Leveraging Multisource Information in Matrix Factorization for Social Collaborative Filtering

Lele Huang[1], Huifang Ma[1,2], Xiangchun He[1], Liang Chang[2]
1 College of Computer science and engineering, Northwest Normal University, Lanzhou Gansu 730070, China
2 College of Computer Science and Information Security, Guilin University of Electronic Technology
Guilin Guangxi 541004, China
Email: {nwnuhuangll, mahuifang}@yeah.net, hxc@nwnu.edu.cn, changl@guet.edu.cn

***Abstract*—Collaborative Filtering is a technique that automatically predicts the interest of an active user by collecting rating information from other similar users or items. However, existing works either leave out the user reliability or cannot capture the implicit correlation between two users who are similar but not socially connected. Besides, they also take the trust value between users is either 0 or 1, thus will degenerate the prediction accuracy. In this work, we propose a social collaborative filtering algorithm via leveraging multisource information in matrix factorization, which aims to solve the above problems via employing both rating records and users' social network information. To this end, we design a novel recommendation framework, which consists of two fundamental components: computation of user reliability and implicit interaction, designing of user-shared feature space. In particular, the latter performs matrix factorization based on user reliability. Finally, a reasonable objective function is designed to synthetically consider the multi-source information of users, such that high-quality recommendation can be produced. Experimental results demonstrate that the proposed model achieves better accuracy than other counterpart recommendation techniques.**

Keywords—collaborative filtering, multi-source information, implicit interactions, user-shared space, matrix factorization

I. INTRODUCTION

Ever since the thriving of the Web, it becomes difficult for online users to find relevant information from massive data. Recommender systems, attempting to tackle the information overload problem by suggesting to online users the information that is potentially of interest, attract more and more attention in recent years [1], [2]. One of the most successful technologies for recommender systems, called collaborative filtering (CF), is effective and widely studied for many commercial web sites [3]. However, most existing CF confronts two main challenges: data sparsity and cold start, which will greatly degenerate its performance [4], [5]. In order to overcome the weaknesses mentioned above, the recommendation algorithms that make full use of the multi-source information are the workhorse methods of choice. Matrix factorization (MF) has recently received greater exposure, mainly as an unsupervised learning method for latent variable decomposition and dimensionality reduction, because of its good characteristics of scalability and flexibility and therefore become an important basis for researchers to construct social recommendation models [6-8].

Several social matrix factorization methods based on the user-shared feature matrix have been proposed for collaborative filtering [9]. For instance, Ma et al. [10] propose a social regularization method (SoRec) by considering the constraint of social relationships. The idea is to share a common user-feature matrix factorized by ratings and by trust. Different from the construction method of SoRec, Yang et al. [11] propose a hybrid method (TrustMF) that combines both a truster model and a trustee model from the perspectives of trusters and trustees, that is, both the users who trust the active user and those who are trusted by the user will influence the user's ratings on unknown items, however, in TrustMF, trust values with either 0 or 1 are factorized to estimate ratings, while we argue that it is more reasonable to take trust degree into account in learning process, also, TrustMF cannot capture the implicit correlation between two users. Our approach builds on TrustMF, through which both the user reliability and the implicit correlation between users are involved to generate predictions. Tang et al. [12] consider both global and local trust as the contextual information in their model, where the global and local perspective reveal the reputation of a user in the whole social network and the correlations between users and their neighborhoods, respectively. However, the existing matrix factorization models for social recommendation exist the following problems: (1) the impact of user reliability is not considered; (2) the trust value always takes either 0 or 1, unable to reveal the strength of trust degrees among users; (3) They cannot capture the implicit correlation between two users who are similar but not socially connected.

Our work is to construct a matrix factorization recommendation model with user reliability and implicit correlation, integrating twofold sparse information, the conventional rating data given by users, and the social network relationship among users. As for the user implicit correlation, users' influence relationship is leveraged to evaluate the correlations between two users by modeling implicit interactions, and ultimately integrate the implicit correlation into the proposed model, which ensures that user-specific vectors can be learned from their social information even if few or no ratings are given. The corresponding graphical illustration of our model is presented in Fig. 1, from which we can see that the model

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

generally contains the computation of user reliability and implicit interaction, and the designing of user-shared feature space $\boldsymbol{Q}$ and $\boldsymbol{U}$.

Our contributions are summarized as follows.

- We utilize the reliability relationship between users to measure the recommendation accuracy.

- A more reasonable user shared feature space based on user reliability is constructed, solving the problem of neglecting the degree of trust between users of factorizing trust values with either 0 or 1.

- Users' influence relationship are exploited as two regularization terms to capture the implicit user correlation, which is an effective way to alleviate the problem of cold start.

The rest of the paper is organized as follows: Section 2 illustrates some preliminaries, including problem definition, the calculation of user reliability, and user influence relationship. Section 3 introduces the proposed social recommendation model with multisource information. Section 4 demonstrates the experiments and results. Section 5 draws the conclusion.

