ciao	1.1324	1.1382	1.0067

As a user is not only a truster but also a trustee in trust relations, we conducted experiments on Epinions and Ciao, to demonstrate the comparison results of CF-TC-Me based on truster-truster similarities, CF-TC-Me based on trustee-trustee similarities and hybrid CF-TC-Me by utilizing least squares method to fuse the ratings predicted by the first two methods. From **Table 8** we can observe that hybrid CF-TC-Me performs better than CF-TC-Me (truster) and CF-TC-Me (trustee). This indicates that the data fusion technique optimizes RMSE and it is possibly useful to utilize information on users' different roles for designing a better method.

To test top-N recommendation, we conducted relevant experiments and compared the results of the above-mentioned first three methods (CF-Me, CF-TM-Me, and CF-TC-Me) on Epinions, as shown in **Figs. 8** and **9**. **Fig. 8** shows the recalls of all comparison partners for trusters, respectively. From this figure, we can see that the recalls of the three methods are increasing as the length of recommendation list, N , increases. This is obviously due to the fact that, a larger value of N makes more items recommended to each truster and then more relevant items are possibly find in the recommendation list. In addition, we can also see that, for different values of N , the recalls of both CF-TC-Me and CF-TM-Me are higher than that of CF-Me. Namely, CF-TC-Me and CF-TM-Me outperform CF-Me in terms of recall. This further proves that, it is necessary and effective to incorporate trust into recommender systems for improving recommendation quality. Furthermore, compared CF-TC-Me to CF-TM-Me, it is significantly observed that, the recall of CF-TC-Me is higher than that of CF-TM-Me for all different values of N . Namely, CF-TC-Me is better than CF-TM-Me in terms of recall. This indicates that, our proposed method can mine more accurate correlations between users with the same role (here, it is truster) and then improve recommendation quality in terms of recall.

**Fig. 8.** Comparison of Recall of the first three methods for trusters on Epinions.

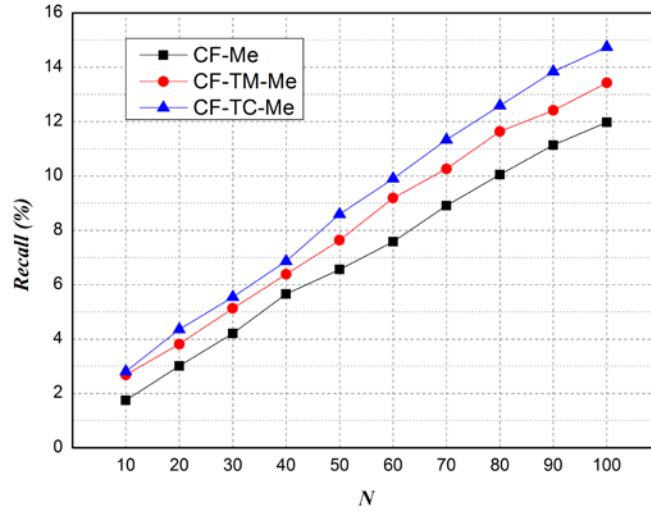


Fig. 9. Comparison of Recall of the first three methods for trustees on Epinions.

Furthermore, Fig. 9 shows the recalls of all comparison partners for trustees, respectively. From this figure, we can see that, for trustees as recommendation targets, CF-TM-Me is better than CF-Me and CF-TC-Me is better than CF-TM-Me in terms of recall. Meanwhile, compared to the results for trusters, the difference between the recalls of CF-TC-Me and CF-TM-Me is smaller. This might be influenced by users' roles and rating information. Also, we conducted relevant experiments on Ciao dataset, as shown in Figs. 10 and 11. We can see that their results on Ciao are similar to those on Epinions. This further improves the effectiveness of our proposed method.

