IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 19, 2018, accepted January 7, 2019, date of publication January 14, 2019, date of current version February 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2893024

A Novel Social Recommendation Method Fusing User's Social Status and Homophily Based on Matrix Factorization Techniques

RUI CHEN[1,2], QINGYI HUA[1], BO WANG[1,3], MIN ZHENG[4], WEILI GUAN[5], XIANG JI[1], QUANLI GAO[6], AND XIANGJIE KONG[7], (Senior Member, IEEE)

[1]School of Information Science and Technology, Northwest University, Xi'an 710127, China
[2]School of Computer Science, Zhengzhou University of Aeronautics, Zhengzhou 450046, China
[3]School of Computer Science, Xi'an University of Posts & Telecommunications, Xi'an 710121, China
[4]School of Information Engineering, Engineering University of People's Armed Police Force, Xi'an 710038, China
[5]Hewlett Packard Enterprise, 918112, Singapore
[6]College of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China
[7]Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, School of Software, Dalian University of Technology, Dalian 116620, China

Corresponding authors: Qingyi Hua (nwuchenrui@126.com and huaqy@nwu.edu.cn) and Xiangjie Kong (xjkong@ieee.org)

This work was supported in part by the National Natural Science Foundation of China under Grant 61272286, in part by the joint funded projects of the Special Scientific Research Fund for Doctoral Program of Higher Education under Grant 20126101110006, in part by the Industrial Science and Technology Research Project of Shaanxi Province under Grant 2016GY-123, and in part by the Science Research Foundation of Northwest University under Grant 15NW31.

ABSTRACT As one of the most successful recommendation techniques, collaborative filtering provides a useful recommendation by associating an active user with a crowd of users who share the same interests. Although some achievements have been achieved both in theory and practice, the efficiency of recommender systems has been negatively affected by the problems of cold start and data sparsity recently. To solve the above problems, the trust relationship among users is employed into recommender systems to build a learning model to further promote the prediction quality and users' satisfaction. However, most of the existing social networks-based recommendation algorithms fail to take into account the fact that users with different levels of trust and backgrounds, that is, user's social status and homophily have different degrees of influence on their friends. In this paper, a novel social matrix factorization-based recommendation method is proposed to improve the recommendation quality by fusing user's social status and homophily. User's social status and homophily play important roles in improving the performance of recommender systems. We first build a user's trust relationship network based on user social relationships and the rating information. Then, the degree of trust is calculated through the trust propagation method and the PageRank algorithm. Finally, the trust relationship is integrated into the matrix factorization model to accurately predict unknown ratings. The proposed method is evaluated using real-life datasets including the Epinions and Douban datasets. The experimental results and comparisons demonstrate that the proposed approach is superior to the existing social networks-based recommendation algorithms.

INDEX TERMS Recommender systems, collaborative filtering, matrix factorization, PageRank algorithm, trust networks.

I. INTRODUCTION

As Internet technology and ubiquitous computing develop rapidly, the geometric growth of the information makes it more and more difficult to find interest items for users to meet their own needs from the vast amount of information in time. This leads to ''information overload'' problem. As a complementary tool for search engine, recommender system (RS) recommends items that may be of interest to users through establishing users' interest models based on analyzing users' historical behaviors. It can provide personalized recommendation services for users without requiring users to provide clear needs [1]. The research of RS began in the 1990s, and a lot of research achievements of related fields are employed into RS, such as cognitive science, information retrieval, machine learning, and data mining. Until now, RS has become a research hotspot, attracted extensive attention from the science and industry communities, and is also widely and successfully applied in industrial community such as mobile-news, e-business services, and e-tourism [2], [3].

As one of the most successful and widely used recommendation technologies, collaborative filtering (CF)-based recommendation methods can filter multimedia information that are difficult for the computer to automatically analyze content [4], such as movies, and music. In CF, products that might be of interest are suggested to customers by analyzing the ratings on those items from his/her neighbors [5].

In recent years, CF-based RS suffers from the following major problems: cold start and data sparsity [4], [5]. CF-based algorithms cannot obtain the accurate similarities between users or between items using sparse rating data, thereby affects the accuracy of CF-based RS [6]. For cold start, there is a lack of ratings on items from newly registered users and the item newly added to the system does not have any ratings from users, thus RS cannot determine the similar neighbor of the user or item, resulting in the inability to provide a personalized recommendation for the newly registered user, or the addition of new items are recommended to interested users [7], [8].

To solve the problems of cold start and data sparsity, more and more recommender systems (RSs) fuse social context information into the recommendation models to promote the prediction quality of RS. Especially, as the basis of interpersonal communication, trust relationship among users play a crucial role in solving information interaction, and experience communication. Some recommendation algorithms for trust relationship-based social networks have appeared recently. Moradi *et al.* [5] propose a novel reliability measure based on trust relationship to evaluate the recommendation performance, and the measure is introduced into trust-based CF approach to promote the prediction quality of the social relationship-based RSs. Ma *et al.* [9] present a trust relationship-based probabilistic graph algorithm, which incorporate the user's hobbies and he/her trusted friends' preferences to optimize the solution. Jamali *et al.* [10] study a recommendation method for matrix factorization (MF) of social relationship networks, and the method of social relationship propagation is employed into the proposed method to promote the recommendation accuracy. Based on research in psychology and sociology, Jiang *et al.* [11] explore various social recommendation methods and present an enhanced recommendation method which fuses the social context information including user's favors and influences between individuals into the MF model. To express the internal relationships of social networks, Zheng *et al.* [12] combine hypergraph theory with probabilistic matrix factorization technology, and propose a novel hybrid recommendation model. To establish a more accurate recommendation model based on trust relationship between users, Pan et al. [13] investigate the different roles that a user as a truster and a trusted person in a social network, and a novel social MF model based on adaptive trust relation training is proposed to exactly reflect social relationships.

However, these recommendation approaches based on social networks have the following problems: (1) It only utilizes the direct social trust relationship among users, but ignores their implicit trust relationships between users. (2) In fact, the trust relationship is sparse and the propagation of social relationships among users is not considered in the model training process. (3) Although we take into account the factor that a user's behavior might be affected by his/ her most trusted and closest friends, ignore the factor that users with different social status and backgrounds have different effects on their friends.

To solve the above problems, inspired by [9], [10], and [14]–[17], we investigate the impacts of user social status and homophily on user trust relationships in social networks and predictions in RS, and analyze users' interactive behaviors from the user's perspective to establish a recommendation model. In this study, we fuse the social status among users, homophily, and the trust propagation into probability matrix factorization (PMF) model, and propose a social recommendation approach based on the MF technology fusing comprehensive evaluation of user social status and homophily (USSHMF). In addition, we propose metrics on user's social status and homophily by using the PageRank algorithm and TF-IDF. Inspired by [9], [10], [14], [16], [22], and [42], by following the impact of users' social relationships on users' behaviors in real life, the metrics on user's social status and homophily are introduced into the MF model to optimize the user latent feature space. Meanwhile, we introduce a social regularization term to constrain user feature space and item's feature space with their similar neighbor users and items, respectively.

The rest of this paper is organized as follows. Section 2 presents the relevant social theory and some typical recommendation techniques. Next, we present a recommendation model and framework based on the MF technology fusing comprehensive evaluation of trust relationship in Section 3. We introduce the evaluation methods and present the experimental results on the real-life datasets in Section 4. The last section draws the conclusions.

II. RELATED WORK

As the online social network services are becoming more popular, RS can obtain rich information from the social networks platforms [16]. Trust relationship between users plays a crucial role in promoting the quality of RS. However, trust is a very complex concept, which might be affected by many factors. Sociological theories from the social sciences including user social status and homophily are helpful to explain the degree of influence on trust relationship between users and make a prediction in RS.

A. SOCIAL THEORY

1) SOCIAL STATUS THEORY

A user's social status represents the importance of this user in the social relationship network, indicating the extent of this user attaching to other users in the network. The relationships between users are usually expressed by a directed graph [13], [17]. Yu *et al.* [16] present an advanced MF method through considering the influence of user social status on users' trust relationships. Wang *et al.* [14] explore the prediction of trust relationship in view of sociology, and propose a novel prediction algorithm of user's trust relationship on the basis of investigating the effect of social prestige theory and homophily theory on trust relationship between users. By introducing user's social status and interest bias into the establishment of social relationship networks, Li and Ma [43] analyze a the influence of social prestige on user social relationship, and present a user's social relationship-based recommendation approach.

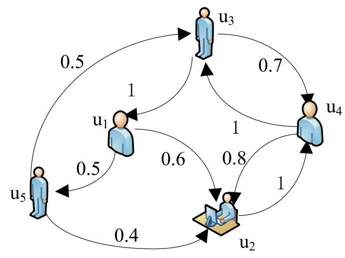
The social status theory is used to explain how the social levels of different users affect the establishment of trust relationship between users. Usually, a high-level user in the social relationship network is considered to be the authority users, and a low-level user is more likely to build trust relationships with a higher-level user. For instance, suppose that user v is an authoritative scholar in literary research, but he/her is a beginner in economics in a social network. Thus the user u is likely to accept the advice from the user v when buying a literary book, but will not accept the advice from the user v when buying an economics book.

2) HOMOPHILY THEORY

Homophily represents the tendency of individuals to relate to similar individuals. It is easier to establish a connection between users with similar characteristics in the real world. That is, users tend to interact with users who are similar to themselves in certain aspects. Homophily factors are mainly composed of the following two factors: similarities of individual characteristics and the similarities of the social environment [14], [42], [49]. Among them, the similarities of individual characteristics include race, gender, age, religion, belief, occupation, educational background, etc. The characteristics of social environment include position, social status, network position, behavior, ability, wishes, etc. Wang *et al.* [14] analyze the effect of homophily on user trust relationship, and construct a novel trust relationship prediction algorithm based on matrix tri-decomposition techniques. Tang *et al.* [42] study the impacts of homophily on trust prediction, and integrate homophily regularization into matrix factorization model to optimize user feature space. Wang *et al.* [41] investigate the impacts of social status and homophily on trust and distrust, and present a novel method of making predictions of trust and distrust relationships between users from different inducing factors through employing multilayer neural network, namely, homophily, emotion tendency, and social status.

3) TRUST METRICS

Trust networks reflect the trust relationships and the levels of trust between users in social relationship networks. Fig. 1 shows a trust network, a node of which denotes a user, and a directed arc denotes a trust relationship of one user to other user. For instance, user u_3 has a trust value of 0.7 for user u_4 and a trust value of 1 for user u_1 in Fig. 1(a), so both u_1 and u_4 are in the trust network of user u_3. Fig. 1(b) is the trust relationship matrix corresponding to the trust relationship network of Fig.1(a).



