Soft Computing (2022) 26:11479–11496
https://doi.org/10.1007/s00500-022-07440-x

DATA ANALYTICS AND MACHINE LEARNING

A comprehensive social matrix factorization for recommendations with prediction and feedback mechanisms 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 Gao[4] · Pu Li[1,2] · Hui Liang[1,2]

Accepted: 30 July 2022 / Published online: 29 August 2022

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, existing recommendation methods have difficulty in accurately obtaining user features and item features, which seriously affects recommendation system performance. To accurately model social relationships and improve recommendation quality, we use both explicit (e.g. user-item ratings, trust relationships) and implicit (e.g. social tags) social relationships to mine users' potential interest preferences; thus, we propose a social recommendation method incorporating trust relationships and social tags. The method maps user features and item features to a shared feature space using the above social relationship, obtains user similarity and item similarity through potential feature vectors of users and items, and continuously trains them to obtain accurate similarity relationships to improve recommendation performance. The experimental results demonstrate that our proposed approach achieves superior performance over 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) has been widely used in RSs due to its strong scalability and high recommendation quality. The social MF model mainly maps multiple social relationships and user preference information among users to the shared user and item feature spaces, realizing user preference prediction to alleviate the problem of data sparsity (Yu and Li 2018). However, in the face of a complex network structure and social relationships, directly measuring social relationships tends to cause the biased error that the estimated user preferences deviate from the real user preference model, degrading the quality of the recommendation (Yang et al. 2014).

A few influencing factors, such as user social relationships, social tags and user personal interests, have been incorporated into recommendation models to improve the quality of recommendations (Zheng and Luo 2018; Bagher et al. 2017; Ricci et al. 2010; Yao et al. 2014). The above

✉ Jian-wei Zhang
zhangjianwei2019@126.com; ing@zzuli.edu.cn

Rui Chen
13382038@qq.com

Jingli Gao
gaojingli.cv@gmail.com

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

2 Henan Joint International Research Laboratory of Computer Animation Implementation Technologies, Zhengzhou, China

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

4 Software College, Pingdingshan University, Pingdingshan 467000, China

methods improve the accuracy of predictions by mapping user ratings and social relationships to shared user feature spaces and item feature spaces. Although the recommendation inaccuracy caused by the sparsity of ratings is alleviated to some extent, user social relationships are not further trained to obtain accurate similarity relationships when neighbour relationships are adopted to obtain user features; in this case, the predicted preference model may deviate from the real user preference model, weakening recommendation accuracy.

Although some prediction methods on user trust relationships have recently been proposed, only few studies have applied user prediction results of social relationships to the recommendation model (Yang et al. 2017; Ma et al. 2009; Wang et al. 2014; Tang et al. 2013; Rafailidis and Crestani 2017; Ahmadian et al. 2018). In (Yang et al. 2017), the MF technique is used to complete the recommendation by integrating approximate ratings for unknown items from the user and his or her friends with similar interests. In (Ma et al. 2009; Wang et al. 2014; Tang et al. 2013), user features are considered to be composed of his or her trustees and trusters, so the trust relationship and the features of the trustees and trusters are integrated to obtain user latent feature vectors. In (Can and Alatas 2019; Portugal et al. 2018; Pan et al. 2018; Nabizadeh et al. 2020), user ratings and social tags are used to map into the shared user and item feature spaces by employing MF technology to obtain more accurate user and item similarity relationships, but these methods generally directly predict users' preferences for items based on the obtained user and item features; however, the complex social relationships among users in social networks have been largely neglected. In (Pan et al. 2018), user implicit similarity is obtained more accurately by indirect feedback from explicit user ratings and trust relationships, thereby improving recommendation accuracy. However, this method does not take into account the impact of a user's implicit comments (such as social tags) on the user's trust relationships or the item's social relationship estimation.

In view of the complexity of social networks and the difficulty of accurately modelling social relationships, inspired by the literature (Yang et al. 2017; Ma et al. 2009, 2008; Wang et al. 2014; Pan et al. 2018; Zhang et al. 2018; Kong et al. 2018; Yan et al. 2017), this paper aims to fuse the prediction of implicit trust relationships between users and social relationships between items into the recommendation method by employing a feedback mechanism to improve recommendation accuracy. This paper comprehensively considers the framework of previous models but introduces the social relationship prediction and feedback mechanism into the social relationship recommendation model, redesigns the social matrix factorization model, and proposes a novel recommendation method based on the social relationship prediction and feedback mechanism. By obtaining more accurate user and item features in complex social networks, the recommendation quality is effectively improved. In other words, social relationships between users or between items are considered a "black box", without regard to multiple direct and indirect social relationships and interactions between users; this not only mitigates data sparsity, cold start and data imbalance but also simplifies the complex social relationship modelling process for users.

Our contributions in this paper are summarized as follows:

(1) We propose a comprehensive social recommendation method integrating trust relationships and social tags.
(2) We propose a recommendation framework based on the feedback mechanism for establishing an accurate user preference model and item features.
(3) We improve recommendation accuracy 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, through continuous training to measure implicit user and item similarities.

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

2 Preliminary and related studies

In this section, we will review the research related to social relationships and the low-rank social MF-based latent factor model because of the importance of social relationships in improving the quality of recommendations (Tang et al. 2013; Rafailidis and Crestani 2017; Can and Alatas 2019; Shokeen and Rana 2020; Li et al. 2017a; Paradarami et al. 2017).

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 (Ricci et al. 2010). In recent years, several trust relationship metrics have been proposed, among which TidalTrust (Ricci et al. 2010; Ahmadian et al. 2020; Tang et al. 2013) and RTCF (Ahmadian et al. 2018; Azadjalal et al. 2017) are the most representative trust relationship calculation methods. TidalTrust is a user trust relationship measurement method based on trust propagation theory, which mainly uses the idea that trust relationships gradually decrease as the distance between users increases. RTCF is proposed to construct a trust network of users based on the reliability measure considering the similarity and trust relationships between users, and the proposed method leads to improved reliability and accuracy of predictions (Yao et al. 2014; Salakhutdinov and Mnih 2008):

It can alleviate the problem of inaccurate recommendations caused by sparse rating data using trust relationships. In the early stage of recommendation systems, a trust propagation mechanism is used in memory-based collaborative filtering (CF) recommendation methods. However, the prediction of the user trust relationship obtained by the trust propagation mechanism will also lead to inaccurate recommendations.

2.2 Matrix factorization model based on social relationships

The social MF has become one of the most widely used methods in RSs due to its accurate prediction and high efficiency. In addition, neighbourhood 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 features and item features (Sambhav and Vikash 2018; Ma et al. 2008; Feng et al. 2017; Chen et al. 2018; Yu et al. 2018; Ahmadian et al. 2020).

2.2.1 Recommendation model based on neighbourhood relationships

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 as follows (Ma et al. 2009; Portugal et al. 2018; Herce-Zelaya et al. 2020).

$$J(R,U,V) = \frac{1}{2}\sum_{u=1}^{N}\sum_{i=1}^{M}\left(r_{ui} - g(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk} U_k^T V_i)\right)^2 + \frac{\lambda_U}{2}\sum_{u=1}^{N} U_u^T U_u + \frac{\lambda_V}{2}\sum_{i=1}^{M} V_i^T V_i \tag{1}$$

where R represents the user-item rating matrix, r_{ui} represents the rating on item i from user u, N_u represents the set of neighbours of user u and S_{uk} denotes the similarity relationship between users u and k.

2.2.2 Recommendation model based on trust relationships

In (Portugal et al. 2018; Guo et al. 2015), user interest preference models were built from the perspectives of users' trustees and trusters. The modelling idea is based on the following considerations: user u's trustee is likely to have the same or similar preference interests as user u, so these trustee users have similar feature vectors as those of user u. Similarly, when some users are trusted by the same user u, these users have similar features as user u. The rating should be determined by the user's truster-specific preferences, as well as trustee-specific preferences. Therefore, the user preference model is described as follows (Yang et al. 2017; Pan et al. 2018):

$$\hat{r}_{ui} = \alpha B_u^T V_i + (1-\alpha) E_u^T V_i \tag{2}$$

where B_u indicates the preference features of user u's trusters, E_u indicates the preference features of the user u's trustees, and V_i indicates the feature vector of the item i.

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 can be obtained as the following loss function (Yang et al. 2017; Ma et al. 2009; Pan et al. 2018):

$$J(T,B,E) = \frac{1}{2}\sum_{u=1}^{N}\sum_{v=1}^{N}(T_{uv} - B_u^T E_v)^2 + \frac{\lambda_B}{2}\sum_{u=1}^{N}\|B\|_F^2 + \frac{\lambda_E}{2}\sum_{v=1}^{N}\|E\|_F^2 \tag{3}$$

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 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 (Rafailidis and Crestani 2017; Yu et al. 2018; He et al. 2016; Gao et al. 2016; Herce-Zelaya et al. 2020; Cohen et al. 2017).

