Research Article

Matrix Factorization Recommendation Algorithm Based on Multiple Social Relationships

Sheng Bin and Gengxin Sun

School of Data Science and Software Engineering, Qingdao University, Qingdao 266071, China

Correspondence should be addressed to Sheng Bin; binsheng@qdu.edu.cn

Received 4 December 2020; Revised 30 January 2021; Accepted 19 February 2021; Published 26 February 2021

Academic Editor: Mustafa Inc

Copyright © 2021 Sheng Bin and Gengxin Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the widespread use of social networks, social recommendation algorithms that add social relationships between users to recommender systems have been widely applied. Existing social recommendation algorithms only introduced one type of social relationship to the recommendation system, but in reality, there are often multiple social relationships among users. In this paper, a new matrix factorization recommendation algorithm combined with multiple social relationships is proposed. Through experiment results analysis on the Epinions dataset, the proposed matrix factorization recommendation algorithm has a significant improvement over the traditional and matrix factorization recommendation algorithms that integrate a single social relationship.

1. Introduction

With the development of big data technology, data redundancy has seriously interfered with obtaining effective information. The recommendation system solves the problem well and becomes a research hotspot in related fields. The recommendation system recommends items or information that may be of interest to users based on their hobbies, demand information, and consumer behavior [1, 2]. At present, the recommender system had been widely used in different industries, such as Amazon product recommendation, iTunes music recommendation, and Netflix movie recommendation because the recommendation algorithm can filter according to the mass of user history information, mine the deep relationship between users and users or items, and produce more accurate personalized recommendation with preference characteristics, which can better meet the needs of users. The algorithms used by the recommendation system consist of three types: collaborative filtering recommendation algorithms [3, 4], content-based recommendation algorithms [5], and hybrid recommendation algorithms [6]. Among them, collaborative filtering recommendation algorithm is currently the most popular, and it consists of three types: item-based collaborative filtering [7], user-based collaborative filtering [8], and matrix factorization collaborative filtering [9]. The matrix factorization collaborative filtering recommendation algorithm has attracted more and more researchers’ attention because of its outstanding performance in the Netflix Prize competition. However, in practical applications, only a small number of users will rate or comment on a small number of items. Therefore, the matrix factorization recommendation algorithm has obvious data sparsity problem and item cold-start problem. The algorithm expresses the user’s scorings of items in matrix form and factorizes the matrix to mine low-dimensional hidden feature space and then get user feature matrix and item feature matrix, finally through inner product operation of the two low-dimensional feature matrixes to describe the relation between users and items. Although the existing recommendation algorithms had obtained a good recommendation result, these traditional recommendation algorithms ignored the influence of social relationships among users on recommendation results. Since social relationships can reflect the similarity between users’ preferences, simply considering the user’s scoring of the items can no longer meet the recommendation needs, so the social recommendation algorithm that introduces social relationship into recommendation algorithm became
current research hotspots in recommendation systems [10]. Hao et al. [11] proposed a weighted bipartite graph recommendation algorithm that used the monotonic saturation function as the weight and used the true value of the number of target users and other items to score relative to the total number of users as the traditional similarity coefficient. Chen et al. [12] calculated the similarity between users and items through similar tags and proposed a joint probability matrix factorization recommendation algorithm that merged the neighbor perception of social tags, effectively using the semantics of the tag improved the recommendation quality. Lin et al. [13] used the principle of trust generation in social psychology, based on the trust extension method of user credibility, to alleviate the sparseness of data, and proposed a matrix recommendation algorithm with enhanced trust. Korpinar et al. [14] based on the shared representation method of user feature matrix, proposed a novel social recommendation model based on shared representation of user feature matrix, which effectively improved the accuracy of recommendation.

Most existing social recommendation algorithms only introduce a social relationship, but each added social relationship affects the recommendation result differently, so introduction of multiple social relationships would definitely improve recommendation accuracy. In this paper, a multiple relationships social network is constructed through a multi-subnet composite complex network model [15], a shared user feature matrix is used to introduce multiple social relationships into the recommendation algorithm, and a matrix factorization recommendation algorithm that integrates multiple social relationships is proposed.

