CrossMark

ISoTrustSeq: a social recommender system based on implicit interest, trust and sequential behaviors of users using matrix factorization

Vahideh Nobahari[1] · Mehrdad Jalali[1] ·
Seyyed Javad Seyyed Mahdavi[1]

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract *Recommender systems* try to propose a list of main interests of an on line social network user based on his predicted rating values. In the recent years, several methods are proposed such as *Interest Social Recommendation method (ISoRec)*, and *Social Recommendation method* based on *trust Sequence Matrix Factorization* which employs matrix factorization techniques to address the trust propagation and sequential behaviors issues. Main drawback of these works is that they ignore implicit interest of users. Therefore, the main goal of this paper is to solve user-item rating based on the trust, sequential interest and the implicit interest of users, simultaneously. In order to solve this problem, our proposed method combines these parameters as its inputs. This method based on matrix factorization named as ISoTrustSeq. Experimental results show higher accuracy of predicted values in compared to the above-mentioned methods. Our results are also much better than these methods in terms of variation in the number of user-items features.

Keywords Online social networks · Recommender system · Matrix factorizations · Trust · Sequential behaviors · Implicit interest

1 Introduction

Recommender systems suggest a list of items that a user may be interested in. These systems help users to easily choosed their suitable items in a very short time. Recent studies

✉ Mehrdad Jalali
Jalali@mshdiau.ac.ir

Vahideh Nobahari
Vnobahari@mshdiau.ac.ir

Seyyed Javad Seyyed Mahdavi
Mahdavi@mshdiau.ac.ir

[1]   Department of Computer, Mashhad Branch, Islamic Azad University, Mashhad, Iran

show that social networks data can be used to improve the accuracy of the recommendations. Generally, there are two different categories of the Recommender systems: Context base and Collaborative filtering (CF) (Bobadilla et al. 2013). Collaborative filtering-based recommender systems are divided into two categories: Matrix Factorization based social recommender methods (SocialMF) and Neighborhood based social recommender methods (Su and Khoshgoftaar 2009). Among the several types of recommender systems, collaborative filtering is one of the most popular methods which is simple to implement and can discover complicated patterns (Zhang et al. 2013).

On-line Social Networks (OSN) provided conditions for the recommender systems. These systems can improve the accuracy of predictions. In social recommender systems, interests and trust between social network people improve the accuracy of traditional recommender system (Yang et al. 2014). In recent years, experts in industry and academia gained satisfactory performance of model-based methods, especially Latent Factor Model (LFM) (Zhang and Liu 2014). This model mapped users and items to latent factor space of them. In the collaborative filtering, high performance of matrix factorization method increases the accuracy of predicted rating in the user-items rating matrix. This performance is in dealing with large dataset (Zhang et al. 2013). Since matrix factorization supposes just some factors which effect on user-item rating performance, so it can be impressive in training. Singular Value Decomposition (SVD) and Expectation Maximization (EM) methods are suggested to minimize the Root Mean Squared Error (RMSE) (Zhang et al. 2013).

Interest Social Recommendation method (ISoRec) (Yu-sheng et al. 2014), is used to solve problem of predicting user rating of items in social recommender systems based on matrix factorization. If the user social communication information is not available, it only uses the implicit user's interest for the social recommendation. The researchers in the scientific field used the similarity of users by Pearson Correlation Coefficient (PCC) (Sections 2 and 3) for achieving implicit user interest matrix (Yu-sheng et al. 2014; Zhang and Liu 2015). If the users' social recommendation data is available, implicit and explicit communication information can be used simultaneously. The latent features matrix (U) of the user and the latent features matrix (V) of the item are obtained by the user-item rating matrix (R) decomposition. Now with users' explicit and implicit information matrix and their decomposition matrix, the relevant conditional probability distribution is created, and then they try to obtain a matrix U and V using a gradient descent method (Bottou 2010) which ultimately can form the predicted rating matrix through transpose multiplication (Koren et al. 2009). This approach ignores the sequence of user behaviors (temporal relations between the items). *Trust Sequence Matrix Factorization-based Social Recommendation (TrustSeqMF)* (Zhang and Liu 2015) creates the trusted user matrix and another matrix containing the relationships between the items according to social relations of users on social networks and their trust in trusted friends. If the user social relationship information is not available, this method does not use users' implicit interests and their effect.

The proposed method in this study is an extension of recent methods and tries to overcome the drawbacks of previous methods. Although *Interest Social Recommendation (ISoRec)* model (Yu-sheng et al. 2014) and *Trust Sequence Matrix Factorization (TrustSeqMF)* method (Zhang and Liu 2015) have tried to solve the problem of predicting user rating for items in social recommender systems, ignoring implicit interest of users has remained as a challenge. Therefore, providing a system which could solve this problem is important. The main objective of this paper is to present a system that increases the prediction accuracy. In doing so, it can solve the problem related to predicting user rating of items, in a way that is closer to the user view of point.

In this paper, it is assumed that every user who is a member of the social network must establish social connections with his trusted friends and people. Also, we assume that each person has rated for at least five items of all the items provided.

The reminder of this paper is organized as follows: In section two, some basic concepts are introduced. In the third section, the proposed method is described in detail. Section four shows the experimental results and the summary and recommendations for feature works will be discussed in the last section.

2 Basic concepts

In this section some concepts to evaluate previous works in the field of recommender systems based on matrix factorization are explained. Some concepts that assist us in understanding algorithms in this field include social network graph, recommender methods based Matrix Factorization and Similarity function which will be discussed in the following.

2.1 Social network graph

Recommendations in the real world include two central elements for social Recommendations, friends in social network and favorite items for them (Ma et al. 2011). This is modeled by user-item matrix (Fig. 1) and social network graph (Fig. 2).

Consider a Social Network directed graph $g(u, \varepsilon)$, so that a set of vertices $U = \{u_i\}_{i=1}^n$ represents all users in social networks, and a set of edges ε represents relationships between users. $R = \{r_{ui}\}$ is an $n \times m$ matrix, where n is the number of users and m number of the items (Fig. 1). $U = \{u_1, u_2, .., u_n\}$ is a set of the users and $I = \{i_1, i_2, i_m\}$ is a set of items. $T = \{t_{jk}\}$ is $n \times n$ matrix of the graph g which in the paper (Ma et al. 2008) is called social network matrix (Fig. 3). For each pair of vectors u_k and u_j, $t_{jk} \in (0, 1]$ is the weight related to the edge that indicates to what extent user j knows user k in the social network and trusts in him. $t_{jk} = 0$ means completely trust and $t_{jk} = 1$ means distrust. Note that, T is an asymmetric matrix in social networks, especially in the trust-based social network (Ma et al. 2011). It should be noted that necessarily user j's trust in user k doesn't represent the k user's trust in the user j.

People always pay attention to their trusted friend to make recommendations. It is more possible that people choose items such as books, movies, music etc. based on the choices of their trusted friend (Yang et al. 2014; Zhang et al. 2016). Researchers believe, although a user doesn't prefer proposed items, but it is possible that this user is affected by their friends' idea about those (Ma et al. 2011).

Fig. 1 Rating matrix of user-item

	i_1	i_2	i_3	i_4	i_5	i_6
u_1		5			3	
u_2	1			3		4
u_3		2			2	
u_4	4			3		
u_5		4	3			5

Fig. 2 The social trust of users graph
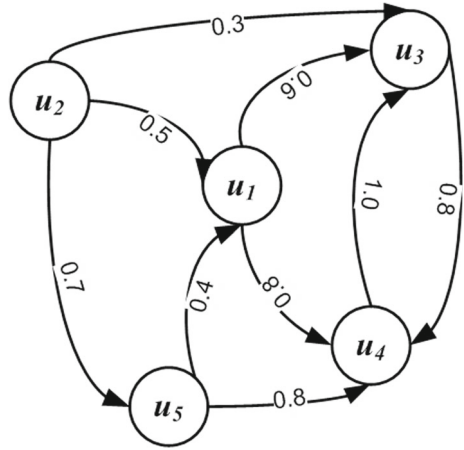
In the real world, we assume that someone wants to watch a film (Item i_1 in Fig. 1), while he does not know anything about it, (like the user u_1 in Fig. 1). What the user would normally do is that he trusts in recommendations by his trusted friends. For example, u_1 trusts to the users' u_2 and u_4 (Fig. 2). The users' u_2 and u_4 have given the point of 1 and 4 to film. u_1 to u_4 (weight 0.8) more trusts in u_2 (weight 0.5) (Fig. 3). According to this information, the probability is very high that u_1 chooses this film based on the u_4 opinion.

2.2 Recommender systems based matrix factorization

Matrix factorization is appropriate for processing large recommender systems data sets and providing scale-up methods (Bobadilla et al. 2013). The main idea of matrix factorization is modeling of user-item interactions which factorization from the latent users and items feature in the system is done by it. Recently, recommender systems based on social matrix factorization is presented to improve the recommendations (Ma et al. 2008, 2011, 2011; Jamali 2010), which will be discussed in Section 3.

This model is trained using available data, and then, is used to predict ratings of the user to items. In the matrix factorization, *Singular Value Decomposition (SVD)* and *Latent Semantic indexing (LSI)* are used. However, the SVD method has a very heavy computational load but provides good results for prediction (Ford et al. 2005). Hence, this decomposition technique is used in the researches (Yu-sheng et al. 2014; Zhang and Liu 2015; Ma et al. 2011; Zhang et al. 2016; Matuszyk et al. 2015; Kannan et al. 2013) carried out in matrix factorization field. An explanation of this method is provided in the following section.

Fig. 3 The trust of user matrix

	u_1	u_2	u_3	u_4	u_5
u_1			0.9	0.8	
u_2	0.5		0.3		0.7
u_3				0.8	
u_4			1		
u_5	0.4			0.8	

2.3 Similarity function

If all users' rating data are available, the similarity between two users is evaluated by calculated rating value of two users. Two common methods in previous studies are *Vector Space Similarity (VSS)* and *Pearson Correlation Coefficient (PCC)* (Kadie et al. 1998). The Vector Space Similarity knows the similarity between two users i and j based on the items that they have rated:

$$sim(i, j) = \frac{\sum\limits_{k \in I(i) \cap I(j)} R_{ik}.R_{jk}}{\sqrt{\sum\limits_{k \in I(i) \cap I(j)} R_{ik}^2} \sqrt{\sum\limits_{k \in I(i) \cap I(j)} R_{jk}^2}} \tag{1}$$

According to the (1), vector space similarity *sim(i, j)* is in *[0,1]*. Plenty of this similarity means that users i and j are similar to each other to much.

