Received July 24, 2017, accepted August 8, 2017, date of publication August 15, 2017, date of current version September 6, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2739752

Social Recommendation With Multiple Influence From Direct User Interactions

CHUNFANG LI AND FEI XIONG

School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

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

This work was supported in part by the National Natural Science Foundation of China under Grant 61401015 and in part by the Fundamental Research Funds for the Central Universities under Grant 2017JBM013.

ABSTRACT With the sustained development of social networks, increasing attention has been paid to social recommender systems. Current studies usually focus on indirect factors such as the similarity between users, but multiple direct interactions, such as mentions, reposts, and comments, are seldom considered. This paper addresses direct connections between users in social recommender systems. We analyze direct interactions to investigate the connection strength between users, and then, user preferences and item characteristics can be better described. Based on the analysis of social influence between users and users' influence over the whole social network, we propose a recommendation method with social influence, which makes full use of information among users in social networks and introduces the mechanisms of macroscopic and microscopic influences. Direct interactions between users are incorporated into a matrix factorization objective function. Real-world microblog data are applied to verify our model, and the results show that the proposed recommendation method outperforms other state-of-the-art recommendation algorithms.

INDEX TERMS Recommender systems, social network, social influence, user interactions.

I. INTRODUCTION

Recommender systems are drawing more attention these days, and the increasing popularity of social networks further accelerates this process, in particular in the big data age [1]. Mass information produced by online users creates new opportunities to help researchers and developers better understand user preferences and provide more accurate recommendations [2].

The successful applications of recommender systems can be found in a number of commercial websites, e.g., Amazon and eBay. One emergent topic raised here is social recommender systems, with the hypothesis that popular items that are adopted by a user's trusted friends can be recommended to the user.

Current studies on social recommender systems focus on the improvement of the accuracy of recommendations, in which, based on user relationship networks, indirect or implicit user connections are emphasized rather than the direct connections between users [3]. Some other algorithms [4], [5] use the trust links between users to improve the accuracy, while the algorithm with social dimensions [6], [7] digs implicit links by detecting the communities, where users share similar interests. Despite tremendous efforts, some situations are still ignored in previous studies. For instance, direct interactions, such as mentions, reposts and comments, have a significant impact in the real world [8]. Furthermore, previous studies on social recommendation simply regard users with links or high similarities as friends, which may lead to a bias evaluation for user preferences.

To solve the problems above, this paper proposes a social recommendation method that restrains invalid links by assigning different weights to different links. Invalid links mean that users recently have no direct interaction with friends. Simultaneously, this paper presents two definitions of influence in social networks, i.e., microscopic social influence (MISI) and macroscopic social influence (MASI). MISI finds which link is valid and assigns high weights for valid links. MASI denotes the strength of a user's influence over the whole social network. The method aims to improve the accuracy of recommendation in those networks with direct user interactions, such as Facebook, Twitter and LinkedIn. In addition to these websites, our method can work well in the websites where users link their accounts to some other active social networks, and its extensions might also adapt to co-author networks or bibliometric networks in broader disciplines [9].

This paper verifies the proposed method by using a public dataset, i.e., KDD Cup 2012 [10], in which there are many kinds of interaction data collected from Tencent microblog (a well-known microblog service in China). Our method generates MISI and MASI, and improves the performance of recommendations based on these two kinds of social influence. Considering this dataset only contains two recommendation results: "1" for acceptance and "-1" for rejection, we utilize the vector space similarity (VSS) [4], [11] to measure the similarities between users. At the same time, we randomly extracted two datasets from Tencent microblog data with different time periods to evaluate our method, in which user interactions are independent.

The contributions of this paper include: (1) We investigate and present two kinds of social influence, i.e., MISI and MASI. (2) A social recommendation method is designed, which addresses direction connections between users, and incorporates macroscopic and microscopic influence. (3) We compare our method with two baseline approaches, i.e., average user rating (UserAvg) and average item rating (ItemAvg), and with several state-of-the-art social recommendation algorithms. The results indicate our method performs better than other algorithms on real-world datasets.

The rest of this paper is organized as follows. In Section II, an overview of recommender systems is provided. Section III analyzes the characteristics of social networks and original data. Section IV details our social recommendation method and proposes a way to extend our method for some other networks. Experimental results are presented in Section V, followed by the conclusion in Section VI.

II. RELATED WORK

Collaborative Filtering (CF) is an important branch of recommender systems in recent studies, which can be divided into two categories: memory-based CF and model-based CF [10], [12]. Memory-based CF approaches [13], [14] use statistical techniques to build neighbourhood relationships between users or items, and predict missing rating based on a weighted sum of ratings from similar users or items. Model-based approaches [15], [16] use user-item ratings to train a model first, and then make recommendations via the model instead of the similarity comparison.

Matrix factorization [17] is a representative of model-based CF, which decomposes the $m \times n$ rating matrix R to a user factor matrix $U \in \mathbb{R}^{m\times k}$ and an item factor matrix $V \in \mathbb{R}^{n\times k}$, shown in Eq. (1).

$$R \approx UV^T \tag{1}$$

where m and n are the number of users and items, and the product is used to construct a matrix approximation. In previous studies [18], [19], minimizing the sum-of-squared-errors objective function is the most popular method to solve recommendation problems. Moreover, to avoid overfitting issues, the regularization terms of matrix U and V are added into the object function as follows:

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}\left(R_{ij} - U_i V_j^T\right)^2 + \frac{\lambda_1}{2}\|U\|_F^2 + \frac{\lambda_2}{2}\|V\|_F^2 \tag{2}$$

where I is a binary function standing for whether user i rates item j, λ_1 and λ_2 are the parameters of regularization terms, and $\|\cdot\|_F^2$ denotes the Frobenius norm [20].

