

A Factored Similarity Model with Trust and Social Influence for Top-N Recommendation

X.F. Zhang, X.L. Chen, D.W. Seng, X.J. Fang

Xuefeng Zhang

School of Computer Science and Technology, Hangzhou Dianzi University
Hangzhou 310018, China

Xiuli Chen

School of Computer Science and Technology, Hangzhou Dianzi University
Hangzhou 310018, China

Dewen Seng*

School of Computer Science and Technology, Hangzhou Dianzi University
Hangzhou 310018, China

*Corresponding author: sengdw@163.com

Xujian Fang

School of Computer Science and Technology, Hangzhou Dianzi University
Hangzhou 310018, China

Abstract: Many trust-aware recommendation systems have emerged to overcome the problem of data sparsity, which bottlenecks the performance of traditional Collaborative Filtering (CF) recommendation algorithms. However, these systems most rely on the binary social network information, failing to consider the variety of trust values between users. To make up for the defect, this paper designs a novel Top-N recommendation model based on trust and social influence, in which the most influential users are determined by the Improved Structural Holes (ISH) method. Specifically, the features in Matrix Factorization (MF) were configured by deep learning rather than random initialization, which has a negative impact on prediction of item rating. In addition, a trust measurement model was created to quantify the strength of implicit trust. The experimental result shows that our approach can solve the adverse impacts of data sparsity and enhance the recommendation accuracy.

Keywords: recommendation system, matrix factorization, trust, social influence, deep learning, top-n recommendation.

1 Introduction

The dawn of the big data era has brought abundant information resources. However, human beings are sometimes provided with too much information, that is, faced with information overload. To solve the problem, many portals and e-commerce systems have adopted recommendation systems to push the desired information to users. These systems learn users' preferences from their historical behaviors, aiming to recommend the items meeting their interests and needs.

CF is one of the most popular and successful personalized recommendation algorithms. The algorithm discovers the preferences of users by mining the data on their historical behaviors, divides the users into different groups based on their preferences, and recommends similar items to users in each group. The recommendation accuracy of the CF hinges on two issues: inferring user preferences from historical data, and identifying similar users or items. The traditional CF algorithm faces two major challenges, namely, data sparsity and cold start. Many methods have been developed to cope with these challenges. Some scholars have introduced various auxiliary

data to improve the CF, including membership [7, 26], trust [2, 8, 14, 23, 27, 28] and other social information [5, 21, 25].

With the development of social networking, it is increasingly convenient to acquire the data on social relationship. Against this background, a series of trust-based approaches have been proposed and applied in various domains, making up an integral part of recommendation algorithms. Trust relationship, closely related to social information, is a hotspot in the research of recommendation systems. In social networks, users always prefer the items recommended by those they trust. Therefore, trust-aware recommendation can greatly improve the recommendation effect of sparse users (i.e. data sparsity), leading to better user experience and enhanced loyalty. Based on trust propagation mechanism, Jamali and Ester [10] designed a MF-based model for recommendation in social rating networks (SocialMF). The potential feature vectors of target users were described as the weighted average of their trustee's feature vectors, which significantly reduced the recommendation error, especially for cold-start users. Guo et al. [8] extended the SVD++ algorithm [14] into the TrustSVD algorithm, which achieves excellent recommendation effect through considering the explicit trust relationship of the target user. The above studies fully demonstrate the promotion effect of the trust relationship on recommendation accuracy. However, these trust relationships are single and incomplete.

The observations on real-world datasets like FilmTrust, Epinions and Ciao reveal two problems of trust relationship: (1) the trust relationship may be implicit due to privacy concerns, resulting in the sparsity of trust information; (2) the trust value is usually binary, i.e. each trusted friend has the same impact on the target user, which goes against the reality.

Therefore, the existing trust information can not fully mine the implicit user preference information in social networks. It is necessary to accurately measure the trust relationship, and examine its role in item recommendation.

The user behavior in the social network is also greatly affected by social influence. The behavior or opinion of a person is easily affected by influential users, who seem to be authoritative. To study user behavior, it is both necessary and difficult to analyze the influence of different users out of the massive information in the social network.

The social influence analysis [18] refers to the evaluation of the impact of each user against his/her information or social behavior by a certain standard, with the aim to identify those with significant impact in his/her group or the social network.

In fact, the analysis of social influence has attracted much attention from scholars engaging in sociology, management, information science, and economics [1, 12, 13, 15, 22]. For example, Kitsak et al. [13] empirically investigated social networks and email networks, and divided network nodes by k-kernel decomposition to different levels from the edge to the core, revealing that the most influential nodes are not necessarily the most connected nodes, i.e. those with high betweenness centrality, but those with high k-shells. Nevertheless, in the field of recommendation system, trust plays an important role, and social influence is neglected by most researchers, although it is also a key factor affecting the behavior of social network users.

Our research is mainly motivated by two issues. On the one hand, the existing trust-based recommendation methods usually use the binary trust relationship of social networks directly to improve the recommendation quality, seldom considering the difference and potential impact of trust intensity among users, and also can not fully tap the implicit user preference information in social networks, which can not achieve good recommendation effect in the case of sparse trust data.

On the other hand, the existing recommendation systems rarely utilize social influence, despite its significant impact on user behavior in the social network. To overcome these two defects, this paper fully excavates the implicit social relationships among users in the social network, and propose a top-N item recommendation algorithm which fusion of trust and social influence, called

Factored user and item Similarity Model with Trust and social Influence based on Deep learning (FSTID). The architecture of our recommendation algorithm is shown in Fig. 1. Finally, the proposed algorithm was proved more accurate and effective than other recommendation methods on three datasets (Epinions, Ciao and FilmTrust).

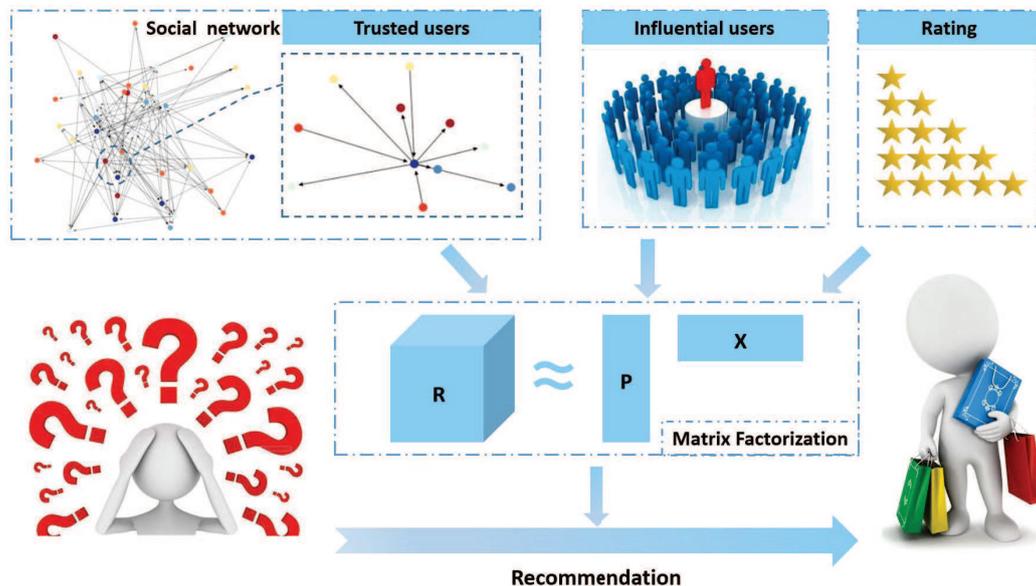


Figure 1: The architecture our recommendation algorithm

There are four main contributions of this research:

- (1) A novel trust measurement model was conducted to quantify the implicit trust in social network. The trust value covers three aspects, namely, user interaction, item rating and user preference. The combination of the two kinds of relationship between trust and similarity are compact, which are effective to solve the extreme data sparsity of recommendation system. In addition, our model has played a significant positive role correspond to accurate positioning neighbor users.
- (2) The ISH method was proposed to pinpoint the key nodes, i.e. the most influential users, in the social network.
- (3) Considering the influence of user trust and social influence on recommendation, an innovative top-N recommendation model with the incorporation of user and item similarity, trust and social influence is proposed. Besides, Deep Autoencoder (DAE) technology was employed to optimize the initial values of the latent features for MF.
- (4) Many top-N recommendation algorithms were compared through experiments on three real-world datasets. The comparison shows that our approach can alleviate the sparse problem of rating data to a certain extent and obtain more reliable recommendations.

The remainder of this paper is organized as follows: Section 2 details the construction of our algorithm; Section 3 compares our algorithm with other top-N recommendation algorithms through experiments; Section 4 puts forward the conclusions and foretells the future research.

2 Model construction

This section introduces the overall framework of the FSTID and details the four key stages of the method in seven steps. Before presenting the FSTID, a new trust measurement model was first established, the ISH method was employed to identify the most influential nodes in the social network, and the DAE was adopted to extract user and item feature vectors.

Let $U = \{u_1, u_2, \dots, u_m\}$ be the set of m users, $I = \{i_1, i_2, \dots, i_n\}$ be the set of n items, $R = [r_{u,i}]_{m \times n}$ be the binary matrix of the item ratings by the user ($r_{u,i} = 1$ means user u has purchased or rated item i ; otherwise, user u has not purchased or rated item), $G = (V, E)$ be the graph of social network (V is the set all nodes (users) and E is the set of all edges), and $\langle i, j \rangle$ be the edge from node i to node j , that is, the trust relationship between users i and j .

2.1 Overall framework

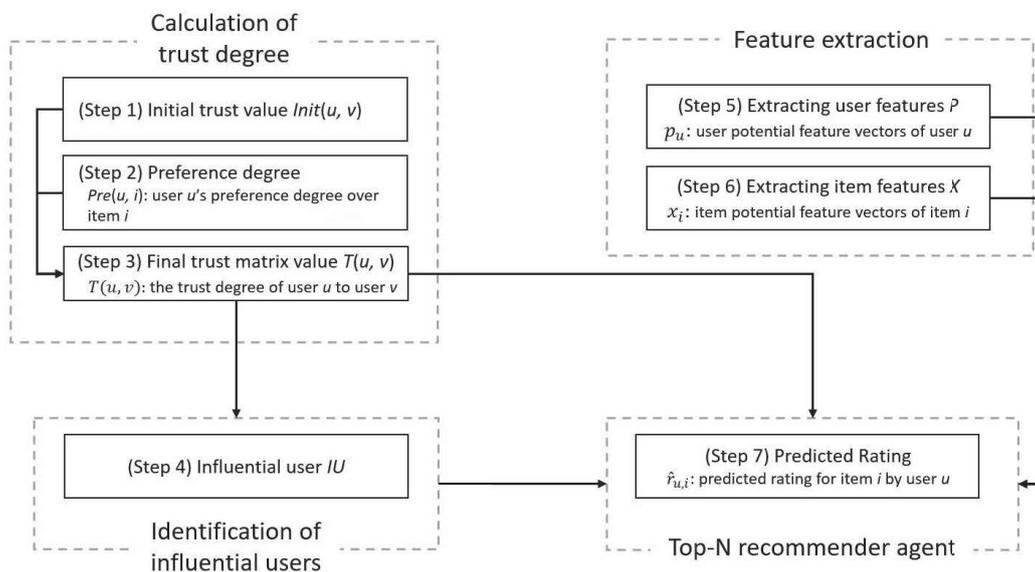


Figure 2: The overall framework of the FSTID

As shown in Fig. 2, the procedure of the FSTID algorithm roughly contains four key stages: the calculation of trust value, the identification of influential users, feature extraction, and top-N recommendation. Firstly, the trust value is computed based on the initial trust value and preference degree of each user, forming the trust network. Then, the ISH is adopted to identify the key nodes in the social network, and allocate them into the set of influential users. After that, the features of users and items are extracted by the DAE.

Finally, the CF is applied to predict the item ratings, and provide personalized recommendation. The four stages are implemented in seven steps:

(1) Calculation of trust value

Step 1: The trust value of each user is initialized based on his/her interaction information with others. The term interaction is defined as two users rating on the same item i .

Step 2: Considering the mutual influence between trust value and interactions, and the users' tendency to trust those who have interacted successfully on their favorite items, the author measured the user preference over different items by preference degree.

Step 3: Based on user preference over item i , different weights are assigned to items in each interaction, then generate the final trust value, and the trust network is formed after filtering

out the weak trust relationship.

(2) *Identification of influential users*

Step 4: The constraint value of each node is computed according to the topology of its neighbors in the directed trust network. Taking the constraint value as the indicator, the importance of each node is evaluated and recorded as "influence". Finally, the most influential users are determined and allocated to the same set.

(3) *Feature extraction*

Steps 5 and 6: By using DAE to learn users' behaviors without supervision so that the high-dimensional, sparse users' behaviors can be compressed into low-dimensional, dense users and items feature vectors. Compared to initial features, these features are no longer sparse, but also more representative.

(4) *Top-N recommendation*

Step 7: Combining the user similarities and item similarities, we introduce trust users and influential users to predict rating and offer personalized recommendation for users

2.2 Calculation of trust value (Steps 1-3)

The trust value can be expressed as $T = [t_{u,v}]_{m \times m}$, where $t_{u,v}$ is a nonzero element indicating the existence of social relationship between users u and v . In the real world, the trust relationship of most online social networks, e.g. Facebook, Epinions and Flixster, is usually represented in binary form (0 or 1), which is not exactly the same as the trust relationship of users.

This subsection sets up a novel trust measurement model, regardless if the trust value obeys explicit distribution. First, two assumptions were put forward: (1) an interaction is defined as two users perform a rating on the same item; (2) the trust value comes from cumulative experience of subjective individuals. Under these assumptions, the trust value can be initialized as:

$$Init(u, v) = \frac{\min(|I_u \cap I_v|, D_u)}{D_u} \quad (1)$$

where $I_u \cap I_v$ is the number of interactions between users u and v , i.e. the number of items rated by both users; $D_u = \sqrt{|I_u|}$ is an adjustable threshold specifying the minimum number of interactions between the two users that fully trust each other. The trust relationship is not a stable factor, which along with influence of interaction experience will be changed. A successful interaction would increase the trust value and an unsuccessful interaction act oppositely. If both users u and v have rated item i , then the interaction is successful if the two ratings are both above or below the average rating of user himself:

$$\begin{cases} \text{successful,} & (r_{u,i} - \bar{u}) * (r_{v,i} - \bar{v}) \geq 0 \\ \text{unsuccessful,} & (r_{u,i} - \bar{u}) * (r_{v,i} - \bar{v}) < 0 \end{cases} \quad (2)$$

where \bar{u} and \bar{v} are the average rating of users u and v on all items, respectively.