In recent years, various social relationships, such as social tags, personal interests and other influencing factors, have been integrated into recommendation models to improve recommendation quality by mapping user rating information and social relationships to shared user feature spaces and item feature spaces (Yang et al. 2017; Rafailidis and Crestani 2017; Ahmadian et al. 2018; Portugal et al. 2018; Guan et al. 2018; Meng et al. 2015; Yu et al. 2018; He et al. 2016, 2018a). Similarly, the loss function can be obtained as follows according to the weight relationships between user u and between the item i for tag t (Tang et al. 2013; Ahmadian et al. 2018).

$$\begin{aligned} J(U,V,L,G,H) = &\frac{1}{2}\sum_{u=1}^{N}\sum_{t=1}^{D}(G_{ut} - g(U_u^T L_t))^2 \\ &+ \frac{1}{2}\sum_{i=1}^{M}\sum_{t=1}^{D}(H_{it} - g(V_i^T L_t))^2 \\ &+ \frac{\lambda_U}{2}\sum_{u=1}^{N}\|U_u\|_F^2 + \frac{\lambda_V}{2}\sum_{i=1}^{M}\|V_i\|_F^2 \\ &+ \frac{\lambda_L}{2}\sum_{t=1}^{D}\|L_t\|_F^2 \end{aligned} \tag{4}$$

where G is decomposed into the user feature matrix U and the tag feature matrix L, respectively, and H is decomposed into the item feature matrix V and the same tag feature matrix L, respectively. G_{ut} and H_{it} indicate the weight relationships between user u and item i for the tag t, respectively.

Although the introduction of implicit social relationships such as social tags can alleviate data sparsity to a certain extent and enhance recommendation accuracy, it lacks further training on user social relationships to obtain accurate similarity relationships when measuring user features through neighbourhood relationships and trust relationships. Recently, some prediction methods of user trust relationships have been proposed, but few studies have applied the prediction results of user social relationships to recommendation models (Yang et al. 2017; Can and Alatas 2019; Portugal et al. 2018; Sambhav and Vikash 2018; He et al. 2016; Li et al. 2018).

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

This paper assumes that user preferences are influenced by users and their neighbours, trusted users, and social tags. It first analyses 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 the two perspectives of explicit and implicit social relationships. At the same time, based on the framework of Yang et al. (2017); Ma et al. 2009; Pan et al. 2018; Meng et al. 2015; Cao et al. 2018; Salakhutdinov and Mnih 2008), the feature matrices of users and items are trained to improve the quality of recommendations.

The ISocialMF recommendation framework proposed in this paper is shown in Fig. 1. The recommendation method is divided into the following processes: (1) a social recommendation model based on explicit interaction information such as user ratings, direct trust relationships and social tags is established; (2) the user's latent feature, the latent feature based on trust relationships and the item's latent feature are mapped to a shared space using MF technology; (3) the user's weight and item's weight based on social tags are established using interaction relationships such as social tags; (4) the implicit feature and the social tag weight information are combined to obtain the implicit similarity between users and between items; (5) the above social recommendation models are combined with the implicit similarity by using the SocialIT framework structure to establish explicit and implicit relationships, and implicit feature matrices of users and items are obtained through the above model learning; (6) the parameters are trained using the above model; and (7) the ratings are predicted, and recommendations are generated.

3.1 Regularization of social relationships based on implicit interaction

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-dimensional shared space, respectively, in this section.

We use the similarity of trust relationships between trustees and trusted users to constrain the trustee and trustee features. 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 (Yang et al. 2017).

$$\frac{\lambda_B}{2}\sum_{u=1}^{N}\sum_{v=1}^{N}P_{uv}^{(B)}\|B_u - B_v\|_F^2 \tag{5}$$

where $P_{uv}^{(B)}$ denotes the similarity between user u and user v based on the trust relationship, which can be obtained from the common trust relationships between users u and v and their common trust users. In addition, the higher the degree of trust of a user by other users, the more users will adopt his or her suggestions and the greater the user's in-degree

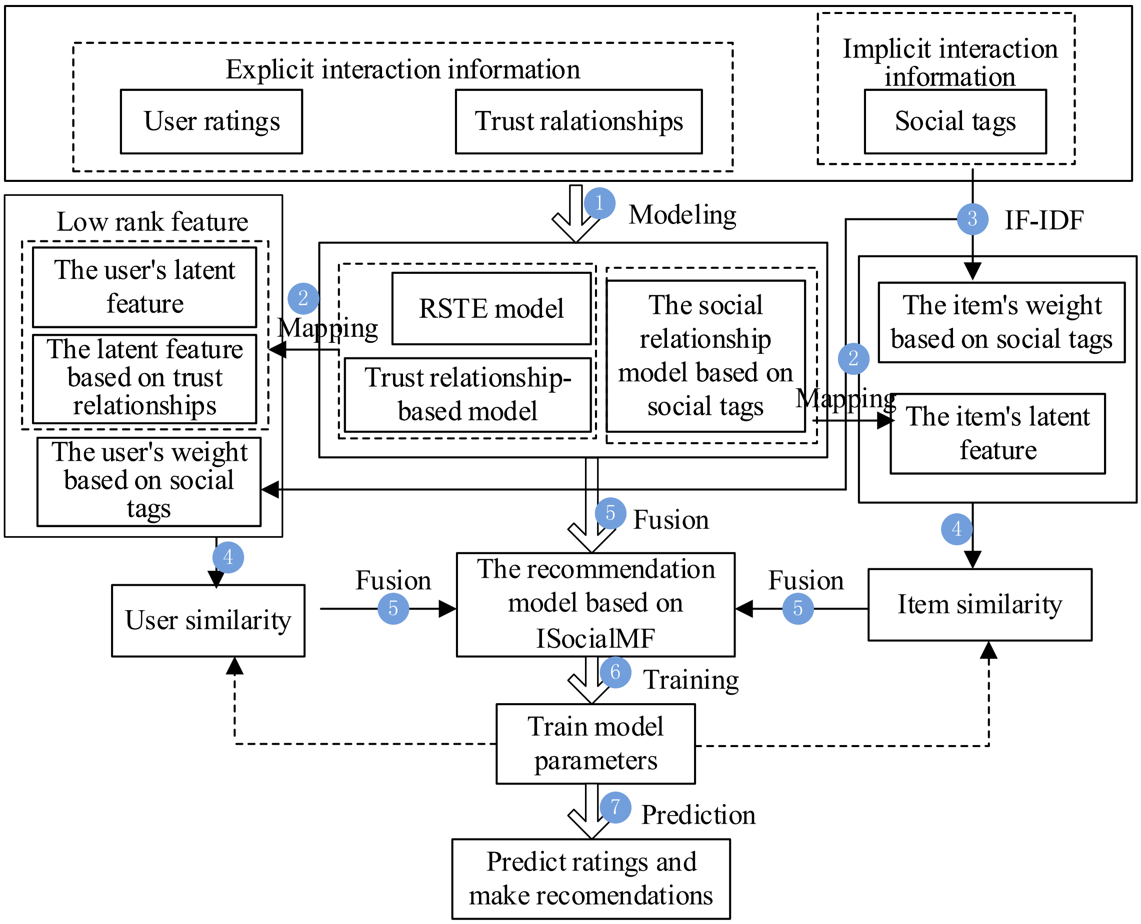


Fig. 1 ISocialMF recommendation framework

will be. According to Yang et al. (2014); Guo et al. 2015; He et al. 2016; Cohen et al. 2017), the improved user similarity is described as follows:

$$P_{uv}^{(B)} = \frac{\sum_{k=1}^{N} \frac{d_k^{(in)}}{d_u^{(out)}+d_k^{(in)}} T_{uk} * \frac{d_k^{(in)}}{d_v^{(out)}+d_k^{(in)}} T_{vk}}{\sqrt{\sum_{k=1}^{N} \left(\frac{d_k^{(in)}}{d_u^{(out)}+d_k^{(in)}} T_{uk}\right)^2} \sqrt{\sum_{k=1}^{N} \left(\frac{d_k^{(in)}}{d_v^{(out)}+d_k^{(in)}} T_{vk}\right)^2}} \tag{6}$$

where $d_k^{(in)}$ and $d_u^{(out)}$ represent the in-degree and out-degree in the trust network, respectively. Similarly, as a trusted user, the closer their features 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 (Yang et al. 2017):

$$\frac{\lambda_E}{2} \sum_{u=1}^{N} \sum_{v=1}^{N} Q_{uv}^{(E)} \| E_u - E_v \|_F^2 \tag{7}$$

where $Q_{uv}^{(E)}$ denotes the similarity of the trust relationship between two trusted persons. According to the literature (Yang et al. 2017; Ma et al. 2009, 2008), the improved user similarity based on the trustee is as follows:

$$Q_{uv}^{(E)} = \frac{\sum_{k=1}^{N} \frac{d_u^{(in)}}{d_k^{(out)}+d_u^{(in)}} T_{ku} * \frac{d_v^{(in)}}{d_k^{(out)}+d_v^{(in)}} T_{kv}}{\sqrt{\sum_{k=1}^{N} \left(\frac{d_u^{(in)}}{d_k^{(out)}+d_v^{(in)}} T_{ku}\right)^2} \sqrt{\sum_{k=1}^{N} \left(\frac{d_v^{(in)}}{d_k^{(out)}+d_v^{(in)}} T_{kv}\right)^2}} \tag{8}$$