(a)

	u_1	u_2	u_3	u_4	u_5
u_1	1	0.6			0.5
u_2		1		1	
u_3	1		1	0.7	
u_4		0.8	1	1	
u_5		0.4	0.5		1

(b)

FIGURE 1. **An example of a trust network. (a) A trust network. (b) A trust relationship matrix between users.**

As the distance of trust between users increases, the degree of trust relationships between users gradually decreases. Therefore, the value of trust relationship of u to v is as follows [4], [51]:

$$t_{uv} = \frac{g_{max} - g_{uv} + 1}{g_{max}} \quad (1)$$

where g_{max} denotes the maximum trust relationship-based distance between two users, and g_{uv} denotes the trust relationship-based distance of u to v. The distance-based trust measurement considers the trust distance value between users as 1. For example, to predict trust values for u_3 to u_1 and u_2, we set 4 as the maximum propagation distance.

According to Eq.(1), the trust distance values of u_3 to u_1 and u_1 to u_2 would be assigned 1, respectively, so the predicted trust value of u_3 to u_1 is $(4 - 1 + 1)/4=1$. The trust distance value of user u_3 to user u_2 would be assigned 2, so the predicted trust value u_3 to u_2 is $(4 - 2 + 1)/4=0.75$.

For datasets that do not have trust relationships, we can calculate the users' trust statements using the users' neighbor relationships in the rating matrix [15].

$$t_{uv} = \frac{\sum_{i \in I_{uv}} f(u, v, i)}{|I_{uv}|} = \frac{\sum_{i \in I_{uv}} (1 - \frac{1}{S}|r_{ui} - r_{vi}|)}{|I_{uv}|} \quad (2)$$

where r_{ui} and r_{vi} denote the ratings on i from u and v, respectively. S denotes the range of the scale, and I_{uv} represents the common item set rated from u and v. For instance, the value of S is 5 for 5-scale integer rating in the Epinions dataset.

B. MATRIX FACTORIZATION MODEL

CF-based recommendation algorithms are mainly classified into following two categories: memory-based CF and model-based CF. Among them, the algorithm of the former is simple and have high recommendation quality. However, as the number of users and items grows rapidly, the online prediction quality of the memory-based CF gradually fails to adapt to the needs of users and platforms. In recent years, because

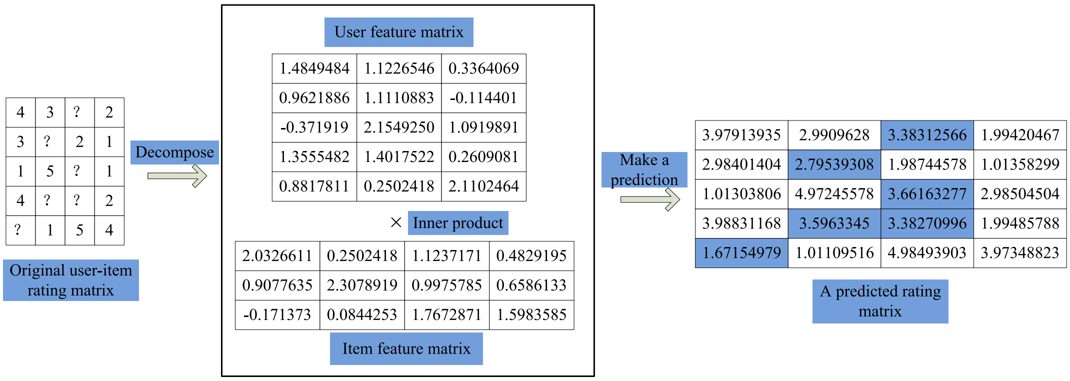


FIGURE 2. The process of rating prediction.

of its good scalability and accurate predictability in processing large-scale data, the model-based CF has been studied extensively, especially the recommendation algorithms based on MF technology have attracted widespread attention from academia and industry.

The MF model projects user's favors and item's features into the same latent factor space, and predicts ratings according to the degree of matching between user's favors and items' features. The MF model is written as follows [1]:

$$R \approx P_{n\times k} Q_{m\times k}^{T} = \hat{R} \tag{3}$$

Here, R indicates the rating matrix of $n \times k$, which is approximately decomposed into following two matrices: $P_{n\times k}$ and $Q_{m\times k}$. Among them, $P_{n\times k}$ represents the characteristic matrix of users and $Q_{m\times k}$ indicates the characteristic matrix of items. n represents the number of users, and m represents the number of items, respectively. The inner product of $P_{n\times k}$ and $Q_{m\times k}$ forms the approximate matrix denoted by $\hat{R}$ of R. k indicates the dimension of P.

As an original rating matrix, R is usually extremely sparse in RS. Our goal is to minimize deviation between $\hat{R}$ and R, obtain the model parameters of P and Q by training the established model, and get a filled user-item rating matrix [7], [10]. The objective function is as follows:

$$L = \min \left\| R - \hat{R} \right\| = \min(\sum\nolimits_{u=1}^{n} \sum\nolimits_{i=1}^{m} (r_{ui} - P_u Q_i^T)^2 + \lambda(\|P\|^2 + \|Q\|^2)) \tag{4}$$

where P_u and Q_i represent the corresponding eigenvectors of low-dimensional matrices P and Q, respectively, and $\lambda(\|P\|^2 + \|Q\|^2)$ indicates the regularization item, which is introduced to avoid overfitting.

The real values P and Q are approximated by a stochastic gradient descent method, which are described as follows:

$$P_u \leftarrow P_u - \eta \frac{\partial L}{\partial P_u} \tag{5}$$

$$Q_i \leftarrow Q_i - \eta \frac{\partial L}{\partial Q_i} \tag{6}$$

where η represents the learning rate. The difference between r_{mi} and $\hat{r}_{mi}$ for an observed pair is minimized by differentiating $(r_{mi} - \hat{r}_{mi})^2$ as follows:

$$\frac{\partial}{\partial p_{uk}} (r_{mi} - \hat{r}_{mi})^2 = -2p_{mk}(r_{mi} - \sum\nolimits_{k=1}^{K} p_{mk} q_{ki}) \tag{7}$$

$$\frac{\partial}{\partial q_{ki}} (r_{mi} - \hat{r}_{mi})^2 = -2q_{ki}(r_{ki} - \sum\nolimits_{k=1}^{K} p_{mk} q_{ki}) \tag{8}$$

The recurrence formulas are obtained according to gradient descent-based updating procedure:

$$p_{uk} = p_{uk} + \eta(q_{ki} \cdot e_{ui} - \lambda p_{uk}) \tag{9}$$

$$q_{ki} = q_{ki} + \eta(p_{uk} \cdot e_{ui} - \lambda q_{ki}) \tag{10}$$

where e_{ui} equals to $r_{mi} - \hat{r}_{mi}$. The predicted ratings are obtained as follows:

$$\hat{r}_{ui} = \sum\nolimits_{k=1}^{K} p_{uk} q_{ki} \tag{11}$$

Fig. 2 shows a process of rating prediction. A matrix of 4×5 is decomposed into user and item feature matrices by Eq.(9) and Eq.(10), and a predicted rating matrix is obtained by Eq.(11).

III. THE ALGORITHM AND FRAMEWORK INTEGRATING USERS' SOCIAL NETWORKS AND MATRIX FACTORIZATION TECHNIQUES

In this section, the proposed algorithm for the fusion of users' social networks and MF techniques will be presented. First, the framework of the recommendation approach fusing user social relations and MF techniques is described. Then, the processes of the model establishment and parameter learning are illustrated specifically.

A. THE PROPOSED FRAMEWORK

From the user's perspective, user's decisions would be affected by his/her trusted friends with different social status, and the degree of influence will depend on the extent of his trust in his/her friends. Inspired by the above intuition, we integrate the user's social status, homophily, and trust relationship into the matrix factorization model and propose a

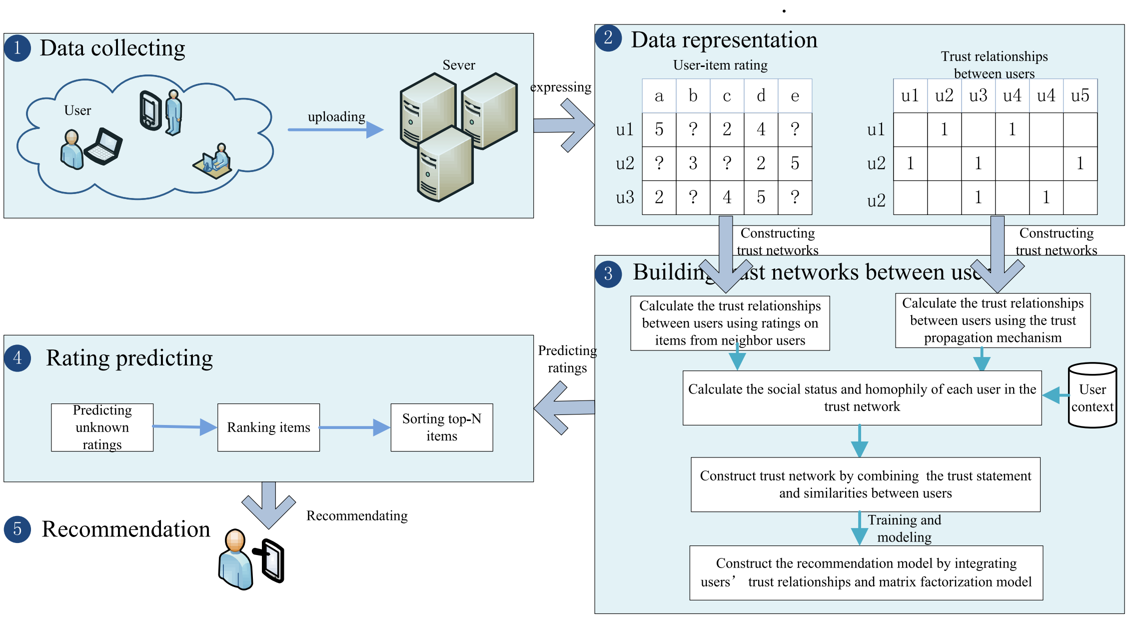


FIGURE 3. The framework of the proposed recommendation algorithm.

framework for a recommendation algorithm that fuses social relationships as shown in Fig. 3.

The implementation of our recommendation approach is divided into the following steps: (1) We build a rating matrix and trust relationship matrix by collecting data; (2) Calculate the implicit the values of social relationships between users using the rating information according to Eq.(2), and explicit trust statements between users using trust propagation based on the trust relationship matrix according to Eq.(1); (3) Calculate users' social status using the PageRank algorithm; (4) Adjust the trust relationships among users using user's social status and homophily characteristics; (5) Establish the matrix factorization trust relationships fusing the users' social status and homophily characteristics.

1) CALCULATE THE SOCIAL STATUS OF USERS IN TRUST NETWORKS

In social networks, users with high social status usually provide valuable information to users with low social status, therefore they have a lot of in-degrees. Correspondingly, users with low social status usually refer to the suggestions of users with high social status, and thus they have more out-degrees [16]. The (user, item) pair in RS is expressed as a bipartite graph, i.e., G (V, E). The two types of vertices are denoted as user u_i and item i_i, respectively, as shown in Fig. 4. Suppose that u_1 clicks on i_1, i_2 and i_3, u_2 clicks on i_1 and i_3, u_3 clicks on i_2 and i_5, and u_4 clicks on i_3 and i_4. Then it can be converted to Fig.4(b), which shows the users and the items they clicked.