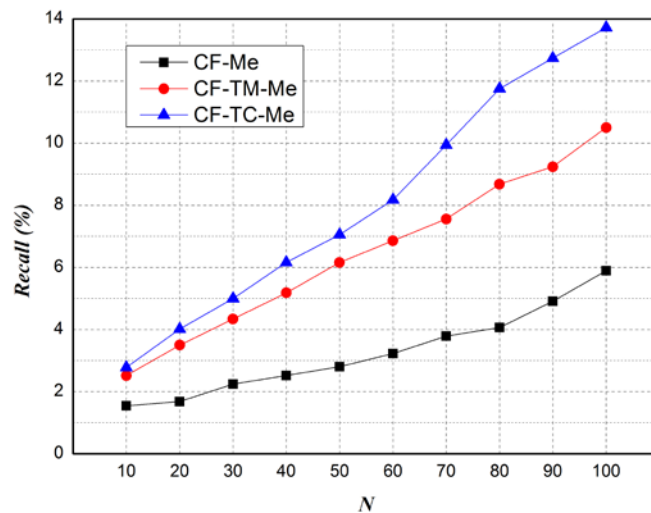


Fig. 10. Comparison of Recall of the first three methods for trusters on Ciao.

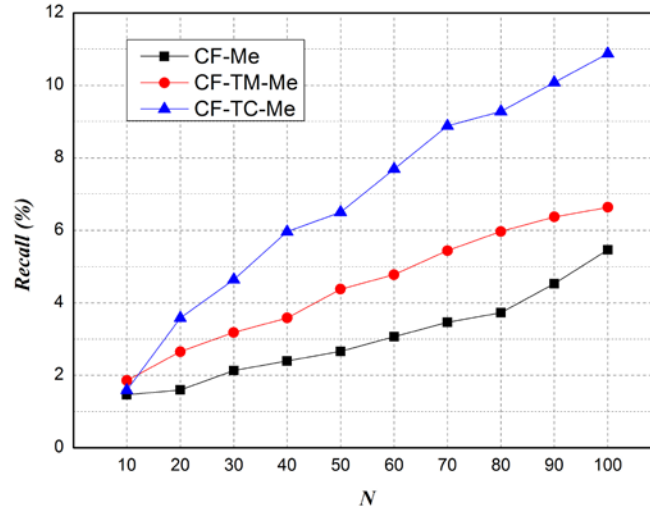


Fig. 11. Comparison of Recall of the first three methods for trustees on Ciao.

Additionally, we compared the results of the abovementioned last three methods (CF-MF, CF-TM-MF, and CFTC-MF) on Epinions and Ciao. In our experiments, we set the values of the parameters θ and λ as 0.01 and 0.1, respectively. We let α equal β and assign 0.001, 0.01, 0.1, 1, 10, 20, 30, 50 to them to test their impacts on CF-TC-MF method in terms of RMSE and MAE on Epinions and Ciao, as shown in **Figs. 12** and **13**. From these two figures, we can see that, when alpha and beta are equal to 1, CF-TC-MF achieves the best values of RMSE and MAE on Epinions and Ciao. Furthermore, we set the values of these parameters α , β , λ , and θ as 1, 1, 0.1, and 0.01, respectively for next comparison. **Table 9** summarizes their performance comparisons when ts is assigned the same value in the previous experiments, i.e., ts is equal to 0.005/0.01 for Epinions/Ciao when computing the similarities between trusters, and ts is equal to 0.005/0.0015 for Epinions/Ciao when computing the similarities between trustees. From **Table 9**, it can be observed that, CF-TCMF and CF-TM-MF achieve smaller values of RMSE and MAE than CF-MF. This proves that trust relations can improve recommendation quality of traditional matrix factorization-based CF method, as demonstrated in other work [18], [31]. We can obviously see that CF-TC-MF performs better than CF-TM-MF in terms of RMSE and MAE. This also indicates that, trust context-based similarity yields more accurate user relations being discovered and exploited when compared to the trust matrix-based matrix factorization method. This also verifies the effectiveness of our proposed method.

Table 9. Performance comparison

Datasets	CF-MF	CF-TM-MF	CF-TC-MF	Metrics
Epinions	2.0434	1.8620	1.7262	RMSE
	1.7058	1.5154	1.3640	MAE
Ciao	1.8501	1.7200	1.6473	RMSE
	1.4959	1.3498	1.2669	MAE

Comparing **Table 9** to **Tables 6** and **7**, we can observe that matrix factorization-based CF outperforms memory-based CF in terms of RMSE and MAE on Epinions and Ciao. In spite of

this, we cannot say that the former is absolutely better than the latter. Actually, they have their own advantages. The former can generate real-time recommendations while the latter can predict a user's rating on an item even though the item is not rated by those who are similar to the user.

