Research Article

A Recommendation System Based on Regression Model of Three-Tier Network Architecture

Wang Bailing, Huang Junheng, Zhu Dongjie, and Hou Xilu

Department of Computer Science and Technology, Harbin Institute of Technology, Weihai 264209, China

Correspondence should be addressed to Zhu Dongjie; zhudongjie@hit.edu.cn

Received 17 November 2015; Revised 21 February 2016; Accepted 9 March 2016

Copyright © 2016 Wang Bailing et al. 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.

The sparsity problem of user-item matrix is a major obstacle to improve the accuracy of the traditional collaborative filtering systems, and, meanwhile, it is also responsible for cold-start problem in the collaborative filtering approaches. In this paper, a three-tier network Architecture, which includes user relationship network, item similarity network, and user-item relationship network, is constructed using comprehensive data among the user-item matrix and the social networks. Based on this framework, a Regression Model Recommendation Approach (RMRA) is established to calculate the correlation score between the test user and test item. The correlation score is used to predict the test user preference for the test item. The RMRA mines the potential information among both social networks and user-item matrix to improve the recommendation accuracy and ease the cold-start problem. We conduct experiment based on KDD 2012 real data set. The result indicates that our algorithm performs superiorly compared to traditional collaborative filtering algorithm.

1. Introduction

A variety of recommendation systems can help people to find useful information from big data sets. In these recommendation systems, the collaborative filtering algorithm (hereinafter referred to as CFA) is the most popular recommendation algorithm for its easy implementation and good expandability [1]. All of these algorithms are facing the following problems: (1) the data sparsity problem, in which the user-item matrix is highly sparse in most cases, leading to the inaccuracy of the user similarity calculated through this matrix; (2) the cold-start problem, in which a new user has not specified enough of his or her product preferences for the system to make effective predictions. Calculating the neighbors of a new user will fail because of the lack of evaluation records. This leads to the result that a new user cannot get effective recommendation. For solving these issues in CFA, Wang et al. have put forward a collaborative filtering algorithm based on both user and item, which brought about improvement to forecast difficulties and inaccuracy stem from data sparseness problem [2].

In early stage, these recommended systems were based on the assumption that users are independent and identically distributed. However, people usually ask for some friends’ advice about some things, and friends’ suggestions often play an important role in the final decision of the individual. With the rapid development of social network services (hereinafter referred to as SNS), such as Facebook (http://www.facebook.com/), Twitter (http://www.twitter.com/), and Tencent (http://www.qq.com/) a good social platform for information exchanging among people is provided. User relationship, user attributes, and item attributes in SNS can provide more available information for recommendation systems. Recently, improving the performance of the recommendation system using social network information has aroused the interest of many scholars [3–6].

The main contributions of this paper include (1) improving the accuracy of user similarity using user friendship and user natures in the SNS; (2) improving the accuracy of item similarity using item natures in the SNS; (3) proposing and constructing a three-tier network Architecture, which includes user similarity network, item similarity network,
and user relationship network; (4) establishing a Regression Model Recommendation Approach (hereinafter referred to as RMRA) to calculate the correlation score between the test user and test item. The correlation score is used to predict the test user preference for the test item. The RMRA fully mines the potential information among SNS and user-item matrix, which can improve the recommendation accuracy and ease the cold-start problem. Experiment based on KDD 2012 real data set is conducted. The RMRA performs superiorly compared to traditional CFA.

2. Related Work

2.1. Prerequisites. In a recommendation system, user set and item set are two basic elements. Let the user set be \( S_U = \{u_1, u_2, \ldots, u_m\} \), where \( m \) is the number of users, and let the item set be \( S_I = \{i_1, i_2, \ldots, i_n\} \), where \( n \) is the number of items.

There are different kinds of relationship between users, such as user topology neighbor relations in the SNS, users similarity relationship based on user tags (user tags are some words that describe the user properties) in SNS, and user similarity relationship based on user-item matrix (user-based CFA is based on this relationship to predict). These relationships among users can constitute an integrated user relationship network (denoted by \( G_U \)). Meanwhile, there are a variety of relationships between items, including item similarity relationship based on user-item matrix (item-based CFA is based on this relationship to predict) and item similarity relationship based on item tags (item tags are some words that describe the item properties) in SNS. These relationships among items can compose an integrated item relationship network (denoted by \( G_I \)). The rating where user rates item in user-item matrix can form a directed bipartite graph (denoted by \( G_{U-I} \)). These three networks are defined as follows.

