A Comprehensive Social Matrix Factorization with Social Regularization for Recommendation Based on Implicit Similarity by Fusing Trust Relationships and Social Tags

Rui Chen (✉ 13382038@qq.com)

Zhengzhou University of Light Industry

Jianwei Zhang

Zhengzhou University of Light Industry

Zhifeng Zhang

Zhengzhou University of Light Industry

Yan-Shuo Chang

Xi'an University of Finance and Economics

Jingli Gao

Pingdingshan University

Pu Li

Zhengzhou University

Hui Liang

Zhengzhou University of Light Industry

Research Article

Keywords: Recommender Systems, Collaborative filtering, Matrix factorization, Social relationships, Social networks

Posted Date: August 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-490657/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A Comprehensive Social Matrix Factorization with Social Regularization for Recommendation Based on Implicit Similarity by Fusing Trust Relationships and Social Tags

Rui Chen^{1,2}, Jian-wei Zhang^{1,2,*}, Zhifeng Zhang^{1,2}, Yan-Shuo Chang^{3,*}, Jingli Gao⁴, Pu Li^{1,2}, and Hui Liang^{1,2}

Received: 8 Jun 2020 / Revised: xx xx 2020 / Accepted: xx xx 2020/

Published online: 12 September 2020

Abstract Social relationships play an important role in improving the quality of recommender systems (RSs). A large number of experimental results show that social relationship-based recommendation methods alleviate the problems of data sparseness and cold start in RSs to some extent. However, since the social relationships between users are extremely sparse and complex, and it is difficult to obtain accurately user preference model, thus the performance of the recommendation system is affected by the existing social recommendation methods. In order to accurately model social relationships and improve recommendation quality, we use explicit social relationships such as user-item ratings, trust relationships and implicit social relationships such as social tags to mine potential interest preferences of users and propose an improved social recommendation method integrating trust relationship and social tags. The method map user features and item features to the shared feature space by using the above social relationship, respectively, and obtains user similarity and item similarity through potential feature vectors of users and items, and continuously trains them to obtain accurate similarity relationship to improve the recommendation performance. Experimental results demonstrate that our proposed approach achieves superior performance to the other social recommendation approaches.

Keywords Recommender Systems; Collaborative filtering; Matrix factorization; Social relationships; Social networks.

1. Introduction

In recent years, the latent factor model based on matrix factorization (MF) is widely used in RSs due to strong scalability and

Rui Chen

13382038@qq.com

Jian-wei Zhang

zhangjianwei2019@126.com

Jingli Gao

gaojingli.cv@gmail.com

¹Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou, 450002, China

²Henan Joint International Research Laboratory of Computer Animation Implementation Technologies

³Institute for Silk Road Research, Xi'an University of Finance and Economics, Xi'an, 710100, China

⁴Computer College, HN-ITSETRC, HN-MDITSETRC, Pingdingshan University, Pingdingshan 467000, China

high recommendation quality. The social MF model mainly maps multiple social relationships and user preference information among users in a social network to shared user and item feature spaces, thereby realizes user preference prediction to alleviate the problem of data sparsity^[1]. However, in the face of complex network structure and social relationships, directly measuring social relationships tends to cause the estimated user preferences to deviate from the real user preference model, so the quality of the recommendation will be seriously affected^[2].

Some influence factors such as user social relationships, social tags, and user personal interests are incorporated into recommendation models to improve the quality of recommendations^[3-6]. The above methods improve the accuracy of the prediction by mapping the user rating and the social relationship to the shared user feature space and the item feature space, although the recommendation inaccuracy caused by the sparsity of the rating is alleviated to some extent, the user's social relationship is not further trained to obtain an accurate similarity relationship when the neighboring relationship is adopted to obtain user feature, thereby the predicted preference model may deviate from the real user preference model, resulting in a limited accuracy of the recommendation.

Although some prediction methods on user trust relationships have been proposed recently, but few studies applied user prediction results of social relationship to the recommendation model^[7-12]. In [7], the MF technique is used to complete the recommendation by integrating the approximate ratings for unknown items from the user and his/her friends with similar interests. In [8-10], the user characteristics are considered to be composed of his/her trustees and trusters, so the trust relationship, and the characteristics of the trustees and trusters are integrated to obtain user latent feature vectors by using the MF technology, so as to the unknown trust relationship is predicted. In [13-16], user ratings and social tags are used to map into shared user and item feature spaces by employing the MF technology to obtain more accurate user and item similarity relationships, and social recommendation methods are proposed, which improve the accuracy of recommendation, but these methods generally directly predict users' preferences for items based on the obtained user and item features, while ignoring the complex social relationships among users in the social networks. In [15], user implicit similarity is obtained more accurately by indirect feedback from explicit user ratings and trust relationships, thereby the recommendation accuracy is improved. However, this method does not take into account the impact of user's implicit comment (such as social tags) on the user's trust relationships and the item's social relationships estimation.

The above methods improve the accuracy of the prediction by mapping the user rating information and the social relationship to the shared user feature space and the item feature space, although to some extent, the recommendation inaccuracy caused by the rating sparsity is alleviated, but the user's social relationship is not further trained for an accurate similarity relationship when user feature is measured according to the neighboring relationship, so that the user preference model obtained may deviate from the real user preference model, resulting in a limited accuracy of the recommendation.

Considering the complexity of social networks and the difficulty of accurately modeling social relationships, inspired by the literature [7-9, 15, 22, 43, 46, 51], the prediction model fusing the implicit trust relationships between users and social relationships between items is integrated into the recommendation method by employing feedback mechanism and recommendation framework from the literature [6, 7, 15], and a comprehensive recommendation method combining trust relationships and social tags is proposed to improve recommendation accuracy. The method comprehensively considers the relationship between explicit and implicit interactions of users, and combines RSTE, SocialIT and RoRec framework structure to mine potential user and item features from explicit relationships such as user-item ratings, trust relationships and social tags, mapping user features and item features to a shared space, respectively, accurate user and item features are obtained through continuous training to measure user and item implicit similarities. In other words, social relationship between users or between items are considered as a "black box", without regard to multiple direct and indirect social relationships and interactions between users, which not only mitigates data sparsity, cold start and data imbalance and also simplifies the complex social relationship modeling process for users.

The rest of this paper is organized as follows: In Section 2, we briefly review the related work. In Section 3, we propose a comprehensive social MF model for recommendation based on implicit similarity by fusing trust relationships and social tags. The experimental results and analysis are demonstrated in Section 4. Finally, conclusions and further study are discussed in Section 5.

2. Preliminaries and related works

In this section, we will review the work related to the social relationship and the low-rank social MF-based latent factor model because of the importance of social relationship in improving the quality of recommendation [10, 11, 13, 26, 29, 50].

2.1 Social relationships

Trust relationships and propagation are used to evaluate indirect social relationships between users, which are used in the establishment of recommendation models [5]. In recent years, several trust relationship metrics have been proposed, among which TidalTrust and RTCF are the most representative trust relationship calculation methods [10, 12]. TidalTrust is a user trust relationship measurement method based on trust propagation theory, which mainly uses the idea that the trust relationship would gradually decrease as the distance between users increases. Its trust relationship is expressed as follows [2, 24]:

$$S_{uv} = \frac{\sum_{w \in F_u^+} S_{uw} S_{wv}}{\sum_{w \in F_u^+} S_{uw}} \quad (1)$$

where F_u^+ denotes the set of users that user u trusts directly, and S_{uw} denotes the degree of user u 's trust in w . RTCF is a measurement method of user's relationship that takes into consideration the similarity and trust relationships between users. The calculation method of the social relationship between the two users is as follows [6, 23]:

$$S_{uv} = \begin{cases} \frac{2 * sim_{uv} * T_{uv}}{sim_{uv} + T_{uv}}, & \text{if } sim_{uv} + T_{uv} \neq 0, \text{ and } sim_{uv} * T_{uv} \neq 0 \\ sim_{uv}, & \text{else if } sim_{uv} \neq 0, \text{ and } T_{uv} = 0 \\ T_{uv}, & \text{else if } sim_{uv} = 0, \text{ and } T_{uv} \neq 0 \\ 0, & \text{else} \end{cases} \quad (2)$$

where sim_{uv} denotes the similarity of preference between users u and v , and T_{uv} denotes the trust degree of user u in user v , which is bound to the range $[0,1]$.

Trust propagation can be used to measure the strength of trust between two users without a direct correlation. For example, Figure 1 (a) describe the process of predicting ratings for items using trust relationships between users. It is known that the user u_2 's rating for the item i_2 is 5 in Figure 1 (a), according to the trust propagation theory we can get the user u_1 's rating for the item i_2 , namely 3 because the user u_1 's trust strength for the user u_2 is 0.6. However, there are complex social relationships between users in social networks and various social relationships among users affect each other, thus it is difficult to measure accurately the social relationship between users. Figure 1 (b) shows the complex social relationships between users.