Based on matrix factorization, much influential information is incorporated into recommender systems, which can relieve the cold-start problem [21]–[23]. For example, user links that come from trust networks are utilized frequently. A number of recommender systems have already been proposed based on the characteristics of trust networks. One effective algorithm is the Recommendation with Social Trust Ensemble (RSTE) [24], which is a method of probabilistic factor analysis utilizing the social trust ensemble term. RSTE is more realistic to model the real-world recommendation processes. Recently, trust-ware recommender systems focus on both online and offline trust [25]. A new trust-ware recommendation algorithm, named USBN [26], synthesizes both online and offline social trust to reinforce the performance of personal recommendation. In addition, Bayesian network probabilistic inferences are also used to deduce trusted degree or accredited degree and obtain time sequences to track the changes of user interests [27].

Social recommender systems have drawn increasing attention. Differing from trust networks, friendships in social networks are bidirectional, and the interests between friends can be different [28]. In addition, implicit interactions existing in social networks are emphasized. Some related work can be outlined: An efficient method of social recommender systems is SoReg [4], which involves both average-based regularization and individual-based regularization. SoReg adds a regularization term generated by social links to a matrix factorization framework in the model. Community detection technology [6] is also used in social recommender systems, which improves the performance of recommendation because users in the same group have similar preferences. Further, some studies consider both local information and global information and explore recommendation methods in signed social networks [29].

III. NETWORK AND DATA ANALYSIS

This paper aims to construct a recommender system approach based on user relationship networks. The characteristics of user relationships will be extracted as features to improve the performance of recommendation. To determine which feature is significant, the network and data characteristics are analyzed first.

A. NETWORK CHARACTERISTICS

Many kinds of relationships and interactions exist in real-world social networks [30]–[32]. For example, in a microblog network, a user can interact with others by three ways:

mentioning other users (@username), reposting others' posts and commenting on posts. One hypothesis here is: if a user does not interact with friends in a long time, they would hardly affect each other in making decisions. On the contrary, recent interactions with friends would have a significant impact. Therefore, we only use one-month data in the dataset and the number of mentions, comments and reposts are calculated.

In addition to the influence between individual users, a user may also affect the whole social network through information propagation. According to the result of the influence from heterogeneous social networks [33], the number of posts and followers can determine a user's influence towards the whole social network. Those two features related to either posts or followers can represent the weight of a user's interest in the network. Here, we do not take timelines and the number of posts into consideration. (1) No matter when followers contact with a user, each post of the user will be pushed to its followers' homepages and the post can be read by followers. (2) Despite the fact that the number of posts can indicate the activity of a user in the social network, this feature is not precise enough, since users may remove some posts on purpose, which will result in the dynamics of the number of posts.

In summary, we select four user features, i.e., the number of mentions, the number of reposts, the number of comments and the number of followers, to generate social influence for enhancing the accuracy of recommendation.

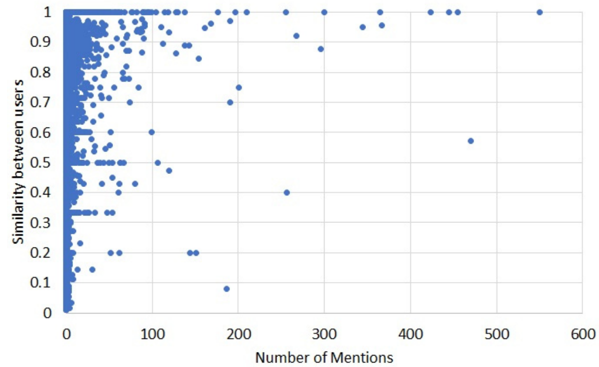


FIGURE 1. Scatter plot of similarity between users versus the number of mentions.

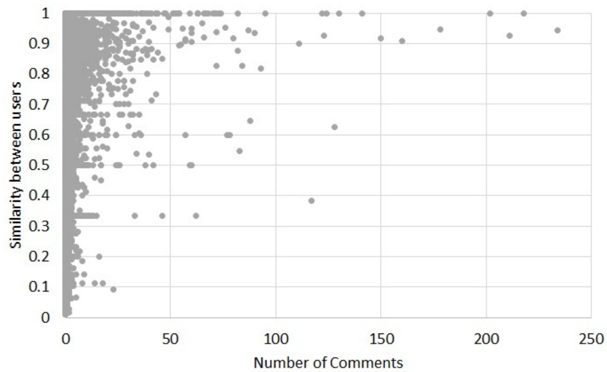


FIGURE 2. Scatter plot for similarity between users versus number of comments.

B. DATA CHARACTERISTICS

The dataset collected in the recent month contains: recommendation logs (including user-item ratings), user profile data, user action data (e.g., the number of mentions, the number of reposts and the number of comments), and user SNS data (i.e., following relationships). Here, we use the whole Tencent microblog data in KDD Cup for analysis.

Firstly, we remove repeated recommendation logs and the users that do not have any action record. Based on user-item ratings, we calculate the interest similarities between friends by VSS. Although the interactions between friends are unidirectional, similarities are bidirectional. Therefore, the similarity between couples of friends only needs to be calculated once.