2. Methodology

2.1. Traditional Matrix Factorization Model. Assuming that the recommendation system includes m customers and n commodities, \( R_{\text{mon}} = [R_{ij}]_{\text{mon}} \) represents the customer-commodity scoring matrix, which is shown as Figure 1. \( R_{ij} \) represents the rating of customer i to item j, where \( R_{ij} \in [1, 5] \). Usually, there are many empty elements in \( R_{\text{mon}} \), and it will cause a sparse matrix of customer-commodity scoring.

In social networks, as shown in Figure 2, the social relationship between customers can be represented by a matrix \( C = [C_{ik}]_{\text{mon}} \), the value of \( C_{ik} \) is 0 or 1, and if \( C_{ik} = 0 \), it means that there is no social relationship between customers.

The flowchart of traditional matrix factorization algorithm [8] is shown in Figure 3.

Matrix \( R_{\text{mon}} \) can be factorized into user feature matrix \( U_{\text{mon}} \) and item feature \( V_{\text{mon}} \), respectively. \( k \) represents dimension of vector, in general, it is much smaller than \( m \) and \( n \), and then dimensionality reduction can be realized. \( U_i \) and \( V_j \) represent the potential feature spaces of corresponding users \( u_i \) and items \( v_j \), respectively. The null value in scoring matrix can be forecast through \( U_i^T V_j \), and then the prediction scoring matrix could be obtained.

For the convenience of research, the function \( f(a) = 1/R_{\text{max}} \) is used to map the customer’s scoring of the commodity to \([0,1]\) interval, where \( R_{\text{max}} \) represents the customer’s maximum scoring of commodity. The traditional matrix factorization only uses a simple linear model \( R = U^T V \), and the obtained results will be too fitted to the scoring matrix, resulting in the prediction scores deviating too much from the real data, and the final prediction results are distorted. Therefore, the logistic function \( g(a) = 1/(1 + e^{-a}) \) is used in this paper, so that the customer’s scoring of commodity is defined in the range of \([0, 1]\). The observed conditional probability distribution can be defined as

\[
p(R|U, V, \sigma^2) = \prod_{i=1}^{m} \prod_{j=1}^{n} \mathcal{N}\left([r_{ij}], g(U_i^T V_j), \sigma_i^2\right)^{I_{ij}^R},
\]

where \( \mathcal{N}(x|\mu, \sigma^2) \) indicates that \( x \) follows a Gaussian distribution whose mean is \( \mu \) and variance is \( \sigma^2 \). \( I_{ij}^R \) represents the null value in the scoring matrix.

![Figure 1: Customer-commodity scoring matrix.](image1)

![Figure 2: Customer social relationship matrix.](image2)

![Figure 3: Flowchart of probabilistic matrix factorization.](image3)
an index function, and if user \( i \) has a score of item \( j \), its value is 1; otherwise, it is 0. Suppose \( U \) and \( V \) obey the spherical Gaussian distribution prior to mean 0:

\[
p(U|\sigma^2_U) = \prod_{j=1}^m \mathcal{N}(U_j|0, \sigma^2_U I),
\]

\[
p(V|\sigma^2_V) = \prod_{j=1}^n \mathcal{N}(V_j|0, \sigma^2_V I).
\]

(2)

Then, through Bayesian inference, the posterior probability distribution of \( U \) and \( V \) could be obtained as follows:

\[
p(U, V|R, \sigma^2_U, \sigma^2_R, \sigma^2_V, \sigma^2_C) \propto p(R|U, V, \sigma^2_R) p(U|\sigma^2_U) p(V|\sigma^2_V)
\]

\[
= \sigma \prod_{i=1}^m \prod_{j=1}^n \mathcal{N}(r_{ij}|g(U_j^T V_j), \sigma^2_R) | I^m_u \times \prod_{i=1}^m \mathcal{N}(U_i|0, \sigma^2_U I) \times \prod_{j=1}^n \mathcal{N}(V_j|0, \sigma^2_V I).
\]