Figure 1. The complex interaction relationship between users in social networks

In recent years, various social relationships such as social tags, personal interests and other influencing factors are integrated into the recommendation model to improve the recommendation quality by mapping user rating information and social relationships to shared user feature space and item feature space [7, 11, 12, 14, 17, 19, 32, 34, 38]. Although the problem of inaccurate

recommendation caused by sparse rating data is alleviated to some extent, it lacks further training on user social relationship to obtain accurate similarity relationship when measuring user characteristics through neighborhood relationships and trust relationships. So that the user preference model obtained by the nearest neighbor or trusted users may deviate from the real user preference model, resulting in the limited range of improving the accuracy of recommendation. Recently, some prediction methods of user trust relationship have been proposed, but few researches have applied the prediction results of user social relationship to recommendation model [7, 13, 14, 21, 34, 37].

2.2 Matrix factorization model based on social relationships

The social MF has been become one of the most widely used methods in RSs due to accurate prediction and high efficiency. In addition, neighborhood relationships and trust relationships are two of the most commonly used social relationships, which are integrated into the MF process to obtain more accurate user feature and item feature [21, 22, 24, 32, 54].

2.2.1 Recommendation model based on neighborhood relationships

The MF approach can be depicted as a probabilistic graphical model, as shown in Figure 2. The MF method uses the potential relationship between users and items to decompose the user-item rating matrix into two low-dimensional matrices: users' preference feature matrix and items' attribute feature matrix. The two matrices are mapped into the same latent factor space, and the unknown ratings are predicted according to the degree of matching between users' preferences and items' attributes.

For the user-item rating matrix R , MF techniques can be used to map them to the user and item latent feature spaces, respectively. The rating information R is composed of inner product of the user and the neighbor user and the item feature, and the probability graph model is shown in Figure 2. In order to obtain a more accurate personalized preference, it is assumed that each user obeys different priori variances. The more a user's rating is, the more accurate his estimated preference is, and vice versa, the greater the uncertainty of his preference is. Then the Gaussian prior distribution of the user features is as follows [8, 14, 39].

$$p(U | \sigma_F^2) = \prod_{u=1}^N N(U_u | 0, \frac{\sigma_F^2}{n_{u_u}} I) \quad (3)$$

where n_{u_u} denotes the number of ratings from the user u . According to Bayesian inference, the feature matrices U and V can be obtained by minimizing the following formula [39].

$$\begin{aligned} L(R, U, V) = & \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M (r_{ui} - g(\alpha U_u^T V_i + (1-\alpha) \sum_{k \in N_u} S_{uk} U_k^T V_i))^2 \\ & + \frac{\lambda_U}{2} \sum_{u=1}^N n_{u_u} U_u^T U_u + \frac{\lambda_V}{2} \sum_{i=1}^M n_{v_i} V_i^T V_i \end{aligned} \quad (4)$$

where N_u represents the set of neighbors of the user u , and S_{uk} denotes the similar relationship between users u and k .

2.2.2 Recommendation model based on trust relationships

In [14, 28], user interest preference models are built from the perspectives of users' trustees and trustees, respectively. The modeling idea is based on the following considerations: a user u 's trustee is likely to have the same or similar preference interests as the user u , so these trustee users have similar feature vectors as the user u . Similarly, when some users are trusted by the same user u , these users have similar characteristics as the user u . Therefore, the user preference model is described as follows [7]:

$$\hat{r}_{ui} = \alpha B_u^T V_i + (1-\alpha) E_u^T V_i \quad (5)$$

where B_u indicates the preference characteristics of the user u 's trustees for the item i , E_u indicates the preference characteristics of the user u 's trusted person for the item i , and V_i indicates the feature vector of the item i .

Figure 2 Probability graph model for recommendation based on social relationship enhancement

The mutual trust relationship between the two users can be represented by the normal distribution of the user's truster feature vector B_u and the trusted user's feature vector E_v , and the truster and the trusted user's feature vector are assumed to obey Gauss distribution with a mean of 0 and variances of σ_B^2 and σ_E^2 , respectively [7, 8]. According to the Bayesian formula, we can get the following loss function [7, 8, 15]:

$$L(T, B, E) = \frac{1}{2} \sum_{u=1}^N \sum_{v=1}^N (T_{uv} - B_u^T E_v)^2 + \frac{\lambda_B}{2} \sum_{u=1}^N n_{b_u} \|B\|_F^2 + \frac{\lambda_E}{2} \sum_{v=1}^N n_{e_v} \|E\|_F^2 \quad (6)$$

where T_{uv} represents the trust relationship between users u and v .

2.2.3 Recommendation model based on social tags

If there is no explicit trust relationship between the users, since the user's annotation and comment information on the item reflects the user's preference for the item to some extent, the tag weight information of users and items can be used to express implicit social relationships between the users and between the items. In order to better measure the social tag weights to reflect the user's personalized preference characteristics, each tag is considered to be subject to a different a priori variance. The Gaussian prior distribution of the tag feature is as follows:

$$p(L | \sigma_L^2) = \prod_{t=1}^D N(L_t | 0, \frac{\sigma_L^2}{n_{l_t}} I) \quad (7)$$

where n_{l_t} indicates the number of tags that are marked by the user u . G_{up} and H_{ip} indicate the weight relationships between the user u and between the item i for the tag p , respectively, which can be obtained by TF-IDF as follows [10, 12]:

$$G_{up} = \frac{c_{up}}{\sum_{p \in K} c_{up}} * \log_2 \frac{N_t}{c_p^{(u)} + 1} \quad (8)$$

$$H_{ip} = \frac{c_{ip}}{\sum_{p \in K} c_{ip}} * \log_2 \frac{N_t}{c_p^{(i)} + 1} \quad (9)$$

where c_{up} denotes the times that the user u selects the tag p , $c_p^{(u)}$ denotes how many users have chosen the tag p , c_{ip} indicates the times the tag p appears in the item i , and $c_p^{(i)}$ indicates how many items the tag p appears in. The more times a user selects the tag, and the few users use the tag, indicating that the tag should have a greater weight for the user. The more times a tag appears in an item set, the few times appear in other item set, and for the item set the more important the tag is, indicating the tag should have a greater the weight for the item.

Inspired by [11, 32, 34, 39, 41], the features of the user and the item can be obtained using the latent semantic model, respectively, and the conditional probabilities of the user-tag and item-tag weights based on the user-tag weight and the item-

tag weight, respectively. The conditional probability distributions are as follows:

$$p(G | U, L, \sigma_G^2) = \prod_{u=1}^N \prod_{t=1}^D N(G_{ut} | g(U_u^T L_t), \sigma_G^2) \quad (10)$$

$$p(H | V, L, \sigma_H^2) = \prod_{i=1}^M \prod_{t=1}^D N(H_{it} | g(V_i^T L_t), \sigma_H^2) \quad (11)$$

where G is mapped to the user feature space U and the tag feature space L , respectively, and H is mapped to the item feature space V and the tag feature space L , respectively. According to the Bayesian formula, the posterior probability of G and H is described as follows:

$$\begin{aligned} & p(U, V, L | G, H, \sigma_L^2, \sigma_G^2, \sigma_U^2, \sigma_V^2) \\ & \propto p(G | U, L, \sigma_G^2) p(H | V, L, \sigma_H^2) p(U | \sigma_U^2 / n_{u_u}) p(V | \sigma_V^2 / n_{v_i}) p(L | \sigma_L^2 / n_{l_t}) \\ & = \prod_{u=1}^N \prod_{t=1}^D N(G_{ut} | g(U_u^T L_t), \sigma_G^2) \times \prod_{i=1}^M \prod_{t=1}^D N(H_{it} | g(V_i^T L_t), \sigma_H^2) \\ & \quad \times \prod_{u=1}^N N(U_u | 0, \frac{\sigma_U^2}{n_{u_u}}) \times \prod_{i=1}^M N(V_i | 0, \frac{\sigma_V^2}{n_{v_i}}) \times \prod_{t=1}^D N(L_t | 0, \frac{\sigma_L^2}{n_{l_t}}) \end{aligned} \quad (12)$$

After solving the above formula, it is equivalent to minimizing the following loss function.

$$\begin{aligned} L(U, V, L, G, H) = & \frac{1}{2} \sum_{u=1}^N \sum_{t=1}^D (G_{ut} - g(U_u^T L_t))^2 + \frac{1}{2} \sum_{i=1}^M \sum_{t=1}^D (H_{it} - g(V_i^T L_t))^2 \\ & + \frac{\lambda_U}{2} \sum_{u=1}^N n_{u_u} \|U\|_F^2 + \frac{\lambda_V}{2} \sum_{i=1}^M n_{v_i} \|V\|_F^2 + \frac{\lambda_L}{2} \sum_{t=1}^D n_{l_t} \|L\|_F^2 \end{aligned} \quad (13)$$

3. The proposed recommendation framework based on similarity feedback on user and item characteristics

This paper is based on the assumption that user preferences are influenced by users and their neighbors, trusted users, and social tags, first analyzes the impact of social tags on user preferences, and builds a comprehensive recommendation model using rating information, trust relationships, and social tag weights to obtain user and item features from two perspectives of explicit and implicit social relationships. At the same time, based on the framework of the literature [7, 8, 15, 19, 20, 23], the feature spaces of users and items are trained to improve the quality of recommendations.