Indeed, calculation of the vector space similarity doesn't consider the differences between the users which may have different ratings. Some potential users may give high scores to all products, while some other members may tend to give a lower rating. Therefore, to solve this problem, Pearson correlation coefficient in (2) is suggested (Hofmann 2003):

$$sim(i, j) = \frac{\sum\limits_{k \in I(i) \cap I(j)} (R_{ik} - \overline{R_i}).(R_{jk} - \overline{R_j})}{\sqrt{\sum\limits_{k \in I(i) \cap I(j)} (R_{ik} - \overline{R_i})^2} \sqrt{\sum\limits_{k \in I(i) \cap I(j)} (R_{jk} - \overline{R_j})^2}} \tag{2}$$

I(i) is a set of items which rated by the user u_i. k belongs to the subset of items, which user i and user j have rated it. R_{ik} is a rating value that user i gives to the item k and $\overline{R_i}$ is the average scores of the user u_i. According to this definition, the similarity between two users i and j (*sim (i,j)*) is in the range of [1,-1]. When this value is larger, means user u_i and u_j are very similar to each other. A mapping function $f(x) = (x + 1)/2$ can be used to limit the similarity value *PCC* in the interval of [0,1].

3 Social recommender models

Some researchers suggested social recommender system methods base on matrix factorization. In the following, most important algorithms with their challenges are investigated.

3.1 User-item matrix factorization models

Salakhutdinov and Mnih (2008) have proposed a probability matrix factorization algorithm (PMF). The graphical model in Fig. 4 shows an expression of this algorithm. $U \in R^{K \times N}$ and $V \in R^{K \times M}$ are latent features Matrix of the user and item is individually. Column vector of U_u and, V_i represent a k-dimensional feature vectors for the user-specific and item-specific. In this paper, zero mean Gaussian on the user-item feature vectors can be used:

$$P(U|\sigma_U^2) = \prod_{u=1}^{n} N(U_u|0, \sigma_U^2 I) \tag{3}$$

$$P(V|\sigma_V^2) = \prod_{i=1}^{m} N(V_i|0, \sigma_V^2 I) \tag{4}$$

Fig. 4 The graphical model of PMF (Salakhutdinov and Mnih 2008)
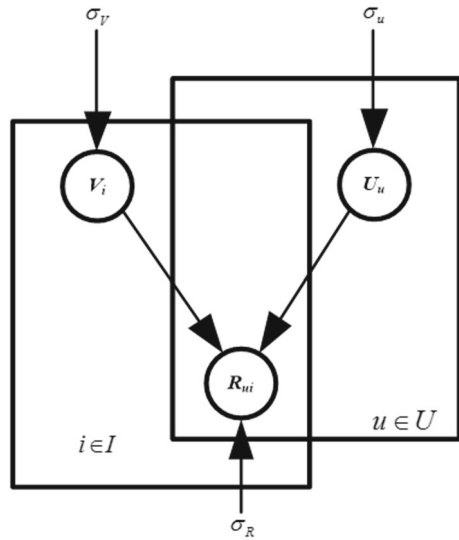
PMF model based collaborative filtering algorithm identifies user-item feature vectors, and then predicts unknown points. User rating for items is a combination of a set of probabilities and conditional distribution as follows (Yang et al. 2014; Zhang et al. 2013):

$$P(R|U, V, \sigma_R^2) = \prod_{u=1}^{n} \prod_{i=1}^{m} \Big[N(R_{ui}|g(U_u^T V_i), \sigma_R^2) \Big]^{I_{ui}^R} \tag{5}$$

$N(x|\mu, \sigma_R^2)$ indicates a Gaussian distribution with mean μ and variance σ_R^2, and I_{ui}^R is an indicator function which if the user has rated its value is equal to one, and otherwise, its value is equal to zero. Logical function is $g(x) = 1/(1 + exp(-x))$, which limits the values of $U_u^T V_i$ to the interval [0,1] (Yang et al. 2014; Zhang et al. 2013).

Based on the above, Prior probability of the latent features U and V is obtained by Bayesian inference:

$$\begin{aligned} P(U, V|R, \sigma_R^2, \sigma_U^2, \sigma_V^2) &\propto P(R|U, V, \sigma_R^2)P(U|\sigma_U^2)p(V|\sigma_V^2) \\ &= \prod_{u=1}^{n} \prod_{i=1}^{m} \Big[N(R_{u,i}|g(U_u^T V_i), \sigma_R^2) \Big]^{I_{ui}^R} \times \prod_{u=1}^{n} N(U_u|0, \sigma_U^2 I) \times \prod_{i=1}^{m} N(V_i|0, \sigma_V^2 I) \end{aligned} \tag{6}$$

The algorithm PMF detects the user-item latent features vectors only through user-item rating matrix, and also user-item social relations are not considered. Therefore, the accuracy of the suggestions presented in this method requires further improvement.

3.2 Social network matrix factorization (SoRec)

The idea raised in the paper (Ma et al. 2008), entitled "Social network matrix factorization (SoRec)" model is to extract the social network graph analysis-based L-dimensional features. $U \in R^{l \times m}$ and $Z \in R^{l \times m}$ are the latent user and decomposition feature matrix with column vectors U_i and Z_k which indicate user-specific and feature-specific latent feature vectors (Yang et al. 2014; Ma et al. 2008). In order to reflect the impact of users' social communication s on the users' interests' items, the Social recommendation is shown by using

the graphical model Fig. 5. This model combines social network graph and user-item rating matrix with each other (Yang et al. 2014; Ma et al. 2008).

Conditional distribution on the relational model of social network is defined as follows:

$$P(C|U, Z, \sigma_C^2) = \prod_{i=1}^{m} \prod_{k=1}^{m} \left[N(C_{ik}|g(U_i^T Z_k), \sigma_C^2) \right]^{I_{ik}^C} \tag{7}$$

Probability density function of Gaussian distribution is N, and $g(x)$ is logic function as mentioned in the previous method. I_{ik}^C is an index function, which if the user i knows the user k, or trusts him, its value is equal to one, otherwise, is equal to zero. Mean-zero Gaussian in the user matrix and decomposition feature vectors are obtained according to the (3) and (4). Therefore, using Bayesian inference, we have:

$$P(U, Z|C, \sigma_C^2, \sigma_U^2, \sigma_Z^2) \propto P(C|U, Z, \sigma_C^2)P(U|\sigma_U^2)p(Z|\sigma_Z^2)$$
$$= \prod_{i=1}^{n} \prod_{k=1}^{m} \left[N(C_{ik}|g(U_i^T Z_k), \sigma_C^2) \right]^{I_{ik}^C} \times \prod_{u=1}^{m} N(U_i|0, \sigma_U^2 I) \times \prod_{k=1}^{m} N(Z_k|0, \sigma_Z^2 I) \tag{8}$$

In online social networks, the value of C_{ik} by user i is associated with the user k. In trust-based social network, if the user i trust in many users, the trust value for C_{ik} increases. Since, it has disorders so is not able to describe the graph structure data on social networks.

The proposed method by Ma et al. (2011), like the previous methods has not considered social communication s and users' trust, as well as the time sequential of relations between items, and their effects on the selection of item by individuals are ignored.

3.3 Social matrix factorization (SocialMF)

Jamali (2010) introduced Social Matrix Factorization (SocialMF). The rating value R_{ui} is achieved by combining U and V matrices. $U \in R^{K \times N}$ and $V \in R^{K \times M}$ are latent features Matrix of the user and item is individually. Column vector of U_u and, V_i represent a k-dimensional feature vectors for the user-specific and item-specific. Prior distribution for U, V and R matrices are σ_U , σ_V and σ_R respectively. The social trust graph shows the each
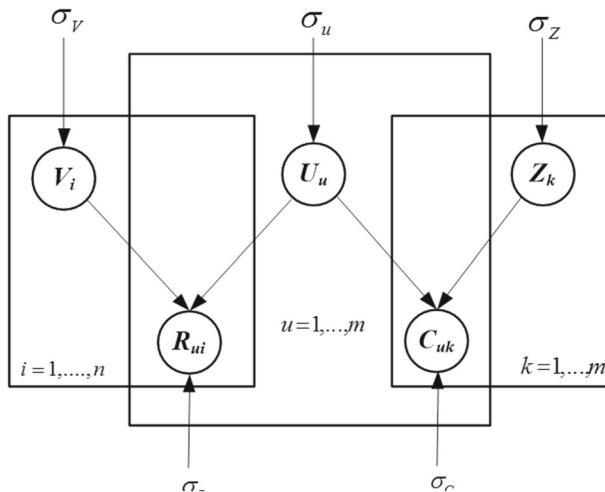

Fig. 5 Graphical model of social network (Ma et al. 2008)

user trust to other users. It should be noted that, the trust of user u_i in user u_j is not based on the trust of user u_j in user u_i. SocialMF Graphical model is shown in the Fig. 6.

Prior distribution according to the Fig. 6 is as follows: (Jamali 2010; Yang et al. 2014)

$$
\begin{aligned}
P(U, V | R, T, \sigma_R^2, \sigma_U^2, \sigma_T^2, \sigma_V^2) &\propto P(R|U, V, \sigma_R^2) P(U|T, \sigma_U^2, \sigma_T^2) P(V|\sigma_V^2) \\
&\propto P(R|U, V, \sigma_R^2) P(U|T, \sigma_T^2) P(U|\sigma_U^2) P(V|\sigma_V^2) \\
&= \prod_{u=1}^{n} \prod_{i=1}^{m} N\left[(R_{ui}|g(U_u^T V_i), \sigma_R^2) \right]^{I_{u,i}^R} \\
&\times \prod_{u=1}^{n} N(U_u | \sum_{u \in N_u} T_{u,v} U_v, \sigma_T^2 I) \times \prod_{u=1}^{n} N(U_u|0, \sigma_u^2 I) \times \prod_{i=1}^{m} N(V_i|0, \sigma_V^2 I)
\end{aligned}
\tag{9}
$$

In the Social Matrix Factorization (SocialMF) (Jamali 2010), whereas the relationship of users' social trust has done through social trust matrix factorization, but the time sequential data has been ignored between items. This method doesn't consider relationship between the items that is an important factor in deciding the users.