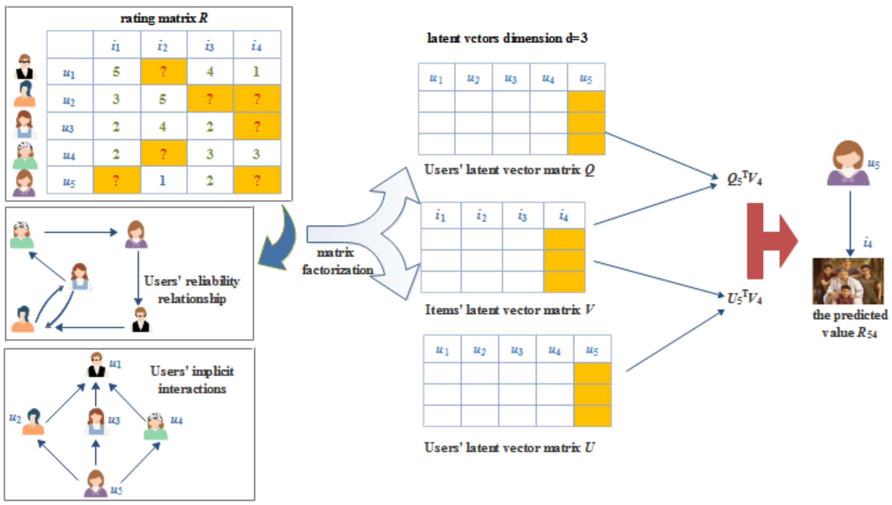

Fig.1. Framework of the proposed matrix factorization recommendation model.

Taking the rating process of user 5 for item 4 as an example

II. PRELIMINARIES

A. Problem Definition

We first introduce the notations used in this paper. Suppose that a recommender system includes a set of m users $U=\{u_1,u_2,\ldots,u_m\}$ and a set of n items $I=\{i_1,i_2,\ldots,i_n\}$. Let $\boldsymbol{R}=[R_{ij}]_{m\times n}$ denote the user-item rating matrix, where each element R_{ij} represents the rating of item j given by user i, which is generally an integer number from 1 to 5. The predicted values are given in a matrix $\boldsymbol{P}=[P_{ij}]_{m\times n}$, which are real numbers within the interval [1,5]. We define a reliability matrix $\boldsymbol{S}=[S_{ik}]_{m\times m}$, S_{ik} denotes the value of reliability among users as a real number in [0,1], where '0' indicates user i is unreliable to user k, whereas '1' indicates user i is absolutely reliable to user k. Note that the matrix $\boldsymbol{S}$ is not symmetric. $\boldsymbol{T}=[T_{ij}]_{m\times m}$ is a transition matrix for influence propagation in the social network, t_{ij} represents the propagation probability from user i to user j, if there is an edge from j to i in the network (i.e., j trusts i), then $t_{ij}>0$, otherwise, $t_{ij}=0$. We define a user influence relationship matrix $\boldsymbol{f}$, $\boldsymbol{f}_i$ is an influence vector for node i, $f_{i\to j}$ denotes the influence from user i to user j.

B. User Reliability

The reliability of a user refers to the accuracy of its recommendation to an active user, i.e., to what extent a user's recommendations to another user are accurate. For example, the more frequently the user recommends items you are favorite, the more convinced you are that the user's future recommendations are reliable[13]. Therefore, it is reasonable to evaluate the reliability by the accuracy of user recommendation.

We consider j as the unique user who recommends items to user i, CI_{ij} is the common scoring rated by users i and j. the prediction value P_{ih} for user i on item h is computed as:

$$P_{ih}=\overline{R}_i+\frac{(R_{jh}-\overline{R}_j)\times sim(u_i,u_j)}{|\,sim(u_i,u_j)\,|} \qquad (1)$$

$sim(u_i, u_j)$ denotes the rating similarity between users u_i and u_j. $\overline{R}_j$ denotes the average rating of user u_j. The recommendation ability of user j to user i on item h is calculated as:

$$pr_{ij}^{h}=1-\frac{|P_{ih}-R_{ih}|}{P_{\max}} \qquad (2)$$

$P_{\max}$ measures the maximum deviation between a predicted rating and the user's true rating. The reliability of user j to user

i on item h, i.e., the recommendation accuracy of user j to user i, is defined as:

$$S_{ij} = PR(u_i, u_j) = \frac{\sum_{h=1}^{|CI_{ij}|} pr_{ij}^h}{|CI_{ij}|} \tag{3}$$

Thus, the user reliability matrix $\boldsymbol{S}$ can be obtained.

C. Implicit Interactions

Existing approaches consider only a single link for inferring user tastes, such as TrustMF, with an inherent limitation of disregarding the implicit correlated users. Two users are similar if they affect many common users, that is, they share many out-links in the user influence relationship matrix $\boldsymbol{f}$. Similarly, two users are similar if they are affected by many common users, that is, they share many in-links in the user influence matrix $\boldsymbol{f}$. By taking into influence relationship instead of relying on one single link such as trust relationship (e.g., trust value is always 0 or 1), we can obtain more robust and accurate correlation between two users even they are not explicitly connected. The influence calculation method between users will be detailed next.