The ISocialMF recommendation framework presented in this paper is shown in Figure 3. The recommendation method is divided into the following processes: (1) a social recommendation model based on explicit interaction information such as user's user ratings, direct trust relationship, and social tags is established; (2) user features, trustee users and item features are mapped to a shared space respectively using MF technology; (3) user weights and item weights are established using interaction relationships such as social tags; (4) the implicit feature space and the social tag weight information are combined to obtain the implicit similarity between users and between items, respectively; (5) the above social recommendation model is combined with the implicit similarity by using the SocialIT framework structure to establish an explicit and implicit relationships, and implicit feature spaces of users and items are obtained from the above model learning training, respectively; (7) the ratings are predicted and recommendations are generated.

3.1 Regularization of social relationships based on implicit interaction

In order to accurately estimate the similarity of users and the similarity of items, the social MF framework with explicit and implicit interactions is proposed by combining various social factors to map user and item features to a low-rank share space respectively in this section.

Figure 3. ISocialMF recommendation framework

We use the similarity of trust relationships between trustees and trusted users to constrain the trustee and trustee characteristics. As two trusted users, the closer the features of the two users are, the more similar their preferences are. Therefore, two truster feature vectors can be constrained by the feature similarity of the user's trustee as follows^[7].

$$\frac{\lambda_B}{2} \sum_{u=1}^N \sum_{v=1}^N P_{uv}^{(B)} \|B_u - B_v\|_F^2 \quad (14)$$

where $P_{uv}^{(B)}$ denotes that the similarity between the user u and the user v based on the trust relationship, which can be obtained from the common trust relationships between users u and v with their common trust user k . In addition, the higher the degrees of trust of a user by other users are, the more users will adopt his/her suggestion, and the greater the user's in-degree will be. According to [2, 28, 34, 41], the improved user similarity is described as follows:

$$P_{uv}^{(B)} = \frac{\sum_{k=1}^N \frac{d_k^{(in)}}{d_u^{(out)} + d_k^{(in)}} T_{uk} * \frac{d_k^{(in)}}{d_v^{(out)} + d_k^{(in)}} T_{vk}}{\sqrt{\sum_{k=1}^N \frac{d_k^{(in)}}{d_u^{(out)} + d_k^{(in)}} T_{uk}^2} \sqrt{\sum_{k=1}^N \frac{d_k^{(in)}}{d_v^{(out)} + d_k^{(in)}} T_{vk}^2}} \quad (15)$$

where $d_k^{(in)}$ and $d_u^{(out)}$ represent in-degree and out-degree in the trust network, respectively. Similarly, as a trusted user, the closer their characteristics are, the more similar the preferences of the two users are. Therefore, the constraint terms of the two trusted user feature vectors are as follows^[7]:

$$\frac{\lambda_E}{2} \sum_{u=1}^N \sum_{v=1}^N Q_{uv}^{(E)} \|E_u - E_v\|_F^2 \quad (16)$$

where $Q_{uv}^{(E)}$ denotes that the similarity of the trust relationship between two trusted persons. According to the literature [7, 8, 22], the improved user similarity based on the trustee is as follows:

$$Q_{uv}^{(E)} = \frac{\sum_{k=1}^N \frac{d_u^{(in)}}{d_k^{(out)} + d_u^{(in)}} T_{ku} * \frac{d_v^{(in)}}{d_k^{(out)} + d_v^{(in)}} T_{kv}}{\sqrt{\sum_{k=1}^N \frac{d_u^{(in)}}{d_k^{(out)} + d_u^{(in)}} T_{ku}^2} \sqrt{\sum_{k=1}^N \frac{d_v^{(in)}}{d_k^{(out)} + d_v^{(in)}} T_{kv}^2}} \quad (17)$$

In summary, considering the contribution of user ratings and trust relationships, the improved similarity between users is expressed as follows:

$$sim(u, v) = \begin{cases} \frac{2*(P_{uv}^{(B)} + Q_{uv}^{(E)})*sim_{uv}}{P_{uv}^{(B)} + Q_{uv}^{(E)} + 2*sim_{uv}}, & \text{if } P_{uv}^{(B)} \neq 0, Q_{uv}^{(E)} \neq 0, \text{ and } sim_{uv} \neq 0 \\ s_{uv}, & \text{else if } P_{uv}^{(B)} = 0, Q_{uv}^{(E)} = 0, \text{ and } sim_{uv} \neq 0 \\ \frac{2*P_{uv}^{(B)}*Q_{uv}^{(E)}}{P_{uv}^{(B)} + Q_{uv}^{(E)}}, & \text{else if } P_{uv}^{(B)} \neq 0, Q_{uv}^{(E)} \neq 0, \text{ and } sim_{uv} = 0 \\ 0, & \text{else} \end{cases} \quad (18)$$

Inspired by the literature [7], if a user trusts another user or a user is trusted by another, their features will be very close. Therefore, the truster and the trusted user have similar feature vectors respectively. So a regularization term is obtained as follows:

$$\frac{\beta}{2} \sum_{u=1}^N sim(u_{I_u^{(B)}}, u_{I_u^{(E)}}) \|B_u - E_u\|_F^2 \quad (19)$$

where B_u and E_u represent the user's set of trusted and trusted users, respectively.

3.2 Implicit similarities between users and between items

In order to accurately describe the degree of similarity between users, based on the literature [14, 27, 36], we analyze the similarity relationships among item similarity and social tag comprehensively, through the user's explicit interaction with the item (Information such as ratings and direct trust relationships) and implicit interactions (users' tag information for items) are optimized to obtain user and item low-rank feature vectors to establish user-preference relationships and item similarity relationships.

3.2.1 Improved user implicit similarity

According to [7, 8, 15, 23], user feature vectors, trust relationships and social tags can reflect the similarity between users to some extent. In [13, 103], user characteristics are measured directly, while the contribution of similarity from other social relationships is ignored. In [45], user characteristics and trust relationships are employed to measure similarity, but the similarity just add up linearly, without considering the impact of data imbalance and the weight of the different influencing factors. From [7, 8, 31], the comprehensive trade-off user characteristics, trust relationship characteristics and social label weight relationships can accurately measure user similarity. Among them, the user's rating similarity is calculated by the user's rating vector as follows [7, 8, 15].

$$S_{uv}^{(UR)} = sim(U_u, U_v) = U_u^T U_v \quad (20)$$

If the two trustee feature vectors are similar, their preferences are similar. Considering trusters and trusted users in a comprehensive way can avoid user feature deviation caused by data sparseness to more accurately reflect the user's preference characteristics. Therefore, the average similarity of preferences based on trust relationship is as follows:

$$S_{uv}^{(T)} = \frac{B_u^T B_v + E_u^T E_v}{2} \quad (21)$$

The implicit similarity between users is considered as a normal distribution consisting of user feature similarity, trust relationship implicit similarity and tag similarity. The conditional probability distribution is as follows:

$$p(S^{(U)} | S_{uv}^{(UR)}, S_{uv}^{(T)}, S_{uv}^{(UL)}, \sigma_S^2) = \prod_{u=1}^N \prod_{v=1}^N N[S_{uv}^{(U)} | g(f(S_{uv}^{(UR)}), S_{uv}^{(T)}, S_{uv}^{(UL)})), \sigma_S^2] \quad (22)$$

where the variance σ_S^2 represents the noise condition of the estimate, and the function $f(x)$ is an improved user similarity function. The relationship $\hat{S}_{uv}^{(U)}$ between users u and v can be expressed as follows:

$$\hat{S}_{uv}^{(U)} = \begin{cases} S_{uv}^{(UL)} \frac{2 * S_{uv}^{(R)} * S_{uv}^{(T)}}{S_{uv}^{(R)} + S_{uv}^{(T)}}, & \text{if } \text{relation}(u, v) \neq 0 \\ \frac{2 * S_{uv}^{(UL)} * S_{uv}^{(R)}}{S_{uv}^{(UL)} + S_{uv}^{(R)}}, & \text{else} \end{cases} \quad (23)$$

where $\text{relation}(u, v)$ indicates a direct trust relationship between users u and v . It can avoid the inaccuracy of user feature description caused by data sparseness and imbalance, and also solve the problem of deviating from real user characteristics caused by linear superposition of user similarity, without explicit user rating and trust relationship data. For user implicit preference similarity based on tags, it can be obtained by comparing the tag weights of the above two users to the item [7, 20].