The trust value differs with the preference degree of each user. Let $Pre(u, i)$ be the preference degree of user u over item i , U_i be the set of users that have rated item i , and o be a user in the set U_i . Then, the similarity between users u and o can be calculated using the Pearson correlation coefficient (PCC):

$$sim(u, o) = \left(\frac{1}{2} + \frac{\sum_{i \in I_u \cap I_o} (r_{u,i} - \bar{u})(r_{o,i} - \bar{o})}{2 * \sqrt{\sum_{i \in I_u \cap I_o} (r_{u,i} - \bar{u})^2} \sqrt{\sum_{i \in I_u \cap I_o} (r_{o,i} - \bar{o})^2}} \right) * \frac{|I_u \cap I_o|}{|I_u|} \quad (3)$$

$$Pre(u, i) = \frac{\sum_{o \in U_i} sim(u, o)}{|U_i|} \quad (4)$$

Based on the user preferences in successful or unsuccessful interactions, different weights were assigned to different items. On this basis, the final trust value $T(u, v)$ can be obtained as:

$$T(u, v) = \frac{\sum_{i \in \text{successful}} \text{Pre}(u, i) - \sum_{i \in \text{unsuccessful}} \text{Pre}(u, i)}{\sum_{i \in \text{successful}} \text{Pre}(u, i) + \sum_{i \in \text{unsuccessful}} \text{Pre}(u, i)} * \text{Init}(u, v) \quad (5)$$

2.3 Identification of influential nodes (Step 4)

In the social network, each user collects and disseminates information as a node. Some of them are more influential than others. These key nodes are similar to Structural Hole (SH) occupants in the competitive group theory proposed by Burt [3]. The SH occupants play a key role in information exchange between the local groups and enjoys more opportunities and options than other group members. Therefore, the author improved the SH method to integrate the impacts of in- and out- degrees of neighbor nodes in the directed trust network. The ISH is illustrated with an instance in Fig. 3 and Table 1. Here, the in-degree is defined as the number of nodes pointing towards a node, and the out-degree has the opposite meaning.

The SH is generally evaluated by two indices: the betweenness centrality and the index given by Burt himself. The former refers to Freeman's betweenness centrality [6] for the overall network and its extended form. If a node is on the shortest path of many pairs of other nodes, then the node has a high betweenness centrality, and a high probability to be a SH occupant. The latter involves four aspects: effective size, efficiency, constraint and hierarchy. Constraints refer to the node's ability to use SH in its own networks, which is more important than the other three aspects. The calculation formula of constraint is as follows, describing the degree to which a node in a network is directly or indirectly connected to other nodes:

$$C(i) = \sum_{j \in \Gamma(i)} \left(p(j, i) + \sum_{q \in \Gamma(i)} p(j, q) * p(q, i) \right)^2, i \neq q \neq j \quad (6)$$

where $\Gamma(i)$ refers to the nodes directly connected to node i in the undirected graph; $p(j, i)$ is the proportion of the energy that node i devotes to maintain its relationship with node j to the total energy of node i . Since the trust network is a directed graph, $\Gamma(i)$ was replaced with T_i^- to describe the set of nodes pointing towards node i . The value of $p(j, i)$ can be calculated by:

$$p(j, i) = \frac{Z_{ji}}{\sum_{j \in T_i^-} Z_{ji}} \quad (7)$$

where $Z_{ji} = 1$ if $\langle i, j \rangle \in E$ (otherwise, $Z_{ji} = 0$); $p(j, i) + \sum_{q \in \Gamma(i)} p(j, q) * p(q, i)$ represents the time and energy that node i invests directly or indirectly to maintain its connection with node j , when node j is the only node pointing to node i , the item reaches a maximum value of 1; this item reaches a minimum when there is no bridge node between node i and node j .

Based on the $C(i)$ value obtained from Eq. (6), the number and closeness of neighbors can be evaluated in a comprehensive manner. The $C(i)$ value is negatively correlated with the number of nodes pointing towards the target node, and positively with the closeness between these nodes. If closely distributed, the nodes are less likely to acquire new information. On the contrary, nodes with small constraint coefficients can promote information propagation.

Table 1 lists the constraint values of all nodes in Fig.3 (a), respectively calculated by the SH and the ISH. During the calculation, any node with no node pointing towards it was neglected. As shown in the table, nodes u_4 and u_7 have the same constraint value, i.e. the same influence. The information flow theory [20] holds that the information on the Internet often propagates to the opinion leader, before spreading to the others distributed in a wide range. In this sense,

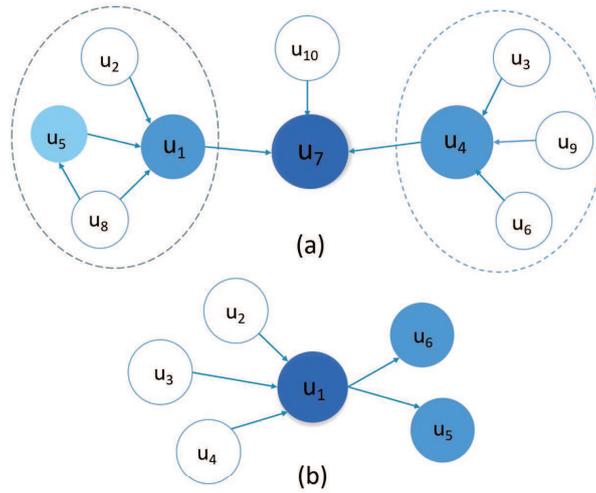


Figure 3: Two toy examples of trust network topological graph

Table 1: Constraint values of all nodes in Fig. 3(a) obtained by the SH and the ISH

$C(i)$	u_1	u_4	u_7	u_8
SH	0.6667	0.3333	0.3333	1
ISH	0.5555	0.3333	0.0625	1

a node connected to multiple opinion leaders is likely to become an SH occupant. Taking the in-degree of node as the evaluation standard for opinion leader, nodes u_1, u_4 and u_7 with the highest in-degrees must be opinion leaders, and play a greater role than the other nodes (e.g. u_8 and u_9) in information dissemination. Among them, node u_7 are connected to two opinion leaders at the same time, indicating that it is more conducive to information dissemination than nodes u_1 and u_4 . In other words, node u_7 is more likely to occupy an SH than the other two nodes. The SH results in Table 1 are obviously unreasonable, as many important nodes are not found. This is because the traditional SH only measures the relationship between a node and its nearest neighbor, without considering that between the node and its two-hop neighbors.