Users in social networks may have an impact on users who have direct or indirect social relationships with them, and the influence can be generated by user influence propagation. We denote the influence from user i on user j by $f_{i\to j}$, then [16]:

$$\begin{aligned} f_{i \to j} &= \frac{1}{1+\lambda}\Big(\sum_{k \in N_j} t_{kj} f_{i \to k} + v_{i \to j}\Big) \\ s.t., f_{i \to i} &= \alpha_i, \quad for \;\; j = i \; and \; \alpha_i > 0 \end{aligned} \tag{4}$$

Where $N_j = \{j_1, j_2, \ldots, j_m\}$ is j's neighbor set, t_{kj} represents the propagation probability from user k to user j. If there is an edge from j to k in network (i.e., j trusts k), then $t_{kj} > 0$, otherwise $t_{kj} = 0$. Since learning the nonzero t_{kj} is beyond the scope of this paper, we assume it is known and usually $\sum_{j=1}^{m} t_{kj} \le 1$ [14]. In this definition, each user i is assigned a constraint value α_i, which is learned from prior content or domain knowledge. Specifically, if i shows full confidence or preference to the information, this value should be the maximum (e.g., 1). In another direction, if i becomes of no interest at all, it will be 0. Parameter λ is the damping coefficient of user j for the influence propagation. It locates in range $(0, +\infty)$, and the smaller the λ is, the less the influence will be blocked.

The influence spread vector $\mathbf{f}_i = [f_{i\to 1}, f_{i \to 2}, \ldots, f_{i \to m}]^{\mathrm{T}}$ for each user i could be represented by the following equation:

$$\mathbf{f}_i = \boldsymbol{A}\mathbf{v}_i \tag{5}$$

In the equation, we denote $m \times m$ matrix $\boldsymbol{A}$ equals $(\mathbf{I} + \lambda \mathbf{I} - \mathbf{T})^{-1}$, As $\boldsymbol{v}_i$ is a vector with only $v_{i,i}$ being nonzero, the value of $v_{i,i}$ could be calculated by equation (6):

$$v_{i,i} = \frac{\alpha_i}{a_{ii}} \tag{6}$$

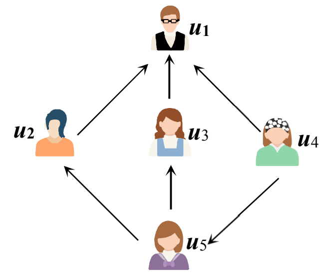

Fig. 2. Trust relationship among users

Thus, a user influence matrix $\boldsymbol{f}$ can be obtained. Fig. 2 depicts an example graph with five users, edges represent the trust relationship among users, if u_2 trusts u_1, then u_1 will generate an influence on u_2. Then the influence from u_1 on u_5 can be calculated as:

$$\begin{aligned} f_{1\to 5} &= \frac{1}{1+\lambda}\Big(\sum_{k\in N_5} t_{k5} f_{1\to k} + v_{1\to 5}\Big) \\ &= \frac{1}{1+\lambda}\big[(f_{1\to 2} \times t_{25} + f_{1\to 3} \times t_{35}) + v_{1\to 5}\big] \end{aligned}$$

III. MATRIX FACTORIZATION FUSING USER RELIABILITY AND IMPLICIT INTERACTIONS

A. Matrix Factorization of the User Reliability

We map each user i as two distinct latent feature vectors, depicted by multi-affect specific feature vector $\boldsymbol{Q}_i$ and multi-affected specific feature vector $\boldsymbol{U}_i$. $\boldsymbol{Q}_i$ and $\boldsymbol{U}_i$ characterize the behaviors of 'to affect others' and 'to be affected by others', respectively. $\boldsymbol{Q}_i$ and $\boldsymbol{U}_i$ depict 'what types of items user i recommends' and 'what types of items user i prefers' in terms of the same d latent features, respectively. Given such vectors, one can model the reliability value S_{ik} as the inner product of $\boldsymbol{Q}_i$ and $\boldsymbol{U}_k$.

1) Multi-affect Regularization: In this subsection, we model the implicit interactions between two users by incorporating out-links structure of the user influence matrix as a regularization term to constrain the objective function. Two users are similar if they affect many same users. Therefore, in order to capture the similarity between them based on the structure of out-links in matrix $\boldsymbol{f}$, we adopt the following metric:

$$C_{ij}^{Q} = \frac{\sum_{k=1}^{m} f_{i,k} \bullet f_{j,k}}{\sqrt{\sum_{k=1}^{m}\sum_{i=1}^{m} f_{i,k}^2} \bullet \sqrt{\sum_{k=1}^{m}\sum_{j=1}^{m} f_{j,k}^2}} \tag{7}$$

The superscribe $\boldsymbol{Q}$ indicates the similarity between two "multi-affect" users. With the similarity between them, multi-affect regularization is to minimize the following term:

$$\sum_{i=1}^{m}\sum_{j=1}^{m} C_{ij}^{Q} \left\| Q_i - Q_j \right\|_2^2 \tag{8}$$

A large value of C_{ij}^{Q} indicates that user i and j share many out-links, and thus we force their preference vectors should be as close as possible. A small value of C_{ij}^{Q} tells that the distance between two preference vectors should be large. By introducing the structure-based similarity, multi-affect specific preference vectors are constrained in the learning process.

