IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON ADVANCED DATA MINING METHODS FOR SOCIAL COMPUTING

Received December 30, 2019, accepted January 9, 2020, date of publication January 20, 2020, date of current version February 4, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968102

Mining Heterogeneous Influence and Indirect Trust for Recommendation

ZHAOYI LI[1], FEI XIONG[1], XIMENG WANG[1,2], ZHE GUAN[3], AND HONGSHU CHEN[4]

[1]School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
[2]Centre for Artificial Intelligence, Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia
[3]China Electronics Technology Group Corporation, Beijing 100846, China
[4]School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

Corresponding author: Fei Xiong (xiongf@bjtu.edu.cn)

This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2019JBM013.

ABSTRACT Relationships between users in social networks have been widely used to improve recommender systems. However, actual social relationships are always sparse, which sometimes bring great harm to the performance of recommender systems. In fact, a user may interact with others that he/she does not connect directly, and thus has an impact on these users. To mine abundant information for social recommendation and alleviate the problem of data sparsity, we study the process of trust propagation and propose a novel recommendation algorithm that incorporates multiple information sources into matrix factorization. We first explore heterogeneous influence strength for each pair of linked users and mine indirect trust between users by using trust propagation and aggregation strategy in social networks. Then, explicit and implicit information of user trust and ratings are incorporated into matrix factorization, and the influence of indirect trust is considered in the recommendation process. Experimental results show that the proposed model achieves better performance than some state-of-the-art recommendation models in terms of accuracy and relieves the cold-start problem.

INDEX TERMS Recommender systems, indirect trust, heterogeneous influence, matrix factorization.

I. INTRODUCTION

Recommender systems provide users opportunities to filter information so that solve the information overload problem. Collaborative filtering (CF) techniques [1], [2] are widely used in e-commerce websites because their efficiency. CF recommendation algorithms only use user-item ratings to mine the users' preference so that make recommendation for the active user which results that CF algorithms encounter data sparsity problem [3], [4] and cold start problem [5], [6]. The former is caused by that ratings collected by the recommender system are few, so it is difficult to model user preferences. The latter is that for a new user or item with no ratings, CF algorithms cannot perform satisfactory recommendation.

One potential way to improve the recommendation accuracy of CF techniques and solve these problems is to incorporate social information provided by users into recommendation models [7]–[14]. The homophily theory [15] indicates that people are more likely to socialize with people who are similar with themselves. Social influence theory [16] indicates that social linked people will influence each other and share more similarities. In real world, people always seek advice from their friends before making decisions and are more likely to accept recommendations from their friends. Therefore, in addition to the rating information, the social relationships among users can also applied in recommendation process. Social relationships provide an opportunity to address cold-start problem [17] and improve the performance of recommender systems.

Although the social information is complementary to the rating information, the direct social relationship is also very sparse which is similar to user-item ratings. To identify more information used for predicting users' preference, some researchers take advantage of the implicit feedback information [18]–[23]. For example, the ratings on items and the trust values between users are explicit information, and the rated items, trusted friends and trusting friends of users are implicit information. In other words, who rates what and who trusts whom are implicit information that can be inferred from the user-item rating matrix and user-user trust matrix. The implicit feedbacks of users can reveal their interests. Implicit feedbacks are precious when explicit information

The associate editor coordinating the review of this manuscript and approving it for publication was Shirui Pan.

is not available. When explicit information is available, the additional implicit feedbacks can also be used to predict users' preferences more accurately. TrustSVD model [12] incorporates the explicit and implicit influence of user ratings and user trust. Comprehensive experimental results show that TrustSVD outperforms both trust-based and rating-based methods in prediction accuracy. Although the above implicit information for recommender system can improve the prediction accuracy, it also directly uses the two matrices, i.e., user-item rating matrix and user-user trust matrix and does not inferr further information from the foundational information.

Based on the propagation of trust in social networks, one can predict the trustworthiness of unconnected users and infer the indirect trust relationships between users. Then, the problem of data sparsity can be effectively alleviated, especially when users do not have enough direct relationships. In addition, users often have heterogeneous influence on their friends. Social recommendation models that consider user influence uniformly may not adequately characterize the dependence of users' preferences on their friends. In this paper, we propose a social recommendation method with heterogeneous influence and indirect trust mining (ReHI). We first investigate heterogeneous influence strength between connected users in terms of affinity and node reputation. Then, we mine indirect trusted users for every active user, and the influence of indirect trusted users can be naturally incorporated by extending user modelling. Within this method, the main contributions of this work are summarized as follows:

(1) We propose a method of calculating indirect trust values between users, and then infer indirect trusted friends for each active user.

(2) We incorporate the influence of indirect trusted friends into social recommendation. Combining indirect trust relationships and heterogeneous influence strength, we propose a recommendation model ReHI.

(3) Extensive experiments demonstrate that our proposed recommendation model reduces rating prediction errors, especially for cold-start users.

The remainder of this paper is organized as follows. In Section 2, we provide an overview of several recommendation methods. In Section 3, we describe the proposed recommendation method. The experimental results and analysis are presented in Section 4. In Section 4, we give the conclusions and suggestions for future work.

II. RELATED WORK

With the rapid development of social networks, social recommendation that uses the social relationships among users to infer their preferences and make recommendations has emerged and been intensively researched in recent years. Social recommendation algorithms based on matrix factorization [8], [9], [14], [24], [25] have received more attention than neighbor based recommendation algorithms because the former have high prediction accuracy and good scalability.