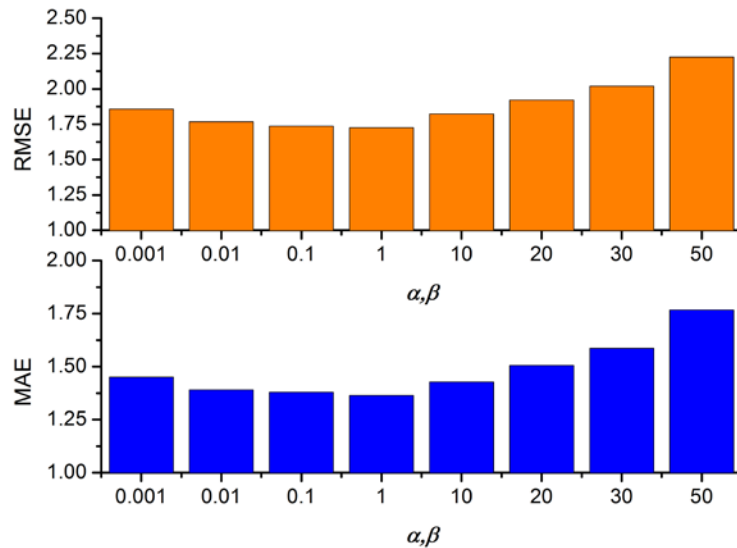


Fig. 12. RMSE and MAE of CF-TC-MF method for different values of α and β on Epinions ($\alpha = \beta$).

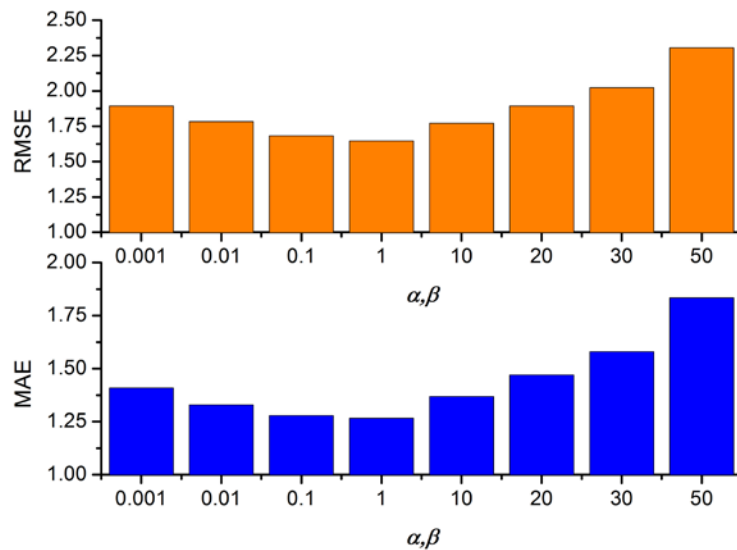


Fig. 13. RMSE and MAE of CF-TC-MF method for different values of α and β on Ciao ($\alpha = \beta$).

5. Conclusion

In this paper, we have investigated how to mine implicit correlation between users with the same role for improving the accuracy of trust-based recommendation. The rationale underlying our method is that two users with the same role (truster or trustees) are considered

to be similar to some extent if they are significantly co-occurring with other same-role users. Co-occurrence relation between two same-role users is determined through an association mining technique based on trust context. Based on obtained co-occurrence relation, each user is represented by user vector and the similarities between any two users with the same role are computed by comparing their vectors. We fused these computed similarities into two traditional collaborative filtering methods. Through relevant experiments on Epinions and Ciao, we show that our proposed CF-TC method outperforms other baseline methods in terms of RMSE, MAE, and recall.