3.4 Social Trust Ensemble (STE)

In matrix factorization and memory-based methods, trust publishing has been widely used in offering a recommendation. Other Recommendations-based Social Networks Matrix factorization method is presented called Social Trust Ensemble (STE) (Ma et al. 2011). This method is a linear combination of matrix factorization method and the social network-based method. Figure 7 shows a graphical model for the proposed method. Predicted rating of the user u for the item i as follows (Yang et al. 2014; Ma et al. 2011).

$$
\hat{R_{ui}} = \alpha U_u^T V_i + (1 - \alpha) \sum_{v \in N_u} T_{uv} U_v^T U_v^T V_i
\tag{10}
$$

The parameter α, controls the impact of neighborhood on the estimated rating. This method is used as a comparative partner in the experiments; therefore it is called Social Trust Ensemble (STE) (Ma et al. 2011).
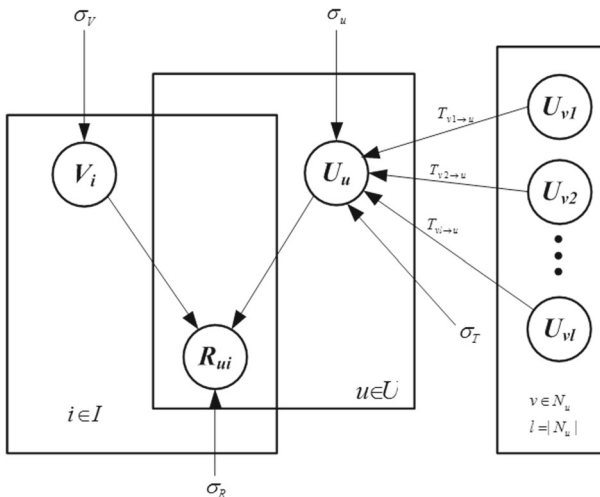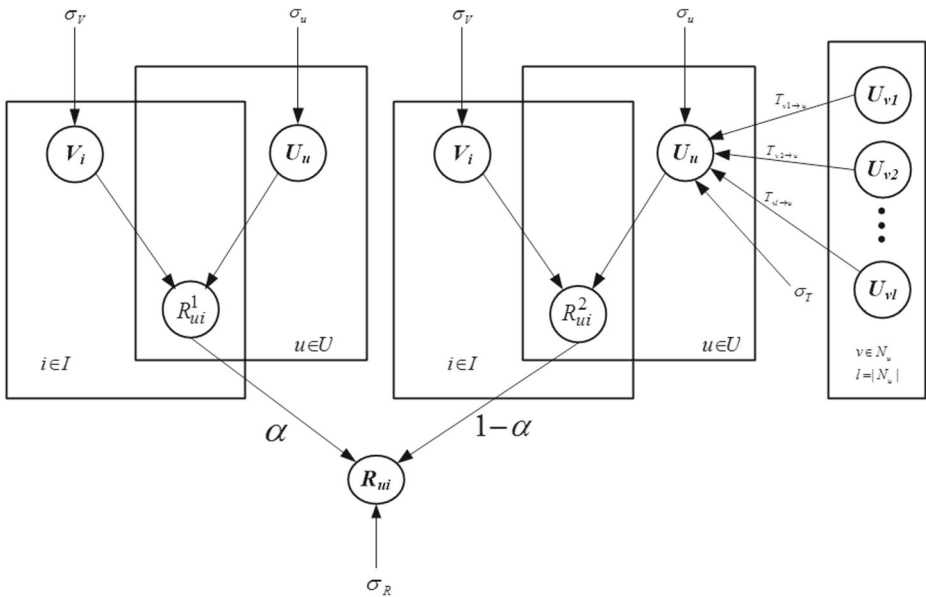

Fig. 6 The graphical model of SocialMF (Jamali 2010)

Fig. 7 The graphical model of Social Trust Ensemble (STE) (Ma et al. 2011)

Social Trust Ensemble (STE) (Ma et al. 2011) with parameter α, combines the effect of the two methods performance. It also controls the effect of Neighborhood on the estimated score. Yu-sheng et al. (2014) and Jamali (2010) who have referred to this method, unfortunately have drawn the Social Trust Ensemble graph incorrectly due to errors in understanding the algorithm and consequently, they have it checked incorrectly in Fig. 8. The criticism related to the misunderstanding of these researchers is that, the direct Neighborhood feature vectors u, have an impact on rating u instead of influencing feature vectors
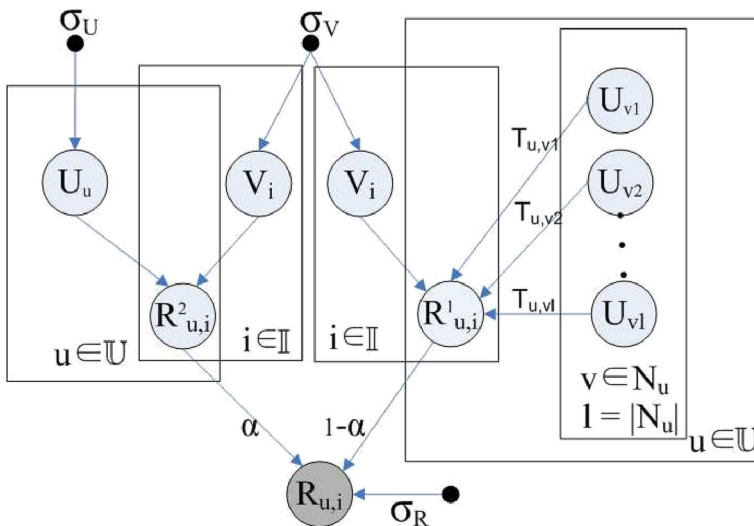

Fig. 8 Incorrect flowchart of STE by Yu-sheng et al. (2014) and Jamali (2010) articles

u. In addition, this model doesn't control the trust publishing. The true image of the graph which has been proposed in the study by HAO in 2011 (Ma et al. 2011) is shown in the Fig. 7. However, this method by integration and the impact of social trust relations has tried to improve the previous methods, but there is not any effect of implicit interest and relationships between the items.

3.5 Interest social recommendation

When information of implicit social relationship is accessible, traditional social recommenders use this information directly. Whereas social relationship information has a lot of reason, such as kinship, colleague and etc. But all of us know recommender systems based on user social relationship means they have similar interest. In the paper (Yu-sheng et al. 2014), if users have the same or similar interests, implicit user's interest communication information is used to obtain social information. The proposed method in this paper, *Interest Social Recommendation (ISoRec)* is used when the information on available user' social communications is available or not. When the users' social communication information is not available, only the implicit user's interest communication information is used for social recommendation. While, when the users' social communication information is available, implicit and explicit communication information is used at the same temporal. The matrix factorization for user-item matrix factorization, the social communication matrix, and implicit interest's matrix is used to learn the users' feature of the user-item matrix. The matrix factorization extracts L-dimensional features matrix of the users U_i and items V_j based on the analysis of matrices R, S and C. The latent user feature matrix U and latent item feature matrix V are used to make the results of recommendations (Yu-sheng et al. 2014). $U \in R^{l \times m}$ and $Z \in R^{l \times m}$ are the latent user and decomposition feature matrix C_{uk} with column vectors U_u and Z_k which indicate user -specific and feature-specific latent feature vectors (Yang et al. 2014; Ma et al. 2008).

A graphical model describing of Fig. 9 shows the impact of social communication and explicit user' interests on the user's judgment to choose the item (Yu-sheng et al. 2014). Similarly, what was said in the previous methods, the conditional distribution on the element observed in related to explicit users social communication of matrix S and implicit user's interests social communication matrix C is defined as follows:

$$P(S|U, W, \sigma_s^2) = \prod_{i=1}^{m} \prod_{k=1}^{m} N\left[(S_{ik}|g(U_i^T W_k), \sigma_S^2)\right]^{I_{ik}^S}$$
$$P(C|U, Z, \sigma_C^2) = \prod_{i=1}^{m} \prod_{t=1}^{m} N\left[(C_{it}|g(U_i^T Z_t), \sigma_C^2)\right]^{I_{it}^C} \qquad (11)$$

With calculation of mean zero Gaussian, the Bayesian inference is used for calculating posterior distribution U and W in (12) and U and Z in (13):

$$P(U, W|R, \sigma_U^2, \sigma_W^2, \sigma_S^2) \propto P(S|U, W, \sigma_S^2)P(U|\sigma_U^2)p(W|\sigma_W^2)$$
$$= \prod_{i=1}^{m} \prod_{k=1}^{m} \left[N(S_{i,k}|g(U_i^T W_K), \sigma_S^2)\right]^{I_{ik}^S} \times \prod_{i=1}^{m} N(U_i|0, \sigma_U^2 I) \times \prod_{k=1}^{m} N(W_k|0, \sigma_W^2 I) \qquad (12)$$

$$P(U, Z|C, \sigma_U^2, \sigma_Z^2, \sigma_C^2) \propto P(R|U, Z, \sigma_C^2)P(U|\sigma_U^2)p(Z|\sigma_Z^2)$$
$$= \prod_{i=1}^{m} \prod_{t=1}^{m} \left[N(C_{i,t}|g(U_i^T Z_t), \sigma_C^2)\right]^{I_{it}^C} \times \prod_{i=1}^{m} N(U_i|0, \sigma_U^2 I) \times \prod_{t=1}^{m} N(Z_t|0, \sigma_Z^2 I) \qquad (13)$$
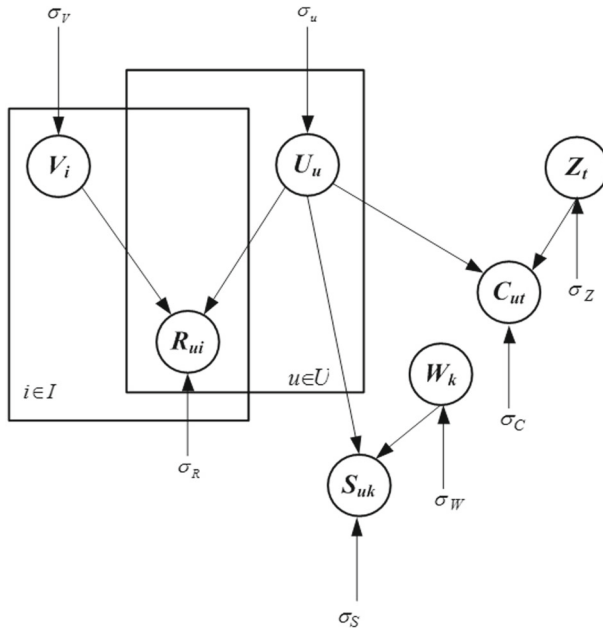
Fig. 9 The graphical model of Interest Social Recommendation (ISoRec) (Yu-sheng et al. 2014)

Maximizing the logarithm of function in above is equal to minimize the objective functions of sum of *Root Mean Squared Error (RMSE)*. The objective function can be minimized by implementing gradient descent on *V*, *U* and *Z* in the (14).

$$
\left.
\begin{aligned}
\frac{\partial L}{\partial U_u} &= \sum_{i=1}^{m} I_{ui}^{R} V_i \acute{g}(U_u^T V_i)(g(U_u^T V_i) - R_{ui}) + \lambda_U U_u \\
&\quad + \lambda_C \sum_{t=1}^{m} I_{ut}^{C} Z_t \acute{g}(U_u^T Z_t)(g(U_u^T Z_t) - C_{ut}) \\
&\quad + \lambda_S \sum_{k=1}^{m} I_{uk}^{S} W_k \acute{g}(U_u^T W_k)(g(U_u^T W_k) - S_{uk}) \\
\frac{\partial L}{\partial V_i} &= \sum_{u=1}^{m} I_{ui}^{R} U_u \acute{g}(U_u^T V_i)(g(U_u^T V_i) - R_{ui}) + \lambda_V V_i \\
&\quad + \lambda_C \sum_{t=1}^{m} I_{ti}^{C} Z_t \acute{g}(Z_t^T V_i)(g(Z_t^T V_i) - C_{ti}) \\
\frac{\partial L}{\partial W_k} &= \lambda_S \sum_{u=1}^{m} I_{uk}^{S} U_u \acute{g}(U_u^T W_k)(g(U_u^T W_k) - S_{uk}) + \lambda_W W_k \\
\frac{\partial L}{\partial Z_t} &= \lambda_C \sum_{u=1}^{m} I_{ut}^{C} U_u \acute{g}(U_u^T Z_t)(g(U_u^T Z_t) - C_{ut}) + \lambda_Z Z_t
\end{aligned}
\right\} \tag{14}
$$

The proposed method by Li Yu-sheng et al, *Interest Social Recommendation model (ISoRec)*, considers only explicit and implicit users' communications. Whereas, indecision-making, the relationship between items and influence of their relationships aren't considered, as well as trusted relationships between users in the social network.