Suppose we want to suggest items he/she might be interested in for the user u_1. Let's start at the vertex that corresponds to the user u_1, and select a path randomly with the probability of φ from the outgoing edge of u_1 to reach the

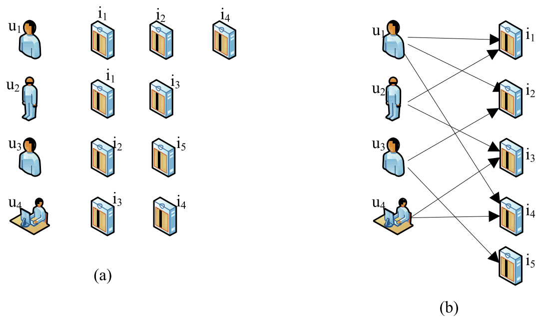


FIGURE 4. The representation of a bipartite graph. (a) (user, item) pair. (b) bipartite graph.

next vertex, say vertex v_1. The starting point u_1 is returned from the vertex v_1 with the probability of 1-φ, and the outgoing edges of the vertex v_1 walk with the probability of φ. After a lot of rounds of travel, it is found that the importance of each vertex would converge, and the probability of the vertex corresponding to each user is the social status value of user in social relationship network. The value of PR_u can be obtained by iteration. For PageRank algorithm, since the initial access probability of each node is the same, the initial access probability of each vertex is 1/N. Through the above analysis, the PageRank algorithm is employed to compute the social status of each user in social networks as follows [16]–[18], [53]–[55]:

$$\mathrm{PR_u} = \varphi \sum\nolimits_{\mathrm{v} \in \mathrm{C_u}} \frac{\mathrm{PR_v}}{|\mathrm{C_u}|} + \frac{1}{\mathrm{N}}(1-\varphi) \tag{12}$$

where PR_u indicates the value of the user u's PageRank in each cluster, and the user social status refers to the local relative value of the user in each social network, which reflects the authoritativeness of the user with similar preferences to a certain item. C_u indicates the set of closest friends of u, the number of users is denoted as N, and φ represents the probability of jumping out of the current trust network, which is range of [0, 1].

In real life, the higher the social status of a person in a field is, the stronger his/her influence will be, and the more likely others are to accept his/her advice. If the social status of user u is higher than user v, the possibility that the user u accepts the suggestions of the user v is relatively small. Otherwise, u is more likely to accept the suggestions of v. Motivated by this intuition, the adjusted weighted trust relationship with social status between u and v namely W_{uv} is as follows:

$$W_{uv} = \begin{cases} PR_v * t_{uv} & PR_v > PR_u \\ \dfrac{PR_v}{PR_u} * t_{uv} & \text{otherwise} \end{cases} \tag{13}$$

The higher the status of the user v in social networks, the higher the credibility of the user v. For instance, if the values of PR_u, PR_v and PR_w are 0.7, 0.9 and 0.5, respectively, the user u is likely to accept the recommendation of v rather than w. Otherwise, the possibility of user v accepting u's suggestion is less than that of user u's accepting v's suggestion. It is noticed that PR_u and PR_v are normalized values of social status.

The degree of user u's trust in user v is not only affected by the social status of the user, but is also relevant to the amount of commonly rated items of two users. Motivated by the intuition, the common ratings are considered as an interaction between two users. In general, the more products two users rated commonly, the more similar their ratings are, and the closer their preferences are. In addition, the interests of u and v on the same product change with time. Therefore, it is necessary to introduce a time decay factor to compute the interest similarity between u and v. If the time interval for two users to rate the same item is far apart, the similarity of two users' interest similarity should be reduced. A trust network can be established according to the combination of the similarities and the social trust relationships between users. The adjusted weighted trust relationship with common interests of users is computed as follows:

$$W^*_{uv} = \begin{cases} \dfrac{|I_{uv}|}{|I_u| + |I_v|} * \dfrac{\sum_{i\in I_{uv}} 1/(1+\exp(\omega|t_{ui} - t_{vi}|)}{|I_{uv}|} \\ \quad * sim_{uv} * PR_v * t_{uv} \quad PR_v \geq PR_u \\ \dfrac{|I_{uv}|}{|I_u| + |I_v|} * \dfrac{\sum_{i\in I_{uv}} 1/(1+\exp(\omega|t_{ui} - t_{vi}|)}{|I_{uv}|} \\ \quad * sim_{uv} * \dfrac{PR_v}{PR_u} * t_{uv} \quad \text{otherwise} \end{cases} \tag{14}$$

where I_u and I_v indicate the sets of items rated from u and v, respectively. ω represents the parameter of time attenuation. I_{uv} indicates the set of common items rated from users u and v. t_{ui} and t_{vi} indicate the times of item i rated from u and v, respectively. sim_{uv} represents the preference similarity between u and v that is computed as follows [1], [2]:

$$sim_{uv} = \frac{\sum_{i\in I_{uv}} (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i\in I_{uv}} (r_{ui} - \bar{r}_u)^2}\sqrt{\sum_{i\in I_{uv}} (r_{vi} - \bar{r}_v)^2}} \tag{15}$$

where $\bar{r}_u$ and $\bar{r}_v$ denote the average ratings on all items from u and v, respectively. The ratings on all the items from u are written as $r_u = \{r_{u1}, r_{u2}, \ldots, r_{un}\}$. The active user's rating for specified item can be calculated according to the closest k users' ratings as follows [46]:

$$\hat{r}_{ui} = \bar{r}_u + \frac{\sum_{v\in N_u} sim_{uv}(r_{vi} - \bar{r}_v)}{\sum_{v\in N_u} |sim_{uv}|} \tag{16}$$

where N_u denotes the closest neighbors set of u.

Fig. 5 shows a trust network with user's social status. The trust network is composed of 6 nodes and 10 directed arcs. A node denotes a user, and an arc, i.e., an association between two users expresses his/her degree of trust in other user. For instance, user u_1 has a trust relationship of 0.6 to u_2, and user u_1 has a social status value of 0.3.

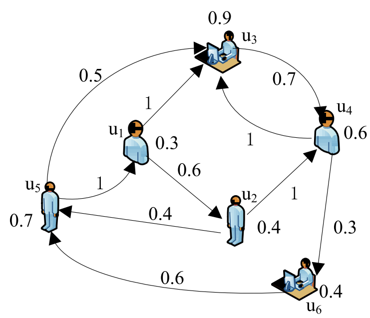


FIGURE 5. A trust network with user's social status.

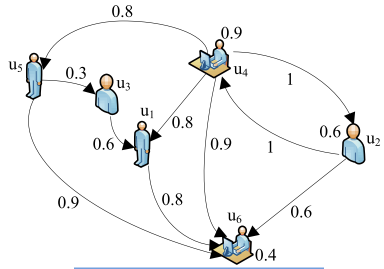
2) CALCULATE THE HOMOPHILY OF USERS IN TRUST NETWORKS

In social networks, homophily is also an important factor affecting the trust relationship between users. In general, users with similar characteristics have similar behaviors and they are more likely to establish trust relationships. The preference similarity of two users is obtained through computing the label similarity and individual characteristics similarity [14], [46]. TF-IDF can be used to calculate to the weight of the label. All labels are considered as a set of documents, and the weight of u_i to label b_k is computed as follows [45]–[47]:

$$tw_{uk} = tf(u,k) \times \log_2\left(\frac{N_t}{d_{uk}+1}\right) \tag{17}$$

where tf(u,k) represents the times that label b_k appears in the label set, d_{uk} represents the amount of users tagging the product b_k, and N_t represents the total amount of users tagging products.

According to the label weight vector of users, their preference similarity between users is obtained as



Evaluated matrix

4.993	1.999	1.920	2.990	4.954	2.999	4.417	3.681
3.983	3.954	0.941	1.924	2.006	1.848	3.304	1.689
4.001	0.961	1.981	1.625	3.348	2.553	1.999	3.954
0	0	0	0	0	0	0	0
4.964	2.012	1.996	2.252	4.001	2.971	3.103	3.869
4.003	2.997	1.100	2.982	3.904	2.091	4.955	2.000

(b) the predicted rating matrix by PMF

Original user-item rating matrix

5	2	0	3	5	3	0	0
4	4	0	0	2	0	0	0
4	0	2	0	0	0	2	4
0	0	0	0	0	0	0	0
5	2	2	0	4	3	0	0
4	3	0	3	0	0	5	2

User feature matrix

0.0569	-0.1632	0.6020	0.3945	0.6269
0.1118	-0.1619	0.2626	-0.0470	0.7410
-0.0984	-0.3580	0.3358	0.3598	0.3204
0.0659	-0.0970	0.3919	0.1123	0.6229
-0.0555	-0.0355	0.6198	0.2242	0.6691
0.2402	0.1028	0.5709	0.1494	0.7975

×

-0.5003	-0.6486	0.5392	0.2052	0.8241
0.2943	-0.4628	-0.2071	-0.4077	0.9472
-0.1918	-0.3529	0.2427	0.2427	0.2520
-0.0211	-0.1498	0.3506	0.1470	0.4909
-0.1961	-0.5698	0.6428	0.8389	0.2811
-0.2390	-0.2441	0.3914	0.1793	0.4279
0.1047	0.0879	0.5351	0.0327	0.8057
-0.3321	-0.8677	0.3456	0.5807	0.3556

Item feature matrix

Evaluated matrix

4.998	2.003	2.232	3.000	4.880	3.004	4.159	3.914
3.959	3.872	1.374	2.354	2.040	2.121	3.667	2.152
4.002	1.119	1.974	1.918	4.155	2.219	2.039	3.911
3.888	2.637	1.505	2.364	2.818	2.239	3.568	2.422
4.912	2.071	1.983	2.926	4.029	2.955	4.346	3.158
4.045	2.997	1.467	2.966	3.054	2.545	4.936	1.994

(a) the predicted rating matrix by the proposed social MF

FIGURE 6. A simple example for social MF. (a) the predicted rating matrix by the proposed social MF. (b) the predicted rating matrix by PMF.

follows [14], [45]:

$$S_{uv} = \frac{\sum_{k=1}^{n} tw_{uk} tw_{vk}}{\sqrt{\sum_{k=1}^{n} tw_{uk}^2}\sqrt{\sum_{p=1}^{n} tw_{vk}^2}} \tag{18}$$

where S_{uv} represents the preference similarity between u_u and u_v based on label weight. In addition, inspired by [14], [45]–[47], [50], and [52], user individual characteristics, as another factor that affects user's decision, can be divided into three categories [13], [52], [58]: the same, similar, and dissimilar.

$$h(ve_1, ve_2) = \begin{cases} (ve_1, ve_2) \in \text{same} & \text{if } ve_1 = ve_2 \\ (ve_1, ve_2) \in \text{similar} & \text{else if } |ra_1 - ra_2| \\ (ve_1, ve_2) \in \text{disimilar} & \text{else} \end{cases}$$
$$\leq \frac{\sum_{u_i,u_j \in U} \sum_{u_i,u_j \in U} |ve_i - ve_j|}{|U| \cdot |U-1|} \tag{19}$$

where user individual characteristics can be defined as a vector $F = (ft_1, ft_2, \ldots, ft_m)$, ve_1 and ve_2 are two values of the same characteristic ft_i $(1 \leq i \leq m)$, and ra_1 and ra_2 represent the average preferences with the characteristics ve_1 and ve_2, respectively. If the two user's attribute values are the same, the two user preferences are the same. If the two user attributes are in the deviation range, the two user's preferences are similar; otherwise, the two user's preferences are dissimilar. Inspired by [10], [14], [16], and [22], user's homophily can be computed as follows:

$$H_{ij} = \frac{|s_a| + |s_s|}{|s_a| + |s_s| + |s_d|} \frac{\sum_{p=1}^{n} tw_{ip} tw_{jp}}{\sqrt{\sum_{p=1}^{n} tw_{ip}^2}\sqrt{\sum_{p=1}^{n} tw_{jp}^2}} \tag{20}$$

where s_a, s_s, and s_d are sets of user's homophily with the same, similar, and dissimilar attributes, respectively.