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

$$sim(u,v)=\begin{cases}\dfrac{2*(P_{uv}^{(B)}+Q_{uv}^{(E)})*sim_{uv}}{P_{uv}^{(B)}+Q_{uv}^{(E)}+2*sim_{uv}}, & if\ P_{uv}^{(B)}\neq 0, Q_{uv}^{(E)}\neq 0,\ \text{and}\ sim_{uv}\neq 0\\ s_{uv}, & else\ if\ P_{uv}^{(B)}=0, Q_{uv}^{(E)}=0,\ \text{and}\ sim_{uv}\neq 0\\ \dfrac{2*P_{uv}^{(B)}*Q_{uv}^{(E)}}{P_{uv}^{(B)}+Q_{uv}^{(E)}}, & else\ if\ P_{uv}^{(B)}\neq 0, Q_{uv}^{(E)}\neq 0,\ \text{and}\ sim_{uv}=0\\ 0, & else\end{cases} \tag{9}$$

Inspired by the literature (Yang et al. 2017; Sa et al. 2021; Sedhain et al. 2017; Shneiderman 2020; Wei et al. 2017; Zhang et al. 2017; Gupta and Kant 2020), if a user trusts another user or a user is trusted by another, their features will be very similar. Therefore, the truster and the trusted user have similar feature vectors. Therefore, a regularization term is obtained as follows:

$$\frac{\beta}{2}\sum_{u=1}^{N} sim(u_{I_u^{(B)}}, u_{I_u^{(E)}})\|B_u-E_u\|_F^2 \tag{10}$$

where $I_u^{(B)}$ and $I_u^{(E)}$ represent the sets of user u's trusters and trustees, respectively.

3.2 Implicit similarities between users and between items

To accurately describe the degree of similarity between users, based on the literature (Portugal et al. 2018; Pereira et al. 2019; Aghdam 2019; Gong et al. 2019; Gupta and Kant 2020), we comprehensively analyse the similarity relationships among item similarity and social tags through the user's explicit interactions 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 Yang et al. (2017); Ma et al. 2009; Pan et al. 2018; Salakhutdinov and Mnih 2008), user feature vectors, trust relationships and social tags reflect the similarity between users to some extent. In (Can and Alatas 2019), [103], user features are measured directly, while the contributions of similarity from other social relationships are ignored. In (Azadjalal et al. 2017), user features and trust relationships are employed to measure similarity, but the similarity just adds up linearly, without considering the impact of data imbalance or the weight of the different influencing factors. From (Yang et al. 2017; Ma et al. 2009; Seo et al. 2017), the comprehensive trade-off user features, trust relationship features and social label weight relationships accurately measure user similarity. Among them, the user's rating similarity is calculated by the user's rating vector as follows (Yang et al. 2017; Ma et al. 2009; Pan et al. 2018; Li et al. 2017b; Lu and Guo 2019).

$$S_{uv}^{(UR)}=sim(U_u,U_v)=U_u^T U_v \tag{11}$$

If the two trustee feature vectors are similar, their preferences are similar. Here, U_u and U_v are normalized feature vectors. 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 features. Therefore, the average similarity of preferences based on the trust relationship is as follows:

$$S_{uv}^{(T)}=\frac{B_u^T B_v+E_u^T E_v}{2} \tag{12}$$

Here, B_u, B_v, E_u, and E_v are normalized feature vectors. The implicit similarity between users is considered a normal distribution consisting of user feature similarity, trust relationship implicit similarity and tag similarity. The conditional probability distribution is as follows:

$$p\left(S^{(U)}|S_{uv}^{(UR)},S_{uv}^{(T)},S_{uv}^{(UL)},\sigma_S^2\right)=\prod_{u=1}^{N}\prod_{v=1}^{N}N\left[S_{uv}^{(U)}|g\left(f\left(S_{uv}^{(UR)},S_{uv}^{(T)},S_{uv}^{(UL)}\right)\right),\sigma_S^2\right] \tag{13}$$

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:

$$\hat{S}_{uv}^{(U)}=\begin{cases}S_{uv}^{(UL)}\dfrac{2*S_{uv}^{(UR)}*S_{uv}^{(T)}}{S_{uv}^{(R)}+S_{uv}^{(T)}}, & if\ relation(u,v)\neq 0\\ \dfrac{2*S_{uv}^{(UL)}*S_{uv}^{(UR)}}{S_{uv}^{(UL)}+S_{uv}^{(UR)}}, & else\end{cases} \tag{14}$$

where relation(u,v) indicates a direct trust relationship between users u and v. It avoids the inaccuracy of user feature descriptions caused by data sparseness and imbalance and solves the problem of deviating from real user features caused by the linear superposition of user similarity, without explicit user rating or trust relationship data. User implicit preference similarity based on tags can be obtained by comparing the tag weights of the above two users to the item (Yang et al. 2017; Cao et al. 2018; Panagiotakis et al. 2021; Rezaeimehr et al. 2017).

$$S_{uv}^{(UL)} = \frac{\sum_{k=1}^{D} G_{uk} * G_{vk}}{\sqrt{\sum_{k=1}^{D} G_{uk}^2}\sqrt{\sum_{k=1}^{D} G_{vk}^2}} \tag{15}$$

Taking the logarithm of the posterior probability of Eq. (13), the objective function is obtained as follows:

$$\begin{aligned} J(S,U,B,G) &= \sum_{u=1}^{N}\sum_{v\in N_u}(S_{uv}^{(U)} - g(f(S_{uv}^{(UR)},S_{uv}^{(T)},S_{uv}^{(UL)}))^2 \\ &= \begin{cases} \sum\limits_{u=1}^{N}\sum\limits_{v\in N_u}(S_{uv}^{(U)} - g(S_{uv}^{(UL)}\dfrac{2*S_{uv}^{(UR)}*S_{uv}^{(T)}}{S_{uv}^{(UR)}+S_{uv}^{(T)}}))^2, & \mathit{if\ relation}(u,v)\neq 0 \\ \sum\limits_{u=1}^{N}\sum\limits_{v\in N_u}\left(S_{uv}^{(U)} - g\left(\dfrac{2*S_{uv}^{(UR)}*S_{uv}^{(UL)}}{S_{uv}^{(UR)}+S_{uv}^{(UL)}}\right)\right)^2, & \mathit{else} \end{cases} \end{aligned} \tag{16}$$

3.2.2 Item implicit similarity

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

$$p\left(S^{(I)}|S_{ij}^{(IR)},S_{ij}^{(IL)},\sigma_S^2\right) = \prod_{i=1}^{M}\prod_{j=1}^{M} N\left[S_{ij}^{(I)}|g\left(f\left(S_{ij}^{(IR)},S_{ij}^{(IL)}\right)\right),\sigma_S^2\right)\right] \tag{17}$$

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

$$J(S,V,H) = \sum_{i=1}^{M}\sum_{j\in N_i}\left(S_{ij}^{(I)} - g(\hat{S}_{ij}^{(I)})\right)^2 \tag{18}$$

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

$$\hat{S}_{ij}^{(I)} = \frac{2*S_{ij}^{(IL)}*S_{ij}^{(IR)}}{S_{ij}^{(IL)}+S_{ij}^{(IR)}} \tag{19}$$

Here, the similarity between items based on social tags is as follows (Ahmadian et al. 2020).

$$S_{ij}^{(IL)} = \frac{\sum_{k=1}^{D} H_{ik} * H_{jk}}{\sqrt{\sum_{k=1}^{D} H_{ik}^2}\sqrt{\sum_{k=1}^{D} H_{jk}^2}} \tag{20}$$

Likewise, the similarity based on item features is calculated as follows:

$$\mathrm{S}_{\mathrm{ij}}^{(IR)} = sim(V_i,V_j) = V_i^T V_j \tag{21}$$

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 Yang et al. (2017); Ma et al. 2009; Pan et al. 2018; Cao et al. 2018; Salakhutdinov and Mnih 2008; Li et al. 2017a; Liu et al. 2017; Aghdam 2019), the effect of user ratings, this paper first analyses explicit and implicit interactions such as social tags and user trust on recommendation quality in depth. It also integrates some classic recommendation frameworks such as SocialIT, RSTE and RoRec into the recommendation process and proposes a comprehensive social recommendation algorithm, namely ISocialMF, to fuse social tags and trust relationships. The method maps the rating information, social tags and user trust relationships to the low-dimensional 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 is 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, not only mitigating the inaccuracy of recommendations caused by sparse data and imbalance but also improving the diversity of recommendations.