To some extent, the similarity between two users shows the strength of their connection. Therefore, comparing interaction data with similarities can be efficient in data analysis. There are three kinds of interactions in user action data, including the number of mentions, number of reposts and number of comments. The following scatter plots of these interactions versus similarities obviously reflect their characteristics.

As shown in Fig. 1, the envelope of user similarity increases slower as the number of mentions increases. The envelope of scatters in Fig. 1 resembles a logarithm function. However, it is obvious that the curve of the generation function crosses the zero point, and has the maximal value of 1. Therefore, the logarithm function is not suitable to fit the envelope of the Fig. 1 and we intend to use a hyperbolic tangent function to instead it.

Fig. 2 shows user similarity versus the number of comments. Similar to Fig. 1, the envelope of the similarities in Fig. 2 is also like a hyperbolic tangent function. Moreover, less exceptional points are found in this figure, compared with Fig. 1. Thus, the influence generated by the number of comments is more convincing than that of the number of mentions.

Fig. 3 shows the relation between user similarity and the number of reposts. User similarity is distributed within a larger range. Compared with Fig. 1 and Fig. 2, the envelope of similarities does not have a clear evolutionary trend. Therefore, this feature can be omitted when calculating the influence.

In addition to interpersonal influence, a global influential feature of social recommendation is to measure how a user's interest can influence the whole social network. This kind of influence is different from those features above. The global influential feature originates from information propagation, which may be affected by the number of followers and the number of users' posts. In a common sense, the number of followers is an effective feature to judge how a user's interest is to affect the whole social network. However, the number of users' posts may not be very appropriate. For instance,

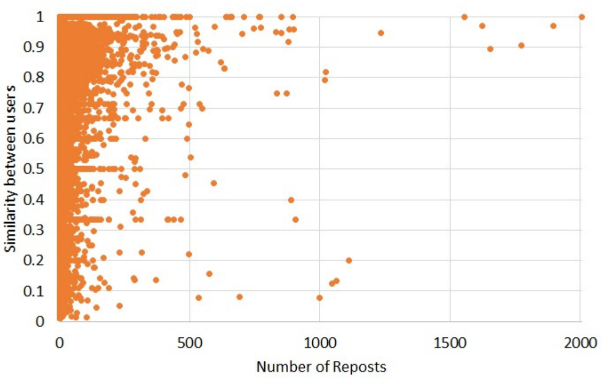


FIGURE 3. Scatter plot of similarity between users versus the number of reposts.

if a user does not publish posts in a long time, its neighbors in the social network may be influenced little by it. Therefore, we suggest using the number of followers as a global influential feature to calculate the user's global influence.

IV. THE PROPOSED RECOMMENDATION FRAMEWORK

Based on the results of the data analysis above, it is obvious that the number of mentions and the number of comments are effective to measure the influence between friends, and the number of followers contributes to generate the global influence of the whole social network. These two kinds of influence are essential for our recommendation method, which is given as follows.

A. SOCIAL INFLUENCE MODEL

Generally, different users will affect each other at different time. In majority cases, the influence between two users only takes effect in a microscopic scale [34]. We define this type of influence as microscopic social influence (MISI). In some other cases, the influence would reflect the impact on all users in social networks, which can be defined as macroscopic social influence (MASI).

1) MICROSCOPIC SOCIAL INFLUENCE

In social networks, MISI arises from a pair of user interactions, and it is often one-way influence. Given the circumstance, the impact of authoritative users is much stronger than normal users' impact. We have two features to characterize user interactions, i.e., the number of mentions and the number of comments, which are defined as NoA_{if} and NoC_{if} respectively. The variables reflect the MISI of a friend f on user i. The generation functions of MISI are listed as follows:

$$EoA_{if} = \tanh\left(NoA_{if}/\alpha_A\right) \quad (3)$$

$$EoC_{if} = \tanh\left(NoC_{if}/\alpha_C\right) \quad (4)$$

where EoA_{if} and EoC_{if} mean the impact of user f aimed at user i in terms of mentions and comments, respectively. If user i mentions and comments on f more frequently, i is affected more strongly by f, leading to larger EoA_{if} and EoC_{if}. The parameter α is used to control the nonlinearity degree of the generation functions.

According to Fig. 1 and Fig. 2, the observations include: (1) The envelope scales of two features are different. (2) The impacts of different generation functions on MISI are also different. (3) The generation functions are independent. Therefore, we use the weighted sum to calculate the MISI with EoA_{if} and EoC_{if}, namely

$$MISI_{if} = \theta_A \times EoA_{if} + \theta_C \times EoC_{if} \quad (5)$$

where the parameters θ_A and θ_C are weights for generation functions. For convenience in following computation, we use $\theta_A + \theta_C = 1$ to limit $MISI_{if}$ in the range [0, 1].

One specific issue raised in real-world social networks is that some close-connected users may have completely opposite interests [35]. To alleviate this problem, the similarity between users is involved to improve the expression of MISI, shown in Eq. (6),

$$MISI_{if} \cdot Sim_{if} \quad (6)$$

where Sim_{if} is the VSS shown in Eq. (7),

$$Sim_{if} = \frac{\sum_{v_j \in \mathcal{G}_i \cap \mathcal{G}_f} R_{ij} \cdot R_{fj}}{\sqrt{\sum_{v_j \in \mathcal{G}_i \cap \mathcal{G}_f} R_{ij}^2}\sqrt{\sum_{v_j \in \mathcal{G}_i \cap \mathcal{G}_f} R_{fj}^2}} \quad (7)$$

where $\mathcal{G}_i$ is the set of ratings from user i, and $v_j \in \mathcal{G}_i \cap \mathcal{G}_f$ stands for the set of mutual ratings from user i and its friend f.

