

## Towards Social Recommendation based on Probabilistic Matrix Factorization

Wei Luo<sup>1,2,a</sup>, Zhihao Peng<sup>1,2,b</sup> and Ansheng Deng<sup>1,c,\*</sup>

<sup>1</sup>*Faculty of Information Science and Technology, Dalian Maritime University,  
Dalian 116024, China*

<sup>2</sup>*Department of Computer Science, Dalian Neusoft Institute of Information,  
Dalian 116626, China*

<sup>a</sup>*luowei@neusoft.edu.cn*, <sup>b</sup>*pengzhihao@neusoft.edu.cn*, <sup>c</sup>*ashdeng@dlmu.edu.cn*

### Abstract

*As an important tool to help users filter Internet information, recommender system has played a very important role wherever in academia or in industrial area. During the past years, different recommendation approaches based on the social network have been proposed with the rapid development of online social networks. Different from the traditional ones which assume all the users are independent and identically distributed, these approaches follow the intuition that a person's implicit or explicit social network will affect his behaviors on the Web. In this paper, on the basis of the existing work, we fuse a baseline predictor model with an improved social recommendation model and propose a social recommendation algorithm based on probability matrix factorization. The experimental result shows that our method outperforms the existing approaches in accuracy.*

**Keywords:** *Social Network, Recommender Systems, Graphical Model, Probabilistic Matrix Factorization*

### 1. Introduction

Due to the popularity of the Internet and information explosion, it is becoming much more difficult to find the information needed efficiently. And this is called information overload. As an effective solution to information overload, recommendation system has been widely used in fields including e-commerce, music, video, social networks, location-based services, and advertising *etc.*

Recommendation algorithm based on collaborative filtering has been used widely in Industry. It can be divided to two categories: memory-based and model-based. The former implements by the user similarity or item similarity, but the sparsity of the user-item rating matrix can reduce the accuracy. So the data sparsity affects its application in real scenario. The latter builds a predefined model by training datasets.

Although the recommendation technology based on collaborative filtering mechanism has gradually matured and been widely used in areas such as Amazon , GroupLens, CDNow and so on, unfortunately, key issues like cold start, data sparsity, system scalability, real-time recommendation, quality of the recommended results, evaluation methods and many others are still waiting to be solved.

In recent years, recommendations based on social networks have gradually been proposed with the growth of Facebook, Twitter and other online social networks. Apart from the traditional recommender systems assuming all the users are independent and identically distributed, these methods support that a user's social network will affect this user's behaviors on the Web directly or indirectly, so we can effectively solve the new

---

\* Corresponding Author

user's cold start and improve the accuracy of forecasting if the user's social relations are combined reasonably.

In this paper, we propose a new social recommendation method based on probability matrix factorization. The method combines the social network with the score record to improve the accuracy of prediction and address the problem of user's cold start. Inspired by previous works, we put forward a new social recommendation algorithm-BISMF, which integrated the baseline forecast model with the improved social recommendation model. The experimental result shows that the BISMF algorithm performs better than the state-of-the-art algorithms and can further improve the quality of the recommendation results.

In Section 2, we provide an overview of several major algorithm for recommender systems and some related work. Section 3 carries on the definition and the problem description, Section 4 presents the three kinds of recommendation algorithms introduced in this article. The results of an experimental analysis are presented in Section 5, followed by the conclusions and future work in the last section.

## 2. Related Work

At present, the matrix factorization algorithm has been widely used in the field of recommendation system, its main idea is to factorize the product of a high-dimension matrix to low-dimension ones to realize computational complexity so that characteristics of high dimensional data can be studied in low dimensional space. Literature[1] in the early stage introduced the trust relationships between users into the recommendation system framework and proposed a collaborative filtering recommendation algorithm based on trust perception. Based on the literature[2], a new algorithm is proposed to improve the accuracy of the recommendation system. In literature[3], the probabilistic matrix factorization (PMF) model is proposed to discuss the concept of matrix factorization from the point of view of probability, the three algorithms proposed in this article are based on the matrix factorization model. Literature[4] is a kind of social recommendation algorithm based on matrix factorization proposed by Epinions data set to solve the problem of score prediction, following the intuition that a user's social network will affect this user's behaviors on the Web. It has a relatively high degree of similarity of latent feature vector between users by fusing the existing probability matrix factorization model and regularization terms so as to solve the new user's cold start problem. Literature [5] deals the label content with LDA algorithm and proposes a social recommendation algorithm based on tags after combining the social recommendation algorithm in literature[4].

Literature [6] considers that a user has similar preference with his trusting users and following the intuition, proposes a social recommendation algorithm in a way of linear combination, largely improving the accuracy of the algorithm proposed by paper [4]. On the basis of the existing work, the MF Social algorithm is proposed in Literature[7] to improve the quality of the recommendation results by fusing the integration of the trust mechanism and the social recommendation model. Literature [8] introduces social network to the regularization term of the matrix factorization model, and Literature [9] further improves the quality of the recommendation results in a certain extent by the integration of the Bayes formula into the probabilistic matrix factorization model. Literature [12-13] choose relatively new geography-based social network, such as Gowalla, Brightkite and so on, as the experimental materials. They have done relative studies of the recommending-area and achieved.

Compared with the above researches, from the perspective of probability, this paper presents a fusion of the baseline forecast model mentioned in Literature [10], [11] and BISMF algorithm of the existing probability matrix factorization model.

What is more, this paper proposed an improved model of social recommendation algorithm –ISMF compared with Social MF model in Literature[7]. Last but not the least, simple fusion of these two algorithms, we proposed the baseline prediction model and an improved social model BISMF recommendation algorithm.

### 3. Problem Description

For convenient exposition, the main symbols in this article are showed in Table 1 below.