Some work has been done to use direct social connections in addition to rating information for inferring user preference and recommending items. In [8], Ma *et al.* proposed a social recommendation model SoRec based on probabilistic matrix factorization (PMF) [26]. SoRec algorithm integrates trust relationships among users into PMF model. Specially, SoRec model firstly calculates the trust values between social linked users and then simultaneously factorizes the user-item rating matrix and user-user trust matrix to learn user feature vectors and item feature vectors more accurately. Compared to PMF model, SoRec has greatly improved the accuracy of rating prediction. However, experiments have shown that jointly matrix factorization is suitable for relationships such as membership but not suitable for friendship [27]. In [9], Ma *et al.* proposed the Recommendation with Social Trust Ensemble algorithm (RSTE) according to the recommendation process in real world, that is, the final rating of a user on an item is a comprehensive result of his/her own preference and his/her friends' opinions. RSTE model combines the preferences of the target user and that of his/her friends, modelling the final preference of the target user, thereby predicting the missing scores. References [8] and [9] are famous efforts to simultaneously use direct social links among users and user-item ratings for rating prediction in recommender systems, and such efforts have achieved great performance improvement of recommender systems. However, the sparsity of the observed social relationships limits the further improvement. Thus, it is necessary to investigate other information such as implicit feedbacks or mine further social information from the observed social context

In recent years, several works have been done to either use implicit feedbacks [21]–[23] or trust propagation [7], [28]–[30] in recommender systems. A well-known recommendation model that incorporates both the explicit and implicit influence of user trust and of item ratings is TrustSVD [12]. TrustSVD is built on the top of SVD++ [31] model. SVD++ model takes into account user bias, item bias and the implicit influence of rated items besides user/item-specific vectors on rating prediction. The rating for user on item is predicted by Eq. (1) in SVD++ model.

$$\hat{r}_{u,i} = b_u + b_i + \mu + (p_u + |I_u|^{-\frac{1}{2}} \sum\nolimits_{i \in I_u} y_i) q_i^T \tag{1}$$

where b_u and b_i represent the rating bias of user u and item i, respectively, μ is the average value of the observed ratings, y_i represents the implicit influence of the item i that has been rated by u on the ratings of unknown items, I_u represents the set of items rated by user u.

TrustSVD extends SVD++ model by further considering both the explicit and implicit influence of user trust. Specially, TrustSVD considers the implicit influence of trustees and trustors of the target user on his/her ratings of unknown items. However, direct trust relationship used in TrustSVD model is also very sparse and the TrustSVD model does not consider the trust propagation. Trust propagation in social networks indicates that indirect trusted friends can also considered be

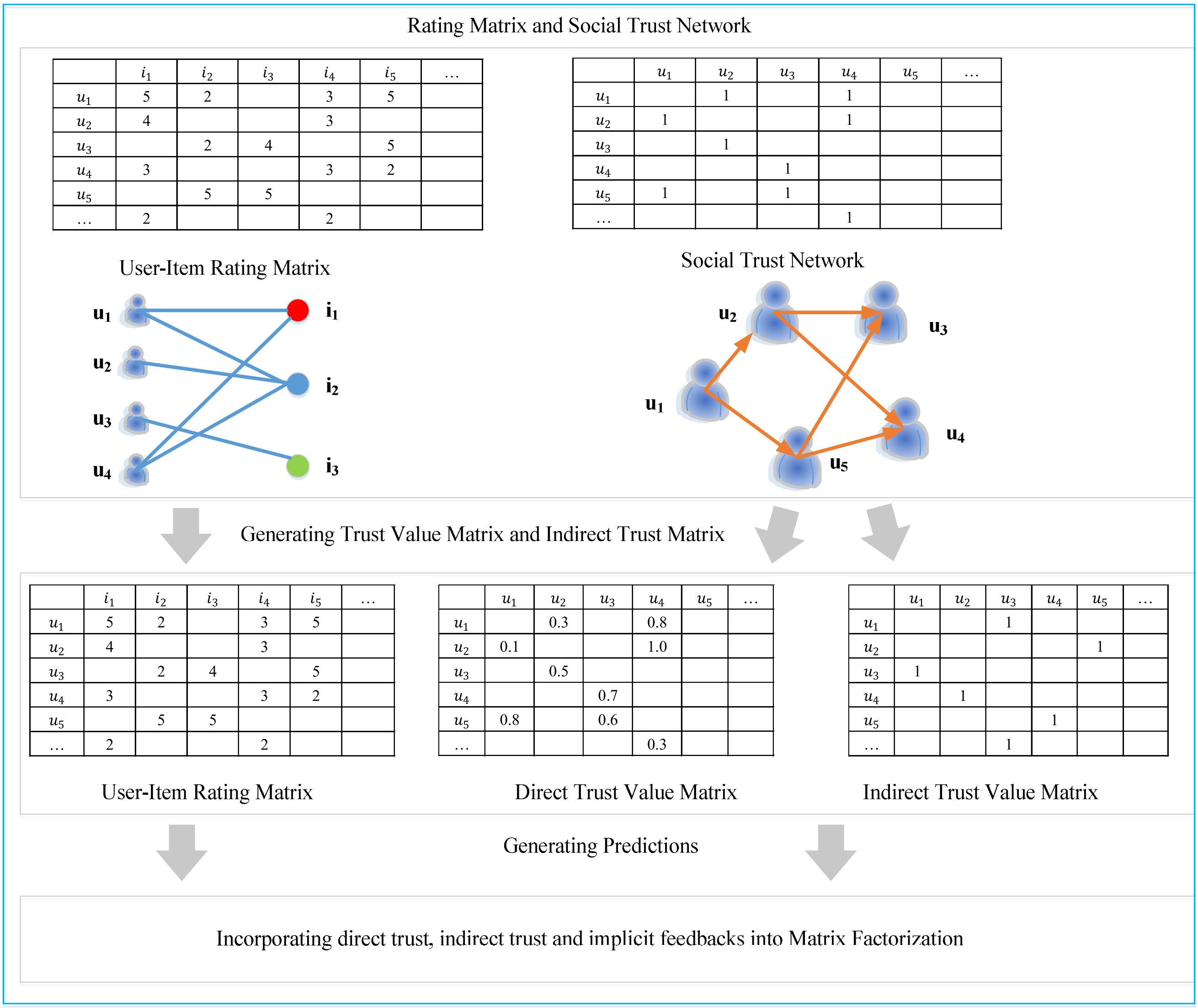


FIGURE 1. **Illustration of the proposed method.**

predictors in recommendation process [7]. Other works mentioned above either only consider implicit feedbacks or trust propagation but do not consider both.

Our ReHI model effectively combines the advantages of indirect trust mining and implicit feedbacks, capturing more information from the data. We propagate trust relationships and use implicit feedbacks to alleviate the data sparsity and cold-start problems to improve the rating prediction accuracy.

III. THE PROPOSED METHOD

In this section, we present a recommendation method with indirect trust mining. We first propose a method of calculating influence strength from social networks. Then, we give the details of indirect trust mining. Last, we introduce the recommendation model ReHI. The illustration of our method is shown in Figure 1.