B. THE PROPOSED ALGORITHM MODEL

By integrating the trust relationships with user's social status into the MF algorithm, we will construct an enhanced recommendation algorithm to solve the problems of cold start and data sparsity in RS. In real life scenarios, the personalized recommendation algorithm includes two aspects as follows: (1) A user' decision is usually influenced by his/her closest friends, therefore, the ultimate rating for an item should be a combination of his/her preferences and his/her closest friends' tastes, and W^*_{uv} denotes the degree of interpersonal trust. (2) A user's favors are similar to that of a user who has the same or similar background, and H_{uv} indicates the extent of preference similarity between u and v based on user's homophily.

1) A SIMPLE EXAMPLE

An example is used to explain the distinction between our proposed social recommendation approach and the traditional PMF recommendation algorithm. A typical rating prediction process based on the social trust network is described in Fig. 6. There are 6 users in total with 10 relationships between users in the trust relationship graph, each user is associated with a social status value v_{ij}, and each arc is in connection with a weight w_{ij}, which is range of [0, 1] to represent the degree of u_i trusts u_j. Our task to accomplish is how to predict the unknown ratings in the rating matrix accurately and quickly. For ease of description, Fig.6 uses only two source, the rating matrix and trust relationships between users to make a prediction. Inspired by the intuition that the social relationships between users will influence the user's behaviors, thus, the rating information and the trust relationship are factorized simultaneously using P^TQ, where P and Q indicate the low-dimensional latent feature matrices of users and items, respectively [9], [22], [48], [49]. The latent eigenvectors of user u relies on the latent eigenvectors of his/her closest friends $v \in N_u$, and this influence is described as follows [10]:

$$\hat{P}_u = \frac{\sum_{v\in N_u} T_{uv}P_v}{\sum_{v\in N_u} T_{uv}} \tag{21}$$

Among them, $\hat{P}_u$ indicates the approximate latent eigenvector of user u. Each row of the user's trust relationship matrix can be normalized as $\sum_{v=1}^{N} T_{uv} = 1$, so the user u's estimated latent eigenvector is obtained by his/her closest friends v as follows [10]:

$$\hat{P}_u = \sum_{v\in N_u} T_{uv}P_v \tag{22}$$

where T_{uv} is calculated according to Eq.(14). We conduct a social MF to make a recommendation (k=5), and obtain $P_{6\times 5}$ and $Q_{8\times 5}$ as Fig.6. Then we can predict unknown ratings in the matrix using P^TQ. It can be seen from the predicted results that even for a cold start user, such as u_4 does not rate any item, the available items can be recommended for user u_4 by utilizing the trust relationship between users. Otherwise, we cannot predict the missing values for user u_4 by the traditional PMF.

2) ADD THE INTERPERSONAL TRUST WITH USERS' SOCIAL STATUS

According to [27], we assume that the latent eigenvectors of users and items obey the Gaussian prior distribution:

$$p\left(P\,|\,\sigma_P^2\right) = \prod_{u=1}^{N} N\left(P_u\,|\,0,\sigma_P^2\right) \tag{23}$$

$$p\left(Q\,|\,\sigma_Q^2\right) = \prod_{i=1}^{M} N\left(Q_i\,|\,0,\sigma_Q^2\right) \tag{24}$$

Inspired by [9], [10], and [27], by adding the trust networks to the user eigenvectors, the user u's conditional distribution can be obtained according to the user eigenvectors of his/her closest neighbors as follows:

$$p\left(P\,|\,W,\sigma_W^2\right) = \prod_{u=1}^{N} N\left(P_u\,|\sum_{v\in N_u} W^*_{uv}P_v^T,\sigma_W^2\right) \tag{25}$$

According to [10], we can get the posterior probability of P and Q by a Bayesian inference:

$$\begin{aligned}
&p\left(P,Q\,|\,R,W,\sigma_R^2,\sigma_W^2,\sigma_P^2,\sigma_Q^2\right)\\
&\quad\propto p\left(R\,|\,P,Q,\sigma_R^2\right)p\left(P\,|\,T,\sigma_P^2,\sigma_W^2\right)p\left(Q\,|\,\sigma_R^2\right)\\
&\quad=\prod_{u=1}^{N}\prod_{i=1}^{M}\left[N\left(R_{ui}\,|\,g\left(P_u^TQ_i\right),\sigma_r^2\right)\right]^{I_{ui}^R}\\
&\quad\quad\times\prod_{u=1}^{N} N\left(P_u\,|\sum_{v\in N_u} W^*_{uv}P_v,\sigma_W^2 I\right)\\
&\quad\quad\times\prod_{u=1}^{N} N(P_u|0,\sigma_P^2 I)\\
&\quad\quad\times\prod_{i=1}^{M} N(Q_i|0,\sigma_Q^2 I)
\end{aligned} \tag{26}$$

where I_{ui}^R represents the indicator function, which equals to 1 when the item v_i has been rated by the user u_u and equals to 0 otherwise.

3) ADD THE INFLUENCE OF INTERPERSONAL HOMOGENEITY BETWEEN USERS

Similar to Eq.(25), motivated by [9] and [32], we can obtain the conditional distribution according to the user latent feature.

$$p\left(P\,|\,H,\sigma_H^2\right) = \prod_{u=1}^{N} N\left(P_u\,|\sum_{v\in N_u} H_{uv}P_v^T,\sigma_H^2\right) \tag{27}$$

According to [10] and [12], we can get the following posterior probability of P and Q by using the rating matrix, social trust matrix with user's social status, and user's homophily:

$$\begin{aligned}
&p\left(P,Q\,|\,R,W,H,\sigma_R^2,\sigma_T^2,\sigma_P^2,\sigma_Q^2\right)\\
&\quad\propto p\left(R\,|\,P,Q,\sigma_R^2\right)p\left(P\,|\,W,\sigma_P^2,\sigma_T^2\right)p\left(Q\,|\,\sigma_R^2\right)\\
&\quad=\prod_{u=1}^{N}\prod_{i=1}^{M}\left[N\left(R_{ui}\,|\,g\left(P_u^TQ_i\right),\sigma_r^2\right)\right]^{I_{ui}^R}\\
&\quad\quad\times\prod_{u=1}^{N} N\left(P_u\,|\sum_{v\in N_u} W^*_{uv}P_v,\sigma_W^2 I\right)\\
&\quad\quad\times\prod_{u=1}^{N} N\left(P_u\,|\sum_{v\in N_u} H_{uv}P_v,\sigma_H^2 I\right)\\
&\quad\quad\times\prod_{u=1}^{N} N(P_u|0,\sigma_P^2 I)\\
&\quad\quad\times\prod_{i=1}^{M} N(Q_i|0,\sigma_Q^2 I)
\end{aligned} \tag{28}$$

4) ADD THE INFLUENCE OF FEATURE VECTORS BETWEEN USERS

In real life, friends trusted by user u may have different interests and hobbies. Some friends have similar interests and hobbies, and other friends may have different hobbies with user u. If the user u's interests differ too much from those of the trusted friends, it will lead to the problem of reducing the accuracy of the user feature vector P_u of user u. To deal with the above problem, regularization items are proposed in the literature [32], [48]–[50], therefore, we employ the following regularization item into the recommendation model:

$$\frac{\beta}{2}\sum_{u=1}^{N}\sum_{t\in T^{+}(u)} sim_{ut}||P_u-P_t||_F^2 \tag{29}$$

where $T^{+}(u)$ represents the set of closest friends of u_i.

5) ADD THE INFLUENCE OF FEATURE VECTORS BETWEEN ITEMS

Inspired by [32], [41], and [49], we propose to impose a social regularization term to constrain item's feature vector Q with its similar neighbor items. In real life, people will consider similar products as substitutes when purchasing products. For example, when a user purchases a book on data structure book, he/she may also choose to purchase an algorithm book.

$$\sum_{i=1}^{M}\sum_{j\in N_i} sim_{ij}||Q_i-Q_j||_F^2 \tag{30}$$

where sim_{ij} is the similarity between items i and j, which can be calculated as follows:

$$sim_{ij}=\frac{\sum_{u\in U_{ij}}(r_{ui}-\bar{r}_i)\left(r_{uj}-\bar{r}_j\right)}{\sqrt{\sum_{u\in U_{ij}}(r_{ui}-\bar{r}_i)^2}\sqrt{\sum_{u\in U_{ij}}(r_{uj}-\bar{r}_j)^2}} \tag{31}$$

where U_i and U_j denote the sets of users who rated items i and j, respectively, and U_{ij} denotes the set of users who rated both items i and j. $\bar{r}_i$ and $\bar{r}_j$ indicate the average ratings on i and j in U_{ij}, respectively.