For convenience, it is assumed that a key node j can boost the influence of the node i it points towards. The nodes pointing away from node j were also taken into account. As shown in Fig.3(b), the influence of node u_6 is reduced due to the fact that the only node pointing towards it u_1 has a node pointing away from it u_5 . Therefore, the ISH integrating the impacts of in- and out-degrees of neighbor nodes can be expressed as:

$$C(i) = \sum_{j \in T_i^-} \frac{|T_j^+|}{|T_j^+| + |T_j^-|} * \left(p(j, i) + \sum_{q \in T_i^-} p(j, q) * p(q, i) \right)^2, i \neq q \neq j \tag{8}$$

where T_j^+ is the set of nodes pointing away from node j .

The validity of the ISH was tested with the instance in Fig.3(a). The test results are recorded in Table 1. It can be seen that node u_7 , with the lowest constraint value, is the most influential node. Both u_1 and u_4 had an in-degree of 3, yet u_8 , a node pointing towards u_1 , has a connection with u_5 , which reduces the influence of u_1 on u_8 . Thus, the influence of u_4 is higher than that of u_1 .

After the constraint values of all nodes were obtained by Eq. (8), the top $k\%$ of users with the smallest constraint values were selected as global influential users IU .

2.4 Feature extraction (step 5 and step 6)

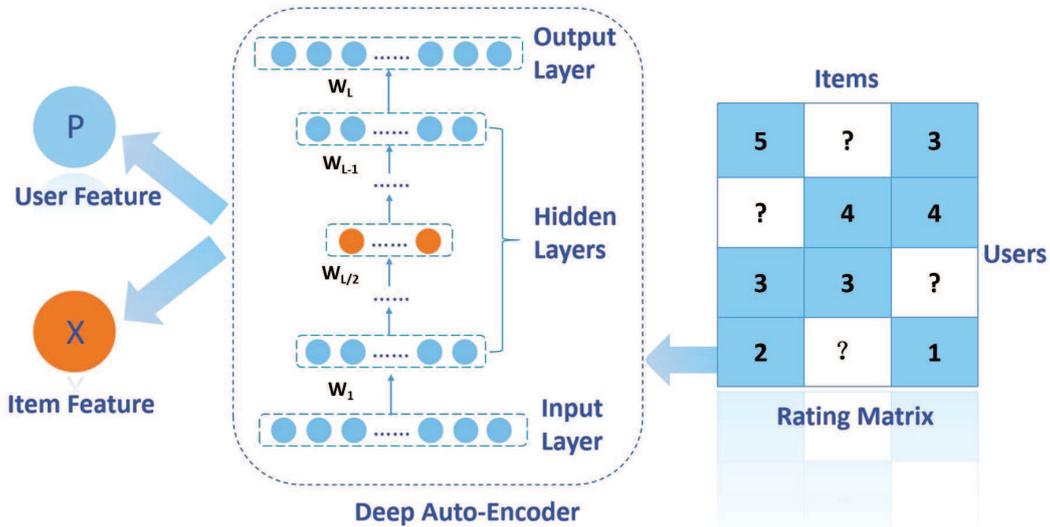


Figure 4: The flowchart of feature extraction.

Considering the sheer number of users and items, the user vectors and item vectors were respectively mapped by the MF to d -dimensional joint potential factor spaces $P_{d \times m}$ and $X_{d \times n}$. Explain intuitively that matrix P represents the preferences of m users on d topics, and X represents the d topics which are concerned by users behind n items, then the rating matrix R can be approximately expressed as the product of two matrices: $R = P^T X$. Most MF-based recommendations initialize P and X randomly and iteratively output local optimal solutions, which depends heavily on the initial values [4]. This paper attempts to make more accurate recommendations by mining out accurate feature vectors from the implicit features of the MF. To this end, the DAE, a deep neural network, was adopted to compress high-dimensional, sparse user behaviors into low-dimensional, dense feature vectors of users and items. This approach provides good initial values to the network through unsupervised layer-by-layer pre-training, and adjust the network weights by supervised fine-tuning training, making it possible to extract key information from the data and form features [24].

The primary goal is to design an item-based (user-based) DAE to map the observed $U_i(I_u)$ into a low-dimensional potential (implicit) space, and then reconstruct $U_i(I_u)$ in the output space to learn advanced, abstract features. The entire workflow is illustrated in Fig.4. Let $S = [I_1, \dots, I_i, \dots, I_m]$ be the input. The specific steps of the DAE to extract user feature matrix P are given below:

Step 1: Given an input S , let $W_i \in \mathbb{R}^{d \times m} (i \in [1, L - 1])$ be a weight matrix, $b_i \in \mathbb{R}^d$ be the bias vectors and $\sigma(x) = \frac{1}{1+e^x}$ be the sigmoid formula. Then, the d -dimensional implicit feature h_i can be expressed as:

$$h_1 = \sigma(W_1 S + b_1) \tag{9}$$

$$h_i = \sigma(W_i h_{i-1} + b_i) \tag{10}$$

Step 2: Given the weight matrix $W_L \in \mathbb{R}^{m \times d}$ between hidden and output layers, and bias

vector $b_L \in \mathbb{R}^m$, the DAE can reconstruct the original data \hat{S} from the hidden layer h_{L-1} by:

$$\hat{S} = \sigma(W_L h_{L-1} + b_L) \quad (11)$$

Step 3: The pretraining program of the DAE intends to minimize the objective function by adjusting weight matrices W and bias vectors b :

$$L = \frac{1}{2m} \sum_{i=1}^m |\hat{S}_i - S_i|^2 + \frac{\lambda}{2} |W_1|^2 + \dots + \frac{\lambda}{2} |W_L|^2 \quad (12)$$

where \hat{S}_i and S_i represent the i -dimensional vectors of \hat{S} and S , respectively; $|\cdot|^2$ is the squared L2 norm of vectors or matrices, i.e. the sum of squared dimensional values; the first factor in objective function denotes the error which is employed to minimize the error of reconstructed data \hat{S}_i and original data S_i ; the other factors are regular terms used to prevent overfitting.

Step 4: Update weight matrix W in each iteration by:

$$W = W - l \times \frac{\partial L}{\partial W} \quad (13)$$

where l refers to learning rate. According to the above formula, bias vectors b update process act in the same way.

The d -dimensional feature vector of the users and the items were obtained by repeating the above steps to continuously optimize the model until the end of training.

2.5 Top-N recommendation agent (step 7)

Our approach FSTID, is based on the model proposed by Guo called Factored user and item Similarity model with social Trust (FST) [9], which predict user u 's rating on item i through four parts: (1) item bias b_i ; (2) the similarity between user u and another user v who has rated item i : $p_v^T q_u$, where p_v, q_u are the implicit feature vectors of users v and u , respectively; (3) the similarity between item i and another item j which has been rated by user u : $x_j^T y_i$, where x_j and y_i are the implicit feature vectors of items j and i , respectively; (4) influence of any user w 's trusted user u on targeted item i : $p_w^T y_i$. The rating prediction formula can be expressed as:

$$\hat{r}_{u,i} = s|U_{i-u}|^{-\beta} \sum_{v \in U_{i-u}} p_v^T q_u + (1-s)|I_{u-i}|^{-\alpha} \sum_{j \in I_{u-i}} x_j^T y_i + |T_u|^{-z} \sum_{w \in T_u} p_w^T y_i + b_i \quad (14)$$

where U_{i-u} refers to the set of users having rated item i except user u ; I_{u-i} refers to the set of items having been rated by user u except item i ; T_u is the set of users trusted by user u ; $s \in [0, 1]$ is the relative importance of user similarity; α, β, z are parameters related to the number of rated items, similar users and trusted users, respectively. Taking β for instance, $\beta = 1$ refers to the mean user similarity, item i will get a high rating only if all users in U_{i-u} are similar to user u ; when $\beta = 0$, this item calculates the sum of the similarities between user u and user in U_{i-u} , which means that item i will get a high rating even if only a few users are similar to user u . In conclusion, the probability of obtaining a high predictive rating decreases with the increase of parameter β .