A. INFLUENCE STRENGTH BASED ON AFFINITY AND NODE WEIGHT

In recommender systems, the user-item rating matrix and the trust matrix are defined as $R \in R^{m\times n}$ and $T \in R^{m\times m}$, respectively, where m is the number of users and n is the number of items. Matrix factorization maps the rating matrix into low-rank user-specific and item-specific matrices. The rating $r_{u,i}$ of user u on item i can be approximated by the inner product of latent user preference vector p_u and item feature vector q_i, i.e., $r_{u,i} \approx p_u q_i^T$. p_u and q_i both have the same dimensionality d.

According to the social correlation theory, social connected friends are likely to share similar behaviors. Each friend of an active user has different influence strength on him/her, so we cannot consider every social link uniformly. In real world, if users u and v have more common friends, there may exist stronger social influence between u and v. Furthermore, the social influence between users u and v should be strong if common friends share similar rating tastes with both users. We define *Affinity* to capture rating similarity and social network structure simultaneously, and then determine the strength of social relationships between users. In directed networks such as trust networks, a user may have both outgoing friends (i.e., trustees) and incoming friends (i.e., trustors).

Hence, *Affinity* consists of two components: $Affinity^+$ and $Affinity^-$. $Affinity^+$ is used to describe the common outgoing friends' impact and $Affinity^-$ is used to describe the common incoming friends' impact. We define $Affinity^+(u, v)$ and $Affinity^-(u, v)$ for users u and v as follows:

$$Affinity^+(u, v) = \frac{\sum_{k \in coee(u,v)} (VSS(u, k) + VSS(v, k))}{\sum_{p \in T^+_{(u)}} VSS(u, p) + \sum_{q \in T^+_{(v)}} VSS(v, q)} \tag{2}$$

$$Affinity^-(u, v) = \frac{\sum_{h \in coor(u,v)} (VSS(u, h) + VSS(v, h))}{\sum_{f \in T^-_{(u)}} VSS(u, f) + \sum_{g \in T^-_{(v)}} VSS(v, g)} \tag{3}$$

where $T^+_{(u)}$ and $T^+_{(v)}$ denote the sets of trustees of users u and v respectively; $T^-_{(u)}$ and $T^-_{(v)}$ are the sets of trustors respectively; $coee(u, v) = T^+_{(u)} \cap T^+_{(v)}$ and $coor(u, v) = T^-_{(u)} \cap T^-_{(v)}$. VSS is the vector space similarity defined as follows:

$$VSS(u, v) = \frac{\sum_{i \in I_{(u)} \cap I_{(v)}} r_{u,i} \cdot r_{v,i}}{\sqrt{\sum_{i \in I_{(u)} \cap I_{(v)}} r^2_{u,i}} \sqrt{\sum_{i \in I_{(u)} \cap I_{(v)}} r^2_{v,i}}} \tag{4}$$

where $I_{(u)}$ and $I_{(v)}$ are the sets of items rated by users u and v respectively; $r_{u,i}$ and $r_{v,i}$ are the ratings on item i given by users u and v respectively. Linear combination is applied to integrate $Affinity^+(u, v)$ and $Affinity^-(u, v)$, and $Affinity(u, v)$ is generated as:

$$Affinity(u, v) = \theta \cdot Affinity^+(u, v) + (1-\theta) \cdot Affinity^-(u, v) \tag{5}$$

where the parameter $\theta \in [0, 1]$ is used to constraint the impact of the common trustees and trustors. For the sake of simplicity, we set $\theta = 0.5$.

For a social link from user u to his/her trusted friend v, if the reputation of user v is relatively large in the social network, user v has great influence on user u. In [11], the PageRank algorithm can be used to determine the reputation rankings of all users. For social networks that contain m users, the ranking of user u's reputation is defined as $r_u \in [1, m]$, and the reputation value rep_u can be calculated according to the rankings by Eq. (6):

$$rep_u = \frac{1}{1 + \log r_u} \tag{6}$$

From Eq.(6), the range of rep_u is [0, 1]. Based on the definitions of *Affinity* and reputation, we define the influence strength that the trusted friend v exerts on user u, as:

$$inf(u, v) = Affinity(u, v) \cdot rep_v \tag{7}$$

B. INDIRECT TRUST RELATIONSHIP MINING BASED ON TRUST PROPAGATION

Because trust is subjective, the trust levels of a user to his/her trusted friends are different in trust networks. It is necessary to calculate trust values between users in social recommendation. In general, the trust value $t_{u,v}$ of user u to friend v drops if user u trusts lots of users, and $t_{u,v}$ increases if user v has high topological status in trust networks. Users' status in social networks can be represented by network centrality indicators such as node degrees, coreness, betweenness, etc. Hence, we calculate the trust value $t_{u,v}$ as:

$$t_{u,v} = tanh(c_v / d^+(u)) \tag{8}$$

where c_v is the coreness of user v in trust networks, and $d^+(u)$ represents the out degree of user u. $tanh(x)$ is the hyperbolic tangent function to limit trust values in [0, 1].

In consideration of trust propagation in social networks, we can infer how much an unknown user is trustworthy according to direct trust information. By predicting the trustworthiness of other users, we find indirect trust relationships between users.

In trust networks, there may be several paths that propagate trust from the same source to the active user, and hence, trust values for different paths should be aggregated so as to accurately estimate indirect trust value between two users. For example, there are several paths that connect user u and user z, and the distances of these paths are not the same. If the length of a path from user u to user z is two, we call user z the 2-hop friend of user u. Because shorter paths can make the indirect trust estimation more accurate [32], in this paper, we only estimate trust values on the paths with the length two, and aggregate these trust values.

$$it_u(z) = \sum_{v \in T^+_{(u)} \cap T^-_{(z)}} t_{u,v} \cdot t_{v,z} \tag{9}$$

where $t_{u,v}$ and $t_{v,z}$ are trust values in Eq. (8).

Considering all the 2-hop friends in rating prediction for the active user will inevitably introduce noise, which will damage the recommendation accuracy. Therefore, we select a certain number of 2-hop friends for an active user as his/her indirect trusted friends according to the following rule. If the number of 2-hop friends of user u is larger than L, we rank all his/her 2-hop friends according to the indirect trust values and select the 2-hop friends with top-L indirect trust values as user u's indirect trusted friends. Else, all 2-hop friends of user u are considered as his/her indirect trusted friends.