2) Multi-affected Regularization: Two users are similar if they are affected by many same users. The similarity between them can be captured by exploiting the structure of in-links as follows:

$$C_{ij}^{U}=\frac{\sum_{k=1}^{m} f_{k,i}\bullet f_{k,j}}{\sqrt{\sum_{k=1}^{m}\sum_{i=1}^{m} f_{k,i}^{2}}\bullet\sqrt{\sum_{k=1}^{m}\sum_{j=1}^{m} f_{k,j}^{2}}} \tag{9}$$

Similar to multi-affect regularization, we propose a multi-affected regularization term to model the assumption as follows:

$$\sum_{i=1}^{m}\sum_{j=1}^{m} C_{ij}^{U}\left\|U_i-U_j\right\|_2^2 \tag{10}$$

B. The Multi-affect(ed) model

The ratings of users will be affected by others who have influence on them and vise versa. Here we first propose a multi-affect model to characterize the first aspect, that is, how a user i affect others to rate item j by means of $Q_i^T V_j$, which is the approximation of real score R_{ij} as well.

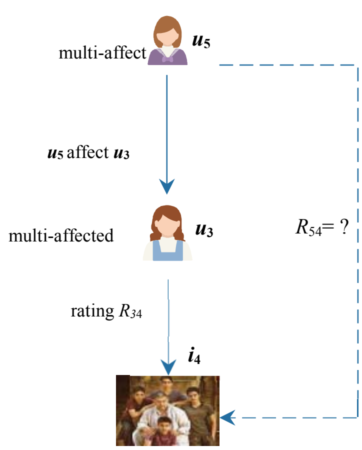

Fig. 3. Multi-affect model $Q_5{}^T V_4$: How user u_5 affect others rating on item i_4

1) Multi-affect Model: Note that *the m* users involved in rating matrix $\boldsymbol{R}$ and reliability matrix $\boldsymbol{S}$ are the same. So, we can associate $\boldsymbol{R}$ and $\boldsymbol{S}$ into one matrix factorization process by sharing a common user-specific latent space. Fig. 3 shows the proposed affect model that is able to characterize how user u_5 to rate item i_4 by means of $Q_5^T V_4$.

In this model, we choose the multi-affect specific feature matrix $\boldsymbol{Q}$ as the latent space commonly shared by $\boldsymbol{R}$ and $\boldsymbol{S}$. the vector $\boldsymbol{Q}_i$ simultaneously characterizes twofold meanings: How a user i affects others and how the same user rates items. The item-specific latent feature vector $\boldsymbol{V}_j$ depicts how an item j is rated by users. Putting together, the $Q_i{}^T V_j$ indicates how user i affects others to rate item j, which is the approximation of real score R_{ij}.

Since the reliability value S_{ik} is between 0 and 1, in order to learn the parameters in a more convenient way, we map the original rating R_{ij} into an interval [0, 1] by employing the function $f(x) = x/R_{\max}$ and adopt a logistic function $g(x) = 1/(1+exp(-x))$ suggested by[15] to bound the inner product of latent feature vectors into the interval [0, 1]. One can get the prediction by $g(Q_i{}^T V_j).R_{\max}$ after training the model. Therefore, we can learn the feature matrices $\boldsymbol{Q}$, $\boldsymbol{V}$, and $\boldsymbol{U}$ simultaneously by minimizing the following objective function:

$$\begin{aligned}L=&\sum_{(i,j)\in\Omega}(g(Q_i^T V_j)-R_{ij})^2+\sum_{(i,k)\in\Psi}(g(Q_i^T U_k)-S_{ik})^2+\\&\alpha_u\Big(\sum_{i=1}^{m}\sum_{j=1}^{m}C_{ij}^{Q}\left\|Q_i-Q_j\right\|_2^2\Big)+\\&\alpha_r\Big(\sum_i(m_{q_i}+n_{q_i})\left\|Q_i\right\|_F^2+\sum_j n_{v_j}\left\|V_j\right\|_F^2+\sum_k n_{u_k}\left\|U_k\right\|_F^2\Big)\end{aligned} \tag{11}$$

The parameter α_u is introduced to control the proportion of effect between rating preference and influence relation when training model. $\left\|\bullet\right\|_F^2$ denote the Frobenius norm, and α_r is a parameter which controls the model complexity to avoid over-fitting, where m_{qi} and n_{qi} denote the number of users who believe user i is reliable and ratings given by user i, respectively, n_{vj} denotes the numbers of ratings given to item j, n_{uk} denotes the number of items user k likes. One can minimize the above objective function by performing the following gradient descents on $\boldsymbol{Q}_i$, $\boldsymbol{V}_j$ and $\boldsymbol{U}_k$ for all users and items.