References

- [1] H.-F. Wang and C.-T. Wu, "A strategy-oriented operation module for recommender systems in e-commerce," *Computers and Operations Research*, vol. 39, no. 8, pp. 1837–1849, 2012. [Article \(CrossRef Link\)](#).
- [2] S. M. Ali, I. Ghani, M. S. A. Latiff, "Interaction-based Collaborative Recommendation: A Personalized Learning Environment (PLE) Perspective," *KSII Transactions on Internet and Information Systems*, Vol 9, No 1, January 2015, pp. 446-465. [Article \(CrossRef Link\)](#).
- [3] N. Aizenberg, Y. Koren, and O. Somekh, "Build your own music recommender by modeling internet radio streams," in *Proc. of the 21st International Conference on World Wide Web*, pp. 1–10, 2012. [Article \(CrossRef Link\)](#).
- [4] P. Knees and M. Schedl, "A survey of music similarity and recommendation from music context data," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 10, no. 1, pp. 1–21, 2013. [Article \(CrossRef Link\)](#).
- [5] W. Carrer-Neto, M. L. Hernandez-Alcaraz, R. Valencia-Garcia, and F. Garcia-Sanchez, "Social knowledge-based recommender system. application to the movies domain," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10 990–11 000, 2012. [Article \(CrossRef Link\)](#).
- [6] A. Azaria, A. Hassidim, S. Kraus, A. Eshkol, O. Weintraub, and I. Netanel, "Movie recommender system for profit maximization," in *Proc. of the 7th ACM Conference on Recommender Systems*, pp. 121–128, 2013. [Article \(CrossRef Link\)](#).
- [7] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, "Mobile recommender systems in tourism," *Journal of Network and Computer Applications*, vol. 39, pp. 319–333, 2014. [Article \(CrossRef Link\)](#).
- [8] J. Borrás, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7370–7389, 2014. [Article \(CrossRef Link\)](#).
- [9] P. Massa and P. Avesani, "Trust-aware recommender systems," in *Proc. of the 2007 ACM Conference on Recommender Systems*, pp. 17–24, 2007. [Article \(CrossRef Link\)](#).
- [10] J. L. Harman, J. O'Donovan, T. Abdelzaher, and C. Gonzalez, "Dynamics of human trust in recommender systems," in *Proc. of the 8th ACM Conference on Recommender Systems*, pp. 305–308, 2014. [Article \(CrossRef Link\)](#).
- [11] W. Yao, J. He, G. Huang, and Y. Zhang, "Modeling dual role preferences for trust-aware recommendation," in *Proc. of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 975–978, 2014. [Article \(CrossRef Link\)](#).
- [12] J. Bobadilla, F. Ortega, A. Hernando, and J. Bernal, "A collaborative filtering approach to mitigate the new user cold start problem," *Knowledge-Based Systems*, vol. 26, pp. 225–238, 2012. [Article \(CrossRef Link\)](#).
- [13] B. Lika, K. Kolomvatsos, and S. Hadjiefthymiades, "Facing the cold start problem in recommender systems," *Expert Systems with Applications*, vol. 41, no. 4, pp. 2065–2073, 2014. [Article \(CrossRef Link\)](#).
- [14] G. Guo, "Resolving data sparsity and cold start in recommender systems," *Lecture Notes in Computer Science*, vol. 7379, pp. 361–364, 2012. [Article \(CrossRef Link\)](#).