2) MACROSCOPIC SOCIAL INFLUENCE

MASI is concerned about a user's influence over the entire social network [36]. With different activities and participants, the influence of users on a network is heterogeneous, which would lead to diverse impacts to the preferences of others.

As mentioned above, the feature to determine each user's MASI is the number of followers. We define NoF_i as the number of followers of user i. Referring to the observations in [33], the global influence increases rapidly at first, with the increase of the number of followers, but it then becomes relatively stable. Some users do not have any follower, but they can still participate in information propagation, and therefore, they can have the global influence as well. Therefore, when $NoF_i = 0$, the MASI of user i is not 0. According to the above characteristics, the expression of MASI is represented in Eq. (8),

$$MASI_i = \gamma \cdot \tanh\left(NoF_i/\alpha_F + b_F\right) \quad (8)$$

where α_F is to control the nonlinearity degree of the generation function, b_F is the offset to make the MASI stay above 0, and γ is to limit the maximal value of MASI below 1.

B. *SoInfRec* RECOMMENDATION APPROACH

With MISI and MASI defined above, a recommendation approach based on social influence (SoInfRec) is proposed. It merges two kinds of social influence to improve the performance of recommendation. Inspired by some previous studies [37], [38], minimizing the sum-of-squared-errors objective function can be used to deal with this recommendation task formulated as :

$$\min_{U,V} \mathcal{L} = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} MASI_i \cdot I_{ij} \left(R_{ij} - U_i V_j^T\right)^2 + \frac{\beta}{2} \sum_{f \in F^+(i)} Sim_{if} \cdot MISI_{if} \cdot \left\|U_i - U_f\right\|_F^2 + \frac{\lambda_1}{2} \|U\|_F^2 + \frac{\lambda_2}{2} \|V\|_F^2 \tag{9}$$

where I_{ij} is the indicator function that is equal to 1 if user i rated item j and equal to 0 otherwise. $F^+(i)$ is the friend set that can influence user i. The Frobenius norm $\|U_i - U_f\|_F^2$ is regarded as a social regularization term [4], and we take advantage of MISI to improve it. β is the parameter of the social regularization term, which is used to control the impact of the MISI regularization on the object function. If a user has large MASI, it affects the preferences of more agents, and should be emphasized in the loss function. Therefore, we use the MASI as a weight for loss function, and it has a global effect in our model. $\|U\|_F^2$ and $\|V\|_F^2$ are quadratic regularization terms of two factor matrices, aiming to avoid overfitting issues. Stochastic gradient descent (SGD) is utilized to find the optimal latent vectors U_i and V_j, shown in Eqs. (10) and (11).

$$\frac{\partial \mathcal{L}}{\partial U_i} = \sum_j MASI_i \cdot I_{ij} \cdot \left(U_i V_j^T - R_{ij}\right) \cdot V_j + \lambda_1 U_i + \beta \cdot \sum_{f \in F^+(i)} Sim_{if} \cdot MISI_{if} \cdot \left(U_i - U_f\right) + \beta \cdot \sum_{g \in F^-(i)} Sim_{gi} \cdot MISI_{gi} \cdot \left(U_i - U_g\right) \tag{10}$$

$$\frac{\partial \mathcal{L}}{\partial V_j} = \sum_i MASI_i \cdot I_{ij} \cdot \left(U_i V_j^T - R_{ij}\right) \cdot U_i + \lambda_2 \cdot V_j \tag{11}$$

where $F^-(i)$ is the friend set that can be influenced by user i.

C. TRAINING AND PREDICTION

Four steps are designed to train our proposed model:

Step 1: To generate the MISI matrix with the dataset of user action data, in which $MISI_{if}$ is calculated using Eq. (5). Note that if a user neither takes any action nor publishes any comment, the $MISI_{if}$ is set to 0. For the MASI matrix, the influence is based on the number of followers, which is given in Eq. (8).

Step 2: To calculate the similarities between pairs of friends.

Step 3: To generate the user-item rating matrix with the microblog data.

Step 4: To use SGD to find the optimal U and V of the objective function in Eq. (9).

The whole algorithm is shown in Algorithm 1.

D. EXTENSION

The proposed social recommendation approach can be used in any social networks with multiple interactions among users, such as Facebook, Twitter, and LinkedIn. The results of recommendation in a majority of social networks are binary,

Algorithm 1 The Proposed Recommendation Framework SoInfRec With Microscopic and Macroscopic Social Influence Model