**Definition 1.** Define user relationship network as \( G_U = (S_U, E(\omega_U)) \), where \( G_U \) is a weighted directed network, \( S_U \) is a set of nodes and one node represents one user, and \( E(\omega_U) \) is a weighted directed edge’s set expressed by \( E(\omega_U) = \{(a, b, \omega_U(a, b)) \mid \forall a, b \in S_U \land a \neq b \land \omega_U(a, b) > 0\} \), where \( \omega_U(a, b) \) is a nonnegative real number and it indicates the strength of association between the node \( a \) and the node \( b \) and \( \omega_U(a, b) = 0 \) indicates no edge between \( a \) and \( b \).

**Definition 2.** Define item similarity network as \( G_I = (S_I, E(\omega_I)) \), where \( G_I \) is a weighted undirected network, \( S_I \) is a set of nodes and one node represents one item, and \( E(\omega_I) \) is a weighted undirected edge’s set expressed as \( E(\omega_I) = \{(a, b, \omega_I(a, b)) \mid \forall a, b \in S_I \land a \neq b \land \omega_I(a, b) > 0\} \), where \( \omega_I(a, b) \) is a nonnegative real number and it indicates the degree of similarity between the node \( a \) and the node \( b \) and \( \omega_I(a, b) = 0 \) represents no edge between \( a \) and \( b \).

**Definition 3.** Define user-item relationship network as \( G_{U-I} = (S_U, S_I, E(\omega_{U-I})) \), where \( G_{U-I} \) is a directed network, \( S_U \) is the same as in Definition 1, and \( S_I \) is the same as in Definition 2. \( E(\omega_{U-I}) \) is a directed edge’s set and \( E(\omega_{U-I}) = \{(a, b, \omega_{U-I}(a, b)) \mid \forall a \in S_U \land b \in S_I \land \omega_{U-I}(a, b) > 0\} \), where \( \omega_{U-I}(a, b) \) is a nonnegative real number and it indicates the rating where user \( a \) rates item \( b \).

A three-tier network Architecture is constructed by integrating networks \( G_U, G_I, \) and \( G_{U-I} \), as shown in Figure 1.

2.2. Construction of an Integrated User Relationship Network. Many commercial websites, including Amazon, Taobao, and JD, provide users with personalized recommendation. The so-called personalized recommendation refers to the fact that commercial sites offer wish lists for people to buy commodities. Most of these sites offer recommendations for clients using CFA [7], since CFA has better scalability and easy implementation. CFA recommendation system uses two strategies, one is based on user [8, 9], and the other one is based on item [10, 11]. The former is mainly based on the user high similarity top \( N \) neighbors by user-item matrix calculation. If the nodes represent users and weighted edges represent similar relationship between the users, then the CFA constitute a user relationship network defined the same as in Definition 1. The user similarity calculated by using the CFA is sometimes inaccurate, since CFA calculates the similarity between users who only use user-item matrix, and the matrix is very sparse. To improve the accuracy of the similarity between users, this paper proposes a comprehensive relationship network with user-item matrix, trust relationship between users, and similarity between users’ tags in SNS.

2.2.1. Calculate the Level of Trust between Users in SNS. Traditional CFA is based on the assumption that users are independent and identically distributed, which ignores the trust relationship between users and does not comply with the phenomenon that people often ask the opinions of friends in real life. In order to improve the accuracy of the recommendation, some recommendation system takes
into account the information in the SNS, known as social recommendation systems [12–14]. Users may trust other users in SNS and the level of trust between users is a good predictor of user preferences. Trust relationship in SNS (e.g., the follower and followee in Twitter or Tencent) forms a directed graph as in Definition 1, where weighted directed edge \((a, b)\) represents the fact that user \(a\) follows user \(b\). The term “user \(a\) follows user \(b\)” means that \(a\) trusts \(b\) to some extent or the interest of user \(a\) and user \(b\) is similar. Meanwhile, to a certain extent, user \(b\)’s preferences can affect user \(a\)’s decision. We use the Gaussian kernel to calculate the degree to which the user \(a\) trusts user \(b\):

\[
T(a, b) = e^{1 - \frac{a}{a}} , \tag{1}
\]

where \(T(a, b)\) indicates the extent to which \(a\) trusts \(b\) and \(L_{ab}\) represents the topological distance between \(a\) and \(b\) based on “follow” relations in SNS. The Gaussian kernel transforms the topological distance between user \(a\) and user \(b\) into the trust level of user \(a\) to user \(b\), and it is obvious that \(T(a, b) \in (0, 1]\).