(3)

The conditional distribution of observed social relationship can be defined:

\[
p(C|U, Z, \sigma^2_C) = \prod_{i=1}^m \prod_{k=1}^n \mathcal{N}(c_{ik}|g(U_i^T Z_k), \sigma^2_C) | I^m_u.
\]

(4)

Suppose that \( U \) and \( Z \) follow the spherical Gaussian prior distribution with mean 0:

\[
p(U|\sigma^2_U) = \prod_{j=1}^m \mathcal{N}(U_j|0, \sigma^2_U I),
\]

\[
p(Z|\sigma^2_Z) = \prod_{j=1}^n \mathcal{N}(Z_j|0, \sigma^2_Z I).
\]

(5)

Then, through simple Bayesian inference, the following results can be obtained:

\[
p(U, Z|C, \sigma^2_C, \sigma^2_U, \sigma^2_Z) \propto p(C|U, Z, \sigma^2_C) p(U|\sigma^2_U) p(Z|\sigma^2_Z)
\]

\[
= \sigma \prod_{i=1}^m \prod_{k=1}^n \mathcal{N}(c_{ik}|g(U_i^T Z_k), \sigma^2_C) | I^m_u \times \prod_{i=1}^m \mathcal{N}(U_i|0, \sigma^2_U I) \times \prod_{k=1}^n \mathcal{N}(Z_k|0, \sigma^2_Z I).
\]

(6)

where \( I^m_u \) is an indicator function, and if user \( i \) has a social relationship with user \( k \), its value is 1; otherwise, it is 0.

The flowchart of matrix factorization recommendation algorithm integrating a social relationship (MFRS1) is shown in Figure 4.

According to the shared user eigenspace, the user’s item scoring matrix is closely related to the social relationship matrix, and the posterior distribution of social recommendations can be obtained by logarithm:
and the maximum posterior distribution function should be equal to the minimum objective function, which is as follows:

\[
\mathcal{L}(R, C_1, C_2, U, V, Z_1, Z_2) = \frac{\lambda_C}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}^R (r_{ij} - g(U_i^T V_j))^2 + \frac{1 - \lambda_C}{2} \sum_{i=1}^{m} \sum_{k=1}^{n} I_{ik}^C (c_{ik}^* - g(U_i^T Z_k))^2
\]

\[
+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2 + \frac{\lambda_Z}{2} \|Z\|_F^2
\]

where \(\lambda_C \in [0, 1]\) is used to adjust the influence proportion of user scoring matrix and social relationship matrix on recommendation result. When \(\lambda_C = 1\), it means that the social relationship between users is not considered, when \(\lambda_C = 0\), it means that the user scoring matrix has a proportion of 0, and the rest means that a social relationship is integrated. \(\lambda_C = \sigma_R^2/\sigma_C^2, \lambda_U = \sigma_R^2/\sigma_U^2, \lambda_V = \sigma_R^2/\sigma_V^2, \lambda_Z = \sigma_R^2/\sigma_Z^2\), \(\|=\|_F\) represents regularization.

2.3. Matrix Factorization Recommendation Algorithm Integrating Multiple Social Relationships. In actual social networks, there are often more than one kind of social relationship between users, and each social relationship has a different impact on the recommendation, so introduction of a kind of social relationship would definitely affect recommendation accuracy. Suppose that there are two kinds of social relationships \(c_1\) and \(c_2\) between users, and the flowchart of matrix factorization recommendation algorithm integrating multiple social relationships (MFRS2) is shown in Figure 5.