$$\begin{aligned}\frac{1}{2}\frac{\alpha L}{\alpha Q_i}=&\sum_{j\in R(i)}g'(Q_i{}^T V_j)(g(Q_i{}^T V_j)-R_{ij})V_j+\\&\sum_{k\in S(i)}g'(Q_i{}^T U_k)(g(Q_i{}^T U_k)-S_{ik})U_k+\\&\alpha_u\Big(\sum_{j=1}^{m}C_{ij}^{Q}(Q_i-Q_j)\Big)+\alpha_r(m_{q_i}+n_{q_i})Q_i\end{aligned} \tag{12}$$

$$\frac{1}{2}\frac{\alpha L}{\alpha U_k}=\sum_{i\in S^+(k)}g'(Q_i{}^T U_k)(g(Q_i{}^T U_k)-S_{ik})Q_i+\alpha_r n_{u_k}U_k \tag{13}$$

$$\frac{1}{2}\frac{\alpha L}{\alpha V_j}=\sum_{i\in R^+(j)}g'(Q_i{}^T V_j)(g(Q_i{}^T V_j)-R_{ij})Q_i+\alpha_r n_{v_j}V_j \tag{14}$$

Where $R(i)$ denotes the set of items which user i has rated, $R^+(j)$ denotes the set of users who have rated item j, $S(i)$ denotes the set of users who believe user i is reliable, $S^+(k)$ denotes the set of users that user k considers to be reliable. $g'(x)=\exp(-x)/(1+\exp(-x))^2$ is the derivative of logistic function $g(x)$.

As the model in Equation (11) uses multi-affect specific feature matrix $\boldsymbol{Q}$ as the commonly shared user latent space, we refer to the learning algorithm as Multi-affect MF, and its pseudocode is given in Algorithm1.

Algorithm1 Multi-affect MF

Input: $\boldsymbol{R}$, $\boldsymbol{S}$, $\boldsymbol{f}$, α_u, α_r, α (learning rate)

1. Initialize $\boldsymbol{Q}$, $\boldsymbol{V}$, $\boldsymbol{U}$ with small random numbers
2. while L has not converged:
3. $V_j \leftarrow V_j - \alpha.\dfrac{\partial L}{\partial V_j},\ \ j = 1,\ldots,n$
4. $Q_i \leftarrow Q_i - \alpha.\dfrac{\partial L}{\partial Q_i},\ \ i = 1,\ldots,m$
5. $U_k \leftarrow U_k - \alpha.\dfrac{\partial L}{\partial U_k},\ \ k = 1,\ldots,m$
6. end while

Output the predicted rating: $P_{ij} \leftarrow g(Q_i^T V_j).R_{\max}$

*2) Multi-**affected Model**:* Fig. 4 shows the proposed multi-affected model that is able to characterize how a user u_5 is affected by the decisions of others by means of $U_5{}^T V_4$. Different from the multi-affect model, this time we choose the multi-affected specific feature matrix $\boldsymbol{U}$ as the latent space commonly shared by $\boldsymbol{R}$ and $\boldsymbol{S}$. In the multi-affected model, the vector $\boldsymbol{U}_i$ simultaneously characterizes twofold meanings: how a user A according to the propagation of such twofold influence among users.

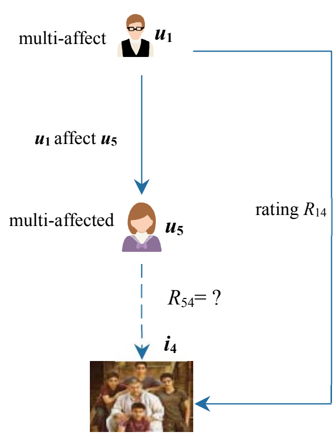

Fig. 4. Multi-affected model: How others affect u_5's rating on item i_4

After independently training multi-affect model and multi-affected model, one can obtain two sets of feature matrices. Let Q_i^t and V_j^t be the multi-affect specific vector and item-specific vector learned by algorithm multi-affect MF. Let U_i^d and V_j^d be the multi-affected specific vector and item-specific vector learned by algorithm multi-affected MF. We suggest the following synthetic strategy to generate the approximation of real rating scores.

is affected by others and how the same user rates items. Again, the item-specific latent feature vector $\boldsymbol{V}_j$ depicts how an item j is rated by users. Putting together, the $U_i^T V_j$ indicates how user i is affected by others to rate item j, which is the approximation of real score $\boldsymbol{R}_{ij}$ as well. Similarly, one can learn the feature matrices $\boldsymbol{U}$, $\boldsymbol{V}$, and $\boldsymbol{Q}$ simultaneously by minimizing the following objective.

$$
\begin{aligned}
L = & \sum_{(i,j)\in\Omega} (g(U_i^T V_j) - R_{ij})^2 + \sum_{(k,i)\in\Psi} (g(Q_k^T U_i) - S_{ki})^2 \\
& + \alpha_u \left(\sum_{i=1}^{m} \sum_{j=1}^{m} C_{ij}^U \left\| U_i - U_j \right\|_2^2\right) \\
& + \alpha_r \left(\sum_i (m_{u_i} + n_{u_i}) \left\| U_i \right\|_F^2 + \sum_j n_{v_j} \left\| V_j \right\|_F^2 + \sum_k m_{q_k} \left\| Q_k \right\|_F^2\right)
\end{aligned} \tag{15}
$$

Where m_{ui} and n_{ui} denote the number of users who user i considers to be reliable and the ratings given by user i, respectively. n_{vj} denotes the number of ratings given to item j, m_{qk} denotes the number of users who believe user k is reliable.