### 2.2.2. Calculate the User Similarity Using User-Item Matrix.

In the user-based CFA, first the similarity between test user and other users is calculated; then, according to the test user similarity neighbor’s preference for test item, the test user preference for test item is predicted. The Pearson correlation coefficient between users is a frequently used method to calculate user similarity [15, 16], as shown in

\[
M(a, b) = \frac{\sum_{j \in S_j} (r_{aj} - \bar{r}_a)(r_{bj} - \bar{r}_b)}{\sqrt{\sum_{j \in S_j} (r_{aj} - \bar{r}_a)^2 \sum_{j \in S_j} (r_{bj} - \bar{r}_b)^2}} , \tag{2}
\]

where \(\bar{r}_a = (1/|S_j|) \sum_{j \in S_j} r_{aj}\) is the average rating value of the user, \(r_{aj}\) is the rating of user \(a\)’s evaluation to the item \(j\), and \(S_j\) is the item set.

### 2.2.3. Calculate the User Similarity Using User Tags in SNS

In SNS, not only are there trust relationships between users, but also some user behavior characteristics are exhibited; for example, users use keywords to represent their self-introduction which reveal their occupation, interest, and viewpoint. A series of those keywords is called user tags.

User tags are the user self-description, which is used to express their standpoint freely. Generally speaking, compared with other information (such as the information via data mining), it is more accurate to obtain user information and express user demand by using the user tags. Therefore, the similarity of user tags represents the user similarity to some degree. Generally, the user tags take the following form: \(U_{ti} = \{\text{keyword}_1, \text{keyword}_2, \ldots, \text{keyword}_n\}\). In this paper, the Jaccard Coefficient is used to calculate the similarity of user tags. Denote \(U_{ta}\) and \(U_{tb}\) as the tag set of users \(a\) and \(b\), respectively; then the Jaccard Coefficient of them is

\[
J(a, b) = \frac{|U_{ta} \cap U_{tb}|}{|U_{ta} \cup U_{tb}|} , \tag{3}
\]

### 2.2.4. Building of an Integrated User Relationship Network.

Due to the sparsity of a single data source, the user similarity calculated with a particular data source is sometimes insufficient and inaccurate. Formula (4) is the combination of formulae (1), (2), and (3) and is used to calculate the weights of the edges in Definition 1:

\[
w_{ij}(a, b) = \frac{\alpha T(a, b) + \beta M(a, b) + \gamma J(a, b)}{\text{Max}_{v \in S_u \land v \neq w} \left(\alpha T(v, w) + \beta M(v, w) + \gamma J(v, w)\right)} , \tag{4}
\]

where \(\alpha\), \(\beta\), and \(\gamma\) represent the proportion of formulae (1), (2), and (3) in (4), respectively. We use formula (4) to establish the users’ similar network \(G_u\).

### 2.3. Construction of an Integrated Item Relationship Network.

In item-based CFA, the similarity between test item and other items is calculated. If each node represents an item and a weighted undirected edge represents similar relationship between a pair of the items, then item-based CFA constitutes an item relationship network as in Definition 2. Item similarity calculated by CFA is inaccurate, since the CFA calculate the similarity between items only using user-item matrix, and the matrix is very sparse, which is the main reason affecting the accuracy of the recommended method. To improve the accuracy of the similarity between items, a comprehensive relationship network has been constructed with user-item matrix and similarity between item tags.

### 2.3.1. Calculate the Item Similarity Using User-Item Matrix.

Item similarity is the basis of the CFA and is calculated with Pearson’s correlation coefficient in this paper:

\[
M(i, j) = \frac{\sum_{a \in S_a} (r_{ai} - \bar{r}_i)(r_{aj} - \bar{r}_j)}{\sqrt{\sum_{a \in S_a} (r_{ai} - \bar{r}_i)^2 \sum_{a \in S_a} (r_{aj} - \bar{r}_j)^2}} , \tag{5}
\]

where \(M(i, j)\) represents the similarity between items \(i\) and \(j\), \(\bar{r}_i = (1/|S_a|) \sum_{a \in S_a} r_{ai}\) is average value that item \(i\) is rated by all users, \(r_{ai}\) is rating that user \(a\) rated item \(i\), and \(S_a\) is the set of users.