$$S_{uv}^{(UL)} = \frac{\sum_{k=1}^D G_{uk} * G_{vk}}{\sqrt{\sum_{k=1}^D G_{uk}^2} \sqrt{\sum_{k=1}^D G_{vk}^2}} \quad (24)$$

Taking the logarithm of the posterior probability of Eq (24), the objective function is obtained as follows:

$$\begin{aligned} L(S, U, B, G) &= \sum_{u=1}^N \sum_{v \in N_u} (S_{uv}^{(U)} - g(f(S_{uv}^{(UR)}, S_{uv}^{(T)}, S_{uv}^{(UL)})))^2 \\ &= \begin{cases} \sum_{u=1}^N \sum_{v \in N_u} (S_{uv}^{(U)} - g(S_{uv}^{(UL)} \frac{2 * S_{uv}^{(UR)} * S_{uv}^{(T)}}{S_{uv}^{(UR)} + S_{uv}^{(T)}}))^2, & \text{if } \text{relation}(u, v) \neq 0 \\ \sum_{u=1}^N \sum_{v \in N_u} (S_{uv}^{(U)} - g(\frac{2 * S_{uv}^{(UR)} * S_{uv}^{(UL)}}{S_{uv}^{(UR)} + S_{uv}^{(UL)}}))^2, & \text{else} \end{cases} \end{aligned} \quad (25)$$

3.2.2 Item implicit similarity

If a user likes an item, the user will often like the item with similar characteristics. Therefore, we introduce the idea to improve the recommendation quality. Specially, it is assumed that item similarity consists of a normal distribution of item feature similarity and social tag relationship similarity as follows:

$$p(S^{(I)} | S_{ij}^{(IR)}, S_{ij}^{(IL)}, \sigma_S^2) = \prod_{i=1}^M \prod_{j=1}^M N[S_{ij}^{(I)} | g(f(S_{ij}^{(IR)}, S_{ij}^{(IL)})), \sigma_S^2] \quad (26)$$

According to Bayesian inference, the following loss function can be obtained as follows:

$$L(S, V, H) = \sum_{i=1}^M \sum_{j \in N_i} (S_{ij}^{(I)} - g(\hat{S}_{ij}^{(I)}))^2 \quad (27)$$

Among them, the similarity between items according to the comprehensive consideration of the item characteristics and social tag factors can be obtained as follows:

$$\hat{S}_{ij}^{(L)} = \frac{2 * S_{ij}^{(IL)} * S_{ij}^{(IR)}}{S_{ij}^{(IL)} + S_{ij}^{(IR)}} \quad (28)$$

Here, the similarity between items based on social tags is as follows [54].

$$S_{ij}^{(IL)} = \frac{\sum_{k=1}^D H_{ik} * H_{jk}}{\sqrt{\sum_{k=1}^D H_{ik}^2} \sqrt{\sum_{k=1}^D H_{jk}^2}} \quad (29)$$

Similarly, the similarity based on item characteristics is be calculated as follows:

$$S_{ij}^{(IR)} = \text{sim}(V_i, V_j) = V_i^T V_j \quad (30)$$

4. Social recommendation algorithm model integrating trust relationships and social tags

To fully exploit and accurately measure the potential complex social relationships between users and items in online social networks, inspired by [7, 8, 15, 20, 23, 29, 33, 36], the effects of user ratings, explicit and implicit interactions such as social tags and user trusts on recommendation quality are analyzed in depth, some classic recommendation frameworks such as SocialIT, RSTE and RoRec are integrated into the recommendation process, and an improved social recommendation algorithm namely ISocialMF fusing social tags and trust relationships is proposed in this paper. The method maps the rating information, social tags and user trust relationships to the low-rank user feature space, the item feature space and the tag feature space, respectively, and uses the optimized feature vector to obtain the user's implicit similarity and item implicit similarity. Then the user and the item feature vectors are trained continually, and the implicit similarity between the users and between the items are optimized by using various influencing factors to improve the accuracy of the recommendation. Social tag information is introduced into the process of establishing the recommendation model, the inaccuracy of recommendation caused by sparse data and imbalance is not only be mitigated, and the diversity of recommendations also can be improved.

4.1 ISocialMF algorithm model

Considering the impact of user trust relationship, rating information and social tags on user preference similarity and item similarity, combined with Eq.(4), Eq.(6), Eq.(13), Eq.(19), Eq.(25) and Eq.(30), the comprehensive social MF recommendation model is shown in Figure 4, and cost function with user and item feature regular terms are obtained as follows:

$$\begin{aligned} L(R, U, V, G, H, S, T) = & \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M I_{ui}^R (r_{ui} - g(\alpha U_u^T V_i + (1-\alpha) \sum_{k \in N_u} S_{uk}^{(U)} U_k^T V_i))^2 \\ & + \frac{1}{2} \sum_{u=1}^N \sum_{t=1}^D I_{ut}^G (G_{ut} - g(U_u^T L_t))^2 + \frac{1}{2} \sum_{i=1}^M \sum_{t=1}^D I_{it}^H (H_{it} - g(V_i^T L_t))^2 \\ & + \frac{\lambda_p}{2} \sum_{u=1}^M \sum_{v \in N_u} (S_{uv}^{(U)} - g(\hat{S}_{uv}^{(U)}))^2 + \frac{\lambda_T}{2} \sum_{u=1}^N \sum_{v=1}^N (T_{uv} - g(B_u^T E_v))^2 \\ & + \frac{\rho}{2} \sum_{u=1}^N \sum_{t \in N_u} \text{sim}(u, t) \|U_u - U_t\|_F^2 + \frac{\gamma}{2} \sum_{i=1}^M \sum_{j \in N_i} \text{sim}(i, j) \|V_i - V_j\|_F^2 \\ & + \frac{\lambda_S}{2} \sum_{u=1}^N \sum_{v=1}^N P_{uv}^{(B)} \|B_u - B_v\|_F^2 + \frac{\lambda_Z}{2} \sum_{u=1}^N \sum_{v=1}^N Q_{uv}^{(E)} \|E_u - E_v\|_F^2 \\ & + \frac{\beta}{2} \sum_{u=1}^M \text{sim}(u_{I_u^{(B)}}, u_{I_u^{(E)}}) \|B_u - E_u\|_F^2 + \frac{\lambda_Q}{2} \sum_{i=1}^M \sum_{j \in N_i} (S_{ij}^{(I)} - g(\hat{S}_{ij}^{(I)}))^2 \\ & + \frac{\lambda_W}{2} \sum_{j=1}^M (V_j - \sum_{i \in N_j} S_{ij}^{(I)*} V_i)^T (V_j - \sum_{i \in N_j} S_{ij}^{(I)*} V_i) + \frac{\lambda_U}{2} \sum_{u=1}^N n_{u_u} \|U\|_F^2 \\ & + \frac{\lambda_V}{2} \sum_{i=1}^M n_{v_i} \|V\|_F^2 + \frac{\lambda_L}{2} \sum_{t=1}^D n_{l_t} \|L\|_F^2 + \frac{\lambda_B}{2} \sum_{u=1}^N n_b \|B\|_F^2 \\ & + \frac{\lambda_E}{2} \sum_{v=1}^N n_{e_v} \|E\|_F^2 \end{aligned} \quad (31)$$

where $\text{sim}(u, t)$ denotes the similarity of users u_u and u_t , see Eq(12), and $\text{sim}(i, j)$ indicates the similarity between items i_i and i_j . n_{b_u} and n_{e_v} denote the number of users trusted by user u_u and the number of users who trust user u_u , respectively. In Figure 4, the user feature space is constrained by the similarity relationship S_{uv} composed of the user feature U , the trust relationship features B and E , and the tag weight relationship G , and the similarity relationship S_{ij} composed of the item feature V and the social tag weight relationship H is used to constrain the items.