Fig. 10 Network graph of item consumption
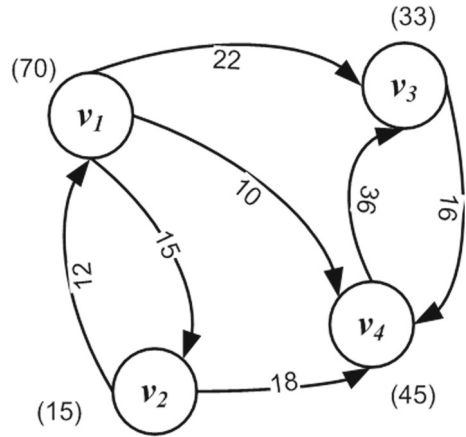
3.6 Trust Sequence Matrix Factorization recommendation (TrustSeqMF)

The proposed method by Zhang et al called "*Trust Sequence matrix factorization social recommendations model (TrustSeqMF)*" (Zhang and Liu 2015). It used time sequential information and trust publishing in the probability matrix factorization-based rating matrix model for integrated social networks recommendations. In this method, in addition to the users' trust network Fig. 2, items consumption network is also considered. In items' consumption network graph, the number in the parentheses indicates the number of users who use the item v_i as in Fig. 10. Each edge has a relevant weight $w_{i \to j}$, so that specifies how many users have used the items v_i and v_j at a same temporal.

Figure 11 shows a scenario of the real-world recommendations including three central elements, users' social trust network, item consumption network and user-item rating matrix. In this scenario, there are four relationships between 4 users (u_1 to u_4) and a relationships between two items (v_1 to v_2).

Prediction of the missing values $r_{u,i}$,(the rating of the user u to the item i) through the user-item rating matrix R, trust matrix T and items consumption matrix S, are the task of recommenders. The graphical model which matched on the (14) is shown in the Fig. 12,
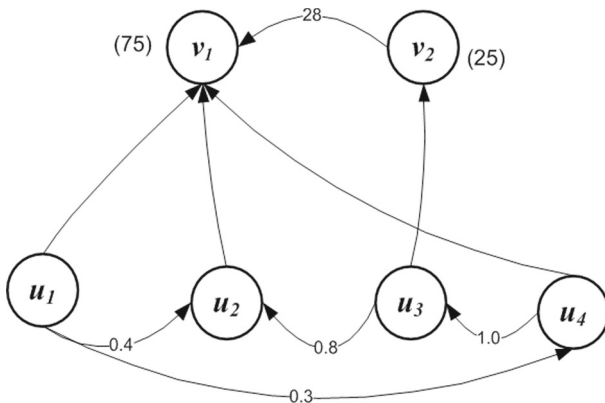

Fig. 11 The graphical model of item consumption and users trust

so that trust publishing combines information of the time sequential of users' behavior and user-item rating matrix.

The prior probability of latent feature vectors is as follows:

$$
\begin{aligned}
&P(U, V | R, T, S, \sigma_R^2, \sigma_T^2, \sigma_S^2, \sigma_U^2, \sigma_V^2) \\
&\propto P(R|U, V, \sigma_R^2) P(U|T, \sigma_U^2, \sigma_T^2) P(V|S, \sigma_V^2, \sigma_S^2) \\
&= \prod_{u=1}^{n} \prod_{i=1}^{m} N\left[(R_{ui}|g(U_u^T V_i), \sigma_R^2)\right]^{I_{ui}^R} \\
&\times \prod_{u=1}^{n} N(U_u | \sum_{u \in N_u} T_{u,v} U_v, \sigma_T^2 I) \\
&\times \prod_{i=1}^{m} N(v_i | \sum_{j \in N_v} S_{j \to i} V_j, \sigma_S^2 I) \\
&\times \prod_{u=1}^{n} N(U_u|0, \sigma_u^2 I) \times \prod_{i=1}^{m} N(V_i|0, \sigma_V^2 I)
\end{aligned}
\tag{15}
$$

Based on Fig. 12, the prior distribution logarithm is calculated. Then the following objective function must be minimized to maximize the mentioned logarithm.

$$
\begin{aligned}
L(R, T, S, U, V) &= \frac{1}{2} \sum_{u=1}^{n} \sum_{i=1}^{m} I_{ui}^R (R_{ui} - g(U_u^T V_i))^2 + \frac{\lambda_U}{2} \sum_{u=1}^{n} U_u^T U_u \\
&+ \frac{\lambda_V}{2} \sum_{i=1}^{m} V_i^T V_i + \frac{\lambda_V}{2} \sum_{u=1}^{n} ((U_u - \sum_{v \in N_U} T_{uv} U_v)^T (U_u - \sum_{v \in N_U} T_{uv} U_v)) \\
&+ \frac{\lambda_S}{2} \sum_{i=1}^{m} ((V_i - \sum_{j \in N_i} S_{j \to i} V_j)^T (V_i - \sum_{j \in N_i} S_{j \to i} V_j))
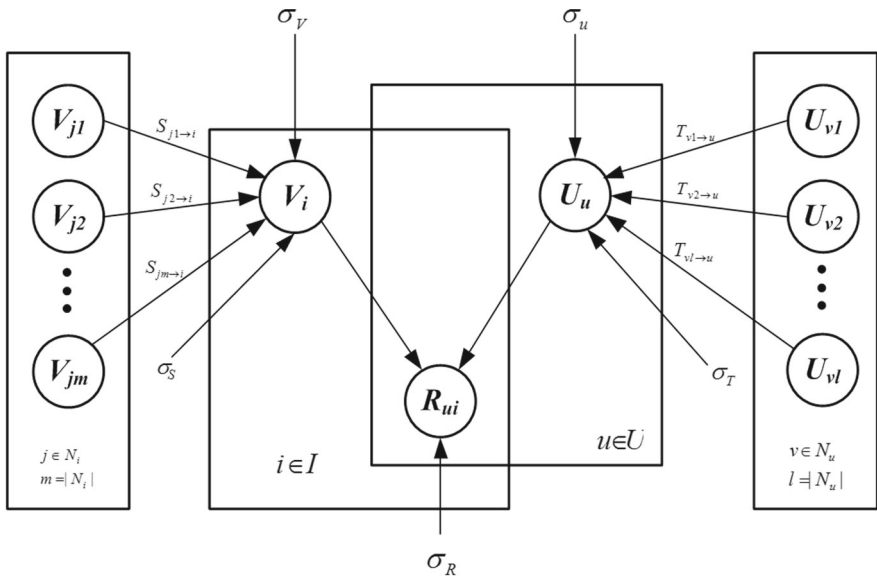\end{aligned}
\tag{16}
$$

Fig. 12 The graphical model of TrustSeqMF (Zhang and Liu 2015)

So that, $\lambda_U = \sigma_R^2/\sigma_U^2$, $\lambda_V = \sigma_R^2/\sigma_V^2$, $\lambda_T = \sigma_R^2/\sigma_T^2$, $\lambda_S = \sigma_R^2/\sigma_S^2$ and by implementing gradient descent on the U_u and V_i, the local minimum of the objective function in the (16), can be found.

$$\frac{\partial L}{\partial U_u} = \sum_{i=1}^{m} I_{ui}^R V_i \acute{g}(U_u^T V_i)(g(U_u^T V_i) - R_{ui}) + \lambda_U U_u$$

$$+ \lambda_T (U_u - \sum_{v \in N_u} T_{uv} U_v) - \lambda_T \sum_{\{v|u \in N_v\}} T_{uv}(U_v - \sum_{w \in N_v} T_{wv} U_w) \qquad (17)$$

$$\frac{\partial L}{\partial V_i} = \sum_{u=1}^{n} I_{ui}^R U_u \acute{g}(U_u^T V_i)(g(U_u^T V_i) - R_{ui}) + \lambda_V V_i$$

$$+ \lambda_S (V_i - \sum_{j \in N_i} S_{j \to i} V_j) - \lambda_S \sum_{\{j|i \in N_j\}} T_{ij} - (V_j - \sum_{k \in N_j} S_{k \to j} V_k) \qquad (18)$$

$\acute{g}(x) = \frac{exp(x)}{(1+exp(x))^2}/$ is derivative of logical function $g(x)$.

Trust Sequence Matrix Factorization (TrustSeqMF) (Zhang and Liu 2015), in 2015, was introduced by Zhang and et al where in addition to considering the relationship of between users' trust, relationships between items are also considered. They didn't consider implicit users' interests in their equation. In order to complete and study the challenge, we propose a new method, In addition to the impact of trusted relationships between users and the relationships between items, implicit users' relationships are also considered. In other words, In addition to the impact of the time sequential in the recommender system, Non-explicit relationship also will be effective.

Prediction of users' rating to items in social recommender systems is closer to the user idea, is ever a challenge. According to studied methods in this section, providing a system which might solve this problem, is very important. This method will be discussed in the next chapter.

4 The proposed model

The concept of "social recommender systems", which are mentioned in this paper, is defined in order to use social network information. It includes comments from friends of a person on social networks to improve the recommendations. However, a lot of studies has been done in matrix factorization-based social recommender systems. Most of them have focused on social matrix factorization (SocialMF) and Social Trust Ensemble (STE), trust sequence matrix factorization social recommendations (TrustSeqMF) on the users' trust. Even some models such as TrustSeqMF are based on integration of the users' trust and relationships between items, but these methods aren't sufficiently accurate. In the presented paper, in addition to considering the relations between users and items, implicit users' interests are also taken into account. Using the implicit user's interests causes more real-world conditions be involved in decision-making and this increases the accuracy of decision-making in recommender systems. Matrix factorization is used to perform this operation.

4.1 The graphical representation

In the proposed method, we use matrix factorization related to the impact of the availability of social communication by using social trust sequence method. Hence, the name of this
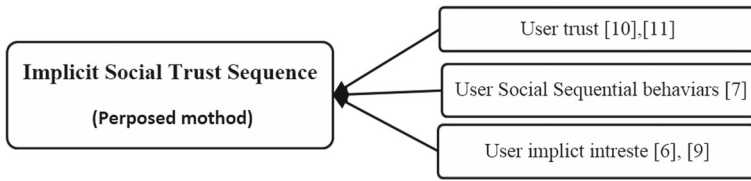
Fig. 13 The component of Implicit Social Trust Sequence (ISoTrustSeq) method

technique is Implicit Social Trust Sequence (ISoTrustSeq). Figure 13 shows a graphical representation of the fundamental elements of this method.

Figure 14 shows a graphical representation of the proposed method.