After selecting the indirect trusted friends for each user, we conduct a social relation matrix, i.e., indirect trust matrix, named IT. The value of $IT(u, z)$ is 1 if user z is user u's indirect trusted friend, and otherwise 0. The notation $B^+_{(u)}$ represents the set of indirect trusted friends of user u.

C. RECOMMENDATION WITH HETEROGENEOUS INFLUENCE AND INDIRECT TRUST MINING

In this section, we propose a social recommendation method with heterogeneous influence and indirect trust mining (ReHI). The sparsity of user-item rating matrix and direct trust relationships motivates us to further consider the implicit influence of users' indirect trusted friends in rating prediction tasks. The implicit influence of indirect trusted friends can be included in modelling users' preferences, as follows:

$$p_u + |I_u|^{-\frac{1}{2}} \sum_{j \in I_u} y_j + \alpha \left|T^+_{(u)}\right|^{-\frac{1}{2}} \sum_{v \in T^+_{(u)}} w_v + (1-\alpha) \left|B^+_{(u)}\right|^{-\frac{1}{2}} \sum\nolimits_{z \in B^+_{(u)}} x_z \quad (10)$$

where $B^+_{(u)}$ is the set of indirect trusted friends of user u; α controls the impact of direct and indirect trusted friends; y_j is the implicit influence of item j that was rated by user u; w_v and x_z are the latent vectors of user u's direct trusted friend v and indirect trusted friend z respectively, which indicate the implicit influence of users v and z.

In ReHI, the explicit influence of direct trust is incorporated from two perspectives: trust relationship factorization and social regularization. Eq. (11) is used to factorize trust relationships into user-specific vectors of trustors and trustees, i.e., p_u and w_v.

$$\frac{\lambda_t}{2} \sum\nolimits_u \sum\nolimits_{v \in T^+_{(u)}} \left(p_u w_v^T - t_{u,v}\right)^2 \quad (11)$$

where λ_t is a parameter that controls the degree of trust regularization, and $t_{u,v}$ is obtained from Eq. (8).

The explicit influence of direct trust can be also incorporated by introducing a social regularization term as shown in Eq. (12).

$$\frac{\beta}{2} \sum\nolimits_u \sum\nolimits_{v \in N(u)} i_{u,v} \left\|p_u - p_v\right\|_F^2 \quad (12)$$

where the definitions of $N(u)$ and $i_{u,v}$ are shown in Eq. (13) and Eq. (14), respectively:

$$N(u) = \begin{cases} T^+_{(u)}, & T^+_{(u)} \neq \emptyset \\ S(u), & T^+_{(u)} = \emptyset \end{cases} \quad (13)$$

$$i_{u,v} = \begin{cases} inf(u,v), & v \in T^+_{(u)} \\ \dfrac{\sum_{k \in T^-_{(v)}} inf(k,v)}{\left|T^-_{(v)}\right|}, & v \in S(u) \end{cases} \quad (14)$$

where $S(u)$ represents the K-nearest neighbors of user u according to the VSS in Eq. (4). The reason for identifying K-nearest neighbors of a user is that there are a large number of users who have no trusted friends in social networks. $i_{u,v}$ represents the influence strength of user v in $N(u)$ on user u. If $v \in T^+_{(u)}$, $i_{u,v} = inf(u,v)$, and if $v \in S(u)$, we define $i_{u,v}$ as the average of influence values that user v exerts on all of his/her trustors.

ReHI integrates indirect trust relationship and heterogeneous influence into the loss function, shown in Eq. (15).

$$L = \frac{1}{2} \sum_u \sum_{i \in I_u} \left(\hat{r}_{u,i} - r_{u,i}\right)^2 + \frac{\lambda_t}{2} \sum_u \sum_{v \in T^+_{(u)}} \left(p_u w_v^T - t_{u,v}\right)^2 + \frac{\beta}{2} \sum\nolimits_u \sum\nolimits_{v \in N(u)} i_{u,v} \left\|p_u - p_v\right\|_F^2 + \frac{\lambda}{2} \begin{pmatrix} \sum_u b_u^2 + \sum_i b_i^2 + \sum_u \|p_u\|_F^2 + \sum_i \|q_i\|_F^2 \\ + \sum_i \|y_i\|_F^2 + \sum_v \|w_v\|_F^2 + \sum_z \|x_z\|_F^2 \end{pmatrix} \quad (15)$$

where q_i is the latent item feature vector of item i, and $\hat{r}_{u,i}$ is the predicted rating that user u gives item i, shown in Eq. (16).

$$\hat{r}_{u,i} = b_u + b_i + (p_u + |I_u|^{-\frac{1}{2}} \sum_{i \in I_u} y_i + \alpha \left|T^+_{(u)}\right|^{-\frac{1}{2}} \sum\nolimits_{v \in T^+_{(u)}} w_v + (1-\alpha) \left|B^+_{(u)}\right|^{-\frac{1}{2}} \sum\nolimits_{z \in B^+_{(u)}} x_z) q_i^T \quad (16)$$

where b_u is the rating bias of user u, and b_i is the rating bias of item i. A local minimization of the loss function can be obtained by performing gradient descent on the variables shown as follows:

$$\frac{\partial L}{\partial b_u} = \sum\nolimits_{i \in I_u} e_{u,i} + \lambda b_u \quad (17)$$

$$\frac{\partial L}{\partial b_i} = \sum\nolimits_{u \in U_i} e_{u,i} + \lambda b_i \quad (18)$$

$$\frac{\partial L}{\partial p_u} = \sum\nolimits_{i \in I_u} e_{u,i} q_i + \lambda_t \sum\nolimits_{v \in T^+_{(u)}} e_{u,v} w_v + \beta \left(\sum\nolimits_{v \in N(u)} i_{u,v} (p_u - p_v) + \sum\nolimits_{g \in N^-_{(u)}} i_{g,u} (p_u - p_g)\right) + \lambda p_u \quad (19)$$

$$\frac{\partial L}{\partial q_i} = \sum\nolimits_{u \in U_i} e_{u,i} \left(p_u + |I_u|^{-\frac{1}{2}} \sum\nolimits_{j \in I_u} y_j + \alpha \left|T^+_{(u)}\right|^{-\frac{1}{2}} \times \sum\nolimits_{v \in T^+_{(u)}} w_v + (1-\alpha) \left|B^+_{(u)}\right|^{-\frac{1}{2}} \sum\nolimits_{z \in B^+_{(u)}} x_z\right) + \lambda q_i \quad (20)$$