4.2 Model learning

For the above objective function, the gradients of U_u , V_i , B_u , E_v , $S_{uv}^{(U)}$, $S_{ij}^{(I)}$ and L_t can be calculated by the stochastic gradient descent method to obtain the local minimum as follows:

$$\begin{aligned}
\frac{\partial L}{\partial U_u} = & \alpha \sum_{i=1}^M I_{ui}^R V_i g'(\alpha U_u^T V_i + (1-\alpha) \sum_{k \in N_u} S_{uk}^{(U)} U_k^T V_i) \\
& \times (g(\alpha U_u^T V_i + (1-\alpha) \sum_{k \in N_u} S_{uk}^{(U)} U_k^T V_i) - r_{ui}) + \lambda_u n_{u_u} U_u \\
& + \rho \sum_{t \in N_u} \text{sim}(u, t) (U_u - U_t) + \lambda_p \sum_{v \in N_u} ((S_{uv}^{(U)} - g(\hat{S}_{uv}^{(U)})) g'(\hat{S}_{uv}^{(U)}) U_u' \\
& + \lambda_p \sum_{\{u|u \in N_v\}} ((S_{vu}^{(U)} - g(\hat{S}_{vu}^{(U)})) g'(\hat{S}_{vu}^{(U)}) U_u' \\
& + \sum_{t=1}^D I_{ut}^G L_t (G_{ut} - g(U_u^T L_t)) g'(U_u^T L_t)
\end{aligned} \tag{32}$$

Figure 4 Probability graph model of ISocialMF

$$\begin{aligned}
\frac{\partial L}{\partial V_i} = & \alpha \sum_{u=1}^N I_{ui}^R g'(\alpha U_u^T V_i + (1-\alpha) \sum_{k \in N_u} S_{uk}^{(U)} U_k^T V_i) \\
& \times (g(\alpha U_u^T V_i + (1-\alpha) \sum_{k \in N_u} S_{uk}^{(U)} U_k^T V_i) - r_{ui}) \\
& \times (\alpha U_u + (1-\alpha) \sum_{k \in N_u} S_{uk}^{(U)} U_k^T) + \gamma \sum_{j \in N_i} \text{sim}(i, j) (V_i - V_j) \\
& + \lambda_W (V_i - \sum_{k \in N_i} S_{ik}^{(I)} U_k) - \lambda_W \sum_{\{k|i \in N_k\}} S_{ik}^{(I)} (V_i - \sum_{l \in N_i} S_{lk}^* U_k) \\
& + \lambda_Q \sum_{j \in N_i} ((S_{ij}^{(I)} - g(\hat{S}_{ij}^{(I)})) g'(\hat{S}_{ij}^{(I)}) V_i') \\
& + \lambda_Q \sum_{\{i|j \in N_j\}} ((S_{ji}^{(I)} - g(\hat{S}_{ji}^{(I)})) g'(\hat{S}_{ji}^{(I)}) V_i')
\end{aligned}$$

$$+ \sum_{t=1}^D I_{it}^H L_t g'(V_i^T L_t) (H_{it} - g(V_i^T L_t)) + \lambda_v n_{v_i} V_i \quad (33)$$

$$\begin{aligned} \frac{\partial L}{\partial B_u} &= \lambda_p (S_{uv}^{(U)} - g(\hat{S}_{uv}^{(U)})) g'(\hat{S}_{uv}^{(U)}) B_u' + \lambda_p (S_{vu}^{(U)} - g(\hat{S}_{vu}^{(U)})) g'(\hat{S}_{vu}^{(U)}) B_u' \\ &+ \lambda_T \sum_{v=1}^N (T_{uv} - g(B_u^T E_v)) g(B_u^T E_v) E_v + \lambda_B n_{b_u} B_u + \beta \text{sim}(u_{I_u^{(B)}}, u_{I_u^{(E)}}) (B_u - E_u) \\ &- \lambda_S P_{uv}^{(B)} (B_u - B_v) \end{aligned} \quad (34)$$

$$\begin{aligned} \frac{\partial L}{\partial E_v} &= \lambda_p (S_{uv}^{(U)} - g(\hat{S}_{uv}^{(U)})) g'(\hat{S}_{uv}^{(U)}) E_v' + \lambda_S (S_{vu}^{(U)} - g(\hat{S}_{vu}^{(U)})) g'(\hat{S}_{vu}^{(U)}) E_v' \\ &+ \lambda_T \sum_{v=1}^N (T_{uv} - g(B_u^T E_v)) g(B_u^T E_v) B_u + \lambda_E n_{e_v} E_v - \lambda_Z Q_{uv}^{(E)} (E_u - E_v) \end{aligned} \quad (35)$$

$$\begin{aligned} \frac{\partial L}{\partial S_{uv}^{(U)}} &= (g(\alpha U_p^T V_i + (1-\alpha) \sum_{k \in T(p)} S_{pu}^{(U)} U_k^T V_i) - r_{ui}) \\ &\times (1-\alpha) \left(\frac{U_k \sum_{k \in N_u} S_{uk} - \sum_{k \in N_u} S_{uk} U_k}{(\sum_{k \in N_u} S_{pu}^{(U)})^2} \right) + \lambda_p (S_{uv}^{(U)} - g(\hat{S}_{uv}^{(U)})) \end{aligned} \quad (36)$$

$$\frac{\partial L}{\partial S_{ij}^{(I)}} = \lambda_Q (S_{ij}^{(I)} - g(\hat{S}_{ij}^{(I)})) + \lambda_W (U_u - \sum S_{ij}^{(I)} U_k) \left(\frac{U_k \sum_{k \in N_i} S_{ik}^{(I)} - \sum_{k \in N_i} S_{ik}^{(I)} U_k}{(\sum_{k \in N_i} S_{ik}^{(I)})^2} \right) \quad (37)$$

$$\begin{aligned} \frac{\partial L}{\partial L_t} &= \sum_{u=1}^N I_{ut}^G g'(U_u^T L_t) (G_{ut} - g(U_u^T L_t)) U_u + \lambda_L n_{l_t} L_t \\ &+ \sum_{i=1}^M I_{it}^H g'(V_i^T L_t) (H_{it} - g(V_i^T L_t)) \end{aligned} \quad (38)$$

Here, U_u' , B_u' , and E_v' are the partial derivatives of U_u , B_u , and B_v , respectively.

Algorithm 1 Social recommendation algorithm that integrates trust relationships and social tags (ISocialMF).

Input : user-item rating matrix R , trust relationship matrix T between users, user-tag and item-tag relationship matrix G and H , respectively, latent feature dimension K , regularization parameters U , V , T , P , Q , W , S , L and learning rate η .

Output : user latent factor feature vector U , item latent factor feature vector V , truster feature vector B , trusted user feature vector E , user-to-user similarity relationship matrix $S^{(U)}$ and item similarity relationship matrix $S^{(I)}$.

- 1 Initialize U , V , B , E , S .
- 2 Construct user and item weight relationship matrices G and H based on the user's annotation behavior for the item.
- 3 Get social tag-based user similarity and item similarity according to Eqs. (26) and (31).

```

4  while  $t < \text{maxIter}$ 
5       $L_{old} \leftarrow L$ ;
6      for  $r_{ui} \in R$ 
7           $U_u \leftarrow U_u - \eta \frac{\partial L}{\partial U_u}, V_i \leftarrow V_i - \eta \frac{\partial L}{\partial V_i}$ 
8           $B_u \leftarrow B_u - \eta \frac{\partial L}{\partial B_u}, E_v \leftarrow E_v - \eta \frac{\partial L}{\partial E_v}$ 
9           $S_{uv}^{(U)} \leftarrow S_{uv}^{(U)} - \eta \frac{\partial L}{\partial S_{uv}^{(U)}}, S_{ij}^{(I)} \leftarrow S_{ij}^{(I)} - \eta \frac{\partial L}{\partial S_{ij}^{(I)}}$ 
10          $L_t \leftarrow L_t - \eta \frac{\partial L}{\partial L_t}$ 
11         Obtain the user feature, trust relationship and social tag according to Eqs.(22),(23),and (26).
12         Obtain the item similarities based on social tag and item feature according to Eqs(31),and (32).
13         Obtain the user and item implicit similarities according to Eqs(27),and (29).
14     end for
15     Update L according to Eq(30);
16     if  $|L - L_{old}| < \text{threshold}$  then
17         break;
18     end if
19 end while
20  $t \leftarrow t + 1$ ;
21  $\hat{r}_{ui} \leftarrow U_u^T V_i$ 

```

5 Experimental results and analysis

In this section, some experiments are performed on the Epinions and Douban datasets, and the performance of the algorithm is compared with mainstream social recommendation algorithms to evaluate the effectiveness of our model.