**Table 1. Symbol Definition**

Symbol	Explanation
$\mathbf{US} = \{u_1, u_2, \dots, u_N\}$	Users Set
$\mathbf{IS} = \{i_1, i_2, \dots, i_M\}$	Items Set
$\mathbf{R} = [\mathbf{R}_{u,i}]_{N \times M}$	User-item rating matrix
$K \in R$	Latent feature space dimension
$\mathbf{U} \in R^{K \times N}$	Users latent feature matrix
$\mathbf{V} \in R^{K \times M}$	Items latent feature matrix
$\mu$	global mean of scores in training set
$\mathbf{Bu} \in R^{1 \times N}$	User's preference vector
$\mathbf{Bv} \in R^{1 \times M}$	Item's preference vector
$\mathbf{T} = [\mathbf{T}_{u,v}]_{N \times N}$	User - user trusting matrix
$\mathbf{S} = [\mathbf{S}_{u,v}]_{N \times N}$	Users - Users interest similarity matrix
$\mathbf{C} = [\mathbf{C}_{u,v}]_{N \times N}$	Users - Users correlation matrix

In the recommendation system, for the given N users  $\mathbf{US} = \{u_1, u_2, \dots, u_N\}$  and given M objects  $\mathbf{IS} = \{i_1, i_2, \dots, i_M\}$ , as well as the user's score matrixes for the items  $\mathbf{R} = [\mathbf{R}_{u,i}]_{N \times M}$ , among which  $r_{u,v}$  says the item's score given by the user  $u$ ,  $r_{u,i}$  can be any real number, usually using integer form of five-under,  $\mathbf{R}_{u,i} \in [1,5]$ . In this paper, without loss of generality, we standardized scores and the scores can be mapped to the range[0,1]. In the social score network, the trust relationships between the users can be represented by the matrix  $\mathbf{T} = [\mathbf{T}_{u,v}]_{N \times N}$ , assume that the user  $u$  has  $|N_u|$  trusting users, for users  $v \in N_u$ ,  $\mathbf{T}_{u,v}$  indicates the degree of trust between the user  $u$  and user  $v$ , within the range  $\mathbf{T}_{u,v} \in [0,1]$ , and larger value indicates the more trusting. Usually,  $\mathbf{T}$  indicates matrix of non-symmetrical form.

In addition to the trust relationship among users, we should also consider the similarity of interests among users. For example, we trust our parents in real life, but we don't necessarily like the things they like. The similarity matrix among the users can be represented by matrix A  $\mathbf{S} = [\mathbf{S}_{u,v}]_{N \times N}$ . After fully considering the trust relationship among users and interest similarity, In this paper, we represents a degree of correlation among the users basing on the user - user correlation matrix  $\mathbf{C} = [\mathbf{C}_{u,v}]_{N \times N}$ , but not just the trust network among users:

$$C_{u,v} = T_{u,v} \oplus S_{u,v}, u, v \in \{1, 2, \dots, N\} \quad (1)$$

Among these, the symbol  $\oplus$  indicates that the trust network between the users and the interest similarity of interests can be expressed by a certain method, with the value  $C_{u,v} \in [0, 1]$ . For simplicity,  $\oplus$  is represented as a left term multiplied by the right term whose value range is constant.

Users may not score all the items, so the user - item rating matrix may become a sparse matrix with a large number of missing values. So the target of the recommendation system based on social network is to predict the unknown user's score by the known user history score and the network of trust network and interest network among users.

## 4. Social Recommendation Algorithm

### 4.1. BPF Model

In the practical the recommendation system, there are different application scenarios, and some of its own attributes have nothing to do with the users and the items, similarly, users also have some properties which are not related with application scenarios and items, and so do the items. Each of the three has independent attributes affecting the final scoring system, that is to say, the final score of the  $i$  th item is affected by  $u$  th user in the application scenario, so in the probability matrix factorization model, we added three factors  $\mu$ ,  $Bu_u$ ,  $Bv_i$  in the probabilistic matrix factorization model to modify the existing probabilistic model, as is shown in Figure 1.

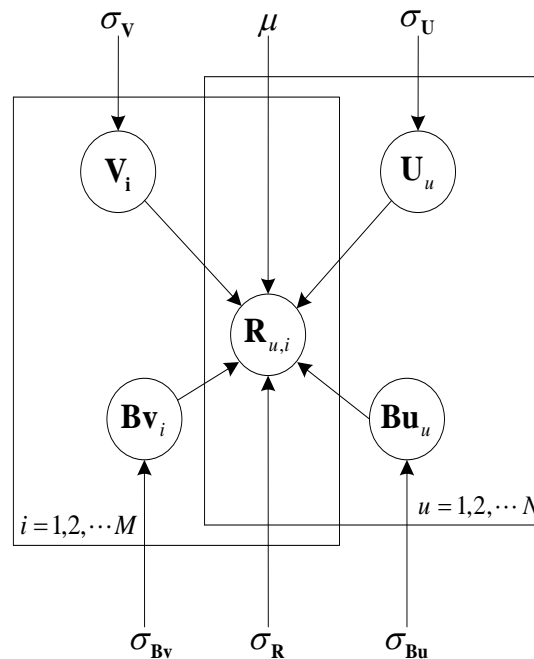


Figure 1. BPF Model

Among these,  $\mu$  indicates the global average of all the scores in the training set, which represents the independent property of the application scene, that is because in different application scenarios, such as clothing sites and books online shopping sites, the overall score distribution may vary from different sites of the positioning and sales of goods, we

use  $\mu$  to show the impact of the application scene itself on the user's score.  $\mathbf{B}\mathbf{u}^u$  means users offset term,  $u$  represents the U-th user's own scoring habits and the application of the scene as well as the object of the object, for instance, as a result of some users with more demanding, the score will be lower, and vice versa, will be high.  $\mathbf{B}\mathbf{v}_i$  indicates that the item reference term, it means that the item has already existed in the user's score. For example, the quality of the goods is relatively high, and the score will be low, and vice versa, will be high.