$$\frac{\partial L}{\partial y_j} = \sum\nolimits_{u \in U_j} \sum\nolimits_{i \in I_u} e_{u,i} |I_u|^{-\frac{1}{2}} q_i + \lambda y_j \quad (21)$$

$$\frac{\partial L}{\partial w_v} = \sum\nolimits_{u \in T^-_{(v)}} \sum\nolimits_{i \in I_u} e_{u,i} \alpha \left|T^+_{(u)}\right|^{-\frac{1}{2}} q_i + \lambda_t \sum\nolimits_{u \in T^-_{(v)}} e_{u,v} p_u + \lambda w_v \quad (22)$$

$$\frac{\partial L}{\partial x_z} = \sum\nolimits_{u \in B^-_{(z)}} \sum\nolimits_{i \in I_u} e_{u,i} (1-\alpha) \left|B^+_{(u)}\right|^{-\frac{1}{2}} q_i + \lambda x_z \quad (23)$$

where $e_{u,i} = \hat{r}_{u,i} - r_{u,i}$ is the rating prediction error; $e_{u,v} = \hat{t}_{u,v} - t_{u,v}$ is the trust prediction error; U_i is the set of users who have rated item i; $B^-_{(u)}$ represents the user set $\left\{k \mid u \in B^+_{(k)}\right\}$.

The pseudocode for training the model is presented in Algorithm 1.

D. COMPLEXITY ANALYSIS

The main cost in learning ReHI model is the calculation of the loss function L and its gradients against the feature vectors of users and items. The computational complexity to calculate the loss function is $O(d\,|R| + d\,|T| + d\,|IT| + mKd)$, where d is the dimensionality of feature space, $|R|$, $|T|$ and $|IT|$ refer to the number of nonzero entries of R, T and IT, respectively, m is the number of users, and K is the number of nearest neighbors. Because of the sparsity of R, T and IT, $|R|$, $|T|$ and $|IT|$ are much smaller than the matrix cardinality. The computational complexities of calculating the gradients $\partial L / \partial b_u$, $\partial L / \partial b_i$, $\partial L / \partial p_u$, $\partial L / \partial q_i$, $\partial L / \partial y_i$, $\partial L / \partial w_v$, $\partial L / \partial x_z$ are

TABLE 1. Algorithm of the proposed method.

```
Algorithm 1: Learning algorithm for ReHI.
Input: rating matrix R, trust matrix T, the dimensionality of feature space
d, learning rate η, the maximum number of indirect trusted friends L, the
number of nearest neighbors K, trust importance parameter α,
regularization parameters β, λ, λ_t.
Output: P, W, Q, Y, X, B_U, B_I
for u = 1:m do
    for each v ∈ T+_(u) do
      calculate inf(u,v)
      calculate t_{u,v}
    end for
end for
for u = 1:m do
    select B+_(u)
    for each z ∈ B+_(u) do
      IT(u,z) = 1
   end for
end for
for u = 1:m do
    select N(u)
    for each v ∈ N(u) do
      calculate i_{u,v}
    end for
end for
while not convergence do
   calculate ∂L/∂b_u and update b_u ← b_u − η · ∂L/∂b_u
   calculate ∂L/∂b_i and update b_i ← b_i − η · ∂L/∂b_i
   calculate ∂L/∂p_u and update p_u ← p_u − η · ∂L/∂p_u
   calculate ∂L/∂q_i and update q_i ← q_i − η · ∂L/∂q_i
   calculate ∂L/∂y_i and update y_i ← y_i − η · ∂L/∂y_i
   calculate ∂L/∂w_v and update w_v ← w_v − η · ∂L/∂w_v
   calculate ∂L/∂x_z and update x_z ← x_z − η · ∂L/∂x_z
end while
```

$O(d\,|R|)$, $O(d\,|R|)$, $O(d\,|R|+d\,|T|+mKd)$, $O(d\,|R|+d\,|T|+d\,|IT|)$, $O(d\,|R|f)$, $(d\,|R|\,t^{+}+d\,|T|\,t^{+})$, $O(d\,|R|\,z^{+})$, where f is the average number of ratings received by an item, t^{+} and z^{+} are the average numbers of direct trusted and indirect trusted friends, respectively. Therefore, the overall computational complexity in one iteration is $O(d\,|R|\,h + d\,|T|\,h + d\,|IT|\,h + mKd)$, where $h = max(f, t^{+}, z^{+})$, $h \ll |R|$, $|T|$, $|IT|$. Since K is always small, $|R|$ is much larger than mK. The overall computational complexity of our method is linear with the numbers of ratings, direct trust and indirect trust relationships. Therefore, ReHI is scalable for large-scale datasets.

IV. EXPERIMENTAL ANALYSIS

In this section, we conduct several experiments to investigate the performance of ReHI in comparison with other state-of-the-art recommendation models.

A. DATASET DESCRIPTIONS

We use two real-world datasets, i.e., Epinions and Ciao, to validate the proposed algorithm. These two datasets contain both item ratings given by users and trust relationships between users. The ratings are expressed by integers from 1 to 5 in both datasets. The statistics of the two datasets are presented in Table 1. The densities of ratings and trust relationships are calculated by $|R|\,\big/(m \times n)$ and $|T|\,\big/(m \times m)$. Table 1 shows that rating data and trust relationship data are both quite sparse in the two datasets.

TABLE 2. Statistics of datasets.

	Epinions	Ciao
# of users	7411	7267
# of items	8728	11211
# of ratings	276,116	149,147
Rating density	0.0043	0.0018
# of trust relationships	52,982	110,755
Trust relationship density	0.00096	0.0021

We split ratings into the training set for learning the model and the test set for validation. Five-fold validation is used in experiments and average results are reported.

B. METRICS AND COMPARISONS

As the task of our proposed model is rating prediction, we choose mean absolute error (MAE) and root mean square error (RMSE) to measure the rating prediction accuracy, which are defined in Eqs. (24) and (25):

$$MAE = \frac{1}{|R_t|} \sum\nolimits_{u,i} \left|\hat{r}_{u,i} - r_{u,i}\right| \tag{24}$$

$$RMSE = \sqrt{\frac{1}{|R_t|} \sum\nolimits_{u,i} (\hat{r}_{u,i} - r_{u,i})^2} \tag{25}$$

where $r_{u,i}$ denotes the rating that user u has given to item i, $\hat{r}_{u,i}$ denotes the predicted rating, and R_t denotes the number of test ratings. From above definitions, we can see that a smaller MAE or RMSE value means more accurate rating prediction.