5.1 Datasets and evaluation indicators

To verify the effectiveness of our algorithm, we select two popular social network datasets, such as Epinions and Douban. For the Epinions dataset, user relationships are directed; for Douban datasets, the relationship between users is undirected. The statistics for our extracted Epinions and Douban datasets are shown in Table 1.

We first use MAE and RMSE evaluation indicators to evaluate the performance of each algorithm. In addition, we also use the evaluation indicators P@N, R@N and NDCG commonly used in the Top-N RS related to the actual scene to evaluate the algorithm more comprehensively.

NDCG is an evaluation index for measuring the quality of recommendation ranking. It considers the relevance and ranking position of all recommended items, and its definition is as follows:

$$NDCG = \frac{1}{N} NDCG(i) = \frac{1}{N} \sum_{i=1}^N \frac{1}{Z_i} \sum_{j \in \Omega(i)} \frac{2^{r_{ij}} - 1}{\log_2(1 + \text{rank}(j))} \quad (39)$$

Among them, $\text{rank}(j)$ represents the position of the list of items v_j in the sorted set $\Omega(i)$, Z_i is a parameter introduced for

standardization $\sum_{j \in \Omega(i)} \frac{2^{r_{ij}} - 1}{\log_2(1 + \text{rank}(j))}$. The value of NDCG is in the interval $[0, 1]$, and the larger the value is, the better

the recommended sorting effect is.

Table 1 Statistics for Epinions and Douban Datasets

Datasets	Epinions	Douban
Number of users	12630	5786
Number of items	3620	26573
Number of ratings	1261218	685936
Number of social relationships	487183	2865
Average number of ratings per user	99.85	118.55
Average number of ratings per item	348.40	25.81
Density of trust network	0.00201	0.00378

5.2 Effect of parameters

To verify the impact of social tags on the performance of the algorithm, we divide the training set into five groups according to the number of tags to be tagged: “=0”, “1~5”, “6~10”, “11~20” and “>20”, then the experiment is conducted. The performance of the ISocialMF algorithm on MAE and RMSE at different numbers of tags is shown in Figure 5.

Figure 5 The MAE performance of the algorithm under different numbers of tags

As can be seen from Figure 5, since the ISocialMF algorithm takes into account the user's tag information for the item, the MAE performance of the recommendation algorithm is significantly improved on both the Epinions and Douban datasets. As the number of tags increases, the performance of the RS increases significantly, and when the number of tagged labels reaches 30, the performance tends to be stable. This phenomenon indicates that when the number of tags is increased to a certain threshold, the user's preferences and item characteristics can be expressed accurately. As the number of tags continues to increase, tag information becomes redundant and there is no significant improvement in the accuracy of the recommendations.

In the ISocialMF algorithm, the parameters α and λ_p are important parameters that affect the recommendation performance. Among them, α controls the degree of contribution of the user and neighboring users to the preference prediction, and the range of which is $[0, 1]$. The user's own behavior is mainly used for prediction when $\alpha = 0$, and the behavior of the neighboring users is mainly used for prediction when $\alpha = 1$. Figure 6 depicts the effect of parameter α on MAE and RMSE for the Epinions and Douban datasets, respectively. The effects of the parameter α on MAE and RMSE are very similar. As α increase, the values of MAE and RMSE decrease first, then gradually increase, and finally become stable. For Epinions and Douban datasets, the prediction error is minimal. This phenomenon indicates that it is more accurate to

consider comprehensively the behavior data of users and their neighbors.

Figure 6 Impact of parameter α on MAE and RMSE

Figure 7 Impact of Parameters λ_p on MAE

Figure 7 shows the effect of parameter λ_p on MAE. The role of λ_p is to control the impact of user implicit similarity on recommendation performance. The larger the value is, the greater the role of neighboring users in predicting the rating process is, and the more the user characteristics depend on neighboring users. On the Epinions and Douban datasets, as λ_p increases, the MAE value decreases first, then increases slowly until stabilizes. The prediction accuracy of ISocialMF on both datasets is optimal.

5.3 Impact of Different Sparse Degrees of Trust Relationship on Recommendation Performance

The sparsity of trust relationships is also a key factor that directly affects the quality of social recommendation algorithms. To verify the robustness of the ISocialMF algorithm when trust relationships are sparse, we first classify all users according to their trust association degree (outdegree + indegree), and then evaluate the MAE performance of the algorithm. The distribution of each group of users and the average number of ratings are shown in Figure 8. The proportion of the number of user relationships with connection relationship range of 0~5 and 5~10 is the highest, the proportion of connection relationships

exceeding 200 is the least, but the number of users who rated items is the largest.

Figure 8 The social relationship distribution of each user group

Figure 9 The MAE performance of users in different groups for each algorithm

Figure 9 shows the MAE performance of each algorithm at different social connection densities. On the Epinions and Douban datasets, ISocialMF performs optimally. The performance trends of several other social recommendation algorithms are the same. As the number of social relationships increases, the recommendation quality also increases. When the number of social relationships increases to a certain extent, the recommendation quality begins to decline and finally become stable. Only the recommendation performance of the PMF algorithm is gradually improved since the algorithm does not consider the influence of social relationships, and the rating data also increases as the number of social relationships increases. Therefore, the performance of the PMF is improving constantly, and finally tends to be stable. Therefore, with the increase of the number of social relationships, although the number of some friends has increased, many of these relationships are only casual. There is no substantial trust relationship and common interest preference between them, these casual social relationship connections will reduce their ability to learn the user's true interest preference characteristics, so the recommendation algorithm does not achieve better recommendation quality.

5.4 Performance comparison and analysis

In order to evaluate the performance of the ISocialMF algorithm proposed in this paper, the algorithm is compared with social recommendation methods such as PMF^[23], TrustPMF^[14], SocialMF^[18], PRM^[31], EnSocialMF^[25], USSHMF^[28], and SocialIT^[3].

5.4.1 Performance of the algorithm on sparse datasets

Table 2 describes the MAE and RMSE performance comparisons of different algorithms on the Epinions and Douban data sets. On both datasets, the performance of the ISocialMF algorithm on MAE and RMSE exceeded that of other algorithms. Compared with other social network-based recommendation algorithms, the traditional PMF algorithm has the worst MAE and RMSE performance. For Douban datasets with relatively dense data, the ISocialMF algorithm has a smaller increase than PMF and other social-based collaborative filtering algorithms. For the Epinions dataset, the association between users is directed. The ISocialMF algorithm can accurately describe the relationship between users. For the Douban dataset, the relationship between users is a two-way friend relationship, so the performance of the ISocialMF algorithm is in Epinions. The increase in the data set is large.

Table 2 MAE and RMSE Performance of Each Algorithm on Epinions and Douban Datasets

Datasets	Metrics	PMF	TrustPMF	SocialMF	PRM	EnSocialMF	USSHMF	SocialIT	ISocialMF
Epinions	MAE	0.816	0.708	0.766	0.742	0.716	0.711	0.698	0.672
	Increase	17.63%	5.08%	12.27%	9.43%	6.15%	5.48%	3.72%	—
	RMSE	1.086	0.873	0.994	0.921	0.895	0.887	0.865	0.836
	Increase	23.02%	4.23%	15.89%	9.23%	6.59%	5.75%	3.35%	—
	MAE	0.665	0.591	0.609	0.605	0.601	0.597	0.582	0.571
Douban	Increase	14.13%	3.38%	6.24%	5.62%	4.99%	4.35%	1.89%	—
	RMSE	0.813	0.742	0.775	0.767	0.761	0.748	0.726	0.718
	Increase	11.69%	3.23%	7.35%	6.39%	5.65%	4.01%	1.10%	—
	MAE	0.665	0.591	0.609	0.605	0.601	0.597	0.582	0.571

Table 3 shows the P@N, R@N, and NDGG performance of each algorithm on the Epinions and Douban data sets. The performance of the ISocialMF algorithm in this paper is significantly improved compared to other social relationship-based recommendation algorithms, which indicates that the algorithm comprehensively considers the similarity of the ratings, the implicit similarity of the trust relationship and the similarity of the label weight to effectively improve the recommendation performance of Top-N. Although the SocialIT algorithm optimizes the user feature space by using feedback from user ratings and trust relationships, it improves the recommendation quality by mining implicit social relationships between users, but this method does not consider the impact of social tag implicit social relationships and item similarity. Compared with other social recommendation algorithms, the USSHMF and EnSocialMF algorithm models improve the recommendation performance to a certain extent, but lack the training of the model because of its direct modeling of social relationships. Their user preferences may deviate from real user preferences, therefore affect the quality of recommendations. The experimental results show that our method can improve the recommendation quality effectively by mining the implicit information of social relationships and social tags to model user preference interests, and training user features and item features for obtaining accurate user similarity and item similarity.