The BPMF model is based on the following assumptions:

Assume that A, B, C are independent and identically distributed.

$$p(\mathbf{U} | \sigma_{\mathbf{U}}^2) = \prod_{u=1}^N N(\mathbf{U}_u | 0, \sigma_{\mathbf{U}}^2 \mathbf{I}) \quad (2)$$

$$p(\mathbf{V} | \sigma_{\mathbf{V}}^2) = \prod_{i=1}^M N(\mathbf{V}_i | 0, \sigma_{\mathbf{V}}^2 \mathbf{I}) \quad (3)$$

$$p(\mathbf{B}\mathbf{u} | \sigma_{\mathbf{B}\mathbf{u}}^2) = \prod_{u=1}^N N(\mathbf{U}_u | 0, \sigma_{\mathbf{B}\mathbf{u}}^2 \mathbf{I}) \quad (4)$$

$$p(\mathbf{B}\mathbf{v} | \sigma_{\mathbf{B}\mathbf{v}}^2) = \prod_{i=1}^M N(\mathbf{V}_i | 0, \sigma_{\mathbf{B}\mathbf{v}}^2 \mathbf{I}) \quad (5)$$

For a given user  $u$ , implied feature vectors  $\mathbf{U}_u$ ,  $\mathbf{V}_i$  for item  $i$ , user  $u$  standardized score of articles  $i$  meet normal distribution who has mean  $g(\mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i + \mathbf{U}_u^T \mathbf{V}_i)$  and variance  $\sigma_{\mathbf{R}}$  and mutually independent. The user - item rating matrix conditional probability distribution is as follows:

$$p(\mathbf{R} | \mathbf{U}, \mathbf{V}, \mathbf{B}\mathbf{u}, \mathbf{B}\mathbf{v}, \mu, \sigma_{\mathbf{R}}^2) = \prod_{u=1}^N \prod_{i=1}^M [N(\mathbf{R}_{u,i} | g(\mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i + \mathbf{U}_u^T \mathbf{V}_i), \sigma_{\mathbf{R}}^2)]^{\mathbf{I}_{u,i}^{\mathbf{R}}} \quad (6)$$

Among these,  $\mathbf{I}_{u,i}^{\mathbf{R}}$  is an indicator function, when the user  $u$  have a score record for item  $i$ , then  $\mathbf{I}_{u,i}^{\mathbf{R}} = 1$ , Otherwise  $\mathbf{I}_{u,i}^{\mathbf{R}} = 0$ ;  $g(x) = \frac{1}{1+e^{-x}}$  is a logistic function, mapping the value of  $\mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i + \mathbf{U}_u^T \mathbf{V}_i$  to range  $[0,1]$ .

A posteriori distribution function of  $\mathbf{U}$ ,  $\mathbf{V}$ ,  $\mathbf{B}\mathbf{u}$ ,  $\mathbf{B}\mathbf{v}$  can be derived from Figure 1, the log function formula of the posterior distribution function can be seen in formula(7):

$$\begin{aligned}
 & \ln p(\mathbf{U}, \mathbf{V}, \mathbf{B}\mathbf{u}, \mathbf{B}\mathbf{v} | \mathbf{R}, \mu, \sigma_{\mathbf{U}}^2, \sigma_{\mathbf{V}}^2, \sigma_{\mathbf{B}\mathbf{u}}^2, \sigma_{\mathbf{B}\mathbf{v}}^2, \sigma_{\mathbf{R}}^2) \\
 &= -\frac{1}{2\sigma_{\mathbf{R}}^2} \sum_{u=1}^N \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} (\mathbf{R}_{u,i} - g(\mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i + \mathbf{U}_u^T \mathbf{V}_i))^2 - \\
 & \frac{1}{2\sigma_{\mathbf{U}}^2} \sum_{u=1}^N \mathbf{U}_u^T \mathbf{U}_u - \frac{1}{2\sigma_{\mathbf{V}}^2} \sum_{i=1}^M \mathbf{V}_i^T \mathbf{V}_i - \frac{1}{2\sigma_{\mathbf{B}\mathbf{u}}^2} \sum_{u=1}^N \mathbf{B}\mathbf{u}_u^2 - \frac{1}{2\sigma_{\mathbf{B}\mathbf{v}}^2} \sum_{i=1}^M \mathbf{B}\mathbf{v}_i^2 - \\
 & \sum_{u=1}^N \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} \ln \sigma_{\mathbf{R}} - K \cdot \sum_{u=1}^N \sigma_{\mathbf{U}} - K \cdot \sum_{i=1}^M \sigma_{\mathbf{V}} - K \cdot \sum_{u=1}^N \sigma_{\mathbf{B}\mathbf{u}} - K \cdot \sum_{i=1}^M \sigma_{\mathbf{B}\mathbf{v}} + C \tag{7}
 \end{aligned}$$

Among these,  $C$  is constant. The maximum formula (7) can be considered as an unconstrained optimization problem, the minimization formula (8) is equivalent to maximizing the formula (7):

$$\begin{aligned}
 L(\mathbf{U}, \mathbf{V}, \mathbf{B}\mathbf{u}, \mathbf{B}\mathbf{v}, \mathbf{R}) &= \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} (\mathbf{R}_{u,i} - g(\mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i + \mathbf{U}_u^T \mathbf{V}_i))^2 + \\
 & \frac{\theta_{\mathbf{U}}}{2} \sum_{u=1}^N \mathbf{U}_u^T \mathbf{U}_u + \frac{\theta_{\mathbf{V}}}{2} \sum_{i=1}^M \mathbf{V}_i^T \mathbf{V}_i + \frac{\theta_{\mathbf{B}\mathbf{u}}}{2} \sum_{u=1}^N \mathbf{B}\mathbf{u}_u^2 + \frac{\theta_{\mathbf{B}\mathbf{v}}}{2} \sum_{i=1}^M \mathbf{B}\mathbf{v}_i^2 \tag{8}
 \end{aligned}$$

$$\theta_{\mathbf{U}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{U}}^2}, \quad \theta_{\mathbf{V}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{V}}^2}, \quad \theta_{\mathbf{B}\mathbf{u}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{B}\mathbf{u}}^2}, \quad \theta_{\mathbf{B}\mathbf{v}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{B}\mathbf{v}}^2}$$