### 2.3.2. Calculate the Item Similarity Using Item Tags in SNS.

In SNS, there are item tags that use a set of keywords to describe the item. The set of keywords is generally written by industry experts and is more accurate than other items of information (such as item information obtained by data mining). Thus, the similarity between item tags is a good complement to the item similarity based on user-item matrix. Item tag set takes the following general form: \(I_{ti} = \{\text{keyword}_1, \text{keyword}_2, \ldots, \text{keyword}_n\}\). In this paper, the Jaccard Coefficient is used to calculate the similarity of item tag. Denote \(I_{ta}\) and \(I_{tb}\) as the tag set of items \(i\) and \(j\), respectively; then the Jaccard Coefficient of them is

\[
J(i, j) = \frac{|I_{ti} \cap I_{tj}|}{|I_{ti} \cup I_{tj}|} , \tag{6}
\]
2.3.3. Construction of an Integrated Item Similarity Network. Formula (7) is the combination of formulae (5) and (6) and is used to calculate the weights of the edges in Definition 2:

$$u_i (i, j) = \frac{\theta M (i, j) + \mu I (i, j)}{\max_{v, w \in S(i) \land v \neq w} (\theta M (v, w) + \mu I (v, w))},$$  

(7)

where $\theta$ and $\mu$ represent the proportion of formulae (5) and (6) in (7), respectively. Formula (7) is used to establish the item’s similar network $G_1$.

2.4. Construction of User-Item Relationship Network $G_{U-I}$. By Definition 3, $G_{U-I} = (S_U, S_I, E(u_{i-j}))$ is a directed bipartite graph, wherein $S_U$ is user node’s set, $S_I$ is item node’s set, and edge’s set $E(u_{i-j})$ is composed of the ratings that users rate items. In this paper, the value of $u_{i-j}(a, b)$ is as follows:

$$u_{i-j} (a, b) = \begin{cases} 
1 & \text{if } a \text{ follow } b \\
0 & \text{otherwise.} 
\end{cases}$$  

(8)

A three-tier network Architecture is constructed by integrating networks $G_U$, $G_I$, and $G_{U-I}$, as shown in Figure 1.

3. The RMRA Based on Three-Tier Network Architecture

The idea of user-based CFA recommendation system is to determine whether a test user $a$ has enough preferences for test item $b$ (whether or not to recommend the item $b$ to the user $a$). Firstly, based on the user-item matrix, the user $a$’s top $N$ most similar neighbors are obtained and let these neighbors constitute a set $Sr$. Next, according to the preferences of users in $Sr$, the test user $a$’s preferences for the test item $b$ are predicted. In the CFA, the similarity between any two users is calculated by the vectors of these two users rate all items, and this algorithm has achieved good results. In view of this, we have reason to believe that any two users’ similarity can be calculated by the similarity of the item that the two users rated. This can be established by user similarity regression model as follows:

$$S_{uai}' = C_u + \sum_{i \in R(u)} \sum_{i' \in R(u')} \eta_{ui} S_{ui'},$$  

(9)

where $S_{uai}'$ is the similarity score between the test user $u$ and other users $u'$, $S_{ui'}$ is the similarity score between item $i$ and item $i'$, $R(u)$ denotes the set of items rated by user $u$, $C_u$ is a constant, and $\eta_{ui}$ is the coefficient of regression model and represents the item $i$’s contribution to users $u$ and $u'$. In general, this model assumes that the similarity of the two users can be interpreted as linear combination of the similarity of items that these two users have rated expressed in formula (9). This model’s idea is consistent with the CFA.

In formula (9), the value $\sum_{i' \in R(u')} S_{ui'}$ is the sum of the similarity scores between item $i$ and the other items rated by user $u'$. Denote

$$\Psi_{ui'} = \sum_{i' \in R(u')} S_{ui'},$$  

(10)

According to formula (10), formula (9) can be rewritten as

$$S_{uai}' = C_u + \sum_{i \in R(u)} \eta_{ui} \Psi_{ui'},$$  

(11)

Formula (11) illustrates the similarity between test user $u$ and other users $u'$ that can be calculated by the similarity between $u'$ and the items that $u$ has rated.