Based on the above intuition, we add the regularization item to the recommendation model, and the original problem is converted to minimize the formula.

$$\begin{aligned}
&L(R,W,H,P,Q)\\
&=\frac{1}{2}\sum_{u=1}^{N}\sum_{i=1}^{M} I_{ui}^{R}\left(R_{ui}-g\left(P_u^TQ_i\right)\right)^2\\
&\quad+\frac{\lambda_U}{2}\sum_{u=1}^{N}P_u^TP_u+\frac{\lambda_V}{2}\sum_{i=1}^{M}Q_i^TQ_i\\
&\quad+\frac{\lambda_W}{2}\sum_{u=1}^{N}((P_u-\sum_{v\in N_u}W_{uv}^{*}P_v)^T\\
&\quad\times(P_u-\sum_{v\in N_u}W_{uv}^{*}P_v))\\
&\quad+\frac{\lambda_H}{2}\sum_{u=1}^{N}((P_u-\sum_{v\in N_u}H_{uv}P_v)^T\\
&\quad\times(P_u-\sum_{v\in N_u}H_{uv}P_v))\\
&\quad+\frac{\beta}{2}\sum_{u=1}^{N}\sum_{t\in T(u)} sim_{ut}||P_u-P_t||_F^2\\
&\quad+\frac{\gamma}{2}\sum_{i=1}^{M}\sum_{j\in N_i} sim_{ij}||Q_i-Q_j||_F^2
\end{aligned} \tag{32}$$

Here, $\lambda_P=\frac{\sigma_R^2}{\sigma_P^2}$, $\lambda_Q=\frac{\sigma_R^2}{\sigma_Q^2}$, $\lambda_W=\frac{\sigma_R^2}{\sigma_W^2}$, $\lambda_H=\frac{\sigma_R^2}{\sigma_H^2}$. To minimize the above objective function, we execute the gradient descent on P_u and Q_i.

$$\begin{aligned}
\frac{\partial L}{\partial P_u}&=\sum_{i=1}^{M} I_{ui}^{R}V_i g'\left(P_u^TQ_i\right)(g\left(P_u^TQ_i\right)-R_{ui})\\
&\quad+\lambda_U P_u+\lambda_T\left(P_u-\sum_{v\in N_u}W_{uv}^{*}P_v\right)\\
&\quad-\lambda_T\sum_{u\in N_v}W_{vu}^{*}\left(P_v-\sum_{w\in N_v}W_{vw}^{*}P_w\right)\\
&\quad+\lambda_H\left(P_u-\sum_{v\in N_u}H_{uv}P_v\right)\\
&\quad-\lambda_H\sum_{u\in N_v}H_{uv}\left(P_v-\sum_{w\in N_v}H_{vw}P_w\right)\\
&\quad+\beta\sum_{t\in T^{+}(u)} sim_{ut}\left(P_u-P_t\right)\\
&\quad+\beta\sum_{s\in T^{-}(u)} sim_{us}\left(P_u-P_s\right)
\end{aligned} \tag{33}$$

$$\begin{aligned}
\frac{\partial L}{\partial Q_i}&=\sum_{u=1}^{N} I_{ui}^{R}P_v g'\left(P_u^TQ_i\right)(g\left(P_u^TQ_i\right)-R_{ui})+\lambda_V Q_i\\
&\quad+\sum_{j\in N_i} sim_{ij}\left(Q_i-Q_j\right)
\end{aligned} \tag{34}$$

where $T^{-}(u)$ represents the set of friends who trust u_u. $g'(a)$ equals to $g'(a)=\frac{e^{-a}}{(1+e^{-a})^2}$, which represents the derivative of $g(a)=\frac{1}{1+e^{-a}}$. sim_{ut} and sim_{us} denote similarities between users u and t, u and s, respectively, which are calculated according to Eq.(15).

Then the latent eigenvectors P_u and Q_i of users and items are updated as follows:

$$P_u^{(t+1)}=P_u^{(t)}-\eta\frac{\partial L(t)}{\partial P_u} \tag{35}$$

$$Q_i^{(t+1)}=Q_i^{(t)}-\eta\frac{\partial L(t)}{\partial Q_i} \tag{36}$$

where η represents the learning rate.

The execution process of USSHMF can be illustrated as follows:

IV. EXPERIMENTS

Several experiments are performed on the Douban and Epinions datasets to contrast the recommendation quality of USSHMF to other state-of-the-art methods. We present the detailed introduction of experimental datasets, parameter settings, metrics, and experimental results in this section [3], [9], [10], [22], [23].

A. EXPERIMENTAL SETUP

1) EXPERIMENTAL DATASETS

Two real-life experimental datasets including Douban[1] and Epinions[2] are used in the experiments [5], [19], [28]: (1) As one of the major datasets for verifying the performance of recommendation algorithms, each user in the Douban dataset can provide ratings for movies, books, and music, and the ratings range from 1 to 5, which represent user preferences for items. The sampled dataset contains 825,058 ratings of 14,682 movies by 3,220 users, and 2245 trust

[1]http://www.douban.com
[2]http://www.epinions.com

Algorithm 1 Algorithm Based on Matrix Factorization Technique Fusing Comprehensive Evaluation of User Social Status and Homophily

Input : an original rating matrix R, a social trust relationship matrix W, a user-tage set K.

initialize $P_u^{(0)}$, $Q_i^{(0)}$, ε, t=0, and number of iterations N.

Output : an estimated rating matrix R, a user eigenvector $P_u^{(*)}$, an item eigenvector $Q_i^{(*)}$.

1 if t>N, then goto step 8;
2 building the trust relationship matrix with user social status.
3 for u=1 to N
4 for v=1 to N
5 if $W_{uv} \notin \phi$ then calculate W_{uv} using Eq.(14).
6 end for
7 end for
8 building the user's homophily matrix.
9 for u=1 to N
10 for v=1 to N
11 if $H_{uv} \notin \phi$ then calculate H_{uv} using Eq.(20).
12 end for
13 end for
14 Calculate $\frac{\partial L(t)}{\partial P_u}$ and $\frac{\partial L(t)}{\partial Q_i}$;
15 Calculate $P_u^{(t+1)} = P_u^{(t)} - \eta\frac{\partial L(t)}{\partial P_u}$ and $Q_i^{(t+1)} = Q_i^{(t)} - \eta\frac{\partial L(t)}{\partial Q_i}$;
16 if L(t)< ε then goto step 14;
17 t++;
18 Output $P_u^{(*)}$, $Q_i^{(*)}$;
19 Predict the unknown ratings based on Eq.(11).

relationships. The sparsity of the rating information is 98.3%, and the sparsity of user's social relationship is 99.995%. (2) Epinions.com is a customer review website designed to promote experiences sharing about products. The extracted dataset is made up of 36,210 different products rated by 3,521 users. There are 68,329 ratings and 42,336 trust statements. The sparsity of the rating information is 99.94%, and the sparsity of user's social relationship is 99.97%. Each user can review the products as an integer from 1 to 5. In addition, each user on the Epinions website has a "trust relationship" list [9], [22].

2) METRICS

We evaluate the performance of recommendation algorithms using the following metrics in this paper: mean absolute error (MAE) and root mean square error (RMSE). They are the most frequently employed metrics to calculate the accuracy of recommendation.

MAE is defined as follows [1]:

$$MAE = \frac{\sum_{(u,i)\in R_t} \left| r_{ui} - \hat{r}_{ui} \right|}{|R_t|} \tag{37}$$

where R_t represents the set of the ratings of tested items. In RS, the smaller MAE value indicates that the prediction is more accurate and the algorithm performance is better.

TABLE 1. Parameter specification and default values.

Symbol	Description	Default value
λ_P	The regularization constant of user feature in [9, 10, 22, 24].	λ_P=0.1
λ_Q	The regularization constant of item feature in [9, 10, 22, 24].	λ_Q = 0.1
λ_T	The tradeoff parameter plays the role of adjusting the effects of interpersonal trust between users in [10]	λ_T = 1
λ_S	The tradeoff parameter is used to adjust the effect of interest similarity value between users in [23].	λ_S = 0.3
λ_W	The tradeoff parameter is used to adjust the effect of interpersonal trust between users.	λ_W = 4
λ_H	The tradeoff parameter is used to adjust the effect of user' homophily.	λ_H = 0.5
α	The tradeoff parameter is used to adjust the effects from recommendations of neighbors and trusted friends in [9].	α=0.7
β	The regularization parameter is used to control the impact of differences of users features in [32, 48, 49, 50].	β=0.01
γ	The regularization parameter is used to control the impact of differences of items features in [21, 49, 57].	γ=0.01
φ	The probability of jumping out of the currently trust network in [16].	0.85
ω	The parameter of time attenuation.	0.06
k	The dimension of latent feature space in [9, 10, 11].	k=10

RMSE is defined as follows [26]:

$$RMSE = \sqrt{\frac{\sum_{(u,i)\in R_t} \left(r_{ui} - \hat{r}_{ui}\right)^2}{|R_t|}} \tag{38}$$

3) COMPARISON

In the experiments, the proposed method (USSHMF) is compared to the PMF [22], [25], NMF [24], RSTE [9], SocialMF [10], and CSIT [27]. PMF is proposed by Salakhutdinov and only uses the user-item rating matrix for recommendations based on PMF. NMF is a constraint MF algorithm that enables the learnt features to more precisely describe the actual meanings such as the user interests and community tendencies. RSTE is a social matrix factorization method that uses the individual trust between users to optimize the solution, and integrate users' preferences and trusted friends' tastes as final ratings. SocialMF is a recommendation algorithm based on social networks proposed by Jamali, which adds a trust propagation mechanism to PMF to improve the accuracy of recommendations. CSIT is another social recommendation algorithm, which integrates the individual trust among users and Gaussian mixture model into the recommendation framework.

4) PARAMETER SETTINGS

We set the parameters of different algorithms by referring to the corresponding literature and experimental results of the comparison algorithms. Under the settings of these parameters, each comparison algorithm achieves optimal performance. Table 1 shows the meanings and the default values of parameters in all algorithms.

B. RESULTS OF EXPERIMENTS

1) RECOMMENDATION ACCURACY

We evaluate the accuracy of these recommendation algorithms in terms of MAE and RSME. Firstly, 80% of the data is selected randomly as a training dataset, and the rest of 20% is as a test dataset. We performed five experiments independently, and the average is taken as the final experimental result. Then serious experiments are performed on the Douban and Epinions datasets, and Table 2 reports the experimental results.

TABLE 2. Performance comparisons on MAE for two datasets.

Methods	Douban		Epinions	
	K=10	K=15	K=10	K=15
PMF	0.668	0.658	0.815	0.823
NMF	0.649	0.641	0.802	0.821
RSTE	0.635	0.619	0.785	0.798
SocialMF	0.626	0.605	0.761	0.783
CSIT	0.613	0.601	0.739	0.763
USSHMF	0.601	0.582	0.706	0.733

For the same dataset, we set the parameters of different algorithms to the same by referring to the corresponding literature [9]–[12], [14], [16], [21], [22], [27], [32], [47], and then we adjust corresponding parameter settings according to experimental results of the comparison algorithms to achieve optimal performance. On the Douban dataset, the number of hidden features, i.e. k is set to 15, $\lambda_P = \lambda_Q = 0.05$, $\alpha = 0.8$, $\lambda_T = 1$, $\lambda_S = 0.2$, $\lambda_W = 5$, and $\lambda_H = 0.4$ for all algorithms. On the Epinions dataset, k=10, $\lambda_P = \lambda_Q = 0.1$, $\alpha = 0.7$,$\lambda_S = 0.3$, $\lambda_T = 5$, $\lambda_W = 4$, and $\lambda_H = 0.7$ for all algorithms.