Note that  $\theta_{\mathbf{U}}, \theta_{\mathbf{V}}, \theta_{\mathbf{B}\mathbf{u}}, \theta_{\mathbf{B}\mathbf{v}}$ . The local minimum of the formula (8) can be obtained by the gradient descent method. The gradient descent formula of the parameter  $\mathbf{U}_u, \mathbf{V}_i, \mathbf{B}\mathbf{u}_u, \mathbf{B}\mathbf{v}_i$  can be seen in formula (9) ~ (12):

$$\frac{\partial L}{\partial \mathbf{B}\mathbf{u}_u} = \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} g'(\mathbf{X}_{u,i}) (g(\mathbf{X}_{u,i}) - \mathbf{R}_{u,i}) + \theta_{\mathbf{B}\mathbf{u}} \mathbf{B}\mathbf{u}_u \tag{9}$$

$$\frac{\partial L}{\partial \mathbf{B}\mathbf{v}_i} = \sum_{u=1}^N \mathbf{I}_{u,i}^{\mathbf{R}} g'(\mathbf{X}_{u,i}) (g(\mathbf{X}_{u,i}) - \mathbf{R}_{u,i}) + \theta_{\mathbf{B}\mathbf{v}} \mathbf{B}\mathbf{v}_i \tag{10}$$

$$\frac{\partial L}{\partial \mathbf{U}_u} = \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} \mathbf{V}_i g'(\mathbf{X}_{u,i}) (g(\mathbf{X}_{u,i}) - \mathbf{R}_{u,i}) + \theta_{\mathbf{U}} \mathbf{U}_u \tag{11}$$

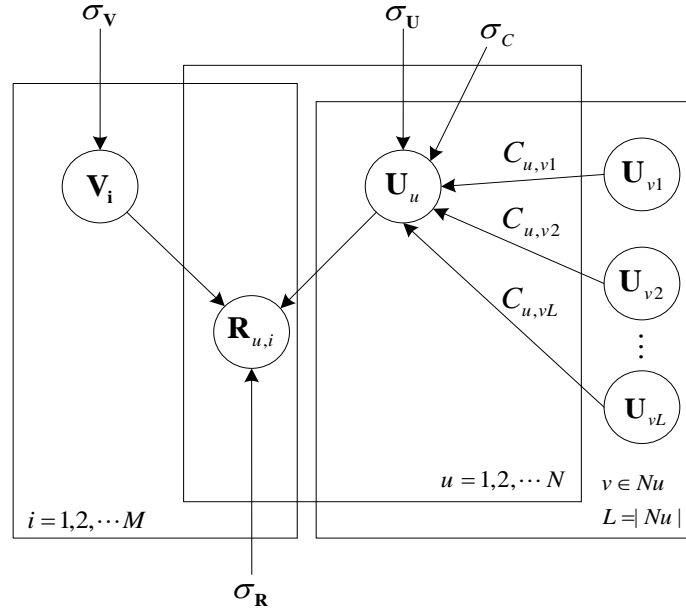
$$\frac{\partial L}{\partial \mathbf{V}_i} = \sum_{u=1}^N \mathbf{I}_{u,i}^{\mathbf{R}} \mathbf{U}_u g'(\mathbf{X}_{u,i}) (g(\mathbf{X}_{u,i}) - \mathbf{R}_{u,i}) + \theta_{\mathbf{V}} \mathbf{V}_i \tag{12}$$

Note that  $g'(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ ,  $\mathbf{X}_{u,i} = \mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i + \mathbf{U}_u^T \mathbf{V}_i$ .

#### 4.2. ISMF Model

ISMF model is an improved social recommendation model of MF Social model proposed by literature [7]. ISMF model uses the formula (1) to modify the deficiency brought by only using trust value to represent the degree of correlation between users in the MF Social model. What is more, Social MF models is designed for the actual situation that different user has various trusting user preferences, which may bring in the problem of information missing, and this will ultimately affect the user's latent feature vector,

reducing the accuracy of prediction to some extent. Therefore, in order to make up for this defect, this situation is corrected by redesigning, the user's potential feature vector of each user's trusting user is computed by the user's correlation matrix  $\mathbf{C}=[\mathbf{C}_{u,v}]_{N \times N}$  to calculate the impact of the potential feature vector generated by the user,  $\mathbf{T}=[\mathbf{T}_{u,v}]$ , following the formula mentioned in the literature [7], thus obtained the ISMF model, as is shown in Figure 2.



**Figure 2. ISMF Model**

The ISMF model is based on the following assumptions:

Assume that  $\mathbf{U}_u$ ,  $\mathbf{V}_i$  are subject to normal distribution and are independent of each other. Formula with (2), (3)

For given user  $u$ , hidden feature vector  $\mathbf{U}_u$ ,  $\mathbf{V}_i$  of items  $i$ , the standard score of item  $i$  given by user  $u$  is  $g(\mathbf{U}_u^T \mathbf{V}_i)$ , the normal distribution of the variance is  $\sigma_R$  and the others are independent of each other.

$$p(\mathbf{R} | \mathbf{U}, \mathbf{V}, \sigma_R^2) = \prod_{u=1}^N \prod_{i=1}^M [N(r_{u,i} | g(\mathbf{U}_u^T \mathbf{V}_i), \sigma_R^2)]^{\mathbf{I}_{u,i}^R} \quad (13)$$

according to Bayesian inference in Figure 2:

$$\begin{aligned} p(\mathbf{U} | \mathbf{C}, \sigma_U^2, \sigma_C^2) &\propto p(\mathbf{U} | \sigma_U^2) \times p(\mathbf{U} | \mathbf{C}, \sigma_C^2) \\ &= \prod_{u=1}^N N(\mathbf{U}_u | 0, \sigma_U^2 \mathbf{I}) \times \prod_{u=1}^N \prod_{v \in Nu} [N(\mathbf{U}_u | \mathbf{U}_v, \sigma_C^2 \mathbf{I})]^{C_{u,v}} \end{aligned} \quad (14)$$