C. Synthetic Influence of Multi-affect and Multi-affected Model

Individuals will affect each other during the process of rating. That is to say, your decision on whether or not to rate something and how to rate them will be influenced by others in social networks. Meanwhile, your decisions will inevitably influence the choices of others. Comprehensively, it is suggested that the observed ratings are actually generated

$$
\hat{R}_{ij} = g\left(\left(\frac{Q_i^t + U_i^d}{2}\right)^T \left(\frac{V_j^t + V_j^d}{2}\right) \right) \cdot R_{\max} \tag{16}
$$

IV. Experiments and Validation

A. Description of the Data set

The data sets used in our experiments are FilmTrust and Epinions. Users in FilmTrust[1] and Epinions[2] can rate products with scores from 1 to 5 and they can also establish social relations (i.e., trust relations) with others [16]. Some statistics of these two datasets are presented in Table I. The rating data is very sparse for both FilmTrust and Epinions.

TABLE I. Statistics of Experimental Datasets

Data Set	FilmTrust	Epinions
# of Users	1641	49 290
# of Items	2 071	139 738
# of Ratings	35 497	664 824
Rating Density(%)	1.044	0.010
# of Social Relations	1 853	487 183
Social Relation Density(%)	0.069	0.020 1

B. Experimental Setup

We use the 5-fold cross-validation for learning and testing. In each time we randomly select 80% of data as training set and the rest of 20% for test. For each experiment discussed below, we conduct five times and take the mean as the final result.

1) Evaluate metrics: We adopt two well-known metrics to evaluate predictive accuracy, namely mean absolute error (MAE) and root mean square error (RMSE)[17], defined by:

$$MAE = \frac{\sum_{i,j} \left| R_{ij} - \hat{R}_{ij} \right|}{N} \tag{17}$$

$$RMSE = \sqrt{\frac{\sum_{i,j} (R_{ij} - \hat{R}_{ij})^2}{N}} \tag{18}$$

Where R_{ij} is the rating in the test set, $\hat{R}_{ij}$ is the predicted rating, and N is the number of ratings in the test set. Smaller values of MAE and RMSE indicate better predictive accuracy.

2) Comparison methods: To comparatively evaluate the performance of our proposed methods, we select six representative CF methods based on user feature space as competitors.

a) we select probabilistic matrix factorization model PMF without using social information to verify the impact of user implicit interaction in social networks on recommendation performance of our proposed model.

b) SoRec, TrustMF and LOCABAL are three social recommendation methods based on user feature matrix representation, SoDimRec and MFC are lastest recommendation methods using indirect social information.

For all comparison methods to be validated, we set respective optimal parameters according to corresponding references. The main parameters of respective methods are given in Table II. Without useful prior knowledge, we could fix α_i to be the same (e.g.,1) for each user i, we set the damping coefficient λ is 0.176 according to reference[14]. For all matrix factorization based methods, we set the dimensionality of latent space as 10 and the learning rate α as 0.001, and adopt the same initialization strategy, which randomly initialize all involved feature matrices with a uniform distribution within the interval [0,1].

C. Experimental Result

1) Validation on FilmTrust and Epinions: We now validate the performance of the proposed model and compare it with its competitors on two data sets. Table III shows that our method performs the best of all. It implies that our proposed multi-affect, multi-affected and synthetic models properly and effectively catch on the users' multisource information on the generation of observed ratings, and hence get a better promotion on the quality of recommendations.

TABLE II. PARAMETER SETTINGS OF DIFFERENT SOCIAL RECOMMENDATION METHODS

Methods	Optimal Parameters
PMF	α_r=0.001
SoRec	α_r=0.001,α_u=1
TrustMF	α_r=0.001,α_u=1
LOCABAL	α_r=0.001,α_u=0.5
SoDimRec [18]	α_r=0.001,α_u=10, α_W=100
MFC [19]	α_r=0.001,α_u=0.001
Our Method	α_r=0.001,α_u=1

2) Validation on cold start users: As mentioned in the introduction section, cold start is one big challenge faced by CF methods. We now evaluate the capabilities of addressing cold start users by respective competitors. Conventionally, those who have rated five or fewer ratings are seemed as cold start users [20]. 5-fold cross-validation is still used in the test but we only care about the accuracy of prediction for cold start users. Table IV shows that our model performs the best once again with respect to addressing cold start users, especially in comparison with PMF. It is shown that the objective function of the implicit interaction between users in the social networks designed in this paper can alleviate the cold start problem to some extent.