It can be seen from the reported results that the proposed method USSHMF outperforms the other methods under the MAE measure for the two datasets. For instance, our model improves the MAE performance of PMF, NMF, RSTE, SocialMF, and CSIT by 11.5%, 9.21%, 5.97%, 3.8%, and 3.16% when k=15 on the Douban dataset, respectively. Our model improves the MAE performance of PMF, NMF, RSTE, SocialMF, and CSIT by 13.4%, 12%, 10.1%, 7.23%, and 4.5% when k=10 on the Epinions, respectively. It is noted that the performance is significantly improved on the Epinions dataset over the Douban dataset. This is because that users' social status and homophily are introduced into our model, and Epinions contain explicit trust relationships between users while Douban dataset do not have explicit trust. relationships between users. However, we establish trust relationships through the user's neighbor relationships in this paper, and the accuracy of the algorithm is also improved compared to other algorithms.

2) PERFORMANCE ON SPARSE DATASETS

a: IMPACT OF DIFFERENT DEGREES FOR SPARSE DATA

To verify the performance of various recommendation approaches when solving sparse data, some new datasets are generated according to the following rules: the number of rated items per user less than 30, 25, 20, 15, 10 and 5 items as the new datasets, and the spasities are 99.80%, 99.82%, 99.86%, 99.89%, 99.93%, and 99.96%, respectively. The accuracy of each model is verified in terms of MAE. Fig. 7 graphically compares the proposed approach with the PMF, NMF, RSTE, SocialMF, and ContextMF approaches on Douban dataset in terms of MAE.

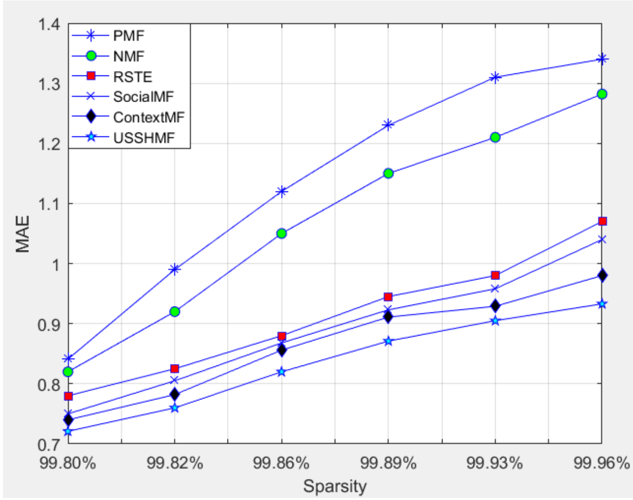


FIGURE 7. The MAE measure on sparse Douban dataset.

In Fig.7, X coordinates indicates sparsity of data, and Y coordinates indicates the measure of MAE. The sparsity of data is defined as $(1 - \frac{\text{the number of ratings}}{|U| \times |I|})$. The reported result from Fig.7 shows that the proposed model outperforms other models compared from 98.8% to 99.96% in term of MAE. It is because USSHMF model fuses the effects of user's social status and homophily on the recommendation accuracy into MF model while other models do not consider these factors, which indicates it is effective to fuse the user's social status and homophily into the USSHMF model.

b: IMPACTS OF PARAMETERS λ_W AND λ_H

In the USSHMF model, parameters λ_W and λ_H are used to control the impacts of social relationship on user behaviors. Fig. 8 and Fig. 9 reveal the effects of the different values of λ_W and λ_H on the MAE measure of the USSHMF. As shown in these figures, the larger the training set is, the higher the accuracy is. As λ_W increases, the quality of the prediction rises too. However, when λ_W exceeds a certain threshold ($\lambda_W = 5$ for the Douban dataset, and $\lambda_W = 4$ for the Epinions dataset), the performance would decrease. The proposed model reaches its best results when $\lambda_W = 5$ and $\lambda_W = 4$ on the Douban and Epinions datasets, respectively. As shown in Fig. 9, the performances of the USSHMF model can achieve their highest levels when $\lambda_H = 0.4$ and $\lambda_H = 0.7$ on the Douban and Epinions datasets.

c: IMPACTS OF PARAMETERS β AND γ

In the USSHMF model, parameters β and γ are both important parameters that influence the performance of the recommendation. They control how the user feature and item feature are adjusted to approximate real user interests. If a very small

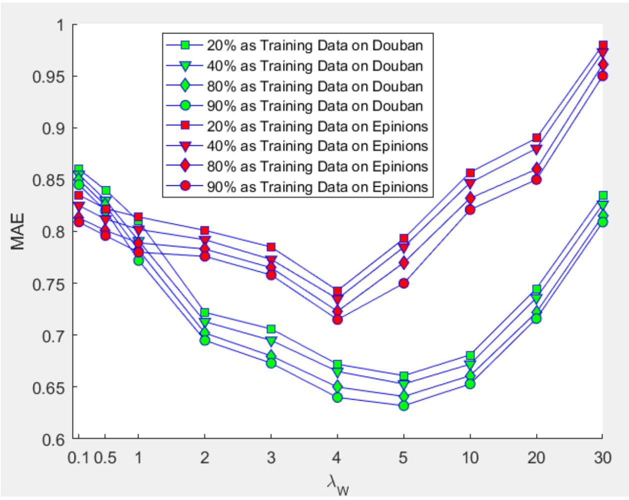


FIGURE 8. The influence of λ_W on the MAE in the Douban and Epinions datasets(dimension=10).

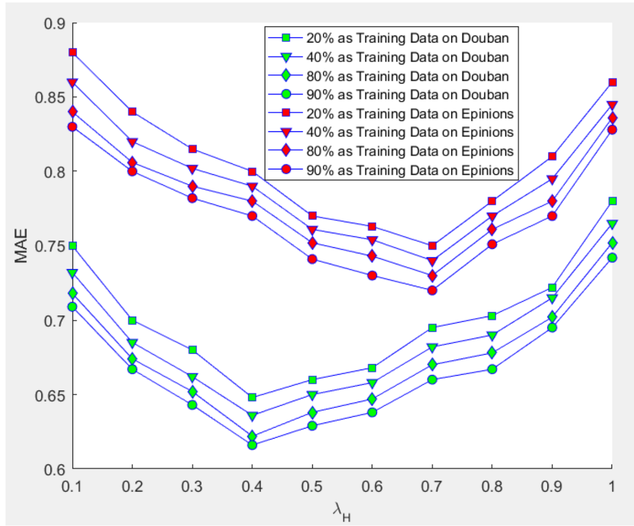


FIGURE 9. The influence of λ_H on the MAE in the Douban and Epinions datasets(dimension=10).

value of β or γ is employed, we only mine the information of the user-item rating matrix, social status, and homophily to constrain the feature vectors of user and item. If a very large value of β or γ is employed, the regularization term will dominate the learning processes. Fig. 10 and Fig.11 show the impacts of β and γ on MAE in our model, respectively. We observe that the value of β influences the recommendation performance significantly, which demonstrates the social network information can improve the recommendation accuracy greatly. As the value of β increases, the MAE value decrease at first, but when the value of β goes below a certain threshold (0.5 for the Douban dataset, and 0.01 for the Epinions dataset), the MAE values increase with further increase of the value of β. The impacts of γ generally share the same trend as the impacts of β on the Douban and Epinions datasets. It can be seen that reasonable adjustment of parameters β and γ can effectively improve the recommendation quality of the USSHMF model.

d: IMPACT OF FEATURE DIMENSIONALITY k

The dimension k of the latent eigenvector is another important parameter that influences the performance of the recommendation for the USSHMF model. Fig. 12 shows the impact of k on MAE for the Douban and Epinions datasets. As k increases, the accuracy of the recommendation also increases firstly, but the prediction accuracy decreases when k surpasses the threshold (15 for the Douban dataset, and 10 for the Epinions dataset), finally the prediction accuracy tends to stable. The phenomenon indicates that although the MF

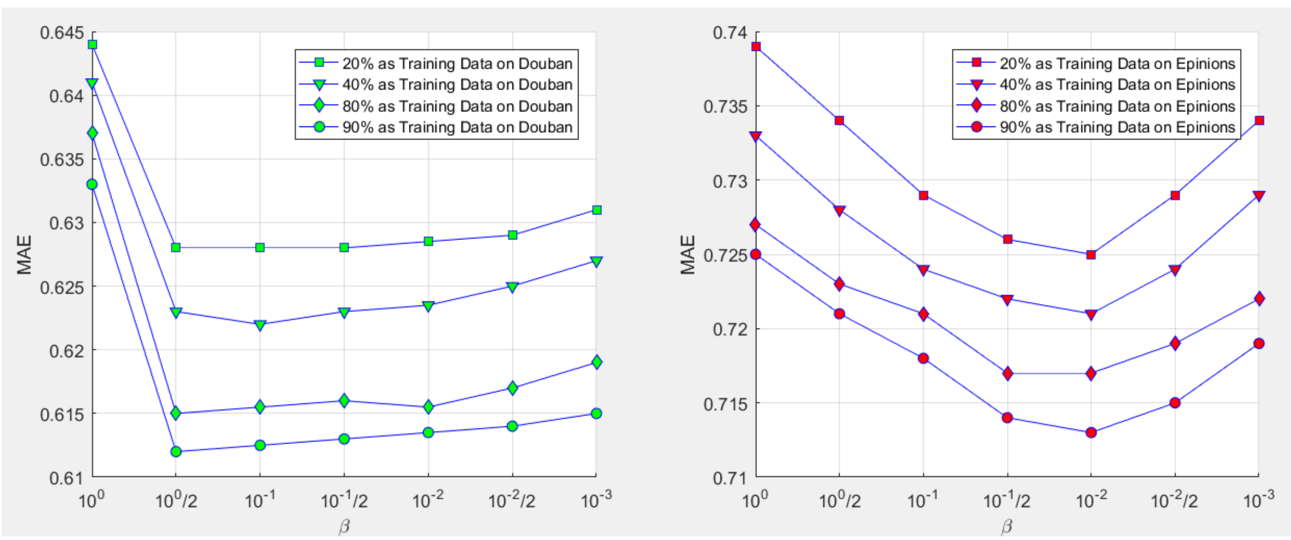


FIGURE 10. The influence of β on the MAE in the Douban and Epinions datasets(dimension=10).