A posterior distribution function of  $\mathbf{U}$ ,  $\mathbf{V}$  is derived, and the formula (15) is obtained as followed:

$$p(\mathbf{U}, \mathbf{V} | \mathbf{R}, \mathbf{C}, \sigma_R^2, \sigma_C^2, \sigma_U^2, \sigma_V^2) \propto p(\mathbf{R} | \mathbf{U}, \mathbf{V}, \sigma_R^2) p(\mathbf{U} | \mathbf{C}, \sigma_U^2, \sigma_C^2) p(\mathbf{V} | \sigma_V^2)$$

$$\begin{aligned}
 &= \prod_{u=1}^N \prod_{i=1}^M [N(\mathbf{R}_{u,i} | g(\mathbf{U}_u^T \mathbf{V}_i), \sigma_{\mathbf{R}}^2)]^{\mathbf{I}_{u,i}^{\mathbf{R}}} \times \prod_{u=1}^N \prod_{v \in Nu} [N(\mathbf{U}_u | \mathbf{U}_v, \sigma_{\mathbf{C}}^2)]^{\mathbf{C}_{u,v}} \\
 &\times \prod_{u=1}^N N(\mathbf{U}_u | 0, \sigma_{\mathbf{U}}^2 \mathbf{I}) \times \prod_{i=1}^M N(\mathbf{V}_i | 0, \sigma_{\mathbf{V}}^2 \mathbf{I})
 \end{aligned} \tag{15}$$

The minimization formula (16) is equivalent to the log function of the maximum formula (15):

$$\begin{aligned}
 L(\mathbf{R}, \mathbf{T}, \mathbf{U}, \mathbf{V}) &= \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} (\mathbf{R}_{u,i} - g(\mathbf{U}_u^T \mathbf{V}_i))^2 + \frac{\theta_{\mathbf{U}}}{2} \sum_{u=1}^N \mathbf{U}_u^T \mathbf{U}_u + \frac{\theta_{\mathbf{V}}}{2} \sum_{i=1}^M \mathbf{V}_i^T \mathbf{V}_i \\
 &+ \frac{\theta_{\mathbf{C}}}{2} \sum_{u=1}^N \sum_{v \in Nu} \mathbf{C}_{u,v} (\mathbf{U}_u - \mathbf{U}_v)^T (\mathbf{U}_u - \mathbf{U}_v)
 \end{aligned} \tag{16}$$

$$\theta_{\mathbf{U}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{U}}^2}, \quad \theta_{\mathbf{V}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{V}}^2}, \quad \theta_{\mathbf{C}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{C}}^2}$$

Note that . The local minimum value of the formula (16) can be solved by the gradient descent method. The gradient descent formula of the parameters  $\mathbf{U}_u$  and  $\mathbf{V}_i$  can be find in formula (17) ~ (18):

$$\begin{aligned}
 \frac{\partial L}{\partial \mathbf{U}_u} &= \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} \mathbf{V}_i g'(\mathbf{U}_u^T \mathbf{V}_i) (g(\mathbf{U}_u^T \mathbf{V}_i) - \mathbf{R}_{u,i}) + \theta_{\mathbf{U}} \mathbf{U}_u \\
 &+ \theta_{\mathbf{C}} \sum_{v \in Nu} \mathbf{C}_{u,v} (\mathbf{U}_u - \mathbf{U}_v) + \theta_{\mathbf{C}} \sum_{\{g|u \in Ng\}} \mathbf{C}_{g,u} (\mathbf{U}_g - \mathbf{U}_u)
 \end{aligned} \tag{17}$$

$$\frac{\partial L}{\partial \mathbf{V}_i} = \sum_{u=1}^N \mathbf{I}_{u,i}^{\mathbf{R}} \mathbf{U}_u g'(\mathbf{U}_u^T \mathbf{V}_i) (g(\mathbf{U}_u^T \mathbf{V}_i) - \mathbf{R}_{u,i}) + \theta_{\mathbf{V}} \mathbf{V}_i \tag{18}$$

### 4.3. BISMF Model

In Section 4.1, combined with the baseline prediction model, in the existing probability model, the independent attribute factors, which are applied to the scene, the user and the item, are added, and the BPMF model is proposed to modify the probability matrix factorization model; in this way, BPMF model is put forward. In Section 4.2, considering correlation between users on trust and interest similarity, ISMF model is presented by amending the latent defect feature vectors for users at the same time; In this section, we fuse the advantages of the above two models and get the BISMF model by using the simple fusion BPMF and ISMF model, as is shown in Figure 3.



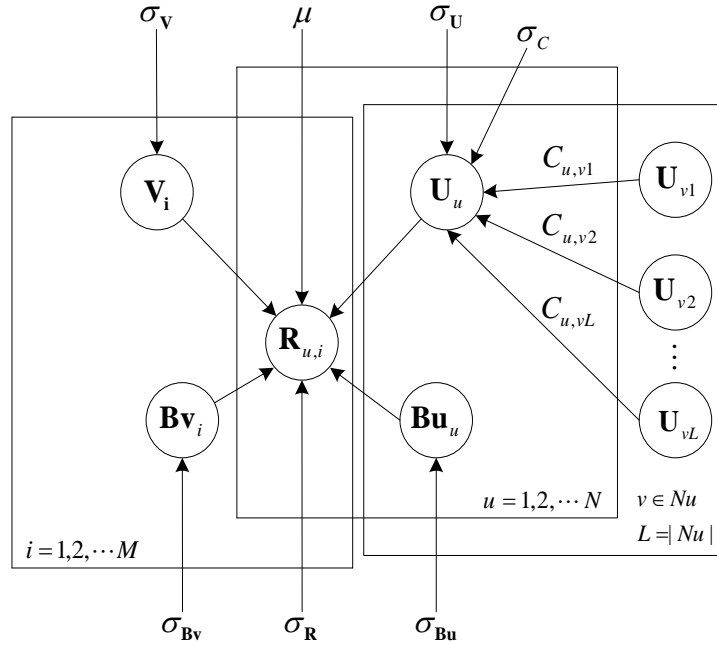


Figure 3. BISMF Model

According to the model, the formula is as follows:

$$\begin{aligned}
 L(\mathbf{R}, \mathbf{C}, \mathbf{U}, \mathbf{V}, \mathbf{B}\mathbf{u}, \mathbf{B}\mathbf{v}) &= \frac{1}{2} \sum_{u=1}^N \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} (\mathbf{R}_{u,i} - g(\mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i \mathbf{U}_u^T \mathbf{V}_i))^2 \\
 &+ \frac{\theta_U}{2} \sum_{u=1}^N \mathbf{U}_u^T \mathbf{U}_u + \frac{\theta_V}{2} \sum_{i=1}^M \mathbf{V}_i^T \mathbf{V}_i + \frac{\theta_{\mathbf{B}\mathbf{u}}}{2} \sum_{u=1}^N \mathbf{B}\mathbf{u}_u^2 + \frac{\theta_{\mathbf{B}\mathbf{v}}}{2} \sum_{i=1}^M \mathbf{B}\mathbf{v}_i^2 \\
 &+ \frac{\theta_C}{2} \sum_{u=1}^N \sum_{v \in Nu} \mathbf{C}_{u,v} (\mathbf{U}_u - \mathbf{U}_v)^T (\mathbf{U}_u - \mathbf{U}_v) \quad (19)
 \end{aligned}$$

Note that  $\theta_U = \frac{\sigma_{\mathbf{R}}^2}{\sigma_U^2}$ ,  $\theta_V = \frac{\sigma_{\mathbf{R}}^2}{\sigma_V^2}$ ,  $\theta_{\mathbf{B}\mathbf{u}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{B}\mathbf{u}}^2}$ ,  $\theta_{\mathbf{B}\mathbf{v}} = \frac{\sigma_{\mathbf{R}}^2}{\sigma_{\mathbf{B}\mathbf{v}}^2}$ ,  $\theta_C = \frac{\sigma_{\mathbf{R}}^2}{\sigma_C^2}$ . The Gradient descent formula of  $\mathbf{B}\mathbf{u}_u$ ,  $\mathbf{B}\mathbf{v}_i$ ,  $\mathbf{V}_i$  are the same as formula (9), (10), (12). The Gradient descent formula of  $\mathbf{U}_u$  is shown in formula (20):

$$\begin{aligned}
 \frac{\partial L}{\partial \mathbf{U}_u} &= \sum_{i=1}^M \mathbf{I}_{u,i}^{\mathbf{R}} \mathbf{V}_i g'(\mathbf{X}_{u,i}) (g(\mathbf{X}_{u,i}) - \mathbf{R}_{u,i}) + \theta_U \mathbf{U}_u \\
 &+ \theta_C \sum_{v \in Nu} \mathbf{C}_{u,v} (\mathbf{U}_u - \mathbf{U}_v) + \theta_C \sum_{\{g|u \in Ng\}} \mathbf{C}_{g,u} (\mathbf{U}_g - \mathbf{U}_u) \quad (20)
 \end{aligned}$$

Among these,  $g'(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ ,  $\mathbf{X}_{u,i} = \mu + \mathbf{B}\mathbf{u}_u + \mathbf{B}\mathbf{v}_i + \mathbf{U}_u^T \mathbf{V}_i$ . For convenience, we set  $\theta = \theta_U = \theta_V = \theta_{\mathbf{B}\mathbf{u}} = \theta_{\mathbf{B}\mathbf{v}} = \theta_C$  in this article.

#### 4.4. Similarity Function

The social recommendation algorithms promoted in Section 4.2 and Section 4.3 need to use the similarity among the users, that is to say, calculation is essential to get the result of  $\mathbf{S} = [\mathbf{S}_{u,v}]_{N \times N}$ . We can calculate the similarity between user and item according to user - item score record of all users, in this paper, the PCC (Pearson Correlation Coefficient) method is used to define the similarity between users  $u$  and users  $v$ , as is shown in formula (21):

$$\mathbf{S}_{u,v} = \frac{\sum_{i \in Z(u) \cap Z(v)} (\mathbf{R}_{u,i} - \bar{\mathbf{R}}_u) \cdot (\mathbf{R}_{v,i} - \bar{\mathbf{R}}_v)}{\sqrt{\sum_{i \in Z(u) \cap Z(v)} (\mathbf{R}_{u,i} - \bar{\mathbf{R}}_u)^2} \cdot \sqrt{\sum_{i \in Z(u) \cap Z(v)} (\mathbf{R}_{v,i} - \bar{\mathbf{R}}_v)^2}} \quad (21)$$

Among these,  $\bar{\mathbf{R}}_u$  indicates average score for user  $u$ ,  $Z(u)$  indicates the collection of items which have been scored by the users, According to this definition,  $\mathbf{S}_{u,v}$  within the range  $[-1,1]$ , the greater the value is, the more similar interests the two users have, in this paper,  $f(x) = (x+1)/2$  is used to measure the similarity within the range  $[0,1]$ .

#### 4.5. Complexity Analysis

In this paper, we analyze the complexity of parameter learning, for the BISMf algorithm, assuming that every user evaluation score is  $\bar{r}$ , the average number of trusting users is  $\bar{t}$ , the complexity of the minimization function is  $O(N\bar{r}K + N\bar{t}K)$ , since the user - item rating matrix and user - user trust matrix are sparse matrix, so the value of  $\bar{r}$  and  $\bar{t}$  is relatively small, therefore, the computational complexity is low as well. The computational complexity of the solution is  $O(N\bar{r}K + N\bar{t}K)$  according to the gradient descent method. Compared with Computational complexity of the Gradient descent equation of Social MF,  $O(N\bar{r}K + N\bar{t}^2K)$ , obviously, the computational complexity of BISMf proposed in this article is much better.