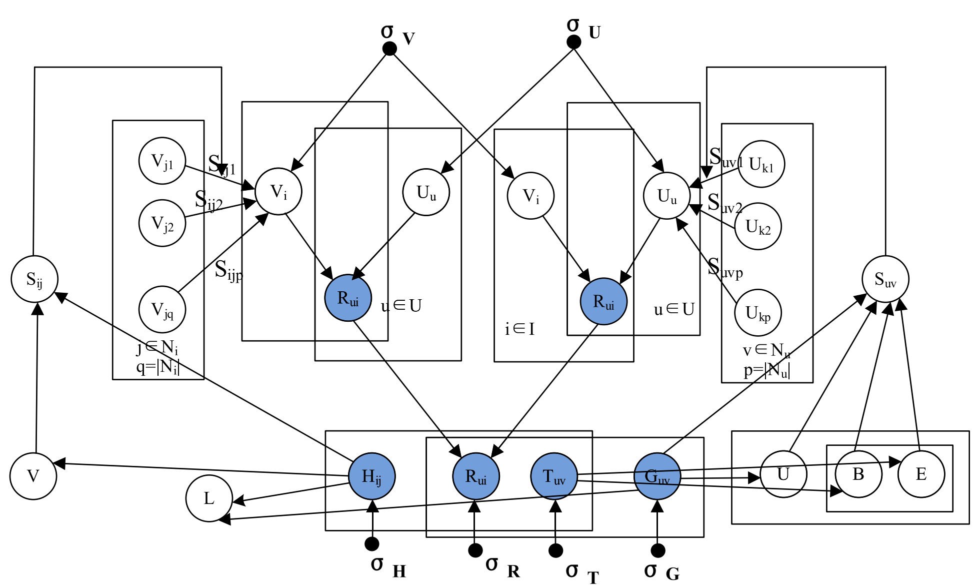
4.1 ISocialMF algorithm model

Considering the impact of user trust relationships, rating information and social tags on user preference similarity and item similarity, combined with Eqs. (1), (3), (4), (10), (16) and (21), the comprehensive social MF recommendation model is shown in Fig. 2, and the cost function with user and item feature regular terms are obtained as follows:

$$
\begin{aligned}
J(R,U,V,G,H,S,T) = {} & \frac{1}{2}\sum_{u=1}^{N}\sum_{i=1}^{M} I_{ui}^{R}\Big(r_{ui} - g\big(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\big)\Big)^2 \\
& + \frac{1}{2}\sum_{u=1}^{N}\sum_{t=1}^{D} I_{ut}^{G}\big(G_{ut} - g(U_u^T L_t)\big)^2 + \frac{1}{2}\sum_{i=1}^{M}\sum_{t=1}^{D} I_{it}^{H}\big(H_{it} - g(V_i^T L_t)\big)^2 \\
& + \frac{\lambda_P}{2}\sum_{u=1}^{M}\sum_{v\in N_u}\big(S_{uv}^{(U)} - g(\hat{S}_{uv}^{(U)})\big)^2 + \frac{\lambda_T}{2}\sum_{u=1}^{N}\sum_{v=1}^{N}\big(T_{uv} - g(B_u^T E_v)\big)^2 \\
& + \frac{\rho}{2}\sum_{u=1}^{N}\sum_{t\in N_u} sim(u,t)\|U_u - U_t\|_F^2 + \frac{\gamma}{2}\sum_{i=1}^{M}\sum_{j\in N_i} sim(i,j)\|V_i - V_j\|_F^2 \\
& + \frac{\lambda_S}{2}\sum_{u=1}^{N}\sum_{v=1}^{N} P_{uv}^{(B)}\|B_u - B_v\|_F^2 + \frac{\lambda_Z}{2}\sum_{u=1}^{N}\sum_{v=1}^{N} Q_{uv}^{(E)}\|E_u - E_v\|_F^2 \\
& + \frac{\beta}{2}\sum_{u=1}^{M} sim(u_{I_u^{(B)}}, u_{I_u^{(E)}})\|B_u - E_u\|_F^2 + \frac{\lambda_Q}{2}\sum_{i=1}^{M}\sum_{j\in N_i}\big(S_{ij}^{(I)} - g(\hat{S}_{ij}^{(I)})\big)^2 \\
& + \frac{\lambda_W}{2}\sum_{j=1}^{M}\Big(V_j - \sum_{i\in N_j} S_{ji}^{(I)} V_i\Big)^T\Big(V_j - \sum_{i\in N_j} S_{ji}^{(I)} V_i\Big) + \frac{\lambda_U}{2}\sum_{u=1}^{N} n_{u_u}\|U\|_F^2 \\
& + \frac{\lambda_V}{2}\sum_{i=1}^{M} n_{v_i}\|V\|_F^2 + \frac{\lambda_L}{2}\sum_{t=1}^{D} n_{l_t}\|L\|_F^2 + \frac{\lambda_B}{2}\sum_{u=1}^{N} n_{b_u}\|B\|_F^2 + \frac{\lambda_E}{2}\sum_{v=1}^{N} n_{e_v}\|E\|_F^2
\end{aligned}
\tag{22}
$$

where sim(u,t) denotes the similarity of users u_u and u_t (see Eq. (12)) and sim(i,j) indicates the similarity between items i_i and i_j. n_{u_u} denotes the number of ratings from the user u, n_{v_i} denotes the number of ratings given to the item i, n_{l_t} denotes the number of tags that are marked by the user u, and 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. I_{ui}^R, I_{ut}^G, and I_{it}^H are the indicator functions. I_{ui}^R is equal to 1 if user u rated item i, and equal to 0 otherwise. I_{ut}^G is equal to 1 if the tag t has been used by user u, and equal to 0 otherwise. I_{it}^H is equal to 1 if the tag t has been used by item i, and equal to 0 otherwise. In Fig. 2, the user feature

Fig. 2 Probability graph model of ISocialMF



matrices 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 J in Eq. (22) with respect to U_u, V_i, B_u, E_v, $S_{\mathrm{uv}}^{(U)}$, $S_{\mathrm{ij}}^{(I)}$ and L_t are presented as follows:

$$
\begin{aligned}
\frac{\partial J}{\partial U_u} =&\alpha \sum_{i=1}^{M} I_{ui}^{R} V_i g\left(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\right)\left(1-g\left(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\right)\right)\\
&\times\left(g\left(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\right) - r_{ui}\right) + \lambda_U n_{u_u} U_u\\
&+\rho\sum_{t\in N_u} sim(u,t)(U_u - U_t) + \lambda_P \sum_{v\in N_u}\left(S_{uv}^{(U)} - g\left(\widehat{S}_{uv}^{(U)}\right)\right) g\left(\widehat{S}_{uv}^{(U)}\right)\left(1-g\left(\widehat{S}_{uv}^{(U)}\right) S_{uv}^{(UR)'} U_u'\right)\\
&+\sum_{t=1}^{D} I_{ut}^{G} L_t (G_{ut} - g(U_u^T L_t)) g(U_u^T L_t)(1-g(U_u^T L_t))
\end{aligned}
\tag{23}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial V_i} =&\alpha \sum_{u=1}^{N} I_{ui}^{R} g\left(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\right)\left(1-g\left(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\right)\right)\\
&\times\left(g\left(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\right) - r_{ui}\right)\\
&\times\left(\alpha U_u + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T\right) + \gamma\sum_{j\in N_i} sim(i,j)(V_i - V_j)\\
&+\lambda_W\left(V_i - \sum_{j\in N_i} S_{ij}^{(I)} V_j\right) - \lambda_W \sum_{j=1}^{M} S_{ij}^{(I)}\left(V_j - \sum_{i\in N_j} S_{ij}^{(I)} V_i\right)\\
&+\lambda_Q \sum_{j\in N_i}\left(g(\hat{S}_{ij}^{(I)}) - S_{ij}^{(I)}\right) g(\hat{S}_{ij}^{(I)})\left(1-g(\hat{S}_{ij}^{(I)})\right) V_i'\\
&+\sum_{t=1}^{D} I_{it}^{H} L_t g(V_i^T L_t)(1-g(V_i^T L_t))(H_{it} - g(V_i^T L_t)) + \lambda_V n_{v_i} V_i
\end{aligned}
\tag{24}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial B_u} =&\lambda_P\left(g\left(\hat{S}_{uv}^{(U)}\right) - S_{\mathrm{uv}}^{(U)}\right) g\left(\hat{S}_{uv}^{(U)}\right)\left(1-g\left(\hat{S}_{uv}^{(U)}\right)\right)\frac{\partial \hat{S}_{uv}^{(U)}}{\partial B_u} + \lambda_P\left(g\left(\hat{S}_{vu}^{(U)}\right) - S_{vu}^{(U)}\right) g\left(\hat{S}_{vu}^{(U)}\right)\left(1-g\left(\hat{S}_{vu}^{(U)}\right)\right) B_u'\\
&+\lambda_T \sum_{v=1}^{N}\left(g(B_u^T E_v) - T_{uv}\right) g(B_u^T E_v) E_v + \lambda_B \mathrm{n}_{b_u} B_u + \beta sim(u_{I_u^{(\mathrm{B})}}, u_{I_u^{(\mathrm{E})}})(B_u - E_u)\\
&-\lambda_{\mathrm{S}} P_{uv}^{(B)}(B_u - B_v)
\end{aligned}
\tag{25}
$$

$$
\begin{aligned}
\frac{\partial J}{\partial E_v} =&\lambda_P\left(g\left(\hat{S}_{uv}^{(U)}\right) - S_{uv}^{(U)}\right) g\left(\hat{S}_{uv}^{(U)}\right)\left(1-g\left(\hat{S}_{uv}^{(U)}\right)\right) E_v'\\
&+\lambda_S\left(g\left(\hat{S}_{vu}^{(U)}\right) - S_{vu}^{(U)}\right) g\left(\hat{S}_{vu}^{(U)}\right)\left(1-g\left(\hat{S}_{uv}^{(U)}\right)\right) E_v'\\
&+\lambda_T \sum_{v=1}^{N}\left(g(B_u^T E_v) - T_{uv}\right) g(B_u^T E_v) B_u + \lambda_E n_{e_v} E_v - \lambda_Z Q_{uv}^{(E)}(E_u - E_v)
\end{aligned}
\tag{26}
$$