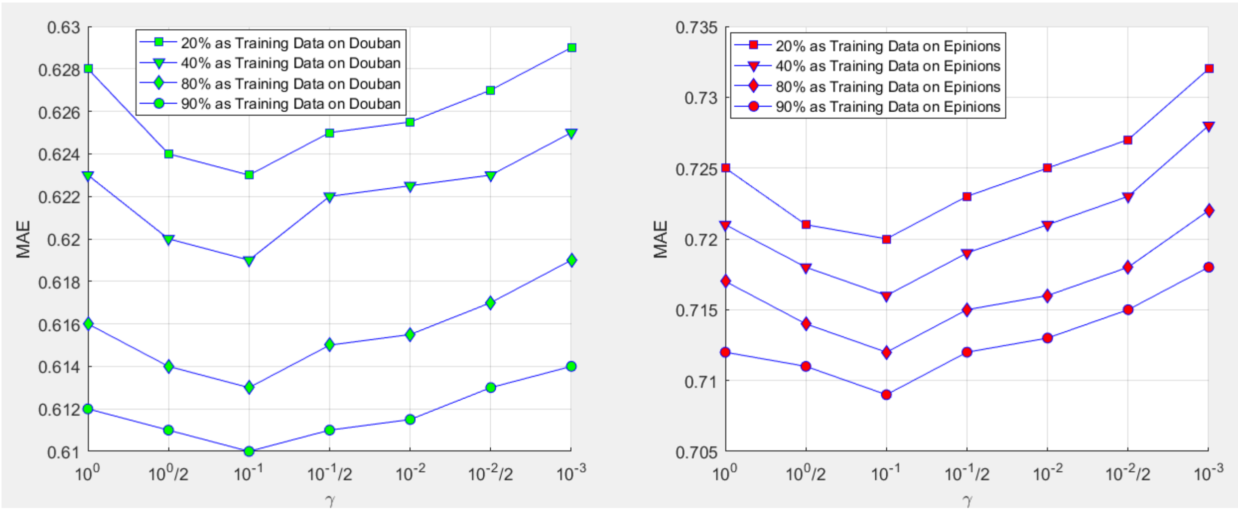


FIGURE 11. The influence of γ on the MAE in the Douban and Epinions datasets (dimension=10).

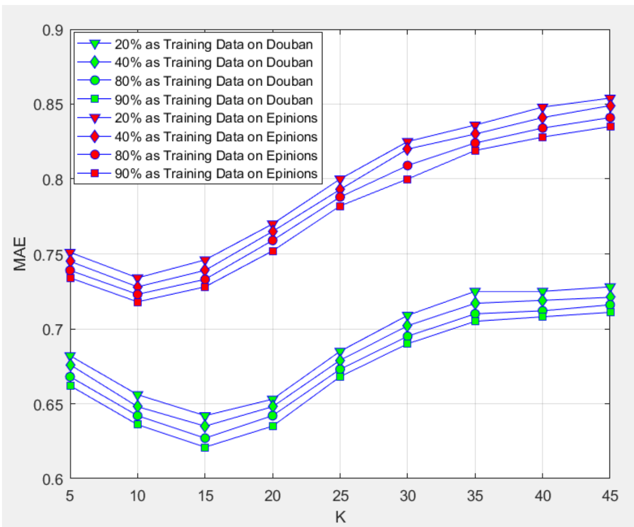


FIGURE 12. The impact of dimension.

model can represent more hidden features when k increases, it also introduces some noise and reduces the accuracy of recommendation. This observation verifies the basic assumptions of the MF model: only a few of latent factors affect the user's favors and characterize the items.

3) PERFORMANCE ON CORD START PROBLEM

Similarly, to demonstrate the prediction quality of the USSHMF algorithm under cold start users, we perform an experiment on cold users on the Douban and Epinions datasets, and Table 3 reports the experimental results of several different algorithms. Here, the users rated no more than 5 products are considered as cold-start users [12], [30].

USSHMF improves the MAE performances of PMF and NMF by more than 20% and by more than 9%, 7%, and 3% against RSTE, SocialMF, and CSIT on the Douban dataset,

TABLE 3. The performance comparisons of cold start users on the Douban and Epinions datasets.

Methods	Douban		Epinions	
	MAE	RMSE	MAE	RMSE
PMF	1.125	1.406	1.215	1.507
NMF	1.116	1.372	1.202	1.489
RSTE	0.933	1.192	1.135	1.409
SocialMF	0.915	1.062	0.961	1.132
CSIT	0.877	1.236	0.943	1.121
USSHMF	0.848	1.051	0.832	1.033

respectively. USSHMF improves the RMSE performances of PMF and NMF by more than 23%, by more than 10% against RSTE and SocialMF, and by more than 14% against CSIT. Similarly, on the Epinions dataset, USSHMF improves the MAE performances of PMF and NMF by more than 30%, by more than 20% against RSTE, by more than 10% against SocialMF and CSIT, respectively. USSHMF improves the RMSE performances of PMF and NMF by more than 30%, by more than 26% against RSTE, by more than 8% against SocialMF and CSIT, respectively. We can clearly observe from the experimental results for cold start users that the performance of the USSHMF model presented in the Epinions dataset is superior to that of the Douban dataset. It is because that the USSHMF model has few trust relationships available on sparse Douban dataset, and it mainly utilizes user's homophily feature, while USSHMF uses the individual trust between users and user's social status influences on the Epinions dataset, in addition to user's homophily feature.

4) PERFORMANCE COMPARISON AND ANALYSIS

We get the following observations by comparing these experimental results on the Douban and Epinions datasets,:

The conventional MF model, i.e., PMF and NMF make recommendations by using the rating information, rather than using additional trust relationship between users, so they have the worst performance of all the models.

The social MF models, i.e., RSTE, SocialMF, and CSIT, are superior to PMF and NMF. This is because they integrate social trust relationships into model training to obtain more accurate user preferences. The experimental results indicate that social trust relationship is useful to alleviate data sparsity and cold start problems.

Our model i.e., USSHMF, can make better use of the trust relationship between users, such as user's social status and homophily feature, to improve the recommendation performance. The experimental results show that USSHMF can further reduce the problems of data sparsity and cold start.

V. CONCLUSIONS AND FUTURE WORK

With the advent of online social networks, it has become increasingly important to exploit the information hidden in the social network to predict the behavior of users. The social network-based recommendation algorithm assumes that social network users' preferences are influenced by their friends and believe that these friends have similar preferences each other. Motivated by the intuition, the existing recommendation algorithms based on social networks such as RSTE, SocialMF, and CSIT integrate individual trust into MF model to improve the performance of recommendation. In real life, each user has different social status in different fields, and users with the same background are likely to have similar preferences. In this paper, we propose a novel recommendation model and a framework based on the MF technique fusing comprehensive evaluation of user social status and homophily, which is named USSHMF. We also conduct extensive experiments on two real-life datasets, and the experimental results show that our model exhibits higher recommendation accuracy over the existing state-of-the-art recommendation models. Although it has some advantages in its recommendation effectiveness, there are still some limitations in the model. For example, the model only considers trust relationships and homophily factors among users, but factors affecting the accuracy of recommendations include mood, geographic location, and the time interval for users to click or purchase items did not consider into our model. These considerations will help alleviate the data sparsity problem and will further potentially increase the prediction accuracy. In addition, deep learning can mine the hidden relationship between users and items through layer by layer unsupervised learning. Thus, deep learning-based recommendation method is our future research direction.

ACKNOWLEDGMENTS

The authors would like to thank Q. Gao from our group and X. Kong for their suggestions on revising this paper and their friendly support of the research. In addition, we would like to thank the anonymous reviewers and editor for their helpful comments.

REFERENCES

[1] F. Ricci, L. Rokach, and B. Shapira, "context-aware recommender systems," in *Recommender Systems Handbook*. New York, NY, USA: Springer, 2010, pp. 217–253.

[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. W. Meng, S. D. Liu, Y. J. Zhang, and X. Hu, "Research on social recommender systems," *J. Softw.*, vol. 26, no. 6, pp. 1356–1372, 2015.

[4] M. M. Azadjalal, P. Moradi, A. Abdollahpouri, and M. Jalili, "A trust-aware recommendation method based on Pareto dominance and confidence concepts," *Knowl.-Based Syst.*, vol. 116, no. 10, pp. 130–143, 2017.

[5] P. Moradi and S. Ahmadian, "A reliability-based recommendation method to improve trust-aware recommender systems," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7386–7398, Nov. 2015.

[6] H. Koohi and K. Kiani, "A new method to find neighbor users that improves the performance of collaborative filtering," *Expert Syst. Appl.*, vol. 83, pp. 30–39, Oct. 2017.

[7] M. Nilashi, O. Ibrahi, and K. Bagherifard, "A recommender system based on collaborative filtering using Ontology and dimensionality reduction techniques," *Expert Syst. Appl.*, vol. 92, pp. 507–520, Feb. 2018.

[8] L.-C. Wang, X.-W. Meng, and Y. J. Zhang, "Context-aware recommender systems," *J. Softw.*, vol. 23, no. 1, pp. 1–20, 2012.

[9] H. Ma, I. King, and M. R. Lyu, "Learning to recommend with social trust ensemble," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2009, pp. 203–210.

[10] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proc. 4th ACM Conf. Recommender Syst.*, 2010, pp. 135–142.

[11] M. Jiang *et al.*, "Social contextual recommendation," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manage. (CIKM)*, 2012, pp. 45–54.

[12] X. Zheng, Y. Luo, L. Sun, X. Ding, and J. Zhang, "A novel social network hybrid recommender system based on hypergraph topologic structure," *World Wide Web*, vol. 21, no. 4, pp. 985–1013, 2017.

[13] Y. Pan, F. He, and H. Yu, "Social recommendation algorithm using implicit similarity in trust," *Chin. J. Comput.*, vol. 41, no. 1, pp. 65–81, 2018.

[14] Y. Wang, X. Wang, and W. Zuo, "Trust prediction modeling based on social theories," *J. Softw.*, vol. 25, no. 12, pp. 2893–2904, 2014.

[15] Z. Zhang, "Research on personalized recommendation models and algorithm in social networks," Ph.D. dissertation, Dept. Manage. Sci. Eng., Shandong Normal Univ., Jinan, China, 2015.

[16] Y. Yu, Y. Gao, and H. Wang, "Integrating user social status and matrix factorization for item recommendation," *J. Comput. Res. Develop.*, vol. 55, no. 1, pp. 113–124, 2018.

[17] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2010, pp. 1361–1370.

[18] P. Fernandez, "Google's pagerank and beyond: The science of search engine ranking," *Math. Intell.*, vol. 30, no. 1, pp. 68–69, 2008.

[19] A. J. B. Chaney, D. M. Blei, and T. Eliassi-Rad, "A probabilistic model for using social networks in personalized item recommendation," in *Proc. ACM Conf. Recommender Syst.*, 2015, pp. 43–50.

[20] F. Wang, T. Li, X. Wang, S. Zhu, and C. Ding, "Community discovery using nonnegative matrix factorization," *Data Mining Knowl. Discovery*, vol. 22, no. 3, pp. 493–521, 2011.

[21] Y. Li and M. Song, "Recommender system using implicit social information," *IEICE Trans. Inf. Syst.*, vol. 98, no. 2, pp. 346–354, 2015.

[22] H. Ma, H. Yang, M. R. Lyu, and I. King, "SoRec: Social recommendation using probabilistic matrix factorization," in *Proc. CIKM*, Napa Valley, CA, USA, Oct. 2008, pp. 931–940.

[23] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user interest and social circle," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 7, pp. 1763–1777, Jul. 2014.

[24] D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

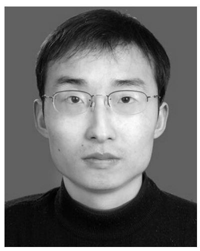
[25] A. Mnih and R. R. Salakhutdinov, "Probabilistic matrix factorization," in *Proc. 21st Annu. Conf. Neural Inf. Proc. Syst. (NIPS)*, 2007, pp. 1257–1264.