The above analysis shows the critical importance of the influential users' ratings on item i . In addition, Guo et al. [9] confirmed that the trustee of user u also affects user u 's rating on item i . Thus, the rating prediction formula of the FST can be modified as:

$$\begin{aligned} \hat{r}_{u,i} = & s|U_{i-u}|^{-\beta} \sum_{v \in U_{i-u}} p_v^T q_u + (1-s)|I_{u-i}|^{-\alpha} \sum_{j \in I_{u-i}} x_j^T y_i + \delta|T_u^+|^{-z} \sum_{o \in T_u^+} p_o^T y_i + \\ & (1-\delta)|T_u^-|^{-z} \sum_{c \in T_u^-} p_c^T y_i + |IU|^{-\mu} \sum_{f \in IU} p_f^T y_i + b_i \end{aligned} \quad (15)$$

The trustees in the FST were divided into a set of trustees T_u^+ and a set of trustors T_u^- , aiming to disclose their impacts on the rating of item i . The parameter $\delta \in [0, 1]$ was adopted to control the weights of the two types of users. When $\delta = 0$, it represents the influence of trustees is totally ignored. To the contrary, $\delta = 1$ denotes only the influence of trustees is considered. $\mu \geq 0$ refers to the parameter of the number of influential users. For each influential user $f \in IU$, its influence over item i is described by the inner product $p_f^T y_i$. In addition, the model parameters b, P, Q, X and Y can be trained by minimizing the objective function below:

$$J = \frac{1}{2} \sum_{u \in U} \sum_{i \in I_u^+, j \in I_u^-} |(r_{u,i} - r_{u,j}) - (\hat{r}_{u,i} - \hat{r}_{u,j})|_F^2 + \frac{\lambda}{2} (|P|_F^2 + |Q|_F^2 + |X|_F^2 + |Y|_F^2 + |b|_F^2) \quad (16)$$

where U is the set of all users; I_u^+ and I_u^- respectively refers to the rated items and unrated items of user u . For simplicity, parameter λ was taken as the regularization term in all cases. The pseudocode for model learning is shown in Algorithm 0, where α, β, z, μ, s and δ are control

Algorithm 1 The learning algorithm of FSTID

Data:

$U \leftarrow$ user set; $R \leftarrow$ user item matrix; $T \leftarrow$ user trust matrix;

$IU \leftarrow$ influential user set; $P \leftarrow$ user feature matrix; $X \leftarrow$ item feature matrix;

Input:

$\alpha, \beta, z, \mu, s, \delta, \rho, \lambda, \eta$ and iteration limitation L ;

Output:

b, P, Q, X, Y

```

1: Initialize bias vector  $b$  with random values in  $(0, 0.01)$ , and set  $Q = P, Y = X$ ;
2: while  $J$  not converged or the number of iterations  $< L$  do
3:   for all  $u \in U$  do
4:     for all  $i \in I_u^+$  do
5:        $Z \leftarrow \text{sample}(\rho, I_u^-)$ 
6:       Compute Loss  $J$  by Eq. (16)
7:       Update  $b_i, b_j, q_u, y_i, y_j, p_v, x_k, p_o, p_c, p_f$  according SGD:
8:        $b_i \leftarrow b_i - \eta \times \frac{\partial J}{\partial b_i}, \quad q_u \leftarrow q_u - \eta \times \frac{\partial J}{\partial q_u}, \quad y_i \leftarrow y_i - \eta \times \frac{\partial J}{\partial y_i}$ 
9:       for all  $j \in Z$  do
10:         $b_j \leftarrow b_j - \eta \times \frac{\partial J}{\partial b_j}, \quad y_j \leftarrow y_j - \eta \times \frac{\partial J}{\partial y_j}$ 
11:         $\forall v \in U_{j-u}, p_v \leftarrow p_v - \eta \times \frac{\partial J}{\partial p_v}$ 
12:       end for
13:        $\forall v \in U_{i-u}, p_v \leftarrow p_v - \eta \times \frac{\partial J}{\partial p_v}, \quad \forall k \in U_{u-i}, x_k \leftarrow x_k - \eta \times \frac{\partial J}{\partial x_k}$ 
14:        $\forall o \in T_u^+, p_o \leftarrow p_o - \eta \times \frac{\partial J}{\partial p_o}, \quad \forall c \in T_u^-, p_c \leftarrow p_c - \eta \times \frac{\partial J}{\partial p_c}$ 
15:        $\forall f \in IU, p_f \leftarrow p_f - \eta \times \frac{\partial J}{\partial p_f}$ 
16:     end for
17:   end for
18:   iteration number++
19: end while
20: return  $b, P, Q, X, Y$ 

```

parameters, λ is the regularization parameter, η is the initial learning rate and ρ is the number of samples. Firstly, ρ unrated items are selected randomly for each user to train the model (Line 6).

Then, the parameters are optimized continuously in the training phase by Stochastic Gradient Descent (SGD) (Lines 7-15) until the loss function converges or the preset maximum number of iterations is reached (Line 3). Finally, the learning vectors and matrices are outputted (Line 20). As for comparative experiments, we will make detailed comparisons and explanations in subsequent experiments to express the advantages of our algorithm.

3 Experiments and results analysis

This section carries out a series of experiments on three real-world datasets, trying to answer the following questions:

- (1) Does the ISH enjoy higher accuracy in identifying influential users than other influence measurement methods?
- (2) How do model parameters like α , β , z and μ affect recommendation accuracy?
- (3) How does the number of influential users affect recommendation accuracy?
- (4) Does the FSTID outperform the other advanced trust-aware recommendation algorithms on typical sparse datasets?

3.1 Datasets and evaluation metrics

There real-world datasets were selected for our experiments, including FilmTrust, Ciao and Epinions, which are currently well-known test data sets. Most trust-based recommendation algorithms are used to test the performance of algorithms because they contain both user ratings and trust relationships. FilmTrust is a dataset from a movie sharing website. The data entries in the dataset are movie ratings based on a four-point scale. Ciao and Epinions are datasets from two famous consumer review websites, on which users rate commodities against a five-point scale and refer to others' ratings before deciding on whether to purchase a commodity. As shown in Table 2, all the three datasets are extremely sparse in nature.