The conditional probability distribution according to the graph Fig. 14 which includes the impact of a sequence of user behaviors, users' trust, and implicit user's interest are as the follows:

$$P(U, V, Z | R, T, S, C, \sigma_R^2, \sigma_T^2, \sigma_S^2, \sigma_C^2, \sigma_U^2, \sigma_V^2) \propto$$
$$P(R|U, V, \sigma_R^2) P(U|T, \sigma_U^2, \sigma_T^2) P(V|S, \sigma_V^2, \sigma_S^2) P(C|U, Z, \sigma_C^2) \tag{19}$$

4.2 The ISoTrustSeq algorithm

At first, we select user-item rating matrix, and users' trust matrix from data set. In the matrix R, we choose some ratings of users as test data, and as part of training data. Test data from user-item rating matrix will be equal to zero and other data as training data remain unchanged. Then we obtain the primitive users' latent features initial matrix U and items' latent features matrix V from item-user rating matrix by using singular value decomposition
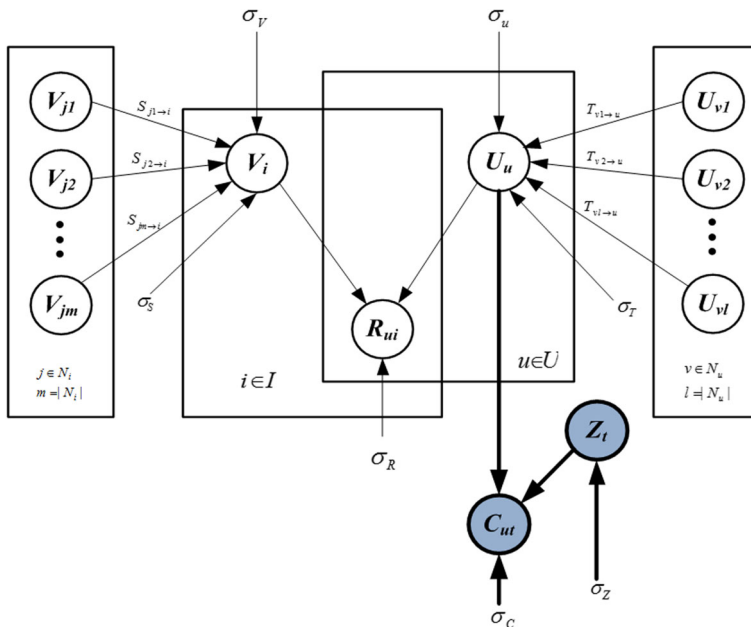

Fig. 14 The graphical model of recommended system ISoTrustSeq

(SVD) method. We make implicit user' relationship matrix C, and the relationships between items matrix S. Now, after derivative of the essential variables of the problem from objective function, the values of gradient descent will achieve. By setting the values in the (36) which is a representation of gradient descent, the new values U, V and Z are obtained using the (37) and the values of matrices $\hat{R}$ and $\hat{C}$. By updating the values of R, Z, V, U and C with the new values, root mean square error is calculated. If the error rate is less than 0.001, this means that the experimental values are close to their original values and the iteration process stop. Otherwise, the procedure is repeated. Figure 15 shows the procedure of the proposed method in the flowchart. The smallest number of features that make slight changes to the system consider as the best number of output so that latent features K are for the next step.

The conditional probability distribution of *ISoTrustSeq* algorithm in (12) contains four concepts A to D as following:

4.3 A: user-item rating matrix

We suppose that a recommender system has N users and M items (Zhang and Liu 2014). $r_{u,i}$ is a rating of user u to the item i. The maximum value of this rating is 5 and includes integer values (Ma et al. 2008). This amount represents the amount of user's interest to the desired item (5, 4, 3, 2, and 1, indicate "very interested", "interested", "Moderate ", "lack of interest", and "hate"). Figure 16 shows the user-item rating Matrix.

The users' set is shown $U = \{u_1, u_2, u_N\}$ and the items' set as $V = \{v_1, v_2, v_M\}$. In other words, $U \in R_{N \times K}$ and $V \in R_{M \times K}$ are users and item latent features matrix. It respectively includes row vector U_u, user-feature vector and V_i a k-dimensional item-feature vector. The
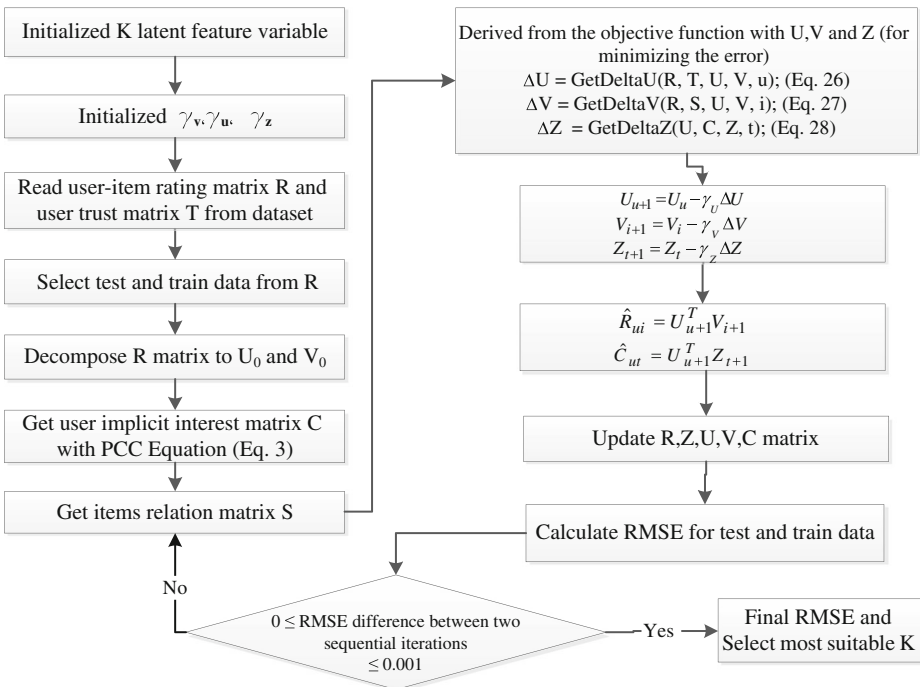

Fig. 15 The flowchart of ISoTrustSeq algorithm

Fig. 16 User-Item rating matrix

	V1	V2	V3
u_1	3	?	?
u_2	?	2	?
u_3	4	?	1
u_4	5	?	?
u_5	?	?	3

values of K is less than the number of users, N and the number of the items, M (Jamali 2010). In fact, K indicates the order of feature matrix. U and V Mean-zero Gaussian on the user-item feature vectors will be as follows:

$$P(U|\sigma_U^2) = \prod_{u=1}^{n} N(U_u|0, \sigma_U^2 I) \tag{20}$$

$$P(V|\sigma_V^2) = \prod_{i=1}^{m} N(V_i|0, \sigma_V^2 I) \tag{21}$$

The objective of using matrix factorization is learning features variables and extraction of them to purpose the recommendations (Qian et al. 2014). Conditional distribution of user-item rating matrix R is defined as follows:

$$P(R|U, V, \sigma_R^2) = \prod_{u=1}^{n} \prod_{i=1}^{m} \left[N(R_{u,i}|g(U_u^T V_i), \sigma_R^2) \right]^{I_{u,i}^R} \tag{22}$$

I_{ij}^R is index variable. Thus, if the user i rates to the item j (row i and column j of the matrix R is a non-zero value), the value of the index variable is equal to one, otherwise zero. $N(x|\mu, \sigma_R^2)$ represents the Gaussian distribution with the average μ and variance σ_R^2. The logical function is $g(x) = 1/(1 + exp(-x))$, so that the values of $U_u^T V_i$ are mapped to interval [0,1]. By substituting a Gaussian normal distribution (Han et al. 2006), in the (22), we have:

$$
\begin{aligned}
P(R|U, V, \sigma_R^2) &= \prod_{u=1}^{n} \prod_{i=1}^{m} \left[N(R_{u,i}|g(U_u^T V_i), \sigma_R^2) \right]^{I_{u,i}^R} \\
&= \prod_{u=1}^{n} \prod_{i=1}^{m} \left[\frac{1}{\sqrt{2\pi\sigma_R^2}} e^{\frac{-(R_{u,i} - g(U_u^T V_i))^2}{2\sigma_R^2}} \right]^{I_{u,i}^R}
\end{aligned}
\tag{23}
$$

4.4 B: users' trust matrix

We assume the Social network graph as $G = (U \cup V, E)$. All vertices $U = \{usr_u\}_{u=1}^{N}$ shows all users on the social network. Sum of vertices $V = \{v_i\}_{i=1}^{M}$ shows all items on the social network and the set of edges E shows the relationship between users and items (Zhang and Liu 2015). Figure 13 shows a scenario of the real world recommendations that includes three

central elements: users' social trust network, items' consumption network, and user-item rating matrix. Figure 17 shows the items' consumption network graph and trust between users. Each node represents a user and each edge represents the trusted relationship between them. As seen in the figure, the graph of trust between users consists of six relationships between five users (u_1 to $u_5 \in U$). $T = \{t_{ij}\}$ shows a $N \times N$ matrix of the user's trust graph and are called social trust matrix. Every edge which is related to the weight t_{ij} in interval $(0, 1]$ is the value of social trust that shown as u_i or u_j. Zero indicates a lack of trust and 1 indicates complete trust. T is an asymmetric matrix, ie the trust of user u_1 in user u_2 is not equal to the trust of user u_2 in user u_1 in the trust social network (Jamali 2010).

Figure 17 user u1 as much as 0.4 trusts in the user u_2, while the user u_2 doesn't trust in the user u_1. According to this description raised, the conditional probability of user features and latent features of its trusted friends are defined as follows:

$$P(U|T, \sigma_U^2, \sigma_T^2) \propto P(U|T, \sigma_T^2)P(U|\sigma_U^2)$$
$$= \prod_{u=1}^{n} N\left(U_u| \sum_{u \in N_u} T_{u,v}U_v, \sigma_T^2 I\right) \times \prod_{u=1}^{n} N(U_u|0, \sigma_U^2 I) \tag{24}$$

By substituting the normal Gaussian distribution, the (25) is by the follows:

$$P(U|T, \sigma_U^2, \sigma_T^2) \propto P(U|T, \sigma_T^2)P(U|\sigma_U^2)$$
$$= \prod_{u=1}^{n} N\left(U_u| \sum_{u \in N_u} T_{u,v}U_v, \sigma_T^2 I\right) \times \prod_{u=1}^{n} N(U_u|0, \sigma_U^2 I)$$
$$= \prod_{u=1}^{n} \frac{1}{\sqrt{2\pi\sigma_T^2}} e^{\frac{-(U_u - \sum_{u \in N_u} T_{u,v}U_v)^2}{2\sigma_T^2}} \tag{25}$$

4.5 C: users' sequential behaviors matrix (relations between the items)

In the items' consumption graph (Fig. 17), there are three relations between three available items in a social network (v_1 to $v_3 \in V$). The number in parentheses close to the node v_i indicates the number of users who use the item v_i (Zhang and Liu 2015). The weight of each
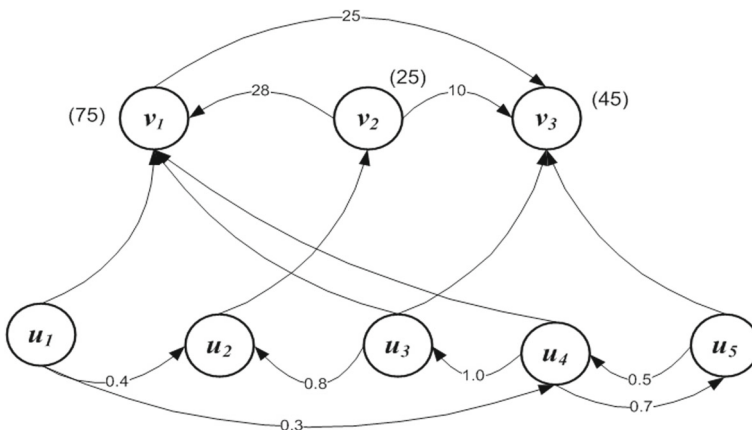

Fig. 17 User trust and item consumption graph

edge $w_{i \to j}$ specifies how many users have used the item v_i and v_j at the same temporal. If item v_i and v_j are used by the same user at a temporal, the weight of $w_{i \to j}$ increases equal to a unit (Zhang and Liu 2015).