[26] X. Luo, M. Zhou, Y. Xia, and Q. Zhu, "An efficient non-negative matrix-factorization-based approach to collaborative filtering for recommender systems," *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 1273–1284, May 2014.

[27] J. Li, C. Chen, H. Chen, and C. Tong, "Towards context-aware social recommendation via individual trust," *Knowl.-Based Syst.*, vol. 127, pp. 58–66, Jul. 2017.

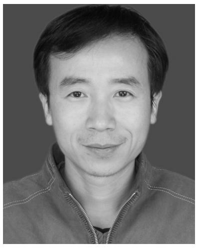
[28] P. Moradi, S. Ahmadian, and F. Akhlaghian, "An effective trust-based recommendation method using a novel graph clustering algorithm," *Phys. A, Stat. Mech. Appl.*, vol. 436, pp. 462–481, Oct. 2015.

[29] Z. Yang, B. Wu, and K. Zheng, "A survey of collaborative filtering-based recommender systems for mobile Internet applications," *IEEE Access*, vol. 5, no. 4, pp. 3273–3287, 2016.
[30] M. Jamali and M. Ester, "TrustWalker: A random walk model for combining trust-based and item-based recommendation," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Pairs, France, 2009, pp. 397–406.
[31] Y. Li, "Research on some key technologies of personalization recommender system," Ph.D. dissertation, School Comput. Sci., Beijing Univ. Posts Telecommun., Beijing, China, 2016.
[32] K. Ji, "Research on hybrid collaborative filtering recommendation algorithm based on context information," Ph.D. dissertation, School Comput. Inf. Technol., Beijing Jiaotong Univ., Beijing, China, 2016.
[33] P. Cremonesi, Y. Koren, and R. Turrin, "Performance of recommender algorithms on Top-N recommendation tasks," in *Proc. 4th ACM Conf. Recommender Syst.*, 2010, pp. 39–46.
[34] Y. Park, S. Park, and W. Jung, "Reversed CF: A fast collaborative filtering algorithm using a k-nearest neighbor graph," *Expert Syst. Appl.*, vol. 42, no. 8, pp. 4022–4028, 2015.
[35] X. Zhou, J. He, G. Huang, and Y. Zhang, "SVD-based incremental approaches for recommender systems," *J. Comput. Syst. Sci.*, vol. 81, no. 4, pp. 717–733, 2015.
[36] 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, 2015.
[37] D. Rafailidis and F. Crestani, "Learning to rank with trust and distrust in recommender systems," in *Proc. 11th ACM Conf. Recommender Syst.*, 2017, pp. 5–13.
[38] M. Aldayel and M. Ykhlef, "A new sentiment case-based recommender," *IEICE Trans. Inf. Syst.*, vol. E-100, no. 7, pp. 1484–1493, 2017.
[39] H. Liu *et al.*, "Context-based collaborative Filtering for citation recommendation," *IEEE Access*, vol. 3, pp. 1695–1703, 2015.
[40] Z. Xu, F. Zhang, W. Wang, H. Liu, and X. Kong, "Exploiting trust and usage context for cross-domain recommendation," *IEEE Access*, vol. 4, pp. 2398–2407, 2016.
[41] X. Wang, Y. Wang, and H. Sun, "Exploring the combination of dempster-shafer theory and neural network for predicting trust and distrust," *Comput. Intell. Neurosci.*, vol. 2016, 2016, Art. no. 5403105, doi: 10.1155/2016/5403105.
[42] J. Tang, H. Gao, X. Hu, and H. Liu, "Exploiting homophily effect for trust prediction," in *Proc. WSDM*, Rome, Italy, Feb. 2013, pp. 53–62.
[43] H. Li and X. Ma, "Research on personalized recommendation algorithms in social network environment," in *Proc. China Univ. Mining Technol.*, May 2016, pp. 57–65.
[44] P. Massa and P. Avesani, "Trust-aware collaborative filtering for recommender systems," in *Proc. Federated Int. Conf. Move Meaningful Internet Syst.*, 2004, pp. 492–508.
[45] Y. Cao, W. Li, and D. Zheng, "An improved neighborhood-aware unified probabilistic matrix factorization recommendation," *Wireless Pers. Commun.*, vol. 102, no. 4, pp. 3121–3140 , 2018.
[46] C. C. Aggarwal, *Recommender Systems*. New York, NY, USA: Springer, 2016.
[47] C. Yin, J. Wang, and J. H. Park, "An improved recommendation algorithm for big data cloud service based on the trust in sociology," *Neurocomputing*, vol. 256, pp. 49–55, Sep. 2017.
[48] H. Ma, H. Yang, I. King, and M. R. Lyu, "Semi-nonnegative matrix factorization with global statistical consistency for collaborative filtering," in *Proc. 18nd Int. ACM Conf. Inf. Knowl. Manage.*, 2009, pp. 767–776.
[49] H. Ma, M. R. Lyu, and I. King, "Learning to recommend with trust and distrust relationships," in *Proc. 3rd ACM Conf. Recommender Syst. (RecSys)*, 2009, pp. 189–196.
[50] H. Li, X. Ma, and J. Shi, "Incorporating trust relation with PMF to enhance social network recommendation performance," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 30, no. 6, p. 1659016, 2016.
[51] P. Massa and P. Avesani, "Trust-aware recommender systems," in *Proc. ACM Conf. Recommender Syst.*, Minneapolis, MN, USA, 2007, pp. 17–24.
[52] Q. Gao, "Research on key issues of context-aware recommender systems," Ph.D. dissertation, School Inf. Sci. Technol., Northwest Univ., Xi'an, China, 2017.
[53] P. Bekhin, "A survey on pagerank computing," *Internet Math.*, vol. 2, no. 1, pp. 73–120, 2005.
[54] Y. Lü, Y.-C. Zhang, C. H. Yeung, and T. Zhou, "Leaders in social networks, the delicious case," *PLoS ONE*, vol. 6, no. 6, p. e21202, 2012.
[55] X. Zhai and T. Xue, *Social Trust: Theory and Application*. Beijing, China: Renmin Univ. Press, 2014, pp. 1–36.
[56] X. Kong, M. Mao, W. Wang, J. Liu, and B. Xu, "VOPRec: Vector representation learning of papers with text information and structural identity for recommendation," *IEEE Trans. Emerg. Topics Comput.*, to be published. [Online]. Available: https://ieeexplore.ieee.org/document/8349947, doi: 10.1109/TETC.2018.2830698.
[57] R. Chen, Q. Hua, Y. Chang, L. Zhang, and B. Wang, "A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks," *IEEE Access*, vol. 6, no. 1, pp. 64301–64320, 2018, doi: 10.1109/ACCESS.2018.2877208.
[58] 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, 2015.
[59] A. Comi, L. Fotia, F. Messina, D. Rosaci, and G. M. Sarné, "A partnership-based approach to improve QoS on federated computing infrastructures," *Inf. Sci.*, vol. 367, pp. 246–258, Nov. 2016.
[60] P. De Meo, F. Messina, F. Rosaci, and G. M. Sarné, "Forming time-stable homogeneous groups into online social networks," *Inf. Sci.*, vol. 414, pp. 117–132, Nov. 2017.
[61] L. Fotia, F. Messina, D. Rosaci, and G. M. Sarné, "Using local trust for forming cohesive social structures in virtual communities," *Comput. J.*, vol. 60, no. 11, pp. 1717–1727, 2017.
[62] P. De Meo, L. Fotia, F. Messina, D. Rosaci, and G. M. L. Sarné, "Providing recommendations in social networks by integrating local and global reputation," *Inf. Syst.*, vol. 78, pp. 58–67, Nov. 2018.
[63] Z. Li, F. Nie, X. Chang, L. Nie, H. Zhang, and Y. Yang, "Rank-constrained spectral clustering with flexible embedding," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 6073–6082, Dec. 2018.
[64] W. Yao, J. He, G. Huang, and Y. Zhang, "Modeling dual role preferences for trust-aware recommendation," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2014, pp. 975–978.



RUI CHEN received the B.S. degree in computer science and technology from the Henan Institute of Science and Technology, Xinxiang, China, in 2004, and the M.S. degree in computer application technology from Northwest University, Xi'an, China, in 2009, where he is currently pursuing the Ph.D. degree in computer application and technology. His research interests include recommender systems, social network analysis, human–computer interaction, and natural language processing.

QINGYI HUA received the B.S., M.S., and Ph.D. degrees in computer software and theory from Northwest University, Xi'an, China, in 1982, 1988, and 2006, respectively. Since 1998, he has been a Professor with the Human–Computer Interaction Laboratory, Northwest University. From 1992 to 1995 and from 2000 to 2004, he was involved on human–computer project research with the National Information Science Research Center, Germany. His research interests include human–computer interaction, recommender systems, and user interface engineering. He is a member of the Editorial Board of the *Chinese Journal of Computer* and Human–Computer Interaction Professional Committee.



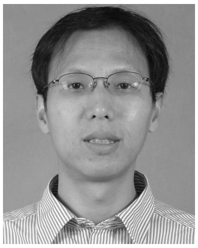
BO WANG received the B.S. degree from the Xi'an University of Architecture and Technology, in 2003, and the M.S. degree from Xidian University, in 2007. He is currently pursuing the Ph.D. degree Northwest University. His current research interests include human–computer interaction and recommender systems.

MIN ZHENG received the B.S. and M.S. degrees from Armed police Engineering University, in 2003 and 2006, respectively. Her research interests include image processing, pattern recognition, and complex system modeling.

WEILI GUAN received the bachelor's degree from Huaqiao University, in 2009, and the Diploma and master's degrees from the National University of Singapore, in 2011 and 2014, respectively. She joined Hewlett Packard Enterprise Singapore as a Software Engineer and worked there for around five years. Her current research interests include employing the multimedia and data mining techniques to boost the business revenue.

XIANG JI received the B.S. degree in computer application from Northwest University, in 2001, and the M.S. degree in computer application technology from Xidian University, in 2005. She is currently pursuing the Ph.D. degree with Northwest University. Her current research interests include human–computer interaction and user interface engineering.

QUANLI GAO received the B.S. degree in information and computer science, in 2010, and the Ph.D. degree from Northwest University, in 2017. He is currently a Lecturer with Xi'an Polytechnic University. He has participated in several national research projects. His research interests include recommender systems and machine learning.



XIANGJIE KONG (M'13–SM'17) received the B.S. and Ph.D. degrees from Zhejiang University, Hangzhou, China. He is currently an Associate Professor with the School of Software, Dalian University of Technology, China. He has published more than 70 scientific papers in international journals and conferences (with 50+ indexed by ISI SCIE). His research interests include intelligent transportation systems, mobile computing, and cyber-physical systems. He is a Senior Member of IEEE and CCF, and a member of ACM. He has served as a Workshop Chair or a PC Member of a number of conferences. He has served as an (Guest) Editor of several international journals.