Let $T_u$ be a vector defined as in (12) that is constituted by the similarity between the test user $u$ and other users:

$$T_u = (S_{uai}_1, S_{uai}_2, ..., S_{uai}_m).$$  

(12)

Let $\Psi_i$ be a vector formed by the tightness between the item $i$ and all users, which is

$$\Psi_i = (\Psi_{i1}, \Psi_{i2}, ..., \Psi_{im}).$$  

(13)

According to formulae (12) and (13), (9) can be extended to the following forms:

$$T_u = C_u + \sum_{i \in R(u)} \eta_{ui} \Psi_i.$$  

(14)

As seen from formula (14), there is a relationship between $T_u$ and $\Psi_i$. Figure 2 shows the relationship between the user $u$ and the item $i$, and Figure 3 shows the correlation between $T_u$ and $\Psi_i$. The correlation between vectors $T_u$ and $\Psi_i$ indicates the test user $u$’s preference for the test item $i$. In this paper, the degree of correlation between vectors $T_u$ and $\Psi_i$ is calculated by the Pearson correlation coefficient and is known as correlation score:

$$CC_{uai} = \frac{\text{cov}(T_u, \Psi_i)}{\sigma(T_u) \sigma(\Psi_i)},$$  

(15)
where \( \text{cov} \) and \( \sigma \) are the covariance and standard deviation, respectively. The correlation score \( CC_{ui} \) can be a good measure of correlation between the test user \( u \) and test item \( i \) and indicates the test user \( u \)'s preference for the test item \( i \). In this paper, the top \( N \) highest score items are recommended to the test user.

4. Algorithm Analysis

The time complexity of RMRA is mainly reflected in formulae (1), (2), (3), (5), and (7). Among them, the time complexity of formulae (2) and (5) is \( O(n^2) \), which is the typical time complexity of the CFA, while the time complexity of formulae (1), (3), and (7) is far less than that of formulae (2) and (5). Therefore, the complexity of RMRA is \( O(n^2) \), which is consistent with that of CFA algorithm.

5. Experimental Design and Analysis

5.1. Data Set. In order to verify the validity of RMRA, we carried out experiments on the KDD CUP 2012 Track 1 data (http://www.kddcup2012.org/). The following is a brief overview of data collection.

The data set of KDD CUP 2012 Track 1 is real sampling user data provided by Tencent (http://www.qq.com/). And the data set includes 2320895 users and 6095 items with a total of 73209277 users' ratings, wherein the data files and data format are as follows: user-item matrix (rec_log_train.txt) file is in a format of (UserID) \( \times \) (ItemID) \( \times \) (Result) \( \times \) (Unix-timestamp). User profile (user_profile.txt) file is in (UserID) \( \times \) (Year-of-birth) \( \times \) (Gender) \( \times \) (Number-of-tweet) \( \times \) (Tag-Ids) format. User relationship data set (user_sns.txt) file is in (Follower-userid) \( \times \) (Followee-userid) format. Item Information Data Set (item.txt) file is in a format of (ItemID) \( \times \) (Item-Category) \( \times \) (Item-KeyWord).

Let \( r_{ui} \) represent ratings of user \( u \) to item \( i \), \( r_{ui} = 1 \) means that user \( u \) follows item \( i \), and \( r_{ui} = 0 \) means that user \( u \) refuses to follow item \( i \). Density calculation of user-item matrix is shown as follows:

\[
\frac{73209277}{2320895 \times 6095} = 0.51\%.
\] (16)

Thus, user-item matrix is pretty sparse. In this paper, the experimental facilities include a computer with CPU 2.6 GHz and RAM 8 G, because the size of the original data set exceeds our computer's processing power in both space and time, and thus we randomly take out 10 million data sets from the original data set as the experimental data set, and, in accordance with the tenfold cross-validation method, the experimental data set is randomly divided into 10 parts, with each part being used as a test set and the remaining 9 parts as the training set.

5.2. Evaluation Criteria. In this experiment, the RMRA's objective is to recommend an item list that the test user is interested in. In order to verify the accuracy of the recommended results, average accuracy (Mean Average Precision at \( N \), which is MAP@\( N \)) is used to evaluate the accuracy of RMRA [17].