Table 2: Specification of the used datasets

Data Set	Users	Items	Ratings	Density
Epinions	40163	139738	664824	0.0118%
Ciao	7375	99746	139738	0.0379%
FilmTrust	1508	2071	40163	1.14%

$$*\text{Density} = \frac{\#Ratings}{\#Users \times \#Items}$$

The 5-fold cross validation was adopted in our experiments. Each dataset was split randomly into 5 equal parts. In each iteration, 1 part was selected as the test set and the other 4 parts as the training set. The mean results of the 5 parts were taken as the final results. Different from the evaluation of rating prediction, the top-N recommendation performance was measured by two metrics: precision and F1-measure. F1-measure is the weighted average of precision and recall rate (the proportion of all recommended items in the items rated by the target user). Let U' be the set of users in the test set and $R_N(u)$ be the top-N items recommended to user u . Then, the precision, recall rate and F1-measure under N recommended items, respectively

denoted as $P@N$, $R@N$ and $F1@N$, can be calculated by:

$$P@N = \frac{1}{|U'|} \sum_{u \in U'} \frac{|R_N(u) \cap I_u|}{N} \quad (17)$$

$$R@N = \frac{1}{|U'|} \sum_{u \in U'} \frac{|R_N(u) \cap I_u|}{I_u} \quad (18)$$

$$F1@N = \frac{2 * P@N * R@N}{P@N + R@N} \quad (19)$$

where $N \in \{5, 10\}$ is the number of items recommended to the target user. The greater the values of $P@N$ and $F1@N$, the more accurate the top-N recommendation.

3.2 Social influence analysis

The Susceptible Infected (SI) epidemic model [15] was selected to evaluate the effects of model parameters on recommendation results. Mimicking the virus propagation process, this classic model is a mature method to simulate the transmission of information. In this model, each network node either exists in the susceptible S status or in the infected I status. Once infected, a susceptible node will irreversibly become an infected node, and transmit virus to its neighbor nodes at the probability of $\gamma \in (0, 1)$. In our experiments, the probability γ was set to 0.001 and the network nodes were initialized randomly. The influence of each network node on the three datasets were evaluated by the parameter S_i at transmission time $t = 10$. The mean value of S_i can be calculated by:

$$\bar{S}_i = \frac{1}{M} \sum_{m=1}^M S_i \quad (20)$$

where M is the repeated times of node i .

Our method ISH was compared with the SH and the Degree Centrality (DC) in terms of the social influence over the three datasets. The experimental results are presented in Fig.5, where the x-axis is the calculated influence of each node and the y-axis is the actual influence of each node.

The purple area in Fig.5 are low-impact nodes. The SH and the DC had almost the same number of low-impact nodes on all three datasets. Comparatively, the SH nodes ranked lower than the DC nodes, indicating that the SH enjoyed better effect than the DC. Meanwhile, the ISH achieved the lowest proportion of low-impact nodes among the three methods. Most ISH nodes concentrated in the second half of the ranking.

The red area in Fig.5 are high-impact nodes. By analysis of top-25 nodes, the number of high-impact nodes calculated by DC, SH, and ISH are 18, 15 and 18 in FilmTrust; the number of high-impact nodes is 8, 9, 11 respectively in Ciao and also is 9, 8, 10 in Epinions. The results are very close, revealing that the node with higher degree is more important.

As shown in Fig.5, the ISH achieved the highest mean value on all datasets, that is, our algorithm outperformed the other methods in identification. To sum up, the ISH can identify the top $k\%$ influential users correctly, without mistaking low-impact users as high-impact users.

3.3 Effects of model parameters

Effect of parameters α , β , z , μ

The parameters α , β , z , μ respectively control the influence of item similarity, user similarity, trustees, and influential users on the item ratings. To save time and space, the values of

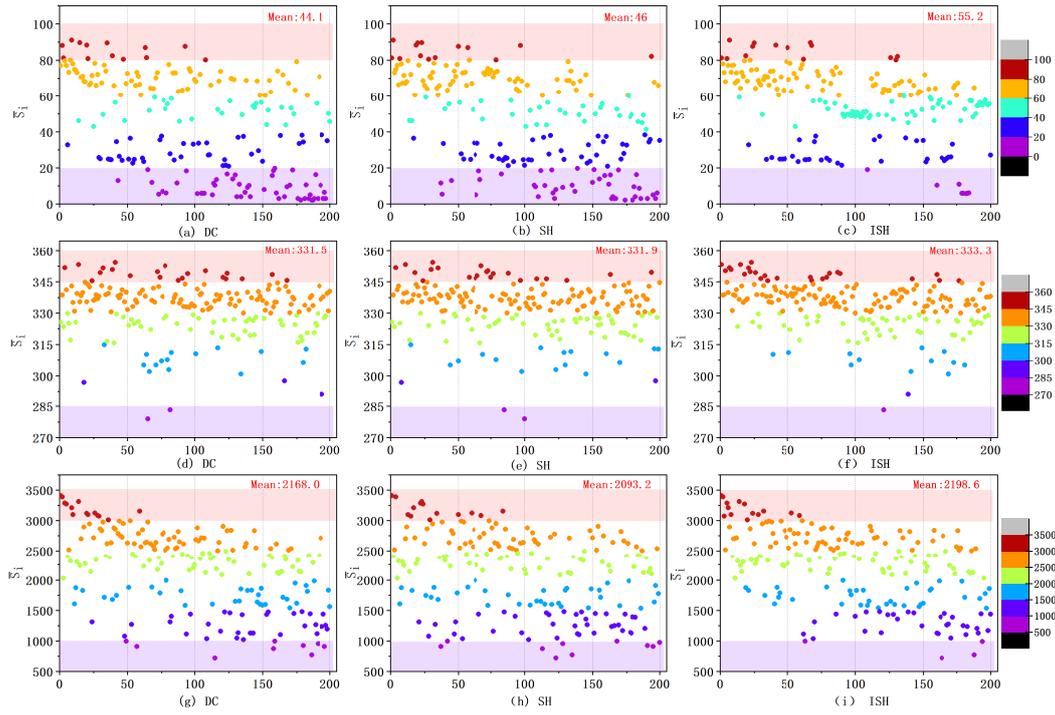


Figure 5: Correlation analysis of different methods and their actual influence. (a)-(c) Results on FilmTrust; (d)-(f) Results on Ciao; (g)-(i) Results on Epinions.

these parameters were limited to a small set: 0.5, 1, 2 while the values of s and δ were fixed at 0.5 in our experiments. For simplicity, only the five best P@10 results on each dataset are listed in Table 3. Obviously, the different parameter configurations led to different results, and the optimal parameters changed from dataset to dataset. The optimal values of α , β , z and μ on Epinions, Ciao and FilmTrust were (0.5, 2, 2, 0.5), (0.5, 2, 0.5, 0.5) and (0.5, 1, 0.5, 0.5), respectively. It can be concluded that the optimal combination is $\alpha = 0.5$, $\beta, z > 1$ and $\mu < 1$, that is, the item similarity and influential users are more important than user similarity and trustees.

Table 3: The five best P@10 results on each dataset.