3) Validation on Users with Different Social Degrees: Another series of experiments are conducted to investigate the performance on users with different social degrees. The social degrees refer to the summation of the number of trusted neighbors specified by a user (i.e., out degree) and the number of trusting neighbors who trust the user (i.e., in degree). Distinct from previous validations focusing on comparing global quality of recommendations in terms of the average accuracy over all users, here we are particularly interested in testing the performance of respective social CF methods in regard to different categories of users. We split the social degrees into seven categories: 1-5, 6-10, 11-20, 21-40, 41-100, 101-500, >500. The results are illustrated as Fig. 5 and Fig. 6.

D. Summary

Major points of the experimental study are summarized as follows.

- The exploitation of improved user reliability according to rating information and implicit correlation between users in social networks are crucial for capturing users' preference. All these factors jointly lead to the superior performance of our model is in the experiments.
- It is more credible to perform matrix factorization based on user reliability (a number of 0 to 1) rather than trust value (either 0 or 1) as most existing methods do, such as TrustMF, Therefore, the rationality of our model is finally proved.

TABLE III. PERFORMANCE COMPARISONS OF DIFFERENT SOCIAL METHODS ON ALL USERS

Data Sets	**Metrics**	PMF	SoRec	TrustMF	LOCABAL	SoDimRec	MFC	Our Method
FilmTrust	MAE	0.735	0.638	0.637	0.641	0.635	0.637	**0.623**
	RMSE	0.968	0.831	0.818	0.829	0.816	0.819	**0.805**
Epinions	MAE	0.838	0.915	0.819	0.817	0.813	0.815	**0.802**
	RMSE	1.115	1.167	1.105	1.094	1.092	1.095	**1.085**

TABLE IV. PERFORMANCE COMPARISONS OF DIFFERENT SOCIAL METHODS ON COLD START USERS

Data Sets	**Metrics**	PMF	SoRec	TrustMF	LOCABAL	SoDimRec	MFC	Our Method
FilmTrust	MAE	0.813	0.672	0.664	0.659	0.651	0.660	**0.644**
	RMSE	1.011	0.901	0.892	0.884	0.852	0.895	**0.830**
Epinions	MAE	1.149	0.961	0.849	0.839	0.852	0.857	**0.827**
	RMSE	1.427	1.179	1.177	1.175	1.174	1.179	**1.166**

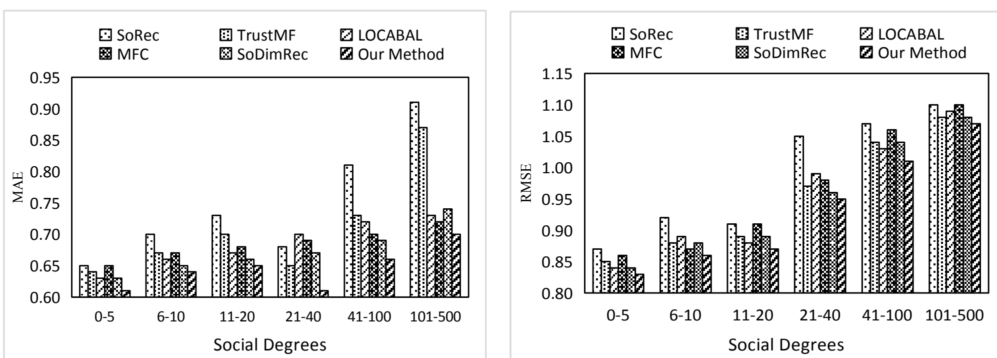

Fig.5. The predict error on users in FilmTrust dataset with different social degrees

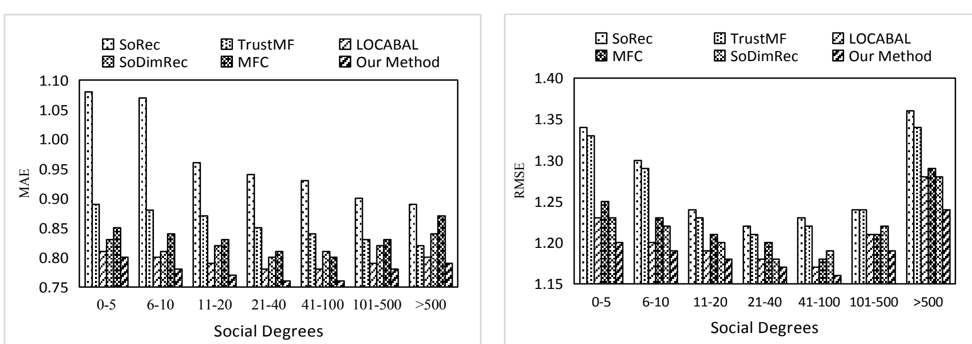

Fig. 6. The predict error on users in Epinions dataset with different social degrees

V. CONCLUSIONS

Collaborative filtering methods are facing the difficulty of data sparsity and cold start. Aiming at addressing them by utilizing additional social data more effectively, we have proposed a novel social CF method, which is motivated by the heuristic that individuals will affect each other in various aspects during the process of interacting. In properly catching on a twofold user influence on the generation of observed opinions, a multi-affect model and a multi-affected model have been proposed to map users into the same latent feature spaces but with different implications that can explicitly describe the feedback how users affect or be affected by others. Moreover, the two models are naturally synthesized to one fusing model simultaneously fitting available ratings and social relations. As has been verified that the proposed model performs better than its competitors.