Input: List of tuples log = (UserID, ItemID, ratings); List of tuples $action$ = (UserID, FriendID, NoAs, NoCs), List of tuples SNS = (UserID, FollowerID), the number of latent factors k, the learning rate μ, regularization parameter λ_1 and λ_2.
Output: $MISI$ matrix, $MASI$ matrix, user factor matrix U and item factor matrix V.

```
0: Building the dictionary from index to UserID
for i ← 1, ..., m do
  dictionary uid = < i : UserID >
end for
1: Generating MISI matrix and MASI matrix
for i ← 1, ..., m do
  for f ← 1, ..., m do
    if i ≠ f then
      if NoA_if exists then
        Calculate EoA_if according to Eq. (3)
      else then
        EoA_if = 0
      if NoC_if exists then
        Calculate EoC_if according to Eq. (4)
      else then
        EoC_if = 0
    Calculate MISI_if according to Eq. (5)
  end for
end for
for i ← 1, ..., m do
  if UserID = uid[i] exists then
    Count rows in sns with UserID = uid[i] as NoF_i
  else then
    NoF_i = 0
  Calculate MASI_i according to Eq. (8)
end for
2: Building the Similarity matrix
for i ← 1, ..., m do
  for f ← 1, ..., m do
    if i ≠ f then
      Calculate Sim_if according to Eq. (7)
  end for
end for
3: Building the Rating matrix
4: Learning user factor matrix U and item factor matrix V
Initialize U and V randomly
while Not convergence do
  Calculate ∂L/∂U according to Eq. (10)
  Calculate ∂L/∂V according to Eq. (11)
  Update U ← U − μ · ∂L/∂U
  Update V ← V − μ · ∂L/∂V
end while
```

including acceptance and rejection; however, in some networks users can assign 5-scale integer ratings (from 1 to 5) to items. For instance, a user in Epinion can give a rating to a product in a range from 1 to 5. At this stage, ratings may be affected by the rating habits of users and then lead to rating bias.

In the situation, the similarity function of VSS may not be accurate. To avoid the bias, the Pearson Correlation Coefficient (PCC) [39] is used to calculate user similarities. Comparing with VSS, PCC deducts the average rating given by each user while calculating similarities, which can avoid the impact of user rating habits.

Besides the similarity function, the generation function of MISI can be extended for different networks. User behaviors are heterogeneous, which indicate the difference of user interactions. Heterogeneous information can be used to generate the MISI as well.

V. EXPERIMENTAL ANALYSIS

In this section, we conducted experiments to compare the performance of SoInfRec with some baselines and several state-of-the-art social recommendation algorithms.

A. EXPERIMENTAL SETTINGS

We extracted two datasets over a period of time randomly from real Tencent microblog data in KDD Cup 2012 [10]. The extractive datasets have 4720 and 7834 users, respectively. Here, posts published by users in the microblog website are regarded as items. We converted user-item ratings to +1 and 0 to make all methods converge more rapidly. The rating 0 represents a user rejects the recommended item.

The basic statistics of the datasets are summarized in Table 1. The density of ratings can be calculated by

$$\frac{\text{the number of ratings}}{\text{the number of users} \times \text{the number of items}}$$

The densities of the user-item rating matrices for two datasets are 2.11%, and 3.29%, respectively. For two datasets, we randomly selected 65%, 75% and 85% data as training data, and treated the remaining data as test data to verify the methods. The random selection is carried out 5 times independently, and we report the average results.

TABLE 1. Statistics of extracted datasets.

	Dataset 1	Dataset 2
User	4720	7834
Item	4303	5910
Rating	429362	1524931
Mentions	10856	12637
Comments	14371	17940
Followers	4576	7025

We used two metrics to evaluate the performance, i.e., Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), which are defined as

$$MAE = \frac{1}{T}\sum\nolimits_{i,j}\left|R_{ij} - \hat{R}_{ij}\right| \tag{12}$$

$$RMSE = \sqrt{\frac{1}{T}\sum\nolimits_{i,j}\left(R_{ij} - \hat{R}_{ij}\right)^2} \tag{13}$$

where T is the number of ratings in the test data and $\hat{R}$ is the predicted ratings. A smaller value of MAE or RMSE means the better recommendation accuracy.

B. COMPARISONS

To demonstrate the performance of SoInfRec, some representative recommendation methods are used for comparison.

UserAvg is a baseline method, which uses the average rating given by each user as the rating of its uncollected items.

ItemAvg is another baseline method, which regards the average rating of each item as the prediction for users who have not collected the item.

MF [17] is a basic matrix factorization method which only uses the user-item matrix to make recommendations.

SocialMF [40] is model-based approach, which incorporates the mechanism of trust propagation into the model.

SoReg [4] is based on matrix factorization, and social regularization is introduced to capture the trust links between users.

RSTE [24] is a probabilistic factor analysis framework with social trust ensemble, which naturally fuses the interests of users and their trusted friends together.

C. PARAMETER SETTINGS

The two datasets have different scales and data characteristics. Thus, different regularization parameters are needed for training. The optimal λ_1 and λ_2 for dataset 1 are 0.001, and those for dataset 2 are 0.0005. We empirically set the number of latent factors at $k = 20$. The number of iterations in each recommendation method is 100.

In SoInfRec, the social regularization parameter is $\beta = 0.0002$ in both datasets. For MISI, the influence from mentions and comments uses the hyperbolic tangent function as the generation function. Parameters in the generation function are set as $\alpha_A = 100$ and $\alpha_C = 50$ for both datasets. The normalization parameters θ_A and θ_C are 0.5. For MASI, parameters in the generation function are set as: $\alpha_F = 2500$, $b_F = 1$, and $\gamma = 0.5$ for both datasets. The learning rate is $\mu = 9.5 \times 10^{-4}$.