Table 3 Comparison of Top-N performance of each algorithm on Epinions and Douban datasets

Datasets	Metrics	PMF	TrustPMF	SocialMF	PRM	EnSocialMF	USSHMF	SocialIT	ISocialMF
Epinions	$P@5$	0.342	0.441	0.389	0.412	0.421	0.436	0.446	0.461
	$P@10$	0.359	0.458	0.418	0.431	0.437	0.451	0.467	0.478
	$R@5$	0.215	0.257	0.227	0.239	0.245	0.258	0.262	0.273
	$R@10$	0.226	0.273	0.239	0.246	0.253	0.261	0.278	0.295
	$NDCG$	0.422	0.485	0.438	0.453	0.462	0.476	0.492	0.508
Douban	$P@5$	0.357	0.407	0.379	0.383	0.386	0.397	0.415	0.432
	$P@10$	0.333	0.375	0.348	0.351	0.357	0.363	0.387	0.405
	$R@5$	0.226	0.281	0.257	0.262	0.269	0.273	0.292	0.313
	$R@10$	0.275	0.323	0.292	0.308	0.313	0.321	0.346	0.367
	$NDCG$	0.424	0.486	0.459	0.466	0.472	0.481	0.498	0.515

5.4.2 Performance of the algorithm on cold start users

Cold start is a challenge for CF-based recommendation algorithms, and we compare our algorithm with other social relationship algorithms in terms of cold start. Here we define users who rate less than 5 items as cold start users. Table 4 shows the recommendation performance of various recommendation algorithms on cold-start users. It can be seen that the performance of the ISocialMF algorithm on the Epinions dataset is better than that on the Douban dataset, and for cold start users, the performance is more significantly improved on the Epinions dataset.

Table 4 MAE and RMSE Performance of Each Algorithm on Cold-Start Users

Datasets	Metrics	PMF	TrustPMF	SocialMF	PRM	EnSocialMF	USSHMF	SocialIT	ISocialMF
Epinions	MAE	1.213	0.823	0.965	0.927	0.827	0.835	0.816	0.787
	Increase	35.12%	4.37%	18.44%	15.10%	4.83%	6.08%	3.55%	—
	RMSE	1.506	1.098	1.281	1.223	1.108	1.130	1.053	0.998
	Increase	33.73%	9.11%	22.09%	18.40%	9.93%	11.68%	5.23%	—
Douban	MAE	1.123	0.845	0.916	0.881	0.842	0.851	0.832	0.813
	Increase	27.60%	3.78%	11.24%	7.72%	3.44%	4.46%	2.28%	—
	RMSE	1.406	1.042	1.147	1.108	1.045	1.058	1.032	1.021
	Increase	27.38%	2.02%	10.98%	7.85%	2.29%	3.50%	1.07%	—

6. Conclusion and future work

Data sparsity and cold start are major challenges of the CF-based RS. This paper further eases the impact of the above problems on RS by introducing trust relationship and user tag information into the MF-based CF method. In this paper, we improve the social relationship measurement method using social tags and trust relationships, and propose a comprehensive social recommendation method, namely ISocialMF. The biggest advantage of ISocialMF is to integrate the social tags and item's social relationship into the MF process, and to train the shared user and item features for more accurate recommendation quality through continuous feedback mechanism in the process of matrix factorization. Finally, the proposed model is applied to different datasets to verify its effectiveness. The experimental results show that the ISocialMF algorithm proposed in this paper is superior to other algorithms in recommendation accuracy, which verifies the correctness of establishing the

recommendation model using explicit and implicit social relationships to train user and item similarities.

Declaration

The content of this article has not been published, nor has it been submitted to other journals. There is no conflict of interest in the content of this article. With the consent of all the authors, this article is authorized to publish. The contributions of each author in this article are as follows: Dr. Rui Chen is responsible for the writing of the article, Prof. Jian-wei Zhang is responsible for the revision of the paper, Prof. Zhifeng Zhang has put forward many valuable suggestions for this article, Dr. Jingli Gao verified the method and experiment, Dr. Pu Li completed the experiment of the paper, and Prof. Hui Liang revised the grammar of the paper. The publication of this article is supported by the following funds.

Acknowledgments

We would like to thank the anonymous reviewers and editor for their helpful comments. This work was supported in part by the National Natural Science Foundation of China under Grants 61672471, 61975187, and 61802352, in part by the Industrial Science and Technology Research Project of Henan Province under Grants 202102210387, 202102210178, 202102210397, 212102210410, 212102210418, 212102310556, and 182102310969, in part by the Research Project of Xi'an "Silk Road" Research Institute under Grant 2019YA07, and in part by the Doctoral Fund of Zhengzhou University of light industry under Grant 2020BSJJ030.

References

1. Yu W, Li S (2018) Recommender systems based on multiple social networks correlation. *Future Generation Computer Systems*, 87:312-327.
2. Yang X, Guo Y, Liu Y, Steck H (2014) A survey of collaborative filtering based social recommender systems. *Computer Communications*, 41:1-10.
3. Zheng X, Luo Y, Sun L (2017) A novel social network hybrid recommender system based on hypergraph topologic structure. *World Wide Web-internet & Web Information Systems*, 2017, (3):1-29.
4. Bagher R C, Hassanpour H, Mashayekhi H (2017) User trends modeling for a content-based recommender system. *Expert Systems with Applications*, 87:209-219.
5. Yu Y, Gao Y, Wang H, et al (2018) Integrating user social status and matrix factorization for item recommendation. *Journal of Computer Research and Development*, 55(1):113-124.
6. Ma H, King I, Lyu R. M., et al (2009) Learning to recommend with social trust ensemble. *Proceedings of 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2009:1-8.
7. Yao W, He J, Huang G, et al (2014) Modeling dual role preferences for trust-aware recommendation. *Proceedings of International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 2014:975-978.
8. Yang B, Yu L, Liu J, et al (2017) Social collaborative filtering by trust. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(8):1633-1647.
9. Wang Y, Wang X, Zuo W (2014) Trust prediction modeling based on social theories. *Journal of Software*, 25(12):2893-2904.
10. Tang J, Gao H, Hu X, et al (2013) Exploiting homophily effect for trust prediction. *Proceedings of ACM International Conference on Web Search and Data Mining (WSDM)*. ACM, 2013:53-62.
11. Rafailidis D, Crestani F (2017) Learning to rank with trust and distrust in recommender systems. *Proceedings of the 11th ACM Conference on Recommender Systems (RecSys'17)*. ACM, 2017:5-13.
12. Ahmadian S, Afsharchi M, Meghdadi M (2018) An effective social recommendation method based on user reputation model and rating profile enhancement. *Journal of Information Science*, 28:180-192, DOI: 10.1177/0165551518808191.
13. Can U, Alatas B (2019) A new direction in social network analysis: Online social network analysis problems and applications. *Physica A Statistical Mechanics and its Applications*, 535:122372.
14. Portugal I, Alencar P, Cowan D (2018) The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems With Applications*, 97:205-227.
15. Pan Y, He F, and Yu H (2018) Social recommendation algorithm using implicit similarity in trust. *Chinese Journal of Computers*, 41(1):65-81.
16. Nabizadeh A, Leal J, Rafsanjani H, et al (2020) Learning path personalization and recommendation methods: A survey of the state-of-the-art. *Expert Systems With Applications*, 2020,159(9):1767-1776. DOI:10.1016/j.eswa.2020.113596
17. Guan J, Xu M, Kong X (2018) Learning social regularized user representation in recommender system. *Signal Processing*, 144(3):306-310.
18. Jamali M, Ester M (2010) A matrix factorization technique with trust propagation for recommendation in social networks. *Proceedings of the 4th ACM Conference on Recommender Systems (RecSys'10)*, ACM, 45:26-30.