![Figure 4: Flowchart of matrix factorization recommendation algorithm integrating a social relationship.](image)

The objective function for minimizing the introduction of two kinds of social relationships is
The gradient descent algorithm can be used to solve the objective function as follows:

\[
\frac{\partial \mathcal{L}}{\partial U_i} = \lambda_C \sum_{j=1}^{m} I_{ij} g'(U_i^T V_j)(g(U_i^T V_j) - r_{ij}) V_j + \beta (1 - \lambda_C) \sum_{j=1}^{m} I_{ik} g'(U_i^T Z_k^1)(g(U_i^T Z_k^1) - C_{ik}) Z_k^1 + (1 - \beta)(1 - \lambda_C) \sum_{j=1}^{m} I_{ij} g'(U_i^T Z_k^1)(g(U_i^T Z_k^1) - C_{ik}) Z_k^1 + \lambda_U U_i,
\]

\[
\frac{\partial \mathcal{L}}{\partial V_j} = \sum_{i=1}^{m} I_{ij} g'(U_i^T V_j)(g(U_i^T V_j) - r_{ij}) U_i + \lambda_V V_j,
\]

\[
\frac{\partial \mathcal{L}}{\partial Z_k^1} = \lambda_C \sum_{i=1}^{m} I_{ik} g'(U_i^T Z_k^1)(g(U_i^T Z_k^1) - C_{ik}) U_i + \lambda_Z Z_k^1,
\]

\[
\frac{\partial \mathcal{L}}{\partial Z_k^2} = \lambda_C \sum_{i=1}^{m} I_{ik} g'(U_i^T Z_k^2)(g(U_i^T Z_k^2) - C_{ik}) U_i + \lambda_Z Z_k^2,
\]
where \( g'(x) = e^x/(1 + e^x)^2 \) represents derivative of logistic function \( g(x) \). For reducing model complexity, the corresponding parameter setting is \( \lambda_U = \lambda_V = \lambda_Z^1 = \lambda_Z^2 \).

### 3. Results and Analysis

#### 3.1. Experimental Data.

In this paper, Epinions is used as the experimental dataset, and it is a knowledge-sharing website where customers could review commodities or give integer scorings between one and five. New customers can use these comments or ratings to determine whether the commodity is worth buying. The Epinions dataset includes the customer’s trust relationship, customer’s scoring for commodities, and comment information of commodities. It includes 49,290 customers, 139,738 commodities, 664,824 comment messages, and 487,181 trust relationships.

In the course of experiment, a five-fold cross-validation method was used to train and verify the proposed recommended algorithm. The Epinions dataset was separated into five equal groups, one group was randomly selected as test set, and the other four groups were acted as training set. Five experiments were conducted to ensure that each test set is tested. The ultimate experimental result was the average of five experiments.

#### 3.2. Evaluation Indicators.

In this paper, two identical evaluation indicators are used to measure the accuracy of recommendations, namely, mean absolute error (MAE) and root mean squared error (RMSE) [16]. These two indicators measure accuracy of recommendation algorithm through calculating the error between the prediction score and the actual score. The smaller their value, the higher the recommendation accuracy. The definitions of MAE and RMSE are as follows:

\[
\text{MAE} = \frac{1}{|E^p|} \sum_{(i,j) \in E^p} |r_{ij} - r'_{ij}| \tag{11}
\]

\[
\text{RMSE} = \sqrt{\frac{1}{|E^p|} \sum_{(i,j) \in E^p} (r_{ij} - r'_{ij})^2},
\]

where \( r_{ij} \) represents the real scoring of customer \( i \) for commodity \( j \), \( r'_{ij} \) represents the prediction scoring of customer \( u_i \) for commodity \( j \), and \( E^p \) represents the test set.