According to the consumption network graph, the weight of the effective equation on the items is calculated through the following equation:

$$S_{i \to j} = \frac{W_{i \to j}}{g(v_i, v_j)} \tag{26}$$

$g(v_i, v_j)$ represents a set of users who use the item A or B at a certain temporal, and $S_{i \to j}$ represents the impact of item v_i on the v_j. Matrix S is the item consumption Matrix, which is asymmetric.

In this proposed method, the task of recommender system is to predict missing values $r_{u,i}$ of the user-item rating matrix through a trust matrix T, item consumption matrix S and implicit interests in the social network. Conditional Probability of the latent features of the items and its neighbors are defined as follows:

$$
\begin{aligned}
p(V|S, \sigma_V^2, \sigma_S^2) &\propto P(V|S, \sigma_S^2) P(V|\sigma_V^2) \\
&= \prod_{i=1}^{m} N \left(V_i \Big| \sum_{j \in N_v} S_{j \to i} V_j, \sigma_S^2 I \right) \times \prod_{i=1}^{m} N(V_i|0, \sigma_V^2 I)
\end{aligned}
\tag{27}
$$

By substituting the normal Gaussian distribution, the (28) is by the follows:

$$
\begin{aligned}
p(V|S, \sigma_V^2, \sigma_S^2) &\propto P(V|S, \sigma_S^2) P(V|\sigma_V^2) \\
&= \prod_{i=1}^{m} N \left(V_i \Big| \sum_{j \in N_v} S_{j \to i} V_j, \sigma_S^2 I \right) \times \prod_{i=1}^{m} N(V_i|0, \sigma_V^2 I) \\
&= \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi \sigma_S^2}} e^{\frac{-(V_i - \sum_{j \in N_v} S_{j \to i} V_j)^2}{2\sigma_S^2}} \times \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi \sigma_V^2}} e^{\frac{-(V_i - 0)^2}{2\sigma_V^2}}
\end{aligned}
\tag{28}
$$

4.6 D: users implicit interest

There are many methods to find similarities between two users. *Pearson Correlation coefficient (PCC)* determines the similarity between two users (2). After calculating all the similarities between pairs of users, matrix similarity is made between users. The similarity matrix is used to make the implicit user's interest matrix C (Yu-sheng et al. 2014; Way 2013).

So that the element C_{ij} is $C_{ij} = Sim(i, j)$, because most of the similarities between users indicate the most fun and interest shared between users in social networks (Zhang and Liu 2015). In this study, Pearson Correlation coefficient which is one of the most common methods for calculating the similarity between users is used to obtain the implicit users' interests (Yang et al. 2013, 2014; Yu-sheng et al. 2014; Way 2013). Conditional distribution of the elements observed in the user's social communication matrix S and the implicit users' interest of social communication matrix is defined as follows:

$$P(C|U, Z, \sigma_C^2) = \prod_{i=1}^{m} \prod_{t=1}^{m} N[(C_{it}|g_(U_i^T Z_t), \sigma_C^2)]^{I_{it}^C} \tag{29}$$

I_{ik}^C is an index function. If the user i, knows the User k or trusts in him, its value is equal to 1, otherwise, it is equal to zero. Gaussian distribution probability density function N and

logical function g are as mentioned in the previous method. By substituting a Gaussian normal distribution, the probability equation is:

$$P(C|U, Z, \sigma_C^2) = \prod_{u=1}^{n}\prod_{t=1}^{n} N\left[(C_{ut}|g(U_u^T Z_t), \sigma_C^2)\right]^{I_{ut}^C}$$

$$= \prod_{i=1}^{m}\prod_{k=1}^{m}\left[\frac{1}{\sqrt{2\pi\sigma_C^2}}e^{\frac{-(C_{ut}-g(U_u^T Z_t))^2}{2\sigma_C^2}}\right]^{I_{ut}^C} \tag{30}$$

For more descriptions of our algorithm, we write the pseudo code of our proposed method.

---

1) Initialized value: $\lambda_U=\lambda_V=\lambda_Z=0.001$ and the learning rate γ equal to 0.001 (In section 5-2-Initial parameters setting.)

2) Read data from dataset.
            R: user-item rating matrix
            T: trust matrix

3) Select test and train data of R matrix

4) C: similarity between user u_i and user u_j means C = PearsonCorrelationCoefficient(R) (In Section 2-3-Similarity function (equation. 2)).

5) First assignment of U and V with Singular Value Decomposition SVD(R)

6) $S : S_{ji}$ denotes which item v_i effects on item v_j
(In part C-Users' Sequential behaviors matrix (equation. 26))

7) Calculate ΔU = GetDeltaU(R, T, U, V, u); (equation. 33)

8) Calculate ΔV = GetDeltaV(R, S, U, V, i); (equation. 34)

9) Calculate ΔZ = GetDeltaZ(U, C, Z, t); (equation. 35)

10) Used descent gradient for updating new value of U, V, Z
        $U_{u+1} = U_u - \gamma_U \Delta U$
        $V_{i+1} = V_i - \gamma_V \Delta V$
        $Z_{t+1} = Z_t - \gamma_Z \Delta Z$
γ values is initialized in step 1 : $\gamma_U = \gamma_V = \gamma_Z = 0.001$

11) Calculate new R ($\hat{R}$) and C matrix with new U, V, Z
        $\hat{R}_{ui} = U_{u+1}^T V_{i+1}$
        $\hat{C}_{ut} = U_{u+1}^T Z_{t+1}$

12) Calculates RMSE value between new R ($\hat{R}$) (equation. 38)

13) If $0 \leq RMSE_2 - RMSE_1 \leq 0.001$
        go to step 6;
else
        The most suitable value k is selected (means the predicted value is closer to the real value)
end

In this pseudo code, some part of algorithm like preparing data, and the stopping condition is shown clearly. The details of every equation is described in the formula definition.

4.7 The objective function

As mentioned above, we proposed the conditional probability distribution in order to reflect the impact of the relationship between implicit user' interests and social trust of the users

on their judgment in relation to the items. First, the logarithm of this equation is calculated in order to perform differentiation operations and then after extending, reach to the (31):

$$
\begin{aligned}
&\ln P(U, V, Z | R, T, S, C, \sigma_R^2, \sigma_T^2, \sigma_S^2, \sigma_C^2, \sigma_U^2, \sigma_V^2) \\
&= -\frac{1}{2\sigma_R^2} \sum_{u=1}^{n} \sum_{i=1}^{m} I_{u,i}^R (R_{u,i} - g(U_u^T V_i))^2 \\
&\quad -\frac{1}{2\sigma_C^2} \sum_{u=1}^{n} \sum_{t=1}^{n} I_{u,t}^C (C_{u,t} - g(U_u^T Z_t))^2 \\
&\quad -\frac{1}{2\sigma_U^2} \sum_{u=1}^{n} U_u^2 - \frac{1}{2\sigma_V^2} \sum_{i=1}^{m} V_i^2 \\
&\quad -\frac{1}{2\sigma_T^2} \sum_{u=1}^{n} (U_u - \sum_{u \in N_u} T_{u,v} U_v)^2 \\
&\quad -\frac{1}{2\sigma_S^2} \sum_{i=1}^{m} (V_i - \sum_{j \in N_v} S_{j \to i} V_j)^2 \\
&\quad -\frac{1}{2} \sum_{u=1}^{n} \sum_{i=1}^{m} I_{u,i}^R \ln (2\pi\sigma_R^2) - \frac{1}{2} \sum_{u=1}^{n} \sum_{t=1}^{n} I_{u,t}^C \ln (2\pi\sigma_C^2) \\
&\quad -\frac{n}{2} [\ln (2\pi\sigma_T^2) + \ln (2\pi\sigma_U^2)] - \frac{m}{2} [\ln (2\pi\sigma_S^2) + \ln (2\pi\sigma_V^2)]
\end{aligned}
\tag{31}
$$

Now, we derive from the (31) of the four last sentence that are without this words V_i, U_u and Z_t, we can remove this form equation. Then, we multiply the resulting equation by σ_R^2 and after substituting the values $\lambda_U = \sigma_R^2/\sigma_U^2$, $\lambda_V = \sigma_R^2/\sigma_V^2$, $\lambda_T = \sigma_R^2/\sigma_T^2$, $\lambda_S = \sigma_R^2/\sigma_S^2$ and $\lambda_C = \sigma_R^2/\sigma_C^2$ can reach the following objective function:

$$
\begin{aligned}
E(R, T, S, C, U, V) &= -\frac{1}{2} \sum_{u=1}^{n} \sum_{i=1}^{m} I_{u,i}^R (R_{u,i} - g(U_u^T V_i))^2 \\
&\quad -\frac{\lambda_C}{2} \sum_{u=1}^{n} \sum_{t=1}^{n} I_{u,t}^C (C_{u,t} - g(U_u^T Z_t))^2 - \frac{\lambda_U}{2} \sum_{u=1}^{n} U_u^2 - \frac{\lambda_V}{2} \sum_{i=1}^{m} V_i^2 \\
&\quad -\frac{\lambda_T}{2} \sum_{u=1}^{n} (U_u - \sum_{u \in N_u} T_{u,v} U_v)^2 - \frac{\lambda_S}{2} \sum_{i=1}^{m} (V_i - \sum_{j \in N_v} S_{j \to i} V_j)^2
\end{aligned}
\tag{32}
$$

Maximizing the logarithm function in (31) is equal to minimizing the objective function in (32). Therefore, with the implementation of the gradient descent and differentiation of objective function in terms of V_i, U_u and Z_t, we obtain the following equations:

$$
\begin{aligned}
\frac{\partial E}{\partial U_u} &= \sum_{i=1}^{m} I_{u,i}^R V_i (R_{u,i} - g(U_u^T V_i)) \acute{g}(U_u^T V_i) \\
&\quad -\lambda_C \sum_{t=1}^{n} I_{u,t}^C Z_t (C_{u,t} - g(U_u^T Z_t)) \acute{g}(U_u^T Z_t) \\
&\quad -\lambda_U U_u + \lambda_T (U_u - \sum_{u \in N_u} T_{u,v} U_v) - \lambda_T \sum_{\{v | u \in N_v\}} T_{u,v} (U_v - \sum_{w \in N_u} T_{W,v} U_w)
\end{aligned}
\tag{33}
$$

$$\frac{\partial E}{\partial V_i} = \sum_{u=1}^{n} I_{u,i}^{R} U_u (R_{u,i} - g(U_u^T V_i)) \acute{g}(U_u^T V_i) + \lambda_V V_i$$

$$+ \lambda_S (V_i - \sum_{k \in N_i} S_{k \to i} V_k) - \lambda_S \sum_{\{j|i \in N_j\}} S_{j \to i} (V_j - \sum_{k \in N_i} S_{k \to j} V_k) \qquad (34)$$

$$\frac{\partial E}{\partial Z_t} = \sum_{u=1}^{n} I_{u,i}^{C} U_u (C_{u,t} - g(U_u^T Z_t)) \acute{g}(U_u^T Z_t) \qquad (35)$$

$\acute{g}(x) = exp(x)/(1 + exp(x))^2$ is a derivative of a logical function $g(x)$. N_u represents a set of neighbors of the user u and N_i is a neighborhood set of item i.