To comparatively evaluate the performance of ReHI, we select the following representative models as comparison methods, including RSTE [9], SocialMF [30], SoReg [10], SVD++ [31], TrustSVD [12] and TrustMF [33].

C. PARAMETER SETTINGS

For ReHI and all the comparison methods, we select optimal parameters for both datasets. To have a fair comparison, we set the dimensionality of all the feature spaces as 20 for all recommendation models. For the Epinions dataset, through cross-validation, we set $\alpha = 0.6$, $\lambda_t = 0.1$, $\beta = 1$, $\lambda = 0.01$, $L = 15$ and $K = 10$. For the Ciao dataset, we set $\alpha = 0.9$, $\lambda_t = 0.01$, $\beta = 1$, $\lambda = 0.1$, $L = 20$ and $K = 10$. We also set the learning rate $\eta = 0.0013$ for both datasets. The optimal parameters are also set for all comparison models.

D. EXPERIMENTAL RESULTS

We validate the performance of ReHI and all the comparison methods on all users and cold-start users. Cold-start users are the users that have rated no more than 10 items. We randomly choose 80% as training set and the remaining 20% as test set. The experimental results are shown in Table 2. The percentages in Table 2 are the improvements of ReHI over the corresponding methods.

TABLE 3. Performance comparisons (MAE and RMSE).

Datasets	Users	Metrics	RSTE	SocialMF	SoReg	SVD++	TrustSVD	TrustMF	ReHI
Epinions	All users	MAE	0.8564	0.8651	0.8232	0.7982	0.8039	0.8890	**0.7877**
		Improve	**8.02%**	**8.95%**	**4.31%**	**1.32%**	**2.02%**	**11.39%**	
		RMSE	1.1475	1.1903	1.0655	1.0414	1.0337	1.1308	**1.0283**
		Improve	**10.39%**	**13.61%**	**3.49%**	**1.26%**	**0.52%**	**9.06%**	
	Cold-start users	MAE	0.8807	0.8953	0.8705	0.8752	0.8688	0.9790	**0.8340**
		Improve	**5.30%**	**6.85%**	**4.19%**	**4.71%**	**4.01%**	**14.81%**	
		RMSE	1.2568	1.3113	1.1372	1.1345	1.1225	1.2606	**1.0930**
		Improve	**13.03%**	**16.65%**	**3.89%**	**3.66%**	**2.63%**	**13.30%**	
Ciao	All users	MAE	0.7786	0.7858	0.7491	0.7411	0.7382	0.7913	**0.7210**
		Improve	**7.40%**	**8.25%**	**3.75%**	**2.71%**	**2.33%**	**8.88%**	
		RMSE	1.0859	1.1230	0.9904	0.9785	0.9652	1.1448	**0.9602**
		Improve	**11.58%**	**14.50%**	**3.05%**	**1.87%**	**0.52%**	**16.13%**	
	Cold-start users	MAE	0.7534	0.7668	0.7550	0.7460	0.7414	0.7686	**0.7200**
		Improve	**4.43%**	**6.10%**	**4.64%**	**3.49%**	**2.89%**	**6.32%**	
		RMSE	1.1370	1.1738	1.0113	0.9749	0.9674	1.1843	**0.9495**
		Improve	**16.49%**	**19.11%**	**6.11%**	**2.61%**	**1.85%**	**19.83%**	

From Table 2, the rating prediction accuracy of ReHI is better than all the comparison counterparts both for all users and cold-start users. Especially for cold-start users, ReHI decreases MAE by 4.01% in contrast to TrustSVD which performs the best among state-of-the-art models. The results indicate that the exploitation of heterogeneous influence and indirect social relationships is effective for improving the recommendation accuracy.

From Eq. (16), the parameter α represents the importance of the direct trustees' implicit feedbacks. To observe the effect of different values of α on rating prediction and determine the optimal value of α, we fix other parameters at the optimal values, and then adjust the value of α to observe the variations of MAE. For each dataset, we select 70% and 80% of data as training set, respectively, and evaluate the performance. The parameters are the same for different training ratios. The results are shown in Figure 2. We observe that the variations of MAE are similar for different training ratios in both datasets. When $\alpha \in [0.1, 0.6]$, the value of MAE for Epinions keeps relatively stable in spite of α, indicating that even a small weight of direct trust takes effect in the modelling of users' preferences, but increasing the importance of direct trusted friends cannot lead to more accurate recommendation. When $\alpha > 0.6$, the value of MAE increases quickly with the increase of α. In Eq. (10), large α means that the implicit feedback of indirect trust plays a small part in recommendation, and direct trust predominates. Therefore, the recommendation performance degrades as a result of sparse direct trust links. For Ciao, the transition point of α is 0.9 despite training ratios.

Another important parameter is λ_t which represents the importance of explicit influence of direct trust. We adjust the value of λ_t while fixing other parameters as the optimal values, and observe the effect of λ_t on recommendation performance. The results are shown in Figure 3. For Epinions, when λ_t is set at 10^{-1}, we get the minimum value of MAE. For $\lambda_t > 10^{-1}$, as the value of λ_t increases, MAE becomes larger rapidly. The reason for the worse recommendation performance is that a too large value of λ_t makes the observed rating information play a very weak role in recommendation, leading to over-fitting. A similar trend appears in the Ciao dataset. Therefore, we set $\lambda_t = 10^{-1}$ and $\lambda_t = 10^{-2}$ for the Epinions dataset and Ciao dataset, respectively.