Table 1 Statistics for the Epinions and Douban datasets

Datasets	Epinions	Douban
Number of users	12,630	5786
Number of items	3620	26,573
Number of ratings	1,261,218	685,936
Number of social relationships	487,183	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

$$\begin{aligned}\frac{\partial J}{\partial S_{uv}^{(U)}} =& g\left(\alpha U_p^T V_i + (1-\alpha)\sum_{k\in N_u} S_{pu}^{(U)} U_k^T V_i\right)\\ &\times\left(1-\left(g\left(\alpha U_p^T V_i + (1-\alpha)\sum_{k\in N_u} S_{pu}^{(U)} U_k^T V_i\right)\right)\right)\\ &\times\left(g\left(\alpha U_u^T V_i + (1-\alpha)\sum_{k\in N_u} S_{uk}^{(U)} U_k^T V_i\right) - r_{ui}\right)\\ &\times (1-\alpha) U_v^T V_i + \lambda_P\left(S_{uv}^{(U)} - g\left(\hat{S}_{uv}^{(U)}\right)\right)\end{aligned} \tag{27}$$

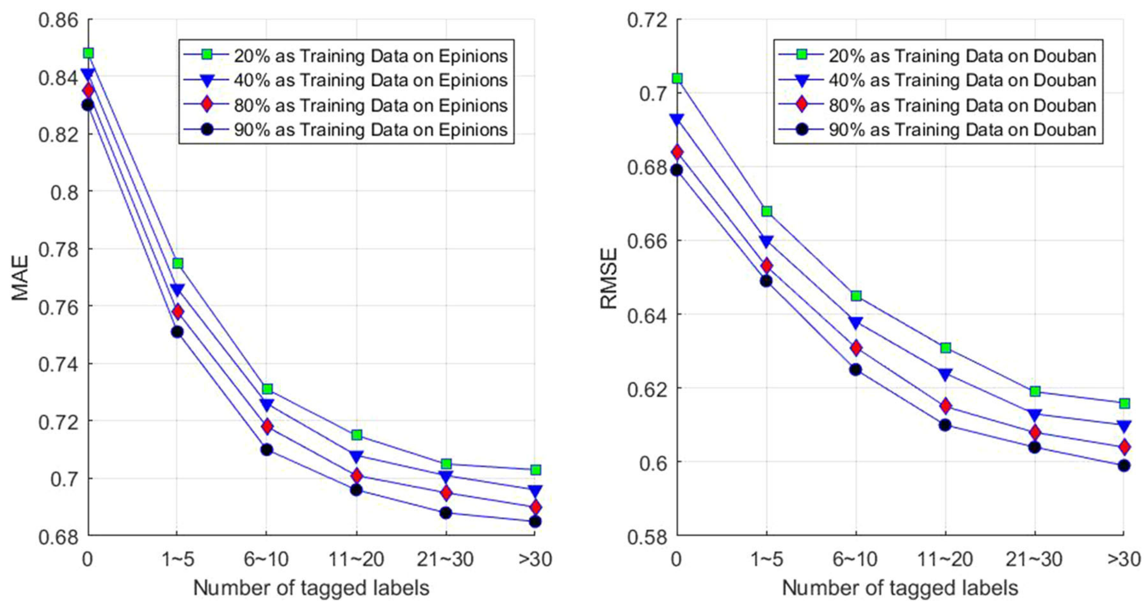


Fig. 3 The MAE performance of the algorithm under different numbers of tags

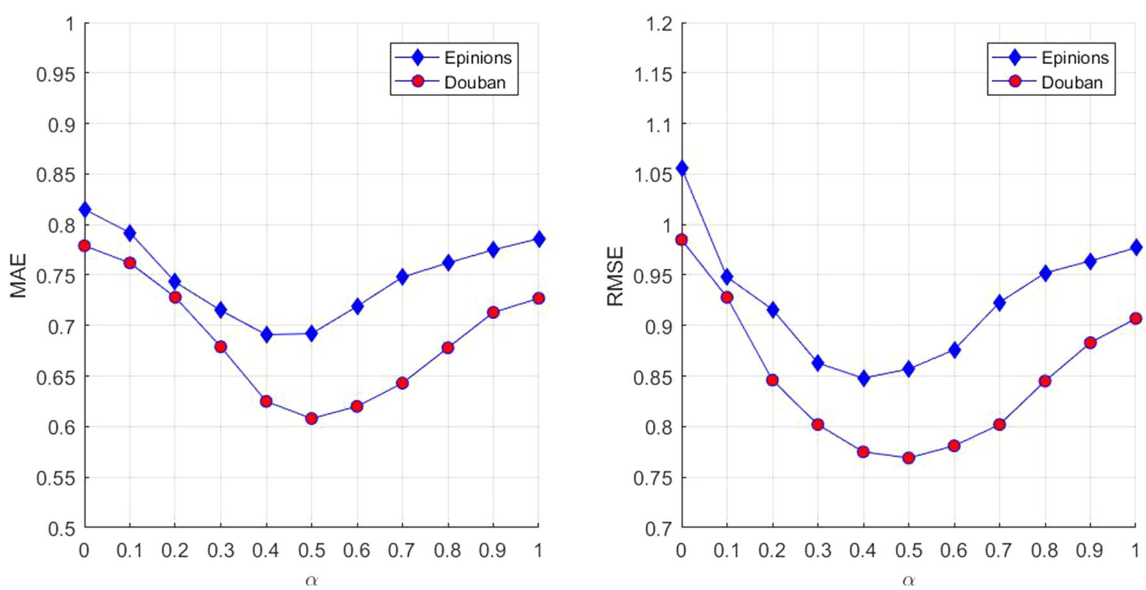


Fig. 4 The effect of parameter α on the MAE and the RMSE

$$\frac{\partial J}{\partial S_{ij}^{(I)}} = \lambda_Q \sum_{j \in N_i} \left(S_{ij}^{(I)} - g\left(\hat{S}_{ij}^{(I)} \right) \right) - \lambda_W V_j \left(V_{\mathrm{i}} - \sum_{j \in N_i} S_{ij}^{(I)} V_j \right) \tag{28}$$

$$\frac{\partial J}{\partial L_t} = \sum_{u=1}^{N} I_{ut}^{G} g\left(U_u^T L_t\right)\left(1 - g\left(U_u^T L_t\right)\right)\left(g\left(U_u^T L_t\right) - G_{ut}\right) U_u + \lambda_L \mathrm{n}_{l_t} L_t + \sum_{i=1}^{M} I_{it}^{H} g\left(V_i^T L_t\right)\left(1 - g\left(V_i^T L_t\right)\right)\left(g\left(V_i^T L_t\right) - \mathrm{H}_{it}\right) \tag{29}$$

Here, U'_u, V'_i, B'_u, E'_v and $S_{uv}^{(UR)'}$ are the partial derivatives of $S_{uv}^{(UR)}, \hat{S}_{ij}^{(I)}, \hat{S}_{uv}^{(U)}$, and $S_{uv}^{(U)}$ over $\underline{U}_u$, V_i, B_u, E_v and $S_{uv}^{(UR)}$, 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 $S^{(I)}$。