Given a test user \( u \) and a recommendatory list \( L \) of items sorted by the correlation score, AP@\( N \) is shown as follows:

\[
\text{AP@} N (u) = \frac{\sum_{k=1}^{N} p(k)}{M},
\] (17)

where \( M \) is the total number of list \( L \)'s items that is followed by the user \( u \), and \( p(k) \) is the accuracy of the \( k \)th position of the list \( L \), which is defined as follows:

\[
p(k) = \begin{cases} 
m(k) \quad \text{the } k\text{th item was followed by user } u \\
0 \quad \text{otherwise}.
\end{cases}
\] (18)

Then, MAP@\( N \) is defined as follows:

\[
\text{MAP@} N (u) = \frac{\sum_{u \in U} \text{AP@} N (u)}{|U|}.
\] (19)

5.3. Experimental Procedure

Step 1. Calculate user similarity based user-item matrix, user tag's similarity, and user relationship topological distance, and then generate user relationship network \( G_{u} \).

Step 2. Calculate item similarity based on user-item matrix and item tag's similarity; then generate item relationship network \( G_{I} \).

Step 3. Generate user-item relationship network \( G_{u-I} \) according to the user-item matrix.

Step 4. Use three networks mentioned above, extract the vector \( T_{u} = (S_{u1},S_{u2},\ldots,S_{u7}) \) and the vector \( \Psi_{i} = (\Psi_{i1},\Psi_{i2},\ldots,\Psi_{in}) \) (obtained from formulae (12) and (13), resp.), and calculate the correlation score between user \( u \) and item \( i \), and then the top \( 3 \) highest score items are recommended to the test user \( u \).
Step 5. Test predictions and calculate MAP@3.

Step 6. Carry on experiment on all 10 data sets. The average of MAP@3 in these experiments is the final MAP@3.

Step 7. Calculate the MAP@3 according to the user-based CFA and item-based CFA. Then compare them to the RMRA.

5.4. Analysis of Results. In order to achieve the best average accuracy, the model parameters were trained. First, for the parameters $\theta$ and $\mu$ in formula (7) of item similarity network, since the two coefficients satisfy $\theta + \mu = 1$, we just need to train $\theta$ and make $\theta$ value increase from 0 to 1 by step size 0.1. Then, according to this, we draw the MAP-$\theta$ curve as in Figure 4 to illustrate the influence of parameter $\theta$ on average accuracy.

In Figure 4, $\theta = 0$ means that only item tags are used, and $\theta = 1$ represents that only the user-item matrix is used in calculating item similarity in the item similarity network. Figure 4 shows that recommendation effect is ineffective when only item tag similarity or item-based similarity is considered. With $\theta$ gradual increase from 0 to 1, it is observed that, starting from $\theta = 0$ to 1, the average accuracy value increases first and then decreases. When $\theta$ equals 0.2, the average accuracy achieves its maximum value. Here $\theta$ is actually a tradeoff of item similarity and user-item matrix similarity. It is observed that item tag similarity solves the data sparseness problem of user-item matrix. When $\theta = 0.2$, tradeoff gets best result and the average accuracy is the highest.

After the parameters of the item similarity network (formula (7)) have been determined, the parameters of user similarity network (formula (4)) need to be trained. Because $\alpha + \beta + \gamma = 1$, we only need to determine two of the parameters. First, we fixed $\gamma = 0.1$ and then $\alpha$ add to 0.9 from 0 by step size of 0.1. According to this, we draw the MAP-$\alpha$ curve as in Figure 5 to illustrate the influence of parameter $\alpha$ on average accuracy.

In Figure 5, $\alpha = 0$ means that only user tags and user SNS relationship are used, and $\alpha = 0.9$ means that only user-item matrix and user SNS relationship are used in calculating user similarity in the user similarity network. It can be seen from Figure 5 that, with the increase of $\alpha$, the curve represents average accuracy value that is in a trend of increases first and then decrease. And the average accuracy achieves the highest value when $\alpha = 0.5$ and $\alpha = 0.6$. Thus, we fix $\alpha$ as 0.5 and, 0.6 respectively, then change the value of $\beta$, and finally get the curve shown in Figure 6, showing the impact of parameters on the average accuracy.

It can be seen from Figure 6 that the peaks of average accuracy in both cases were between 0.0 and 0.2. And in this interval, the curve $\alpha = 0.6$ is generally higher than the curve $\alpha = 0.5$. From this, we know that when $\alpha = 0.6$, $\beta = 0.1$, and $\gamma = 0.3$, the average accuracy achieves its maximum value.