### 5. Results and Analysis

In this section, we analyze some experimental results to compare our proposed recommendation method with the recommendation quality of the existing algorithms.

#### 5.1. Experimental Environment and Data Set

This paper uses a program in a 2.00GB for the GHz, the processor for the dual core 2.93 Intel CPU RAM with Windows 7 operating system on the computer.

In this paper, we use the data set from the Movielens data set and Epinions data set. We use the 1M-version Movielens data set, which has roughly six thousand users, four thousand movies, and one million user rating records to compare the quality of the recommendation of the BPFM with PMF algorithm. We use Epinions data set to compare the recommended quality among the Social MF, ISMF, BISMf, their specific specifications can be seen in Table 2.

**Table 2. Dataset Introduction**

Statistics	Movielens	Epinions
Users Number	6,040	49,290
Item Number	3,900	139,738
Trust State	/	487,181
Ratings Records	1,000,209	664,824

### 5.2. Evaluation Index

We use two evaluation metrics: mean absolute error (MAE) and root mean square error (RMSE) to calculate. Their calculation methods are respectively Formula (22) and (23):

$$RMSE = \sqrt{\frac{\sum_{u,i \in T} (\mathbf{R}_{u,i} - \hat{\mathbf{R}}_{u,i})^2}{|T|}} \quad (22)$$

$$MAE = \frac{\sum_{u,i \in T} |\mathbf{R}_{u,i} - \hat{\mathbf{R}}_{u,i}|}{|T|} \quad (23)$$

Among these, T indicates the test set, for user  $u$  and item  $i$  in the test collection, the actual score of the  $u$  for the item  $i$  is  $\mathbf{R}_{u,i}$ , and the  $\hat{\mathbf{R}}_{u,i}$  is the prediction of the algorithm. but  $\hat{\mathbf{R}}_{u,i}$  is a predictive score for the calculation of the recommendation algorithm. The RMSE does the square for each absolute error, as a result, increasing the penalty squared term, obviously, RMSE is always equal to or greater than MAE, and RMSE has stringenter request for system evaluation than MAE.

### 5.3. Experimental Results

Table 4 and Table 3 are different percentages of training sets when PMF and BPFM each has the percentage of 90%, 80%, 70% respectively, on the condition of K=10 and K=30, And in this experiment, we set  $\theta_u = \theta_v = 0.05$ ,  $\theta_{Bu} = \theta_{Bv} = 0.001$ . Compared with the data in Table 4 and Table 3, in the same conditions, BPFM outperforms PMF in both two evaluation indexes of RMSE and MAE.

**Table 3. Performance Comparison of PMF and BPFM on Movielens Data Sets (K=10)**

Training Data	Evaluation index	PMF	BPFM
90%	MAE	0.6738	0.6664
	RMSE	0.8552	0.8511
80%	MAE	0.6744	0.6685
	RMSE	0.8582	0.8542
70%	MAE	0.6800	0.6747
	RMSE	0.8645	0.8610

**Table 4. Performance Comparison of PMF and BPF on Movielens Data Sets (K=30)**

Training Data	Evaluation index	PMF	BPF
90%	MAE	0.6696	0.6631
	RMSE	0.8504	0.8465
80%	MAE	0.6717	0.6646
	RMSE	0.8537	0.8491
70%	MAE	0.6786	0.6721
	RMSE	0.8599	0.8567

Here are the performance comparisons among Social MF and ISMF, BISMf proposed in this article when  $K = 10$  and  $K = 30$ , respectively, and the experiments were performed on the Epinions data set of 60%, 70%, 80%, 90% training sets, note that,  $\theta_u = \theta_v = 0.05$ ,  $\theta_{Bu} = \theta_{Bv} = 0.001$ ,  $\theta_c = 0.1$ . Obviously, conclusions can be drawn that ISMF has improved the MF Social to a certain extent, and compared with the MF Social, it has achieved more accurate results.

**Table 5. Performance Comparison of Three Methods on Epinions Data Sets (K=10)**

Training Data	Evaluation index	Social MF	ISMF	BISMf
90%	MAE	0.8316	0.8266	0.7844
	RMSE	1.0824	1.0766	1.0361
80%	MAE	0.8394	0.8380	0.7873
	RMSE	1.0922	1.0874	1.0401
70%	MAE	0.8448	0.8396	0.7935
	RMSE	1.0998	1.0923	1.0443
60%	MAE	0.8570	0.8515	0.7999
	RMSE	1.1131	1.1064	1.0530

**Table 6. Performance Comparison of Three Methods on Epinions Data Sets (K=30)**

Training Data	Evaluation index	Social MF	ISMF	BISMf
90%	MAE	0.8308	0.8247	0.7842
	RMSE	1.0809	1.0728	1.0360
80%	MAE	0.8376	0.8326	0.7870
	RMSE	1.0882	1.0813	1.0400
70%	MAE	0.8433	0.8386	0.7932
	RMSE	1.0957	1.0899	1.0442
60%	MAE	0.8534	0.8493	0.7997
	RMSE	1.1087	1.1024	1.0530

#### 5.4. Influence of $\theta_c$

The parameter  $\theta_c$  controls the influence of user's social network to the user's network behavior, the role of social network on the user's network behavior becomes larger as the value of  $\theta_c$  is higher, conversely, the smaller.

Figure 5 and Figure 4 show respectively that when the  $\theta_c$  value varies, the influence on the experimental results in MAE and RMSE. According to Figure 4 and Figure 5, we can see the results of the experiment have the best effect when  $\theta_c=0.1$ .