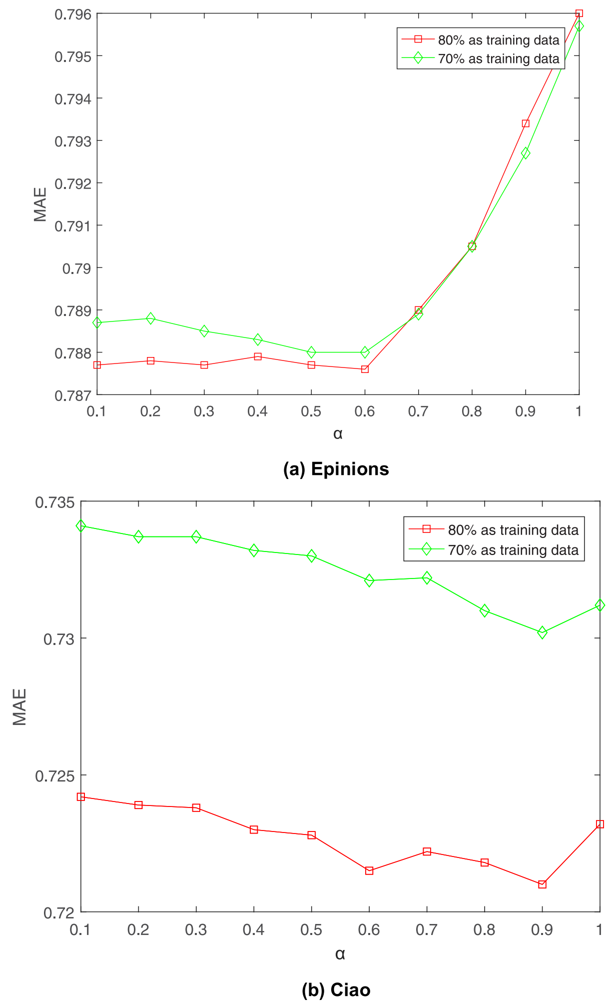


FIGURE 2. Effect of parameter α in different datasets.

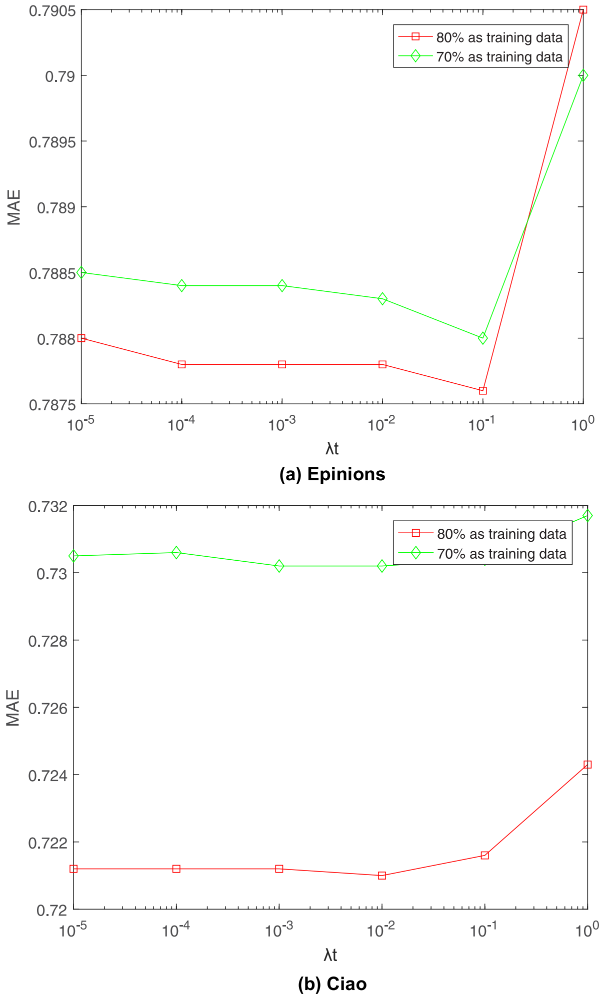


FIGURE 3. **Effect of parameter λ_t on recommendation performance.**

V. CONCLUSION

Previous social recommendation models only use direct social relationships to infer users' preferences. However, direct social data are very sparse and may not help to improve recommendation. Many recommendation methods uniformly consider social links and cannot mine more adequate influence on users' preferences. We addressed these problems by utilizing heterogeneous influence and social relationships. Firstly, we explored the influence strength between connected users according to users' reputation. Then, we mined indirect trusted friends for each active user through trust propagation and aggregation. We considered the influence of indirect trusted friends in recommendation, and proposed a social recommendation model ReHI by combining this type of influence. Experiment results from two real-world datasets demonstrated that the proposed method performs better than state-of-the-art recommendation models, especially for cold-start users.

In this paper, we concentrated on how trust relationships affect users' preferences and how to incorporate trust relationships into CF algorithms. However, distrust relationships between users are also critical in social recommender systems. Even a small number of distrust relationships have great impact on recommendation. Thus, it is worth studying how to fuse distrust relationships in social recommendation in future work.

ACKNOWLEDGMENT

(Zhaoyi Li and Fei Xiong are co-first authors.)

REFERENCES

[1] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," *Commun. ACM*, vol. 35, no. 12, pp. 61–70, Dec. 1992.

[2] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms," in *Proc. 10th Int. Conf. World Wide Web (WWW)*, 2001, pp. 285–295.

[3] Z. Huang, H. Chen, and D. Zeng, "Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 116–142, Jan. 2004.

[4] Y. Chen, C. Wu, M. Xie, and X. Guo, "Solving the sparsity problem in recommender systems using association retrieval," *J. Comput.*, vol. 6, no. 9, pp. 1896–1902, Aug. 2011, doi: 10.4304/jcp.6.9.1896-1902.

[5] X. N. Lam, T. Vu, T. D. Le, and A. D. Duong, "Addressing cold-start problem in recommendation systems," in *Proc. 2nd Int. Conf. Ubiquitous Inf. Manage. Commun.*, New York, NY, USA, 2008, pp. 208–211.

[6] B. Lika, K. Kolomvatsos, and S. Hadjiefthymiades, "Facing the cold start problem in recommender systems," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 2065–2073, Mar. 2014.

[7] P. Massa and P. Avesani, "Trust-aware recommender systems," in *Proc. ACM Conf. Rec. Syst.*, Minneapolis, MN, USA, 2007, pp. 17–24.

[8] H. Ma, H. Yang, M. R. Lyu, and I. King, "SoRec: Social recommendation using probabilistic matrix factorization," in *Proc. 17th ACM Conf. Inf. Knowl. Mining (CIKM)*, Napa Valley, CA, USA, 2008, p. 931.

[9] H. Ma, I. King, and M. R. Lyu, "Learning to recommend with social trust ensemble," in *Proc. 32nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, Boston, MA, USA, 2009, p. 203.

[10] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, "Recommender systems with social regularization," in *Proc. 4th ACM Int. Conf. Web Search Data Mining (WSDM)*, Hong Kong, 2011, p. 287.