#### 3.3. Experimental Results Analysis.

During the experiment, the number of algorithm user features is \( K = 5 \), the number of iterations is 1000, and \( \lambda_U = \lambda_V = 0.001 \). The proportion between social relationship matrix and customer scoring matrix can be adjusted by parameter \( \alpha \), and the proportion between the two social relationships can be adjusted by parameter \( \beta \). The different values of \( \alpha \) and \( \beta \) will directly affect the recommendation result. The value of \( \alpha \) and \( \beta \) is determined by the method of simulation experiment. \( \beta = 1 \) means that only one kind of social relationship is introduced, when \( \alpha \) takes different values, and the change of MAE value and RMSE value in the dataset is shown in Figures 6 and 7.

As shown in Figures 6 and 7, in the Epinions dataset, when \( \alpha = 0.8 \), the values of MAE and RMSE are all the smallest; that is, for one social relationship, when \( \alpha = 0.8 \), the recommendation accuracy rate is the highest.

\( O_u \) and \( O_v \) are used to denote the item sets that customers \( u \) and \( v \) had scored, respectively, the more commodities which customers \( u \) and \( v \) jointly scored, the more likely they are to have the same interest and influence each other, and the specific definitions are as follows:

\[
f_{uv} = \frac{O_u \cap O_v}{O_u \cup O_v}. \tag{12}
\]

When \( f_{uv} > 0.2 \), it means that customers \( u \) and \( v \) have similar interests. Suppose that the relationship between users satisfying this condition is the \( e_2 \) relationship. When \( \alpha \) and \( \beta \) take different values, the changes in MAE and RMSE on the Epinions dataset are shown in Figures 8 and 9.

As shown in Figures 8 and 9, in the Epinions dataset, when the parameters \( \alpha = 0.3 \) and \( \beta = 0.4 \), the value of MAE
is the smallest, that is, the recommended accuracy rate in the proposed algorithm is the highest. Similarly, when $\alpha = 0.7$ and $\beta = 0.5$, the value of RMSE is the smallest, which means that the recommendation accuracy in the proposed algorithm is the highest.

For verifying the performance of MDRS2 algorithm and the impact of various social relationships on the recommendation, we compared MDRS2 algorithm with SocRec algorithm [14], TDSRec algorithm [17], and MDRS1 algorithm on the Epinions dataset. The SocRec algorithm considers the attributes of social relationships between users on the basis of matrix factorization and incorporates a social relationship; the TDSRec algorithm combines the similarity based on user ratings preferences while considering social networks and predicts the values of customer scoring matrix together. The MDRS1 algorithm only considers one kind of social relationship. The MDRS2 algorithm tightly links the user-item scoring matrix and the social relationship matrix through sharing the user’s feature space and integrates multiple social relationships into matrix factorization. The experimental statistical result is shown in Table 1.

As shown in Table 1, in the Epinions dataset, the MAE and RMSE values of the MDRS2 algorithm are smaller than those of other algorithms; that is, the prediction accuracy is higher. It can be seen that the recommendation algorithm that introduces two kinds of social relationships has a higher accuracy rate than other three recommendation algorithms, which indicates that introducing multiple relationships between users would improve recommendation accuracy, and the more relationships between users, the higher the recommendation accuracy.

4. Conclusions

In this paper, through factorizing the user-item scoring matrix, according to multi-subnet composite complex network, a variety of social relationships are integrated into the matrix factorization recommendation algorithm by using the shared user’s potential feature space. Using the matrix factorization recommendation algorithm to the user-item dataset after introducing the social relationship, users’ preferences for items will be accurately obtained. Through experiments on real datasets, it is proved that the proposed matrix factorization recommendation algorithm, which combines multiple relationships, improves the recommendation accuracy. It means that introducing multiple social relationships can better personalize recommendations for users, and the more relationships introduced, the better the recommendation effect. In future research, the user’s indirect relationship and direct relationship can be combined to further study the impact of social relationships on recommendations.

Data Availability

The basic data used in this article are downloaded from the online public dataset: Epinions http://www.trustlet.org/epinions.html.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the Shandong Provincial Natural Science Foundation, China, grant no. ZR2017MG011, and Humanity and Social Science Youth Foundation of Ministry of Education of China, grant no. 15YJC860001.
References