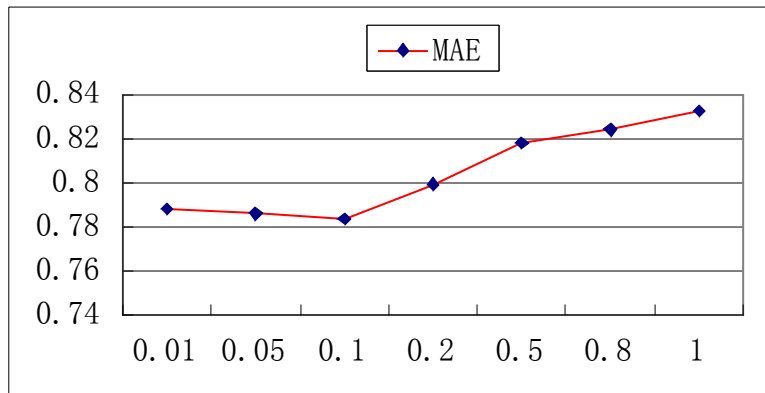


Figure 4. Influence of  $\theta_c$

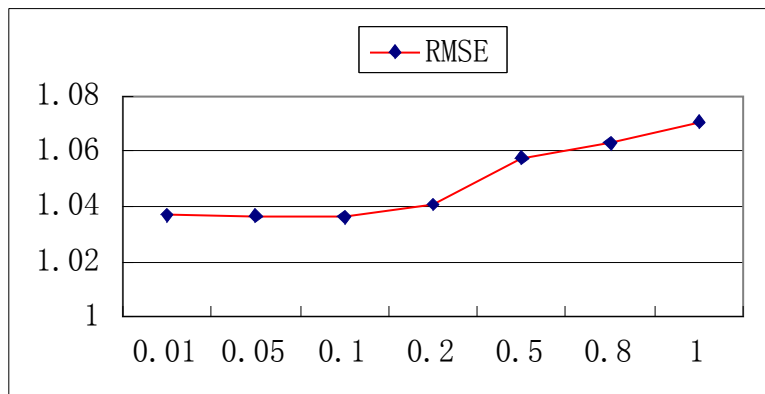


Figure 5. Influence of  $\theta_c$

## 6. Conclusion

This paper mainly studies how to improve the prediction accuracy of recommendation system based on social network, and we put forward three models. The experimental results show that our approach outperforms the other state-of-the-art collaborative filtering algorithms, and can further improve the prediction accuracy of recommendation system.

With the development of the Internet, social networks are becoming more and more complexed, the social networks among users are becoming closer. Social relations of various social networks, such as Facebook based on social graph, Twitter based on interest and Gowalla based on geographic information service will be explored in the coming future.

## Acknowledgment

This work is supported by the National Natural Science Foundation of China under grant No. 61272171.

## References

- [1] P. Massa and P. Avesani, "Trust-aware collaborative filtering for recommender systems", On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE. Springer Berlin Heidelberg, (2004), pp. 492-508.
- [2] S. Ray and A. Mahanti, "Improving prediction accuracy in trust-aware recommender systems", System Sciences (HICSS), 2010 43rd Hawaii International Conference on. IEEE, (2010), pp. 1-9.
- [3] A. Mnih and R. Salakhutdinov, "Probabilistic matrix factorization", Advances in neural information processing systems, (2007), pp. 1257-1264.
- [4] H. Ma, H. Yang and M. R. Lyu, "Sorec: social recommendation using probabilistic matrix factorization", Proceedings of the 17th ACM conference on Information and knowledge management, ACM, (2008), pp. 931-940.
- [5] S. Purushotham, Y. Liu and C. C. J. Kuo, "Collaborative Topic Regression with Social Matrix Factorization for Recommendation Systems", arXiv preprint arXiv:1206.4684, (2012).
- [6] H. Ma, I. King and M. R. Lyu, "Learning to recommend with social trust ensemble", Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, ACM, (2009), pp. 203-210.
- [7] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks", Proceedings of the fourth ACM conference on Recommender systems. ACM, (2010), pp. 135-142.
- [8] H. Ma, D. Zhou and C. Liu, "Recommender systems with social regularization", Proceedings of the fourth ACM international conference on Web search and data mining, ACM, (2011), pp. 287-296.
- [9] R. Salakhutdinov and A. Mnih, "Bayesian probabilistic matrix factorization using Markov chain Monte Carlo", Proceedings of the 25th international conference on Machine learning. ACM, (2008), pp. 880-887.
- [10] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model", Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, (2008), pp. 426-434.
- [11] Y. Koren, R. Bell and C. Volinsky, "Matrix factorization techniques for recommender systems", Computer, vol. 42, no.8, (2009), pp. 30-37.
- [12] J. J. Levandoski, M. Sarwat and A. Eldawy, "Lars: A location-aware recommender system", Data Engineering (ICDE), 2012 IEEE 28th International Conference on. IEEE, (2012), pp. 450-461.
- [13] B. Berjani and T. Strufe, "A recommendation system for spots in location-based online social networks", Proceedings of the 4th Workshop on Social Network Systems. ACM, (2011), pp. 4.
- [14] W. Wu, L. He and J. Yang, "Evaluating recommender systems", ICDIM, (2012), pp. 56-61.

## Authors



**Wei Luo**, he received his M.E. degree from Northeast Dianli University, in 2006. Now, he is working on his doctor degree at Dalian Maritime University. His research interests include Natural Language Processing and Recommendation system.



**Zhihao Peng**, he received his M.E. degree from Dalian Maritime University, in 2007. Now, he is working on his Doctor's degree at Dalian Maritime University. His research interests include Computer Science Reasoning, Formal Concept Analysis and Belief revision *etc.*



**Ansheng Deng**, he received his Undergraduate degree from Peking University, in 1985. He received his M.E. degree from Jilin University, in 1988, and then received Doctor degree in 1995. Now, he is working at Dalian Maritime University. His research interests include Computer Science Reasoning, Formal Concept Analysis and Belief revision *etc.*