- [15] G. Guo, J. Zhang, and D. Thalmann, "Merging trust in collaborative filtering to alleviate data sparsity and cold start," *Knowledge-Based Systems*, vol. 57, pp. 57–68, 2014. [Article \(CrossRef Link\)](#).
- [16] J. Cao, Z. Wu, B. Mao, and Y. Zhang, "Shilling attack detection utilizing semi-supervised learning method for collaborative recommender system," *World Wide Web*, vol. 16, no. 5-6, pp. 729–748, 2013. [Article \(CrossRef Link\)](#).
- [17] G. Noh, Y. Kang, H. Oh, and C. Kim, "Robust sybil attack defense with information level in online recommender systems," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1781–1791, 2014. [Article \(CrossRef Link\)](#).
- [18] H. Liu, F. Xia, Z. Chen, N. Y. Asabere, J. Ma, and R. Huang, "Trucom: Exploiting domain-specific trust networks for multi-category item recommendation," *IEEE Systems Journal*, doi:10.1109/JSYST.2015.2427193, 2015. [Article \(CrossRef Link\)](#).
- [19] J. A. Golbeck, *Computing and applying trust in web-based social networks*, University of Maryland, 2005.
- [20] M. Jamali and M. Ester, "Trustwalker: a random walk model for combining trust-based and item-based recommendation," in *Proc. of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 397–406, 2009. [Article \(CrossRef Link\)](#).
- [21] H. Ma, I. King, and M. R. Lyu, "Learning to recommend with social trust ensemble," in *Proc. of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 203–210, 2009. [Article \(CrossRef Link\)](#).
- [22] J. Tang, H. Gao, X. Hu, and H. Liu, "Exploiting homophily effect for trust prediction," in *Proc. of the Sixth ACM International Conference on Web Search and Data Mining*, pp. 53–62, 2013. [Article \(CrossRef Link\)](#).
- [23] H. Ma, H. Yang, M. R. Lyu, and I. King, "Sorec: social recommendation using probabilistic matrix factorization," in *Proc. of the 17th ACM Conference on Information and Knowledge Management*, pp. 931–940, 2008. [Article \(CrossRef Link\)](#).
- [24] A. K. Menon and C. Elkan, "Link prediction via matrix factorization," in *Lecture Notes in Computer Science*, vol. 6912, pp. 437–452, 2011. [Article \(CrossRef Link\)](#).
- [25] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proc. of the Fourth ACM Conference on Recommender Systems*, pp. 135–142, 2010. [Article \(CrossRef Link\)](#).
- [26] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, "Recommender systems with social regularization," in *Proc. of the Fourth ACM International Conference on Web Search and Data Mining*, pp. 287–296, 2011. [Article \(CrossRef Link\)](#).
- [27] X. Yang, H. Steck, and Y. Liu, "Circle-based recommendation in online social networks," in *Proc. of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1267–1275, 2012. [Article \(CrossRef Link\)](#).
- [28] J. Tang, H. Gao, H. Liu, and A. D. Sarma, "eTrust: Understanding trust evolution in an online world," in *Proc. of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 253–261, 2012. [Article \(CrossRef Link\)](#).
- [29] X. Yang, H. Steck, Y. Guo, and Y. Liu, "On top-k recommendation using social networks," in *Proc. of the Sixth ACM Conference on Recommender Systems*, pp. 67–74, 2012. [Article \(CrossRef Link\)](#).
- [30] A. Mnih and R. Salakhutdinov, "Probabilistic matrix factorization," *Advances in Neural Information Processing Systems*, pp. 1257–1264, 2007.
- [31] G. Guo, J. Zhang, and N. Yorke-Smith, "Trustsvd: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings," in *Proc. of the 29th AAAI Conference on Artificial Intelligence*, 2015.



Haifeng Liu received the B.Sc. and M.Sc. degrees in computer science, in 2009 and 2012, respectively, from the Dalian University of Technology, China, where he is currently working toward the Ph.D. degree in the School of Software. His research interests include recommender systems, mobile computing, and big data.



Zhuo Yang received B.Sc. degree from China Jiliang University and M.Sc. degree from Nanyang Technological University, Singapore. She has been working as a Software Engineer in School of Software, Dalian University of Technology, China since 2009. She was a Software Quality Assurance Engineer at Hundsun Technologies Inc., Hangzhou, China and a Software Validation Engineer in Gemalto R&D, Singapore. Her research interests include mobile computing, recommender systems, and big data.



Jun Zhang received her B.Sc. and M.Sc. degrees in software engineering from Jilin University, Changchun, China. She is currently working toward the Ph.D. degree in Software Engineering in Dalian University of Technology, Dalian, China. Her research interests include computational social science, scientific impact evaluation and scholarly data analysis.



Xiaomei Bai received the B.S. degree in computer and application from University of Science and Technology Liaoning, Anshan, China, in 2000 and the M.S. degree in software engineering from Jilin University, Changchun, China, in 2006. She is currently working toward the Ph.D. degree from the Mobile and Social Computing Laboratory, School of Software, Dalian University of Technology, Dalian, China. Since 2000, she has been working in Anshan Normal University, Anshan, China. Her research interests include computational social science, relation analysis and influence analysis.



Wei Wang received his B.S. degree in Electronic Information Science and Technology from Shenyang University, Shenyang, China, in 2012. He is currently working toward the Ph.D. degree in Software Engineering in Dalian University of Technology (DUT), Dalian, China. His research interests include big scholarly data, social network analysis and computational social science.



Feng Xia received the BSc and Ph.D. degrees from Zhejiang University, Hangzhou, China. He was a Research Fellow at Queensland University of Technology, Australia. He is currently a Full Professor in School of Software, Dalian University of Technology, China. He is the (Guest) Editor of several international journals. He serves as General Chair, PC Chair, Workshop Chair, Publicity Chair, or PC Member of a number of conferences. Dr. Xia has authored/co-authored two books and over 180 scientific papers in international journals and conferences. His research interests include social computing, computational social science, big data, and mobile social networks. He is a Senior Member of IEEE and ACM.