	Epinions	Ciao	FilmTrust
1	0.010486 (0.5, 2, 2, 0.5)	0.02378 (0.5, 2, 2, 2)	0.352258 (2, 1, 1, 0.5)
2	0.010467 (1, 2, 0.5, 0.5)	0.023609 (0.5, 0.5, 0.5, 0.5)	0.351964 (2, 0.5, 2, 2)
3	0.010379 (2, 1, 2, 0.5)	0.023411 (1, 2, 1, 1)	0.351948 (0.5, 1, 0.5, 1)
4	0.010243 (1, 2, 1, 1)	0.023396 (0.5, 2, 1, 0.5)	0.351819 (1, 1, 0.5, 0.5)
5	0.010231 (0.5, 2, 0.5, 2)	0.02336 (2, 0.5, 0.5, 1)	0.351775 (0.5, 0.5, 0.5, 0.5)

Effect of s and δ

The parameters s and δ from Eq.(15) respectively describe the relative importance of user similarity and trustees on the prediction of item ratings. The value of s is positively correlated with the impact of user similarity, while the value of δ is negatively correlated with the impact of trustees. In our experiments, α , β , z , μ were set to the optimal values, while the values of

s and δ were increased at a step of 0.1 in the range $[0, 1]$. Firstly, the value of s was set to 0.5 to obtain a series of results on δ and determine the value of δ . Then, the experiments were conducted with s . According to the experimental results in Fig. 6, the value of P@10 was worse than most other values at $\delta = 0$ or $\delta = 1$. This means the recommendation accuracy can be improved effectively through the proper combination of trustors and trustees. This conclusion also applies to parameter s . The value of s can be set to 0.3, despite the difference in its optimal value on different datasets.

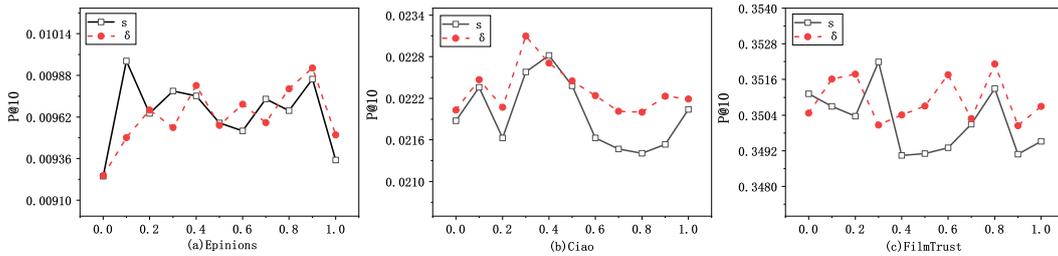


Figure 6: The effects of s and δ on FSTID at P@10

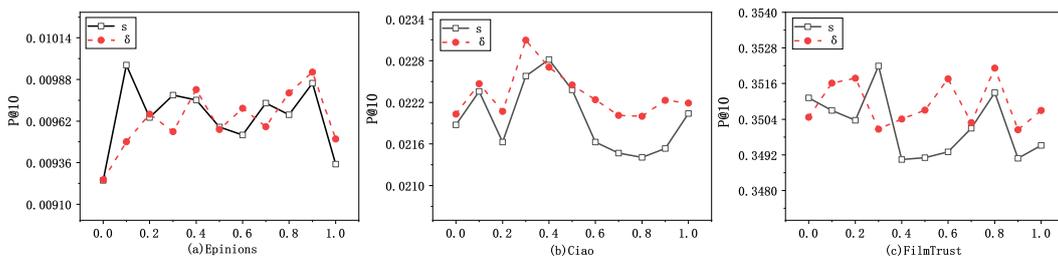


Figure 7: The effects of parameters k on FSTID at P@10

Effect of parameter k

The FSTID takes the top $k\%$ users with higher influence as influential users, as only a tiny fraction of users in the social network have a great influence. To verify the impact of parameter k on the recommendation performance, the value of the parameter was increased at the step of 1 in the interval $[0, 10]$. The experimental results are displayed in Fig. 7. It can be seen that the optimal value of k was 2, 4 and 7 in Epinions, Ciao, and FilmTrust, respectively. Moreover, the poorest result was observed on the three datasets at $k = 0$, i.e. ignoring influential users. Hence, the consideration of influential users can indeed enhance the recommendation effect.

3.4 Comparison with other methods

Several top-N recommendation algorithms were selected to evaluate the effect of our method, such as MostPop, FST, Group Bayesian Personalized Ranking (GBPR), the factored item similarity model (FISM), and the FSTID-. The MostPop is the baseline method that ranks the ratings of an item by popularity, i.e. how frequently the item is rated or consumed by the user. Proposed by Guo et al., the FST [8] takes account of implicit user feedback, similarity and social trust. The GBPR was improved from the Bayesian Personalized Ranking (BPR) by Pan and Chen [16], which introduces the influence of social groups on user preference to enhance the item recommendation quality. FISM is a top-N recommendation method based on item similarity

proposed by Kabbur et al. [11]. The FSTID- removes the DAE pretraining from our approach and randomly sets the values of P, Q, X and Y in the interval of $(0, 0.01)$.

Table 4: Results of different methods at $N=5$ and 10 . The optimal results are in bold characters.

Datasets	N	Metrics	MostPop	GBPR	FISM	FST	FSTID-	FSTID
Epinions	5	P@N	0.01169	0.009353	0.01147	0.01179	0.011888	0.01231
	10	F1@N	0.01298	0.01103	0.01307	0.0133	0.013107	0.01402
	5	P@N	0.009171	0.00756	0.00902	0.009187	0.009486	0.01024
	10	F1@N	0.01305	0.01111	0.01315	0.01328	0.014431	0.01459
Ciao	5	P@N	0.02677	0.02228	0.02704	0.02741	0.027382	0.0283
	10	F1@N	0.02436	0.02063	0.02495	0.02523	0.025442	0.02644
	5	P@N	0.02142	0.01827	0.02141	0.02174	0.022376	0.02329
	10	F1@N	0.02662	0.02116	0.02687	0.0272	0.028402	0.02914
FilmTrust	5	P@N	0.4170	0.4124	0.4171	0.4191	0.418567	0.4198
	10	F1@N	0.4095	0.4051	0.4087	0.4099	0.411364	0.4116
	5	P@N	0.3503	0.3470	0.3503	0.3514	0.352159	0.3532
	10	F1@N	0.4518	0.4458	0.4516	0.4521	0.452825	0.4541

It can be seen from Table 4 that the FSTID achieved better results than the contrastive methods on all datasets, proving that our approach can alleviate data sparsity and make reliable recommendations. Besides, the FSTID, which relies on the DAE to initialize the MF, outperformed the FSTID-, revealing that the features extracted by the DAE are more suitable than randomly generated features. On the other hand, FSTID- because there is no need to utilize DAE for feature initialization, it has lower computational cost than FSTID, and in most experiments it has achieved the best results compared with other approach except the FSTID method, which indicates the effectiveness of the integration of social influences to improve the recommended performance. Furthermore, the FSTID's advantage over the other methods were smaller on FilmTrust than Ciao and Epinions, due to the relatively small scale of FilmTrust. This means the DAE can capture features more effectively in big data.