19. Meng X, Liu S, Zhang Y, et al (2015) Research on social recommender systems. *Journal of Software*, 26(6):1356-1372.
20. Cao Y, Li W, Zheng D (2018) An improved neighborhood-aware unified probabilistic matrix factorization recommendation. *Wireless Personal Communications*, (4):1-20.
21. Sambhav Y, Vikash, Shreyam, et al (2018) Trust aware Recommender System using Swarm Intelligence. *Journal of Computational Science*.
22. Ma H, Yang H, Lyu M R, et al (2008) SoRec: social recommendation using probabilistic matrix factorization. *Proceedings of ACM Conference on Information & Knowledge Management (CIKM'08)*, 2008:931-940.
23. Salakhutdinov R, Mnih A (2008) Probabilistic matrix factorization. In *Proceedings of NIPS*.
24. Chen R, Hua Q, Chang Y, et al (2018) A survey of collaborative filtering-based recommender systems: from traditional methods to hybrid methods based on social networks. *IEEE Access*, 6(1):64301-64320.
25. Luo X, Zhou M, Li S, et al (2018) An inherently nonnegative latent factor model for high-dimensional and sparse matrices from industrial applications. *IEEE Transactions on Industrial Informatics*, 14(5):2011-2022.
26. Jyoti S, Chhavi R (2019) Social recommender systems: techniques, domains, metrics, datasets and future scope. *Journal of Intelligent Information Systems*, 18:83-94. <https://doi.org/10.1007/s10844-019-00578-5>
27. Pereira B L, Ueda A, Penha G, et al (2019) Online learning to rank for sequential music recommendation. *Proceedings of the 13th ACM Conference on Recommender Systems (RecSys'19)*. ACM, Copenhagen, Denmark. New York, NY, USA.
28. Guo L, Ma J, Chen Z, et al (2015) Learning to recommend with social context information from implicit feedback. *Soft Computing*, 19(5):1351-1362.
29. Li Y, Wang D, He H, et al (2017) Mining intrinsic information by matrix factorization-based approaches for collaborative filtering in recommender systems. *Neurocomputing*, 249: pp.48-63.
30. Lu Q, Guo F (2019) Personalized information recommendation model based on context contribution and item correlation [J]. *Measurement*, 142:30-39.
31. Seo Y, Kim Y, Lee E, et al (2017) Personalized recommender system based on friendship strength in social network services. *Expert Systems With Applications*, 69: 135-148.
32. Yu Y, Gao Y, Wang H (2018) Integrating user social status and matrix factorization for item recommendation. *Journal of Computer Research & Development*, 55(1):113-124.
33. Liu H, Jing L, Yu J (2017) Survey of matrix factorization based recommendation methods by integrating social information. *Journal of Software*, 2017:1-24, DOI:10.13328/j.cnki.jos.005391.
34. He C, Parra D, Verbert K (2016) Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. *Expert Systems With Applications*, 56(9): 9-27.
35. Gao Q, Gao L, Fan J, et al (2016) A preference elicitation method based on bipartite graphical correlation and implicit trust. *Neurocomputing*, 237:92-100.
36. Aghdam M H. (2019) Context-aware recommender systems using hierarchical hidden Markov model. *Physica A: Statistical Mechanics & Its Applications*, 518:89-98.
37. Li H, Ma X, Shi J (2018) Incorporating trust relation with PMF to enhance social network recommendation performance. *International Journal of Pattern Recognition and Artificial Intelligence*, 30(6): 113-124.
38. He Y, Wang C, Jiang C J (2018) Correlated matrix factorization for recommendation with implicit feedback. *IEEE Transactions on Knowledge & Data Engineering*, 2018:1-15.
39. Herce-Zelaya J, Porcel C, Bernabe-Moreno J, et al (2020) New technique to alleviate the cold start problem in recommender systems using information from social media and random decision forests. *Information Sciences*, 53:156-170.
40. Ahmadian S, Joorabloo N, Jalili M, et al (2020) A social recommender system based on reliable implicit relationships. *Knowledge Based Systems*, 192:1-17.
41. Cohen D, Aharon M, Koren Y, et al (2017) Expediting exploration by attribute-to-feature mapping for cold-start recommendations. *Proceedings of the 11th ACM Conference on Recommender Systems (RecSys'17)*, 2017:184-192.
42. Zhang Z, Xu G, Zhang P, et al (2017) Personalized recommendation algorithm for social networks based on comprehensive trust. *Applied Intelligence*.
43. Zhang W, Du Y, Yang Y, et al (2018) DeRec: A data-driven approach to accurate recommendation with deep learning and weighted loss function. *Electronic Commerce Research Applications*, 31:12-23.
44. Wei J, He J, Chen K, et al (2017) Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems With Applications*, 69: 29-39.
45. Azadjalal M M, Moradi P, Abdollahpour A, et al (2017) A trust-aware recommendation method based on Pareto dominance and confidence concepts. *Knowledge-Based Systems*, 116:130-143.
46. Kong X, Mao M, Wang W (2018) VOPRec: Vector representation learning of papers with text information and structural identity for recommendation. *IEEE Transactions on Emerging Topics Computing*, DOI: 10.1109/TETC.2018.2830698.
47. Rezaeimehr F, Moradi P, Ahmadian S (2017) TCARS: Time- and community-aware recommendation system. *Future Generation Computer Systems*, DOI:10.1016/j.future.2017.04.003.
48. Sedhain S, Menon A K, Sanner S, Xie L, Brazhunas D (2017) Low-rank linear cold-start recommendation from social data. *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. AAAI Press.
49. Seo Y D, Kim Y G, Lee E, et al (2017) Personalized recommender system based on friendship strength in social network services. *Expert Systems With Applications*, 69:135-148.
50. Paradarami T. K., Bastian N. D., Wightman J. L (2017) A hybrid recommender system using artificial neural networks. *Expert System With Applications*, 83:300-313.
51. Yan S, Lin K J, Zheng X, et al (2017) An approach for building efficient and accurate social recommender systems using individual relationship networks. *IEEE Transactions on Knowledge and Data Engineering*, 29(10):2086-2099.
52. Ahmadian S, Joorabloo N, Jalili M, et al (2020) A social recommender system based on reliable implicit relationship [J]. *Knowledge-Based Systems*, 192:1-17. <https://doi.org/10.1016/j.knosys.2019.105371>

- 53. Feng S, Cao J, Wang J, et al (2017) Recommendations based on comprehensively exploiting the latent factors hidden in items' ratings and content. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 11(3):35-46.
- 54. Ahmadian S , Joorabloo N , Jalili M , et al (2019) A social recommender system based on reliable implicit relationships. *Knowledge-Based Systems*, 192:1-17.
- 55. Gong C , Tao D , Chang X , et al (2019) Ensemble Teaching for Hybrid Label Propagation. *IEEE Transactions on Cybernetics*, 49(2):388-402.
- 56. Gupta S, Kant V (2020) Credibility score based multi-criteria recommender system. *Knowledge-Based Systems*, 2020, 196(1): 1-12.
- 57. Shneiderman B (2020) Human-centered artificial intelligence: reliable, safe & trustworthy. *International Journal of Human - Computer Interaction*, 36: 495-504, DOI:10.1080/10447318.2020.1741118

RUI CHEN received his M.S., and PhD degrees in computer application technology from Northwest University, Xi'an, China, where he is currently a lecturer at Zhengzhou University of Light Industry. His research interests include recommender systems, social network analysis, human-computer interaction, and natural language processing.

JIAN-WEI ZHANG received the B.Sc. degree from Zhengzhou University, Zhengzhou, China, and Ph.D. degree from PLA Information Engineering University, Zhengzhou, China, in 2010. He is currently a professor, master tutor with the Zhengzhou University of Light Industry, Zhengzhou, China. His research interests include broadband information network, distributed system, Information security. He has to take the National Natural Science Foundation (61672471).

ZHI-FENG ZHANG received his B.S.degree in 2001 from XiDian University, and received his M.S. degree in 2006 from Xi'an University of Technology. He is currently an associate professor in Zhengzhou University of Light Industry. His main research interests include big data and deep learning.

YAN-SHUO CHANG received his Ph.D. degree in 2013 from Northwest University. He is currently a Associate Professor at School of Information, Xi'an University of Finance and Economics. His current research interests include machine learning and its applications to multimedia analysis and Quantitative Trading.

JINGLI GAO received his bachelor's degree in computer science and technology and master's degree in applied mathematics from Henan University in 2004 and 2008, and his PH.D degree in control theory and control engineering from Zhejiang University in 2017, respectively. He is currently a full-time associate professor of Pingdingshan University. His research interests include computer vision, information fusion and retrieval.

PU LI received his bachelor and PH.D degrees in computer science and technology from South China Normal University, Guangzhou 510631, China. Now, he is a lecturer in Software Engineering College, Zhengzhou University of Light Industry, Zhengzhou 450000, China. His major research interests include data intelligence, semantic computing, and semantic search, etc. He is a member of China Computer Federation (CCF). His homepage is <http://www.scholat.com/superlipu>.

HUI LIANG received his PH.D degree in communication and information systems from Communication University of China. He is currently a professor at Zhengzhou University of Light Industry, a senior EU "Marie Curie" scholar, a doctoral supervisor at Bournemouth University in the United Kingdom, and the leader of Henan Joint International Research Laboratory of Computer Animation Implementation Technologies. His research interests include computer animation, digital media technology, and virtual reality.