After obtaining the values $\Delta U = \frac{\partial E}{\partial U_u}$, $\Delta V = \frac{\partial E}{\partial V_i}$, $\Delta Z = \frac{\partial E}{\partial Z_t}$ and the new values of V, U and Z can be obtained by (36) (Oh and Han 2015).

$$\left. \begin{array}{l} U_{u+1} = U_u - \gamma_U \Delta U \\ V_{i+1} = V_i - \gamma_V \Delta V \\ Z_{t+1} = Z_t - \gamma_Z \Delta Z \end{array} \right\} \qquad (36)$$

According to singular value decomposition concepts, the new values of matrices C and R can be obtained by multiplying updated values of V, U and Z according to the (37).

$$\left. \begin{array}{l} \hat{R}_{ui} = U_{u+1}^T V_{i+1} \\ \hat{C}_{ut} = U_{u+1}^T Z_{t+1} \end{array} \right\} \qquad (37)$$

In the next section, the results of this proposed method on data set and evaluation of this method with two methods ISoRec (Yu-sheng et al. 2014) and TrustSeqMF (Zhang and Liu 2015) will be examined.

5 Experimental results

Used dataset in this study is Epinions. Epinions.com is a very popular review site that was established in 1999. Users can provide their personal opinions on the topics covered in the site, such as books, music, movies, software and other products. Users' rating values include integers from one to five which not only shows the user interest to the desired item According to type of data, the value but will affect the decision making by other customers.

Epinions dataset contains 22,166 different users who have rated to 27 categories of items. In addition to rating to available items by each user, every user of Epinions keeps a trust list which shows the trusted relationships between users. Epinions dataset rating matrix contains four columns. The content of this column includes the user ID number, item code, the user-item rating and the temporal of rating. Another Matrix of this data set has kept information on the users' trust in itself as trust network. This matrix contains two columns, which shows the user with user ID x trusts to the user with user ID y.

For getting implicit user's interests from dataset, we use Pearson Correlation Coefficient (PCC) (2). The explicit communication information is users' Sequential behaviors matrix which shows relations between the items (19).

5.1 Metric

The most common scale to evaluate the accuracy of recommender The most common scale to evaluate the accuracy of recommender systems, use the root mean square error (RMSE),

(Yu-sheng et al. 2014; Zhang and Liu 2015; Ma et al. 2011; Recht 2013; Salakhutdinov and Mnih 2008; Way 2013; Wu et al. 2012; Li et al. 2015; Yang et al. 2013) which is calculated as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{u,i} (R_{u,i} - \hat{R}_{u,i})^2} \tag{38}$$

$R_{u,i}$ is a score given to item i pair (u, i), and $\hat{R}_{u,i}$ is rating predicted of the user u to item i that is achieved through recommender systems algorithms. N is the number of scores and less level of RMSE shows that the rate of rating predicted is close to the actual value and higher level of RMSE shows that there is a difference between the predicted value and the actual value. In other words, whatever the RMSE is less, the accuracy of the desired algorithm will be greater, and vice versa, as whatever the RMSE is higher the accuracy of the algorithm is less.

As mentioned at the beginning of this section, the scale of evaluation is used in the literature that is active in this area. So we used this metric to compare and evaluate the proposed algorithm.

5.2 Initial parameters setting

In ISoRec and TrustSeqMF methods have considered evaluation criteria based on changes in the number of user and items features (K) in different sets of test data. In this study, also we examine the performance of the proposed algorithm compared with two ISoRec and Trust-SeqMF methods by changes in the number of user and items features. In fact, the aim of all experiments was to get the error RMSE to show the impact of implicit interest on the conditions which have considered the relationships between the users and items on the implicit and non-implicit users' interests (Rana and Jain 2014). Also, as fundamental algorithms, we examined the effect of the user and items features on the accuracy of predictions.

Each stage of the experiment, we test λ_U, λ_V and λ_Z parameters with numeric values 0.1, 0.5, 0.8, 0.01, 0.05, 0.08, 0.001, 0.005, 0.008 for the first temporal with the learning rate γ equal to 0.005 (Salakhutdinov and Mnih 2008), anther with learning rate of 0.01 (Way 2013) then with the learning rate of 0.001 (Yu et al. 2013).

By comparing the results of the above values with results on the fundamental algorithms, the values of $\lambda_U = \lambda_V = \lambda_Z = 0.01$ and learning rate equal to 0.001 was used as the initial parameters in the proposed algorithm. According to the type of data, the value of these parameters are different in the various articles, so we decided to achieve the correct amount of these parameters by conducting several experiments so that these parameters be fitted with the studied data.

5.3 Stopping condition

In the stages of implementing the algorithm which the matrices V, U and Z are learning, the algorithm must be stopped in a step of implementation. As is shown in flowchart Fig. 15, with multiple experiments, it was concluded that when the difference between two continuous stages of RMSE, the error obtained from the experiment predicted data is less than 0.001, the convergence of the predicted values will be closer to actual values. In other words, the low difference between RMSE of two continuous stages shows little changes (and even closer to zero) of the predicted values. Therefore, the convergence is intended as a stop condition (Fig. 15).

5.4 Complexity analysis

In the *TrustSeqMF* algorithm, at first, for extraction social relation, the consumption network graph is computed. Then by effective relationship, an item is suggested to a user. In the fact the objective function and its gradient are computed. The total time complexity of computing this gradient is $O(N\bar{t}^2 K + Nl^2 K + M\bar{r}K + Ml^2 K)$. $\bar{r}$ is the number of rating average for every user, $\bar{t}$ is the number of direct neighbors average for every user and l are the number of neighbors that affect the current user. In *ISoTrustSeq* model, the considering of users implicit interest effect into the time complexity. So the time complexity is linear and similar to *TrustSeqMF* and it is $O(N\bar{t}^2 K + Nl^2 K + M\bar{r}K + Ml^2 K + N\rho_C K)$ in which ρ_C is the numbers of nonzero entries in C. As mentioned in the prior section, according to consider more reality condition due to pervious modes, accuracy of computations are increased and we get an accurate response with less iteration compared with *ISoRec* and *TrustSeqMF*. Although, the time complexity of *ISoTrustSeq* method increase in one iteration by adding implicit interest, but due to the significant decrease of iteration in *ISoTrustSeq* the ultimate time complexity is reduced.

5.5 Experimental results

First a part of the rating matrix scores on the training data (%99, %80, %50, %20, %10 (Yu-sheng et al. 2014)) is considered. %80 of the rating means that from all users rating value, %80 of them predicted training data and the remaining %20 as test data. During the test process, the predicted test data will be compared and the error rate RMSE is calculated for them. The values considered for different K values include 2, 5, 10, 15, 20, 27 (Yu-sheng et al. 2014; Zhang and Liu 2015).

To compare the proposed method *ISoTrustSeq* with *ISoRec* and *TrustSeqMF* methods, we perform the experiment on each algorithms with training data %10, %20, %50, %80 and %99 with different features $K \in \{2, 5, 10, 15, 20, 27\}$ (Yu-sheng et al. 2014; Zhang and Liu 2015). The results of experiments are proposed in Table 1 with $K \in \{2, 5\}$ values, in Table 2 with $K \in \{10, 15\}$ values and in Table 3 with $K \in \{20, 27\}$ values.

The columns of this Table 1 include three algorithms implemented; the proposed algorithm *ISoTrustSeq* and two algorithms evaluated *ISoRec* and *TrustSeqMF* with the number of different features K. Since the columns of the tables show the algorithms above with a number of different features. The rows of the two tables include the percentage of data that are used for learning, which how to adopt training data is explained at the beginning of this section. The content of each of the table cells represents the root mean square error (RMSE).

Figures 18 and 19 show the values of Tables 1, 2 and 3 graphically, in order to make the ease comparison between the different models. These diagrams demonstrate an overview of evaluating the proposed method with some of basic methods with various tests data. The vertical axis shows error rate RMSE and the horizontal axis shows the proposed algorithm and three evaluation algorithms with the number of different user and items features. Both of Figs. 18 and 19 show that RMSE error rate is reduced by increasing the number of training data. Also, in all experiments with number of the specified user and item features K, the proposed algorithm *ISoTrustSeq* has less error rate compared with other algorithms.

5.6 Conclusion and future work

Implicit Social Trust Sequence (ISoTrustSeq) method is presented to predict missing ratings of the users in the user-item rating matrix. In the beginning, we expressed that

Table 1 The comparison of proposed method (ISoTrustSeq) with ISoRec, TrustSeqMF and SocialMF $K \in \{2, 5\}$

Training data	K=2				K=5			
	ISoTrustSeq	ISoRec [6]	TrustSeqMF [7]	SocialMF [11]	ISoTrustSeq	ISoRec [6]	TrustSeqMF [7]	SocialMF [11]
99%	0.7502	0.7537	0.7728	0.9312	0.7471	0.7507	0.7796	0.9245
80%	0.7617	0.7643	0.7880	0.9401	0.7732	0.7790	0.7830	0.9382
50%	0.7726	0.7776	0.7883	0.9523	0.7734	0.7855	0.7879	0.9501
20%	0.7749	0.7856	0.7957	0.9684	0.7775	0.7871	0.7931	0.9599
10%	0.7844	0.7912	0.7962	0.9753	0.7786	0.7902	0.7957	0.9725

Table 2 The comparison of proposed method (ISoTrustSeq) with ISoRec, TrustSeqMF and SocialMF $K \in \{10, 15\}$

Training data	K=10				K=15			
	ISoTrustSeq	ISoRec [6]	TrustSeqMF [7]	SocialMF [11]	ISoTrustSeq	ISoRec [6]	TrustSeqMF [7]	SocialMF [11]
99%	0.7495	0.7517	0.7615	0.9056	0.7367	0.7472	0.7506	0.9001
80%	0.7546	0.7654	0.7676	0.9114	0.7483	0.7582	0.7538	0.9095
50%	0.7643	0.7658	0.7679	0.9203	0.7509	0.7643	0.7562	0.9174
20%	0.7679	0.7687	0.7695	0.9399	0.7538	0.7649	0.7577	0.9301
10%	0.7709	0.7796	0.7755	0.9578	0.7623	0.7787	0.773	0.9456