We also set the optimal parameters for the comparative methods. The learning rates are $\mu = 4.5\times10^{-4}$ and $\mu = 0.09$ for MF and SocialMF, respectively. The social regularization parameter for SocialMF is 0.1. For RSTE, the learning rate is $\mu = 0.001$ and the parameter α is 0.7. For SoReg, the learning rate is $\mu = 4.2\times10^{-4}$ and the social regularization parameter is $\beta = 0.01$.

TABLE 2. Results of recommendation accuracy in dataset 1.

		UserAvg	ItemAvg	MF	SocialMF	SoReg	RSTE	SoInfRec
65%	MAE	0.3485	0.3408	0.3506	0.3147	0.2963	0.2692	**0.2267**
	RMSE	0.5998	0.5965	0.6210	0.5062	0.4620	0.4401	**0.4078**
75%	MAE	0.3456	0.3354	0.3695	0.3253	0.3195	0.2662	**0.2255**
	RMSE	0.5980	0.5902	0.6091	0.5001	0.4815	0.4357	**0.4008**
85%	MAE	0.3230	0.3229	0.3497	0.3357	0.3312	0.2629	**0.2241**
	RMSE	0.5783	0.5736	0.5976	0.5692	0.4881	0.4329	**0.3931**

TABLE 3. Results of recommendation accuracy in dataset 2.

		UserAvg	ItemAvg	MF	SocialMF	SoReg	RSTE	SoInfRec
65%	MAE	0.3295	0.3305	0.3119	0.3104	0.3071	0.2571	**0.2257**
	RMSE	0.5774	0.6335	0.6287	0.5257	0.4996	0.4142	**0.4115**
75%	MAE	0.3288	0.3289	0.3027	0.2983	0.2810	0.2423	**0.2197**
	RMSE	0.5763	0.6023	0.5912	0.5089	0.4974	0.4344	**0.4002**
85%	MAE	0.3285	0.3267	0.2989	0.2810	0.2745	0.2311	**0.2090**
	RMSE	0.5763	0.5983	0.5776	0.4727	0.4598	0. 3866	**0.3689**

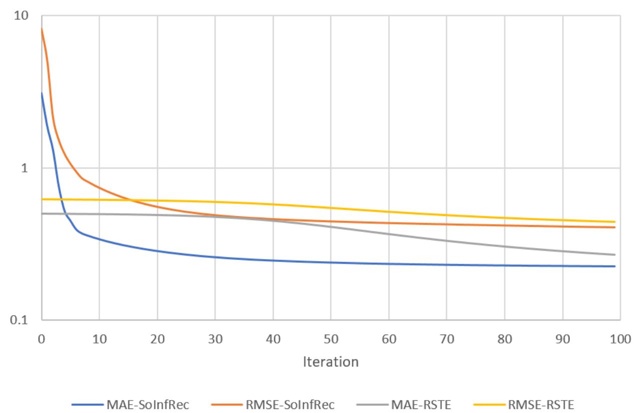


FIGURE 4. Convergence speeds in dataset 1 with 65% data for training.

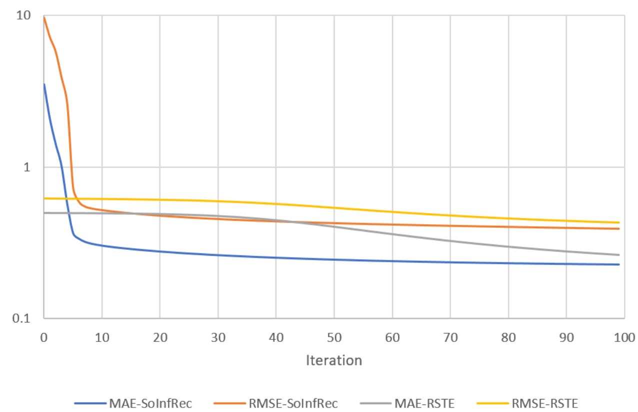


FIGURE 5. Convergence speeds in dataset 1 with 85% data for training.

D. PERFORMANCE COMPARISON

We randomly selected 65%, 75% and 85% of two datasets as training data to verify our proposed method. Results are shown in Table 2 and Table 3.

Based on the tables, two baseline approaches UserAvg and ItemAvg have similar performance, but MF cannot perform better in dataset 1. One possible reason is that the dataset from Tencent microblog has some noises, and MF cannot deal with these noises well. The noises also impact the performance of SocialMF and SoReg, and therefore, the MAE for the two methods in dataset 1 increases when the training data become denser. RSTE receives the best performance among the six comparative methods, which makes a better use of user links and enhances the recommendation with probabilistic factor analysis. According to the comparison, it is clear that SoInfRec works better than the six existing methods in both datasets. In particular, comparing to RSTE, when the proportion of training data is 85%, our method improves RMSE with 9.2% in the dataset 1, and improves RMSE with 4.6% in the dataset 2.

Aiming to compare the iteration process of the state-of-the-art methods with SoInfRec, we reveal the convergence speeds of the methods. Since RSTE performs the best among the six comparative methods, we select RSTE to compare its

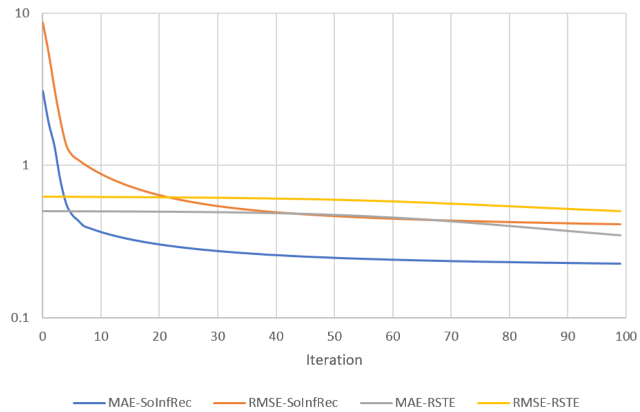


FIGURE 6. Convergence speeds in dataset 2 with 65% data for training.