Initialize U, V, B, E, S。
Construct user and item weight relationship matrices G and H based on the user's annotation behaviour for the item.
Get social tag-based user similarity and item similarity according to Eqs. (15) and (20).
while t<$maxIter$
    $J_{old} \leftarrow J$;
    for $r_{ui} \in R$
        $U_u \leftarrow U_u - \eta \frac{\partial J}{\partial U_u}$, $V_i \leftarrow V_i - \eta \frac{\partial J}{\partial V_i}$
        $B_u \leftarrow B_u - \eta \frac{\partial J}{\partial B_u}$, $E_v \leftarrow E_v - \eta \frac{\partial J}{\partial E_v}$
        $S_{uv}^{(U)} \leftarrow S_{uv}^{(U)} - \eta \frac{\partial J}{\partial S_{uv}^{(U)}}$, $S_{ij}^{(I)} \leftarrow S_{ij}^{(I)} - \eta \frac{\partial J}{\partial S_{ij}^{(I)}}$
        $L_t \leftarrow L_t - \eta \frac{\partial J}{\partial L_t}$
        Obtain user similarities based on user features, trust relationships and social tags according to Eqs.(11), (12), and (15).
        Obtain the item similarities based on social tags and item features according to Eqs (20) and (21).
        Obtain the user and item implicit similarities according to Eqs (16) and (18).
    end for
    Update L according to Eq(22);
    if $|J - J_{old}|$<$threshold$ **then**
        break;
    end if
end while
$t \leftarrow t+1$;
$\hat{r}_{ui} \leftarrow U_u^T V_i$

4.3 Algorithm efficiency analysis

The main cost of the algorithm lies in the objective function J and its gradients against variables. The time complexity of completing an iterative loss function J is $O(\bar{r}K + \bar{t}K + \bar{d}K + \bar{c}K)$, where $\bar{r}$ represents the average number of user ratings, $\bar{t}$ represents the average number of user trust relationships, $\bar{d}$ represents the number of nonzero elements in matrix G, $\bar{c}$ represents the number of nonzero elements in matrix H, and K is the dimension of the eigenvector U_u. The time complexity of gradient learning is O $(\bar{r}K + \bar{t}K + \bar{d}K + \bar{c}K + t^2K)$. Because the user rating matrix and user social relationship matrix are very sparse, the values of $\bar{r}$, $\bar{t}$, $\bar{d}$, $\bar{c}$ and K are small. The time complexity of the algorithm mainly depends on the values of N and M. Therefore, the time complexity of the algorithm is equivalent to that of other models. In the case of the same order of time cost, the recommendation quality is obviously improved.

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, 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 in our experiment are shown in Table 1.

We first use the 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 the recommendation ranking. It considers the relevance and ranking position of all recommended items, and its definition is as follows (Yang et al. 2017; He et al. 2018b):

$$\mathrm{NDCG} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{Z_i}\sum_{j\in\Omega(i)}\frac{2^{r_{ij}}-1}{\log_2(1+rank(j))} \tag{30}$$

where rank(j) represents the position of the list of items v_j in the sorted set $\Omega(i)$ and Z_i Z_i is a parameter introduced for standardization the value of $\sum_{j\in\Omega(i)}\frac{2^{r_{ij}}-1}{\log_2(1+rank(j))}$. The value of NDCG is in the interval [0, 1], and the larger the value, the better the recommended sorting effect is.

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 $\sim$ 5", "6 $\sim$ 10", "11 $\sim$ 20" and " > 20". Then, the experiment is conducted. The performance of the ISocialMF algorithm on the MAE and the RMSE at different numbers of tags is shown in Fig. 3.

As seen from Fig. 3, since the IScocialMF algorithm takes into account the user's tag information for the item, the MAE performance of the recommendation algorithm is significantly improved in 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 features 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 IScocialMF 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 neighbourhood users to the preference prediction, the range of which is [0, 1]. The user's own behaviour is mainly used for prediction when $\alpha = 0$, and the behaviour of the neighbourhood users is mainly used for prediction when $\alpha = 1$. Figure 4 depicts the effect of parameter α on the MAE and the RMSE for the Epinions and Douban datasets, respectively. The effects of the parameter α on the MAE and the RMSE are very similar. As α increases, the values of the MAE and the RMSE first decrease, then gradually increase, and finally become stable. This phenomenon indicates that it is more accurate to comprehensively consider the behaviour of users and their neighbours.

Figure 5 shows the effect of parameter λ_P on MAE for the Epinions and Douban datasets. The role of λ_P is to control the impact of user implicit similarity on recommendation performance. The larger the value, the greater the role of neighbourhood users in predicting the rating process, and the greater user features depend on neighbourhood users. On the Epinions and Douban datasets, as λ_P increases, the MAE value first decreases and then increases slowly until it stabilizes. The prediction accuracy of IScocialMF of both datasets is optimal.

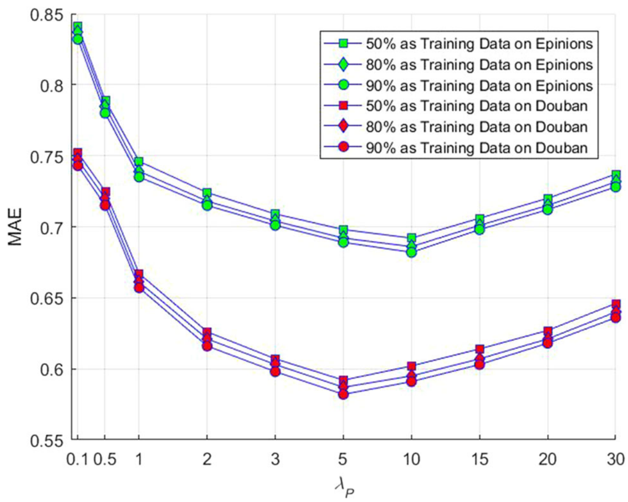


Fig. 5 The effect of parameters λ_P on the MAE for the Epinions and Douban datasets

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 the degree of their trust association (out-degree + in-degree) 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 Fig. 6. The proportion of the number of user relationships with connection relationships ranging from 0–5 and 5–10 is the highest, and the proportion of connection relationships exceeding 200 is the lowest, but the number of users who rated items is the largest.

Figure 7 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 becomes 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 increase as the number of social relationships increases. Therefore, the performance of the PMF is constantly improving and finally tends to be stable. Therefore, as the number of social relationships increases, although the number of some friends has also increased, many of these new relationships are only casual. There is no substantive trust relationship or common interest preference between them. These casual social relationship connections have less ability to learn a user's true interest preference features, so the recommendation algorithm does not achieve better recommendation quality.

5.4 Performance comparison and analysis

To evaluate the performance of the ISocialMF algorithm proposed in this paper, the algorithm is compared with social recommendation methods such as PMF

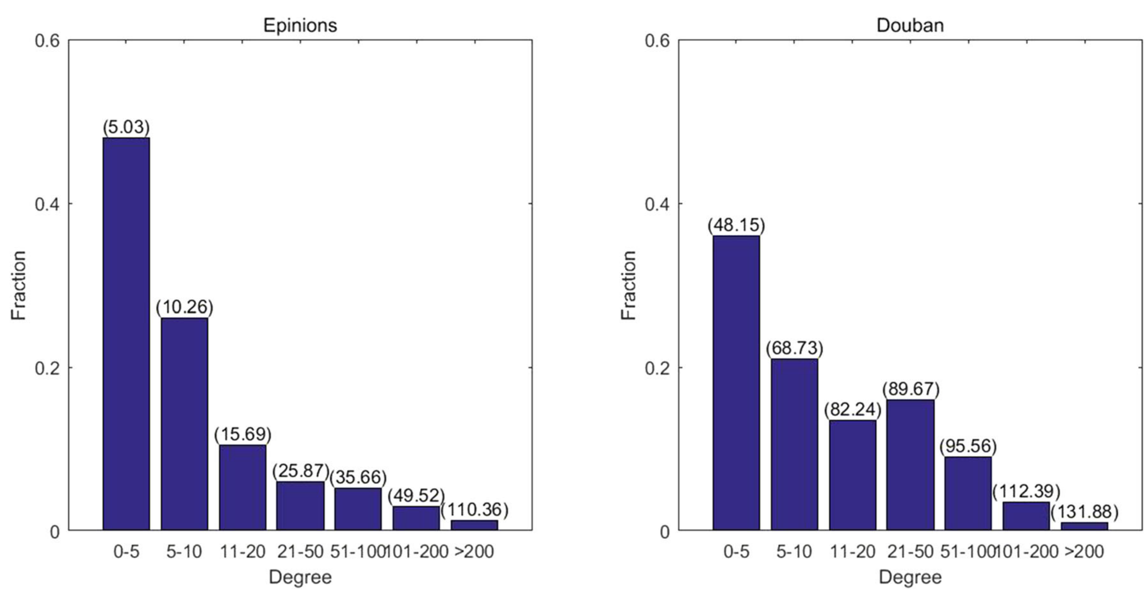


Fig. 6 The social relationship distribution of each user group on the Epinions and Douban datasets