[11] J. Tang, X. Hu, H. Gao, and H. Liu, "Exploiting local and global social context for recommendation," in *Proc. Int. Joint Conf. Artif. Intell.*, Beijing, China, 2013, pp. 2712–2718.

[12] G. Guo, J. Zhang, and N. Yorke-Smith, "A novel recommendation model regularized with user trust and item ratings," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 7, pp. 1607–1620, Jul. 2016.

[13] J. Lee, G. Noh, H. Oh, and C.-K. Kim, "Trustor clustering with an improved recommender system based on social relationships," *Inf. Syst.*, vol. 77, pp. 118–128, Sep. 2018.

[14] Z. Li, F. Xiong, X. Wang, H. Chen, and X. Xiong, "Topological influence-aware recommendation on social networks," *Complexity*, vol. 2019, 2019, Art. no. 6325654.

[15] M. Mcpherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, no. 1, pp. 415–444, Aug. 2001.

[16] M. Ye, X. Liu, and W. C. Lee, "Exploring social influence for recommendation: A generative model approach," in *Proc. 35th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, New York, NY, USA, 2012, pp. 671–680.

[17] L. A. G. Camacho and S. N. Alves-Souza, "Social network data to alleviate cold-start in recommender system: A systematic review," *Inf. Process. Manage.*, vol. 54, no. 4, pp. 529–544, Jul. 2018.

[18] D. W. Oard and J. Kim, "Implicit feedback for recommender systems," in *Proc. AAAI Workshop Rec. Syst.*, vol. 83, 1998, pp. 1–3.

[19] D. Kelly and J. Teevan, "Implicit feedback for inferring user preference: A bibliography," in *Proc. ACM SIGIR Forum*, New York, NY, USA, 2003, pp. 18–28.

[20] C. Davidsson and S. Moritz, "Utilizing implicit feedback and context to recommend mobile applications from first use," in *Proc. Workshop Context-Awareness Retr. Rec.*, New York, NY, USA, 2011, pp. 19–22.

[21] L. Guo, J. Ma, Z. Chen, and H. Zhong, "Learning to recommend with social contextual information from implicit feedback," *Soft Comput.*, vol. 19, no. 5, pp. 1351–1362, May 2015.

[22] R. He and J. McAuley, "VBPR: Visual Bayesian personalized ranking from implicit feedback," in *Proc. 30th AAAI Conf. Artif. Intell.*, 2016, pp. 1–7.

[23] S. Sidana, C. Laclau, and M.-R. Amini, "Learning to recommend diverse items over implicit feedback on PANDOR," in *Proc. 12th ACM Conf. Rec. Syst. (RecSys)*, 2018, pp. 427–431.

[24] F. Xiong, W. Shen, H. Chen, S. Pan, X. Wang, and Z. Yan, "Exploiting implicit influence from information propagation for social recommendation," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/tcyb.2019.2939390.

[25] M. Mao, J. Lu, G. Zhang, and J. Zhang, "Multirelational social recommendations via multigraph ranking," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4049–4061, Dec. 2017.

[26] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2007, pp. 1257–1264.

[27] Q. Yuan, L. Chen, and S. Zhao, "Factorization vs. regularization: Fusing heterogeneous social relationships in top-n recommendation," in *Proc. ACM Conf. Rec. Syst.*, New York, NY, USA, 2011, pp. 245–252.

[28] J. Golbeck and J. Hendler, "FilmTrust: Movie recommendations using trust in Web-based social networks," in *Proc. IEEE Consum. Commun. Netw. Conf.*, Las Vegas, NV, USA, USA, Jan. 2006, pp. 282–286.

[29] X. C. Chen, R. J. Liu, and H. You Chang, "Research of collaborative filtering recommendation algorithm based on trust propagation model," in *Proc. Int. Conf. Comput. Appl. Syst. Model. (ICCASM)*, Taiyuan, China, 2010, pp. 22–24.

[30] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proc. ACM Conf. Rec. Syst.*, New York, NY, USA, 2010, pp. 135–142.

[31] Y. Koren, "Factorization meets the neighborhood: A multifaceted collaborative filtering model," in *Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2008, pp. 426–434.

[32] J. Golbeck and A. Mannes, "Using trust and provenance for content filtering on the semantic Web," in *Proc. MTW*, Edinburgh, U.K., 2006, pp. 3–4.

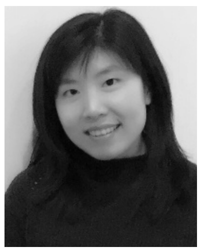
[33] B. Yang, Y. Lei, J. Liu, and W. Li, "Social collaborative filtering by trust," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 8, pp. 1633–1647, Aug. 2017.

ZHAOYI LI received the master's degree in communication and information systems from Beijing Jiaotong University, Beijing, China, in 2019. He is currently an Assistant Engineer with the 54th Research Institute of China Electronics Technology Group. His current research interests include recommender systems and complex networks.

FEI XIONG received the B.E. and the Ph.D. degrees in communication and information systems from Beijing Jiaotong University, Beijing, China, in 2007 and 2013, respectively From 2011 to 2012, he was a Visiting Scholar with Carnegie Mellon University. He is currently an Associate Professor with the School of Electronic and Information Engineering, Beijing Jiaotong University. He has published over 70 articles in refereed journals and conference proceedings. He was a recipient of the National Natural Science Foundations of China and several other research grants. His current research interests include the areas of web mining, complex networks, and complex systems.

XIMENG WANG received the B.E. degree in communication engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2011, and the M.E. degree in software engineering from Beijing Jiaotong University, Beijing, China, in 2013, where he is currently pursuing the Ph.D. degree. Since 2017, he has been a joint Ph.D. student with the University of Technology Sydney. His current research interests include recommender systems, complex networks, and data mining.

ZHE GUAN received the M.E. degree in information network and security from Beijing Jiaotong University, Beijing, China, in 2009. From 2009 to 2011, he was an Assistant Engineer with the China Academy of Electronic Science. He is currently an Engineer with the China Electronic Tec hnology Group Company, Ltd. (CETC).



HONGSHU CHEN received the Ph.D. degrees in management science and engineering from the Beijing Institute of Technology, Beijing, China, in 2015, and in software engineering from the University of Technology Sydney, NSW, Australia, in 2016. She is currently a Lecturer with the School of Management and Economics, Beijing Institute of Technology. Her research interests include bibliometrics, information systems, and text analytics.