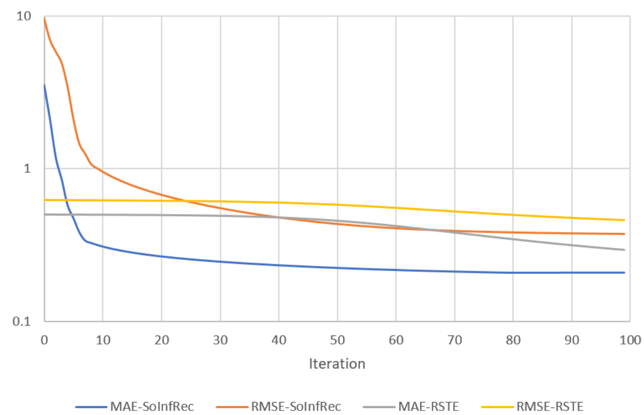


FIGURE 7. Convergence speeds in dataset 2 with 85% data for training.

convergence speed with SoInfRec in Figs. 4–7. According to those figures, SoInfRec always converges to a lower MAE and RMSE within a shorter time. MAE and RMSE for RSTE decrease relatively slow with the increase of the number of iterations. With more training data, small initial fluctuations may exist in the evolution of MAE and RMSE, but SoInfRec achieves a higher accuracy quickly.

VI. CONCLUSION

Direct interactions in social networks can result in social influence on the decision-making process of related friends, and such influence would be more significant in social networks with multiple social activities. However, previous studies on social recommendation simply regard users with links or high similarities as friends, which may lead to a bias evaluation for user preferences. This paper provides a solution to this issue by integrating direct interactions in recommender systems, and a social recommendation approach named SoInfRec is proposed, which uses multiple features in user interactions, such as the number of comments and the number of followers. We define the microscopic and macroscopic social influence to characterize the influence between users and the user's influence over the whole social network. We also introduce the microscopic and macroscopic social influence into matrix factorization to make better recommendations. The experiment results on real-world datasets demonstrate that SoInfRec actually improves the recommendation accuracy.

In future, we will continue our research on improving the performance and effectiveness of recommender systems via exploiting social relations and time information [41]. We hope this paper will inspire readers in this significant direction.

REFERENCES

[1] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey," *Decision Support Syst.*, vol. 74, pp. 12–32, Jun. 2015.

[2] L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, and T. Zhou, "Recommender systems," *Phys. Rep.*, vol. 519, no. 1, pp. 1–49, 2012.

[3] X. Wang, Y. Liu, and F. Xiong, "Improved personalized recommendation based on a similarity network," *Phys. A, Statist. Mech. Appl.*, vol. 456, pp. 271–280, Aug. 2016.

[4] 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*, 2011, pp. 287–296.

[5] G. Guo, J. Zhang, F. Zhu, and X. Wang, "Factored similarity models with social trust for top-*N* item recommendation," *Knowl.-Based Syst.*, vol. 122, pp. 17–25, Apr. 2017.

[6] J. Tang *et al.*, "Recommendation with social dimensions," in *Proc. 30th AAAI Conf. Artif. Intell.*, 2016, pp. 251–257.

[7] H. Wang, J. Wu, S. Pan, P. Zhang, and L. Chen, "Towards large-scale social networks with online diffusion provenance detection," *Comput. Netw.*, vol. 114, pp. 154–166, Feb. 2017.

[8] N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in *Proc. 20th Int. Conf. Mach. Learn.*, 2003, pp. 720–727.

[9] Y. Zhang, G. Zhang, D. Zhu, and J. Lu, "Scientific evolutionary pathways: Identifying and visualizing relationships for scientific topics," *J. Assoc. Inf. Sci. Technol.*, vol. 68, no. 8, pp. 1925–1939, Aug. 2017.

[10] K. Ji and H. Shen, "Addressing cold-start: Scalable recommendation with tags and keywords," *Knowl.-Based Syst.*, vol. 83, pp. 42–50, Jul. 2015.

[11] Y. Zhang *et al.*, "Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research," *Technol. Forecast. Soc. Chang.*, vol. 105, pp. 179–191, Apr. 2016.

[12] G. Zhao, X. Qian, and X. Xie, "User-service rating prediction by exploring social users' rating behaviors," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 496–506, Mar. 2016.

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

[14] G. Karypis, "Evaluation of item-based top-*N* recommendation algorithms," in *Proc. 10th Int. Conf. Inf. Knowl. Manage.*, 2001, pp. 247–254.

[15] J. D. M. Rennie and N. Srebro, "Fast maximum margin matrix factorization for collaborative prediction," in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 713–719.

[16] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. 20th Int. Conf. Neural Inf. Process. Syst.*, 2007, pp. 1257–1264.

[17] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[18] W. Wang, G. Zhang, and J. Lu, "Member contribution-based group recommender system," *Decision Support Syst.*, vol. 87, pp. 80–93, Jul. 2016.