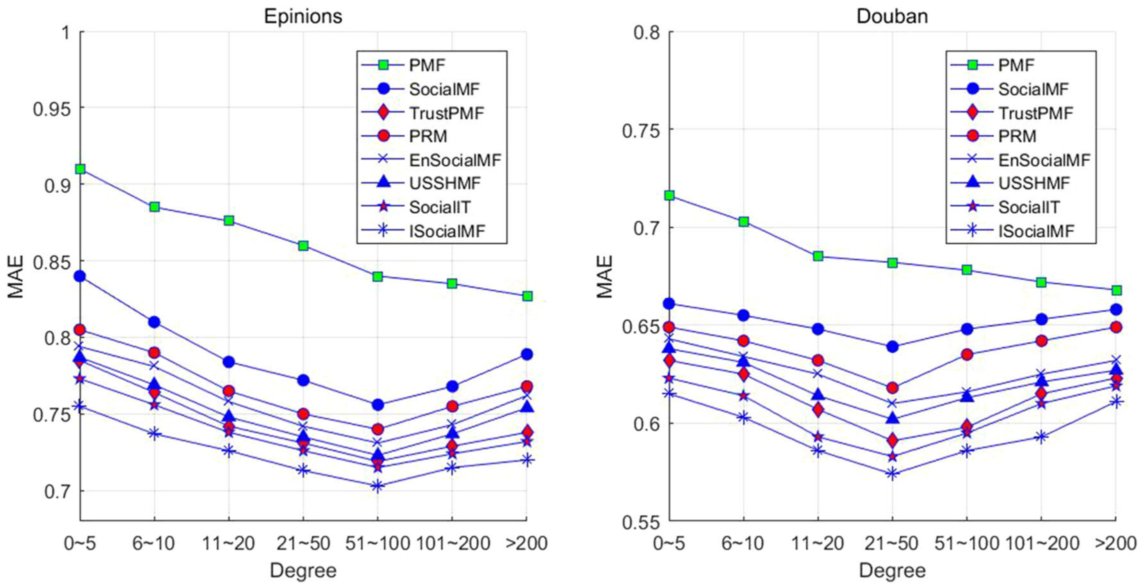


Fig. 7 The MAE performance of users in different groups for each algorithm

Table 2 The MAE and RMSE performance of each algorithm on the 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%	–
Douban	MAE	0.665	0.591	0.609	0.605	0.601	0.597	0.582	0.571
	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%	–

Table 3 Comparison of the top-N performance of each algorithm on the 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

(Salakhutdinov and Mnih 2008), TrustPMF (Portugal et al. 2018), SocialMF (Jamali and Ester 2010), PRM (Seo et al. 2017), EnSocialMF(Luo et al. 2018), USSHMF(Guo et al. 2015) and SocialIT (Zheng and Luo 2018).

5.4.1 Performance of the algorithm on sparse datasets

Table 2 describes the MAE and RMSE performance comparison of different algorithms on the Epinions and Douban datasets. On both datasets, the performance of the ISocialMF algorithm on the MAE and the RMSE exceeded that of other algorithms. Compared with other social network-based recommendation algorithms, the traditional PMF algorithm performs worst on the MAE and the RMSE performance. For the Douban dataset with relatively dense data, the ISocialMF algorithm has a smaller increase than PMF and other social-based CF 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 better in Epinions.

Table 3 shows the P@*N*, R@*N* and NDGG performance of each algorithm on the Epinions and Douban datasets. The performance of the ISocialMF algorithm in this paper is significantly improved compared with other social relationship-based recommendation algorithms; this 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 matrices 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 effect of social tag implicit social relationships or 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 modelling of social relationships. Their user preferences may deviate from real user preferences, therefore affecting the recommendation quality. The experimental results show that our method can improve the recommendation quality more effectively by mining the implicit information of social relationships and social tags to model user preference interests and training user features and item features to obtain accurate user similarity and item similarity.

The foundation of this model is to assume that the training samples and test samples obey the standard normal distribution. We conducted a tenfold cross-validation on the Epinions and Douban datasets to show the effectiveness of our proposed method. SocialIT is the state-of-the-art model among the baselines, so we choose to compare it with the best model. We conduct the two models on one of the test samples to obtain 10 values of top-N performance and then conduct a t test on this basis. When N equals 5, the Top-N Performance comparison of SocialIT and ISocialMF on the Epinions Dataset is shown in Table 4.

For SocialIT, the mean is 0.442, and the standard variance is 0.0024; for IsocialMF, the mean is 0.461, the standard variance is 0.0026, and the p value is 6.676E-08, which is much less than 0.05. Therefore, the result indicates that the differences are statistically significant. When N equals 5, the Top-N Performance comparison of SocialIT and ISocialMF on the Douban Dataset is shown in Table 5.

For SocialIT, the mean is 0.414, and the standard variance is 0.0028; for IsocialMF, the mean is 0.433, the standard variance is 0.0033, and the p value is 5.827E-09, which is much less than 0.05. Therefore, the result indicates that the differences are statistically significant, and the proposed method is effective.

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 fewer than 5 items as cold start users. Table 6 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

Table 4 Top-N Performance comparison of SocialIT and ISocialMF on the Epinions Dataset

Methods	Times										
	1	2	3	4	5	6	7	8	9	10	Standard variance
SocialIT	0.445	0.441	0.442	0.438	0.439	0.441	0.443	0.442	0.446	0.443	0.0024
ISocialMF	0.462	0.463	0.463	0.461	0.465	0.458	0.461	0.457	0.462	0.458	0.0026

Table 5 Top-N Performance Comparison of SocialIT and ISocialMF on the Douban Dataset

Methods	Times										
	1	2	3	4	5	6	7	8	9	10	Standard variance
SocialIT	0.416	0.412	0.413	0.415	0.411	0.419	0.411	0.413	0.418	0.414	0.0028
ISocialMF	0.434	0.436	0.431	0.435	0.429	0.437	0.433	0.428	0.436	0.429	0.0033

Table 6 The 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%	–

Table 7 The average online time performance of several different methods on the datasets of Epinions and Douban

Datasets	K	PMF (ms)	TrustPMF (ms)	SocialMF (ms)	PRM (ms)	EnSocialMF (ms)	USSHMF (ms)	SocialIT (ms)	ISocialMF (ms)
Epinions	5	10.23	11.35	11.32	11.31	11.51	11.23	11.15	11.67
	10	13.25	15.41	13.28	14.52	15.56	13.33	15.36	16.51
	15	18.36	19.28	18.36	18.89	19.92	18.78	18.55	19.27
Douban	5	9.28	10.89	10.38	10.22	10.87	10.56	10.89	11.35
	10	12.63	13.78	12.92	13.3	14.62	12.23	14.31	15.03
	15	16.88	17.92	16.76	16.63	17.86	16.85	17.42	18.36

Douban dataset, and for cold start users, the performance is more significantly improved over the Epinions dataset.

5.4.3 Time efficiency of the algorithm

To evaluate the online recommendation efficiency of the algorithm model proposed in this paper, we set the potential factors to 5, 10 and 15 and compared the online efficiency with other methods on the Epinions and Douban datasets. Table 7 gives the average online time performance of several different methods.

Table 7 shows the time performance of different methods when K is 5, 10 and 15. The average times of ISiocialMF on the Epinions and Doublan datasets are 16.51 ms and 15.03 ms when K = 10, respectively. The average times of the most basic probability matrix factorization algorithm PMF are 13.25 ms and 12.63 ms when K = 10, respectively, and the average times of the classical social matrix factorization algorithm SocialMF are 13.28 ms and 12.92 ms when K = 10, respectively. Therefore, the online running time is acceptable.

6 Conclusion and future work

Data sparsity and cold start are major challenges of CF-based RSs. This paper further eases the impact of the above problems on RS by introducing trust relationships 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 greatest advantage of ISocialMF is that it integrates the social tags and item social relationships into the MF process and trains the shared user and item features for more accurate recommendation quality through a 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.

Acknowledgements 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 212102210410, 212102310556, 202102210387, 202102210178, 212102210418, 222102210031, 222102110045, 222102210323, 222102210030, 222102210024, and 182102310969, in part by the Zhongyuan Science and Technology Innovation Leadership Program under Grant 214200510026, in part by the Natural Science Foundation Projectin of Henan Province under Grant 222300420582, in part by the Blue Book of Science Research Report on the "Belt and Road" Tourism Development Grant 2017sz01, in part by Shaanxi innovation capability support plan under Grant 2018KRM071, in part by the Doctoral Fund Project of Zhengzhou University of Light Industry under Grants 2020BSJJ030 and 2020BSJJ031, and in part by the innovation team of data science and knowledge engineering of Zhengzhou University of Light Industry under Grant 13606000032.

Authors contribution The content of this article has not been published, nor has it been submitted for consideration to other journals. There is no conflict of interest in the content of this article. With the consent of all the authors, this article will be authorized for publication. The contributions of each author in this article are as follows: Dr. Rui Chen wrote the article, Prof. Jian-wei Zhang revised the paper, Prof. Zhifeng Zhang 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.

Funding The authors have not disclosed any funding.

Data availability Enquiries about data availability should be directed to the authors.

Declaration

Conflict of interest The authors have not disclosed any competing interests.