Table 3 The comparison of proposed method (ISoTrustSeq) with ISoRec, TrustSeqMF and SocialMF $K \in \{20, 27\}$

| Training data	K=20				K=27			
	ISoTrustSeq	ISoRec [6]	TrustSeqMF [7]	SocialMF [11]	ISoTrustSeq	ISoRec [6]	TrustSeqMF [7]	SocialMF [11]
99%	0.7077	0.7296	0.7467	0.8927	0.7015	0.7149	0.7373	0.8905
80%	0.7499	0.7528	0.7576	0.9033	0.7366	0.7344	0.7439	0.8983
50%	0.7556	0.7564	0.7581	0.9116	0.7547	0.7658	0.7554	0.9035
20%	0.7578	0.7595	0.7717	0.9235	0.7587	0.7679	0.777	0.9187
10%	0.7647	0.7614	0.777	0.9312	0.7727	0.7712	0.7784	0.9253
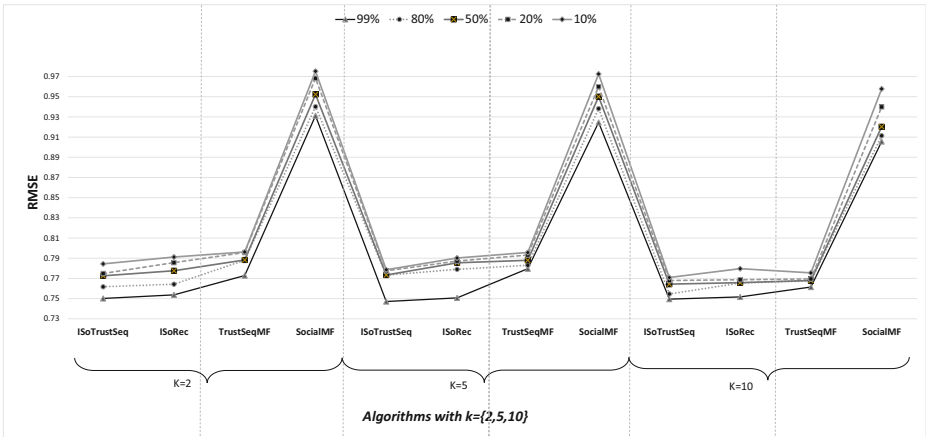
Fig. 18 The overview of the proposed method evaluation with basic methods $K \in \{2, 5, 10\}$

recommender systems attempt to provide the suitable items, in addition to satisfying customers by offering the best things to each person, according to his comments and have adequate profitability for manufactures. Therefore, the aim of this study is to provide a recommender system which can achieve the users, directors and producers' satisfaction by providing some recommendations comply with or close to their views.

A recommender system presented in this study, *ISoTrustSeq*, in addition to considering social trust communication between users, and users' Sequential behaviors, also has considered their implicit interests to solve this problem. The proposed method is trying to solve this problem through the matrix factorization. After obtaining the objective function in accordance with the requirements defined, to update the relevant matrices with gradient descent, the objective function of the basic variables is derived, and the results of derivative after substituting in equations and updating the original variables provide a user-item matrix that has been scored the values that have no rating.
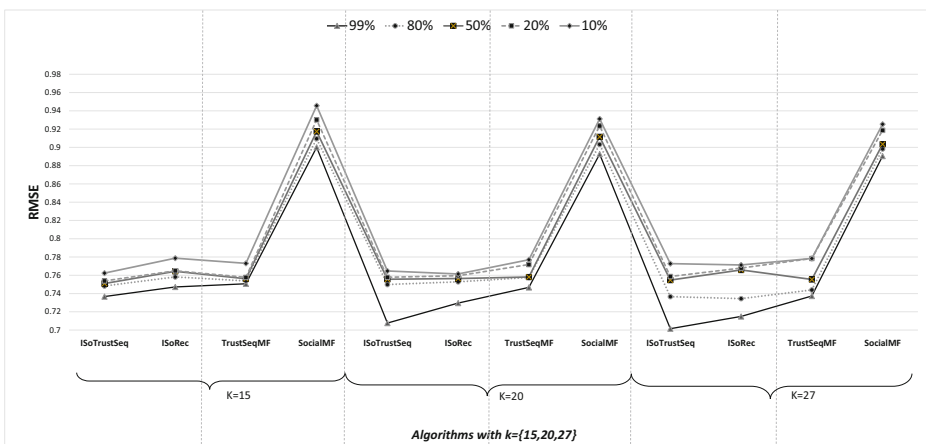

Fig. 19 The overview of the proposed method evaluation with basic methods $K \in \{15, 20, 27\}$

The proposed method has been implemented on Epinions dataset, and the prediction error rate is calculated using RMSE method. The results of the comparisons made with two *ISoRec* and TrustSeqMF methods show that the proposed method increases the prediction accuracy more than other methods. *TrustSeqMF* method also, because of a lack of consideration of the implicit interests compared to *ISoRec* method generates more errors.

The following issues can be mentioned for improving the accuracy of the results of the proposed method in this study:

– As noted, in this study, the effect of trusted relationships between users are examined in the proposed method, and the effect of non-trusted relationship has been ignored. The impact of personal relationships which lack of trust is seen between them can be involved.
– The recommendations are provided in this system based on the implicit interests and sequence of user behavior (choosing the former items) for people who don't establish social trust relationships in this system. To provide the recommendation with high accuracy, for this category of users, first the trusted relationships of these users can be predicted and then the proposed algorithm can be used based on the sequential behavior of users who have similar implicit interests.
– In all the literature in this field, the values of λ_T, λ_S and λ_C are considered the same. The experiments can be performed with different values of these parameters (not equal) to examine the effect of these parameters.

References

Bobadilla, J., Ortega, F., Hernando, A., Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, *46*, 109–132. Elsevier B.V.

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of the International Conference on Computer Statistics*.

Ford, J., Makedon, F., Pearlman, J. (2005). Using singular value decomposition approximation for collaborative filtering. In *7th IEEE International Conference on E-Commerce Technology* (pp. 257–264).

Han, J., Kamber, M., Pei, J. (2006). *Data mining – consepts and techniques, edition S*. San Francisco: Morgan Kaufmann of Elsevier.

Hofmann, T. (2003). Collaborative filtering via gaussian probabilistic latent semantic analysis. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development Informaion Retrieval* (pp. 259–266).

Jamali, M. (2010). A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks Categories and Subject Descriptors. In *The fourth ACM conference on Recommender systems* (pp. 135–142). no. 978–1–60558–906–0.

Kadie, C., Breese, J.S., Heckerman, D. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In *UAI-Fourteenth conference on Uncertainty in artificial intelligence* (pp. 43–52).

Kannan, R., Ishteva, M., Park, H. (2013). Bounded matrix factorization for recommender system. *Knowledge Information Systems*, *39*(3), 491–511.

Koren, Y., Bell, R., Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer (Long Beach California)*, *42*(8), 30–37.

Li, B., Tian, B.-B., Liu, J. (2015). A Novel Probabilistic Matrix Factorization Recommendation Algorithm Using Social Reliable Similarity Propagation. *Intelligence Computing Theory Methodology*, *9226*, 80–91.

Ma, H., Yang, H., Lyu, M.R., King, I. (2008). *SoRec*: Social recommendation using probabilistic matrix factorization. In *Proceeding of the 17th ACM Conference on Information Knowledge Management vol. 08* (pp. 0–9).

Ma, H., King, I., Lyu, M.R. (2011). Learning to Recommend with Social Trust Ensemble. In *SIGIR '09 Proceedings of the 32nd International ACM SIGIR Conference* (pp. 203–210).

Ma, H., Zhou, D., Liu, C., Lyu, M.R. (2011). Recommender systems with social regularization. In *Proceedings of the fourth ACM International Conference on Web Search Data Mining, no. 978–1–4503–0493–1* (pp. 287–296).

Matuszyk, P., Vinagre, J., Spiliopoulou, M., Jorge, A.M., Gama, J. (2015). Forgetting methods for incremental matrix factorization in recommender systems. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing - SAC '15* (pp. 947–953).

Oh, J., & Han, W. (2015). Fast and robust parallel SGD matrix factorization. In *KDD '15 Proceedings of the 21th ACM SIGKDD, The International Conference on Knowledge Discovery and Data Mining* (pp. 865–874).

Qian, X., Feng, H., Zhao, G., Mei, T., Member, S. (2014). Personalized recommendation combining user interest and social circle. *IEEE Transactions on Knowledge and Data Engineering*, *26*(7), 1763–1777.

Rana, C., & Jain, S.K. (2014). An evolutionary clustering algorithm based on temporal features for dynamic recommender systems. *Swarm and Evolutionary Computation*, *14*, 21–30.

Recht, R.C.B. (2013). Parallel stochastic gradient algorithms for large-scale matrix completion. *Mathematical Programming Computation*, *5*, 201–226.

Salakhutdinov, R., & Mnih, A. (2008). Probabilistic matrix factorization. Neural Information Processing System, 1–8.

Su, X., & Khoshgoftaar, T.M. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, *2009*(3), 1–19.

Way, O.M. (2013). An Experimental Study on Implicit Social Recommendation. In *36th international ACM SIGIR conference on Research and development in information retrieval* (pp. 73–82).

Wu, L., Chen, E., Liu, Q., Xu, L., Bao, T., Zhang, L. (2012). Leveraging tagging for neighborhood-aware probabilistic matrix factorization. In *21st ACM international conference on Information and knowledge management* (pp. 1854–1858).

Yang, X., Member, S., Guo, Y., Liu, Y. (2013). Bayesian-inference-based recommendation in online social networks. *IEEE Transactions on Parallel and Distributed Systems*, *24*(4), 642–651.

Yang, X., Guo, Y., Liu, Y., Steck, H. (2014). A survey of collaborative filtering based social recommender systems. *Computer Communications*, *41*, 1–10.

Yu, H.-F., Hsieh, C.-J., Si, S., Dhillon, I.S. (2013). Parallel matrix factorization for recommender systems. *Knowledge Information Systems*, *41*(3), 793–819.

Yu-sheng, L.I., Mei-na, S., Jun-de, S. (2014). Social recommendation algorithm fusing user interest social network. *The Journal of China Universities of Posts and Telecommunications*, *21*, 26–33.

Zhang, Z., & Liu, H. (2014). Application and research of improved probability matrix factorization techniques in collaborative filtering. *International Journal of Control, Automation and Systems*, *7*(8), 79–92.

Zhang, Z., & Liu, H. (2015). Social recommendation model combining trust propagation and sequential behaviors. *Applied Intelligence*, *43*(3), 695–706.

Zhang, C., Zhang, Z., Yu, L., Liu, C., Liu, H. (2013). Information filtering via collaborative user clustering modeling. *Physica A*, *396*, 195–203.

Zhang, Y., Chen, M., Huang, D., Wu, D., Li, Y. (2016). iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization, Future Generation Computer Systems, p. In Press.