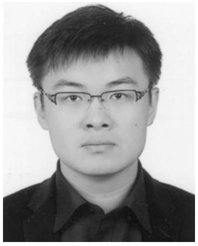
[19] X. Wang, W. Lu, M. Ester, C. Wang, and C. Chen, "Social recommendation with strong and weak ties," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, 2016, pp. 5–14.

[20] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Adv. Large Margin Classifiers*, vol. 10, no. 3, pp. 61–74, 1999.

[21] Z. Miao, J. Yan, K. Chen, X. Yang, H. Zha, and W. Zhang, "Joint prediction of rating and popularity for cold-start item by sentinel user selection," *IEEE Access*, vol. 4, pp. 8500–8513, 2016.

[22] G. Guo, J. Zhang, and D. Thalmann, "Merging trust in collaborative filtering to alleviate data sparsity and cold start," *Knowl.-Based Syst.*, vol. 57, pp. 57–68, Dec. 2014.

[23] S. Deng, L. Huang, G. Xu, X. Wu, and Z. Wu, ''On deep learning for trust-aware recommendations in social networks,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 5, pp. 1164–1177, May 2017.
[24] 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.*, 2009, pp. 203–210.
[25] H. Ma, ''On measuring social friend interest similarities in recommender systems,'' in *Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2014, pp. 465–474.
[26] Y. Cheng, J. Liu, and X. Yu, ''Online social trust reinforced personalized recommendation,'' *Pers. Ubiquitous Comput.*, vol. 20, no. 3, pp. 457–467, Jun. 2016.
[27] H. Ma, H. Yang, M. R. Lyu, and I. King, ''SoRec: Social recommendation using probabilistic matrix factorization,'' in *Proc. 17th ACM Conf. Inf. Knowl. Manage.*, 2008, pp. 931–940.
[28] Q. Zhang, J. Wu, P. Zhang, G. Long, I. W. Tsang, and C. Zhang, ''Inferring latent network from cascade data for dynamic social recommendation,'' in *Proc. 16th Int. Conf. Data Mining*, Dec. 2016, pp. 669–678.
[29] J. Tang, C. Aggarwal, and H. Liu, ''Recommendations in signed social networks,'' in *Proc. 25th Int. Conf. World Wide Web*, 2016, pp. 31–40.
[30] H. Wang, P. Zhang, L. Chen, and C. Zhang, ''SocialAnalysis: A real-time query and mining system from social media data streams,'' in *Proc. Austral. Database Conf.*, 2015, pp. 318–322.
[31] H. Wang, J. Wu, C. Zhou, Z. Ji, and J. Wu, ''Mining subcascade features for cascade outbreak prediction in big networks,'' in *Proc. Int. Joint Conf. Neural Netw.*, Jul. 2016, pp. 3942–3949.
[32] H. Wang *et al.*, ''Incremental subgraph feature selection for graph classification,'' *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 1, pp. 128–142, Jan. 2017.
[33] L. Liu, J. Tang, J. Han, and S. Yang, ''Learning influence from heterogeneous social networks,'' *Data Mining Knowl. Discovery*, vol. 25, no. 3, pp. 511–544, Nov. 2012.
[34] Q. Zhang, J. Wu, H. Yang, W. Lu, G. Long, and C. Zhang, ''Global and local influence-based social recommendation,'' in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, 2016, pp. 1917–1920.
[35] J. Tang, Y. Chang, C. Aggarwal, and H. Liu, ''A survey of signed network mining in social media,'' *ACM Comput. Surv.*, vol. 49, no. 3, pp. 1–37, 2016.
[36] A. Guille, H. Hacid, C. Favre, and D. A. Zighed, ''Information diffusion in online social networks: A survey,'' *ACM Sigmod Rec.*, vol. 42, no. 2, pp. 17–28, 2013.
[37] W. Pan and Z. Ming, ''Interaction-rich transfer learning for collaborative filtering with heterogeneous user feedback,'' *IEEE Intell. Syst.*, vol. 29, no. 6, pp. 48–54, Nov./Dec. 2014.
[38] W. Pan and Q. Yang, ''Transfer learning in heterogeneous collaborative filtering domains,'' *Artif. Intell.*, vol. 197, pp. 39–55, Apr. 2013.
[39] G.-R. Xue *et al.*, ''Scalable collaborative filtering using cluster-based smoothing,'' in *Proc. 28th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2005, pp. 114–121.
[40] M. Jamali and M. Ester, ''A matrix factorization technique with trust propagation for recommendation in social networks,'' in *Proc. 4th ACM Conf. Rec. Syst.*, 2010, pp. 135–142.
[41] H. Wang. *Time-Variant Graph Learning and Classification*. Accessed on 2016. [Online]. Available: http://hdl.handle.net/10453/101943



CHUNFANG LI received the B.S. degree in communication engineering from Beijing Jiaotong University, Beijing, China, in 2017. He is currently pursuing the M.S. degree in computer science with the Institute of Information Engineering, China Academy of Science, Beijing.

His currently research interest is the recommender system.

FEI XIONG received the Ph.D. degree in communication and information system from Beijing Jiaotong University, Beijing, China, in 2013. He was a Visiting Scholar with Carnegie Mellon University from 2011 to 2012. He is currently an Associate Professor with the School of Electronic and Information Engineering, Beijing Jiaotong University.

His current research interests include the areas of web mining, complex networks, and complex systems.