References

Aghdam MH (2019) Context-aware recommender systems using hierarchical hidden Markov model. Physica A 518:89–98

Ahmadian S, Afsharchi M, Meghdadi M (2018) An effective social recommendation method based on user reputation model and rating profile enhancement. J Inf Sci 28:180–192. https://doi.org/10.1177/0165551518808191

Ahmadian S, Joorabloo N, Jalili M et al (2020) A social recommender system based on reliable implicit relationships. Knowl Based Syst 192:105371. https://doi.org/10.1016/j.knosys.2019.105371

Azadjalal MM, Moradi P, Abdollahpouri A et al (2017) A trust-aware recommendation method based on Pareto dominance and confidence concepts. Knowl-Based Syst 116:130–143

Bagher RC, Hassanpour H, Mashayekhi H (2017) User trends modeling for a content-based recommender system. Expert Syst Appl 87:209–219

Can U, Alatas B (2019) A new direction in social network analysis: Online social network analysis problems and applications. Physica A Statal Mech Appl 535(1):1–13

Cao Y, Li W, Zheng D (2018) An improved neighborhood-aware unified probabilistic matrix factorization recommendation. Wireless Pers Commun 102(4):3121–3140

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

Cohen D, Aharon M, Koren Y, et al (2017) Expediting exploration by attribute-to-feature mapping for cold-start recommendations. In: Proceedings of the 11th ACM conference on recommender systems (RecSys'17), pp 184–192

Feng S, Cao J, Wang J et al (2017) Recommendations based on comprehensively exploiting the latent factors hidden in items' ratings and content. ACM Trans Knowl Discov Data 11(3):35–46

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

Gong C, Tao D, Chang X et al (2019) Ensemble teaching for hybrid label propagation. IEEE Trans Cybern 49(2):388–402

Guan J, Xu M, Kong X (2018) Learning social regularized user representation in recommender system. Signal Process 144(3):306–310

Guo L, Ma J, Chen Z et al (2015) Learning to recommend with social context information from implicit feedback. Soft Comput 19(5):1351–1362

Gupta S, Kant V (2020) Credibility score based multi-criteria recommender system. Knowl-Based Syst 196(1):1–12

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 Syst Appl 56(9):9–27

He Y, Wang C, Jiang CJ (2018a) Correlated matrix factorization for recommendation with implicit feedback. IEEE Trans Knowl Data Eng 2018:1–15

He Y, Wang C, Jiang C (2018b) Correlated matrix factorization for recommendation with implicit feedback. IEEE Trans Knowl Data Eng 31(3):451–464

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. Inf Sci 53:156–170

Jamali M, Ester M (2010) A matrix factorization technique with trust propagation for recommendation in social networks. In: Proceedings of the 4th ACM conference on recommender systems (RecSys'10), ACM, pp 26–30

Kong X, Mao M, Wang W (2018) VOPRec: Vector representation learning of papers with text information and structural identity for recommendation. IEEE Trans Emerg Top Comput. https://doi.org/10.1109/TETC.2018.2830698

Li Y, Wang D, He H et al (2017a) Mining intrinsic information by matrix factorization-based approaches for collaborative filtering in recommender systems. Neurocomputing 249:48–63
Li J, Chen C, Chen H et al (2017b) Towards context-aware social recommendation via individual trust. Knowl-Based Syst 2017:58–66
Li H, Ma X, Shi J (2018) Incorporating trust relation with PMF to enhance social network recommendation performance. Int Pattern Recogn Artific Intell 30(6):113–124
Liu H, Jing L, Yu J (2017) Survey of matrix factorization based recommendation methods by integrating social information. J Softw 2017:1–24. https://doi.org/10.13328/j.cnki.jos.005391
Lu Q, Guo F (2019) Personalized information recommendation model based on context contribution and item correlation. Measurement 142:30–39
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 Trans Ind Inf 14(5):2011–2022
Ma H, Yang H, Lyu M R, et al (2008) SoRec: social recommendation using probabilistic matrix factorization. In: Proceedings of ACM conference on information & knowledge management (CIKM'08), pp 931–940
Ma H, King I, Lyu RM, et al (2009) Learning to recommend with social trust ensemble. In: Proceedings of 32nd international ACM SIGIR conference on research and development in information retrieval, pp 1–8
Meng X, Liu S, Zhang Y et al (2015) Research on social recommender systems. J Softw 26(6):1356–1372
Nabizadeh A, Leal J, Rafsanjani H et al (2020) (2020) Learning path personalization and recommendation methods: A survey of the state-of-the-art. Expert Syst Appl 159(9):1767–1776. https://doi.org/10.1016/j.eswa.2020.113596
Pan Y, He F, Yu H (2018) Social recommendation algorithm using implicit similarity in trust. Chinese J Comput 41(1):65–81
Panagiotakis C, Papadakis H, Papagrigoriou A et al (2021) Improving recommender systems via a dual training error based correction approach. Expert Syst Appl 183(5):115386
Paradarami TK, Bastian ND, Wightman JL (2017) A hybrid recommender system using artificial neural networks. Exp Syst Appl 83:300–313
Pereira BL, Ueda A, Penha G, et al (2019) Online learning to rank for sequential music recommendation. In: Proceedings of the 13th ACM conference on recommender systems (RecSys'19). ACM, Copenhagen, Denmark. New York, NY, pp 237–245
Portugal I, Alencar P, Cowan D (2018) The use of machine learning algorithms in recommender systems: a systematic review. Expert Syst Appl 97(1):205–227
Rafailidis D, Crestani F (2017) Learning to rank with trust and distrust in recommender systems. In: Proceedings of the 11th ACM conference on recommender systems (RecSys'17). ACM, pp 5–13
Rezaeimehr F, Moradi P, Ahmadian S (2017) TCARS: Time- and community-aware recommendation system. Futur Gener Comput Syst. https://doi.org/10.1016/j.future.2017.04.003
Ricci F, Rokach L, Shapira B, et al (2010) Recommender systems handbook: context-aware recommender systems. Springer, New York, pp 217–253
Sa A, Nj B, Mj B et al (2021) Alleviating data sparsity problem in time-aware recommender systems using a reliable rating profile enrichment approach. Expert Syst Appl 2021:1–15
Salakhutdinov R, Mnih A (2008) Probabilistic matrix factorization. In Proceedings of NIPS
Sambhav Y, Vikash S et al (2018) Trust aware recommender system using swarm intelligence. J Comput Sci 28(1):180–192. https://doi.org/10.1016/j.jocs.2018.09.007
Sedhain S, Menon A K, Sanner S, Xie L, Braziunas D (2017) Low-rank linear cold-start recommendation from social data. In: Proceedings of the 31st AAAI conference on artificial intelligence. AAAI Press
Seo Y, Kim Y, Lee E et al (2017) Personalized recommender system based on friendship strength in social network services. Expert Syst Appl 69:135–148
Shneiderman B (2020) Human-centered artificial intelligence: reliable, safe & trustworthy. Int J Human-Comput Interact 36:495–504. https://doi.org/10.1080/10447318.2020.1741118
Shokeen J, Rana C (2020) Social recommender systems: techniques, domains, metrics, datasets and future scope. J Intell Inf Syst 54(2):1–35
Tang J, Gao H, Hu X, et al (2013) Exploiting homophily effect for trust prediction. In: Proceedings of ACM international conference on web search and data mining (WSDM). ACM, pp 53–62
Wang Y, Wang X, Zuo W (2014) Trust prediction modeling based on social theories. J Softw 25(12):2893–2904
Wei J, He J, Chen K et al (2017) Collaborative filtering and deep learning based recommendation system for cold start items. Expert Syst Appl 69:29–39
Yan S, Lin KJ, Zheng X et al (2017) An approach for building efficient and accurate social recommender systems using individual relationship networks. IEEE Trans Knowl Data Eng 29(10):2086–2099
Yang X, Guo Y, Liu Y, Steck H (2014) A survey of collaborative filtering based social recommender systems. Comput Commun 41(1):1–10
Yang B, Yu L, Liu J et al (2017) Social collaborative filtering by trust. IEEE Trans Pattern Anal Mach Intell 39(8):1633–1647
Yao W, He J, Huang G, et al (2014) Modeling dual role preferences for trust-aware recommendation. In: Proceedings of International ACM SIGIR conference on research & development in information retrieval. ACM, pp 975–978
Yu W, Li S (2018) Recommender systems based on multiple social networks correlation. Futur Gener Comput Syst 87(1):312–327
Yu Y, Gao Y, Wang H (2018) Integrating user social status and matrix factorization for item recommendation. J Comput Res Dev 55(1):113–124
Zhang Z, Xu G, Zhang P et al (2017) Personalized recommendation algorithm for social networks based on comprehensive trust. Appl Intell 47(3):659–669
Zhang W, Du Y, Yang Y et al (2018) DeRec: a data-driven approach to accurate recommendation with deep learning and weighted loss function. Electron Comm Res Appl 31:12–23
Zheng X, Luo Y (2018) Sun L (2018) A novel social network hybrid recommender system based on hypergraph topologic structure. World Wide Web-Int Web Inf Syst 21:985–1013