ACKNOWLEDGMENT

This research was supported by National Natural Science Foundation of China (61762078, 61363058, 61966009, 61762079, U1711263, U1811264), Guangxi Key Laboratory of Trusted Software (kx202003) and Major project of young teachers' scientific research ability promotion plan (NWNU-LKQN2019-2).

REFERENCES

[1] G. W. Zhao, H.F. Ma, Z.X. Li, X. Ao, N. Li, "SBRNE: An Improved Unified Framework for Social and Behavior Recommendations with Network Embedding," DASFAA, 2019, pp. 555–571.

[2] Z. Li, H. Zhao, Q. Liu, Z. Y. Huang, T. Mei, E. H, Chen, "Learning from History and Present: Next-item Recommendation via Discriminatively Exploiting User Behaviors," KDD, 2018, pp. 1734–1743.

[3] Q.T. Wu, L. Jiang, X. F. Gao, X. C. Yang, G. H. Chen, "Feature Evolution Based Multi-Task Learning for Collaborative Filtering with Social Trust," IJCAI, 2019, pp. 3877–3883.

[4] Y. K. Gao, X. H. Wang, L. Guo, Z. M. Chen, "Learning to recommend with collaborative matrix factorization for new users. Journal of Computer Research and Development," 2017, 54(8), pp. 1813–1823.

[5] J. Chen, D. F. Lian, K. Zheng, "Improving One-Class Collaborative Filtering via Ranking-Based Implicit Regularizer," AAAI, 2019, pp. 37–44.

[6] S. Sedhain, A. K. Menon, S. Sanner, L. Xie, D. Braziunas, "Low-Rank linear cold-start recommendation from social data," In: Proceedings of the 31th AAAI Conference on Artificial Intelligence, 2017, pp. 1502–1508.

[7] A. J. B. Chaney, D. M. Blei, T. Eliassi-Rad, "A probabilistic model for using social networks in personalized item recommendation," In: Proceedings of the 9th ACM Conference on Recommender Systems, New York: ACM Press, 2015. pp. 43–50.

[8] G. Guo , J. Zhang, 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 Conf. on Artificial Intelligence. AAAI Press, 2015, pp. 123–129.

[9] H. F. Liu, L. P. Jing, J. Yu, "Survey of matrix factorization based recommendation methods by integrating social information," Journal of Software, 2018, 29(2), pp. 340-362.

[10] H. Ma, H . Yang, M. R. Lyu, I. King, "SoRec, Social recommendation using probabilistic matrix factorization," In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, 2008, pp. 931–940.

[11] B. Yang, Y. Lei, D. Y. Liu, J. M. Liu, "Social collaborative filtering by trust," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8), pp. 1633–1647.

[12] J. L. Tang, X. Hu, H. J. Gao, H. Liu, "Exploiting local and global social context for recommendation," In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence, 2013, pp. 2712–2718.

[13] R. Q. Wang, Y. L. Jiang, Y. X. Li, J. G. Lou, "A Collaborative Filtering Recommendation Algorithm Based on Multiple Social Trusts," Journal of Computer Research and Development, 2016, 53(6), pp. 1389–1399.

[14] Q. Liu, B. Xiang , N. J. Yuan, E. H. Chen, H. Xiong, Y. Zheng, et al. "An Influence Propagation View of PageRank," ACM Transactions on Knowledge Discovery from Data, 2017, 11(3), pp. 1–30.

[15] Y. Yang, E. Chen, Q. Liu, B. Xiang, T. Xu, S. Shad, "On approximation of real-world influence spread," In proceedings of European Conference on Machine Learning and Knowledge Discovery in Databases, 2012, pp. 548–564.

[16] G. Guo, J. Zhang, N. Yorke-Smith, "A novel Bayesian similarity measure for recommender systems," In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence. AAAI Press, 2013, pp. 2619–2625.

[17] T. T. Pan, F. Wen, Q. R. Liu, "Collaborative filtering recommendation algorithm based on rating matrix filling and item predictability," Acta Automatica Sinica, 2017, 43(9), pp. 1597–1606.

[18] J. L. Tang, S. H. Wang, X. Hu, D. W. Yin, Y. Z. Bi, Y. Chang, et al., "Recommendation with social dimensions," In: Proceedings of the 30th AAAI Conference on Artificial Intelligence, 2016, pp. 251–257.

[19] H. Li, D. M. Wu, W. B. Tang, N. Mamoulis, "Overlapping community regularization for rating prediction in social recommender systems," In: Proceedings of the 9th ACM Conference on Recommender Systems. New York: ACM Press, 2015, pp. 27–34.

[20] M. Jamali, M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," In Proceedings of the Fourth ACM Conference on Recommender Systems, pp. 135–142, Barcelona, Spain, September 2010. Association for Computing Machinery.

[21] H. Fang, Y. Bao , J. Zhang, "Leveraging decomposed trust in probabilistic matrix factorization for effective recommendation," In Proceedings of the 28th AAAI Conference on Artificial Intelligence, 2017, pp. 30–36.