We get these different results of MAP@3 through Step 7 and compare them to the user-based CFA, Item-based CFA, and RMRA in Table 1.
Table 1: Comparison of these models.

| Model       | The user-based CFA | The item-based CFA | RMRA    |
|-------------|---------------------|--------------------|---------|
| MAP@3       | 0.385               | 0.327              | 0.449   |

Table 1 shows that, after combining three networks, the recommendation effect using RMRA has been significantly improved. User tags similarity and user SNS relationship solve a part of data sparsity and cold-start problem.

6. Conclusion

In this paper, first a three-tier network Architecture, which includes user relationship network, item similarity network, and user-item relationship network, is constructed using comprehensive data among the user-item matrix and the social networks. Then, based on these networks, the RMRA was established to calculate the correlation score between the test user and test item. The correlation score is used to predict the test user preference for the test item. The experimental result indicates that our algorithm performs superior than the traditional CFA.

In our future work, we will focus on improving the performance of the recommendation system using a variety of information combination methods. Since the size of raw data sets processed by a recommendation system is very large, parallel algorithms in recommendation system should be used in our future research.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 61371177 and no. 61170262).

References

[1] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 5–53, 2004.

[2] J. Wang, A. P. De Vries, and M. J. T. Reinders, "Unifying user-based and item-based collaborative filtering approaches by similarity fusion," in *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 501–508, ACM, Seattle, Wash, USA, August 2006.

[3] F. Liu and H. J. Lee, “Use of social network information to enhance collaborative filtering performance,” *Expert Systems with Applications*, vol. 37, no. 7, pp. 4772–4778, 2010.

[4] M. Hao, D. Zhou, C. Liu, M. R. Lyu, and I. King, "Recommender systems with social regularization," in *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM ’11)*, pp. 287–296, ACM, Hong Kong, February 2011.

[5] H. Ma, H. Yang, M. R. Lyu et al., “Sorec: social recommendation using probabilistic matrix factorization,” in *Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM ’08)*, pp. 931–940, Napa Valley, Calif, USA, October 2008.

[6] H. Ma, I. King, and M. R. Lyu, "Learning to recommend with social trust ensemble," in *Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’09)*, pp. 203–210, ACM, Boston, Mass, USA, July 2009.

[7] G. Linden, B. Smith, and J. York, “Amazon.com recommendations: item-to-item collaborative filtering,” *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80, 2003.

[8] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, “GroupLens: an open architecture for collaborative filtering of netnews,” in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 175–186, ACM, Chapel Hill, NC, USA, 1994.

[9] R. Jin, Y. J. Chai, and L. Si, “An automatic weighting scheme for collaborative filtering,” in *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 337–344, ACM, 2004.

[10] M. Deshpande and G. Karypis, "Item-based top-N recommendation algorithms," *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 143–177, 2004.

[11] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, “Item-based collaborative filtering recommendation algorithms,” in *Proceedings of the 10th International Conference on World Wide Web*, pp. 285–295, ACM, Hong Kong, May 2001.

[12] H. Ma, “An experimental study on implicit social recommendation,” in *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’13)*, pp. 73–82, ACM, August 2013.

[13] L. Wu, E. Chen, Q. Liu, L. Xu, T. Bao, and L. Zhang, “Leveraging tagging for neighborhood-aware probabilistic matrix factorization,” in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, pp. 1854–1858, ACM, Maui, Hawaii, USA, October 2012.

[14] U. Shardanand and P. Maes, “Social information filtering: algorithms for automating word of mouth,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’95)*, pp. 210–217, ACM Press/Addison-Wesley, Denver, Colo, USA, May 1995.

[15] G. Adomavicius and A. Tuzhilin, “Personalization technologies: a process-oriented perspective,” *Communications of the ACM*, vol. 48, no. 10, pp. 83–90, 2005.

[16] T. Hofmann, “Latent semantic models for collaborative filtering,” *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 89–115, 2004.

[17] R. Kühlen and M. Rittberger, “Hypertext-information-retrieval-multimedia: Synergieeffekte elektronischer Informationssysteme,” in *Proceedings of the (HIM ’95)*, Konstanz, Germany, 1995.
Submit your manuscripts at http://www.hindawi.com