4 Conclusion

This paper proposes a novel factored similarity model for top-N recommendation based on trust and social influence. Firstly, we quantify the trust value between users based on the user's interaction information and similarity, and finally build the user's trust association graph. In the process of calculating the trust degree, the penalty factor based on the common rating item is added to reduce the accidental effect caused by the low number of common rating items, which greatly increases the attack cost of malicious users, makes it difficult to become the neighbors of target users, and effectively enhances the stability of the algorithm. Then, the ISH was utilized to identify the influential users in a complex network. Moreover, the DAE was adopted to compress high-dimensional, sparse user behaviors into low-dimensional, dense vectors of user and item features. Compared with other top-N recommendation approaches, our approach achieved excellent recommendation effects on three real-world datasets. The future research will consider the dynamic changes in user preference and trust with the time of user rating, and try to recommend items preferred by users in a timely manner.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 61473108), the Fundamental Public Welfare Research Program of Zhejiang Province (No. GF19F020052) and the National Experimental Teaching Demonstration Center of Computer-Science in China.

Bibliography

- [1] Anagnostopoulos, A.; Kumar, R.; Mahdian, M. (2008). Influence and correlation in social networks, *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 7–15, 2008.
- [2] Bao, Y.; Fang, H.; Zhang, J. (2014). Leveraging decomposed trust in probabilistic matrix factorization for effective recommendation, *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*, 350: 30–36, 2014.
- [3] Burt, R.S. (2009); *Structural holes: The social structure of competition*, Harvard university press, 2009.
- [4] Deng, S.; Huang, L.; Xu, G. et al. (2017). On deep learning for trust-aware recommendations in social networks, *IEEE Transactions on Neural Networks & Learning Systems*, 28(5), 1164–1177, 2017.
- [5] Deng, X.Y.; Wang, C. (2018). A hybrid collaborative filtering model with context and folksonomy for social recommendation, *Ingenierie des Systemes d'Information*, 23(5), 139–157, 2018.
- [6] Freeman, L.C. (1977); A set of measures of centrality based on betweenness, *Sociometry*, 40(1), 35–41, 1977.
- [7] Guy, I.; Ronen, I.; Wilcox, E. (2009). Do you know?: recommending people to invite into your social network, *Proceedings of the 14th International Conference on Intelligent User Interfaces*, 77–86, 2009.
- [8] Guo, G.; Zhang, J.; Yorke-Smith, N. (2016). A novel recommendation model regularized with user trust and item ratings, *IEEE Transactions on Knowledge & Data Engineering*, 28(7), 1607–1620, 2016.
- [9] Guo, G.; Zhang, J.; Zhu, F. et al. (2017). Factored similarity models with social trust for top-n item recommendation, *Knowledge-Based Systems*, 122, 17–25, 2017.
- [10] Jamali, M.; Ester, M. (2010). A matrix factorization technique with trust propagation for recommendation in social networks, *Proceedings of the fourth ACM conference on Recommender systems*, 135–142, 2010.
- [11] Kabbur, S.; Ning, X.; Karypis, G. (2013). Fism: factored item similarity models for top-n recommender systems, *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 659–667, 2013.
- [12] Kempe, D.; Kleinberg, J.; Tardos, E. (2003). Maximizing the spread of influence through a social network, *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 137–146, 2003.

-
- [13] Kitsak, M.; Gallos, L.K.; Havlin, S. et al. (2010). Identification of influential spreaders in complex networks, *Nature Physics*, 6(11), 888–893, 2010.
- [14] Koren, Y. (2008). Factorization meets the neighborhood: a multifaceted collaborative filtering model, *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 426–434, 2008.
- [15] Li, D.; Luo, Z.; Ding, Y. et al. (2017). User-level microblogging recommendation incorporating social influence, *Journal of the Association for Information Science and Technology*, 68(3), 553–568, 2017.
- [16] Pan, W.; Chen, L. (2013). Gbpr: Group preference based bayesian personalized ranking for one-class collaborative filtering, *Proceedings of The Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI)*, 13, 2691–2697, 2013.
- [17] Pastorsatorras, R.; Castellano, C.; Miegheem, P.V. et al. (2014). Epidemic processes in complex networks, *Review of Modern Physics*, 87(3), 120–131, 2014.
- [18] Peng, S.; Wang, G.; Xie, D. (2017). Social influence analysis in social networking big data: Opportunities and challenges, *IEEE Network the Magazine of Global Internetworking*, 31(1), 11–17, 2017.
- [19] Rendle, S.; Freudenthaler, C.; Gantner, Z. et al. (2009). Bpr: Bayesian personalized ranking from implicit feedback, *Conference on Uncertainty in Artificial Intelligence*, 452–461, 2009.
- [20] Rogers, E.M. (1995). *The Diffusion of Innovations*, Free Press, 1995.
- [21] Sedhain, S.; Menon, A.K.; Sanner, S. et al. (2017). Low-rank linear cold-start recommendation from social data, *Proceedings of the 31th AAAI Conference on Artificial Intelligence (AAAI)*, 1502–1508, 2017.
- [22] Xu, W.; Rezvani, M.; Liang, W. et al. (2017). Efficient algorithms for the identification of top-k structural hole spanners in large social networks, *IEEE Transactions on Knowledge & Data Engineering*, 29(5), 1017–1030, 2017.
- [23] Xu, K.; Zheng, X.; Cai, Y. et al. (2018). Improving user recommendation by extracting social topics and interest topics of users in unidirectional social networks, *Knowledge Based Systems*, 140, 120–133, 2018.
- [24] Yuan, X.; Huang, B.; Wang, Y. et al. (2018). Deep learning based feature representation and its application for soft sensor modeling with variable-wise weighted SAE, *IEEE Transactions on Industrial Informatics*, (99): 1–1, 2018.
- [25] Yang, C. ; Sun, M. ; Zhao; W. X. et al. (2016). A neural network approach to joint modeling social networks and mobile trajectories, *Acm Transactions on Information Systems*, 35(4), 36, 2016.
- [26] Yuan, Q.; Zhao, S.; Chen, L. et al. (2009). Augmenting collaborative recommender by fusing explicit social relationships, *Workshop on Recommender Systems and the Social Web, Recsys*, 2009.
- [27] Zhang, Z., Liu, Y., Jin, Z. et al. (2018). A dynamic trust based two-layer neighbor selection scheme towards online recommender systems, *Neurocomputing*, 285, 94–103, 2018.
- [28] Zhao, H.; Yao, Q.; Kwok, J.T. et al. (2017). Collaborative filtering with social local models, *2017 IEEE International Conference on Data Mining (ICDM)*, 645–654, 2017.

Appendix

In order to facilitate the reader to read better, we have listed the abbreviations used in this article in lexicographic order, as shown in Table 5.

Table 5: Acronyms and Initialisms Dictionary.

Abbreviation	Full name
CF	Collaborative Filtering
DAE	Deep Autoencoder
DC	Degree Centrality
FISM	Factored Item Similarity Model
FST	Factored user and item Similarity model with social Trust
FSTID	Factored user and item Similarity Model with Trust and social Influence based on Deep learning
GBPR	Group Bayesian Personalized Ranking
ISH	Improved Structural Holes
MF	Matrix Factorization
SGD	Stochastic Gradient Descent
SH	Structural Hole
SI	Susceptible Infected