Applying Bi-directional Link Mining in Personalized Recommendation

Hong Wang [1,2]
[1]Institute of Information Science and Engineering, Shandong Normal University, Jinan, China
[2]Shandong Provincial Key Laboratory for Distributed Computer Software Novel Technology, Jinan, China
wanghong106@163.com

Yanshen Sun *
Department of Earth Sciences, ZheJiang University, Hangzhou, China
sunyanshen@163.com
*Corresponding Author

***Abstract*—Recently, many researchers have been attracted in link prediction which is an effective technique to be used in graph based models analysis. By using link prediction method we can understand associations between nodes. To the best of our knowledge, most of previous works in this area have not explored the prediction of links in dynamic Multi-dimension Networks and have not explored the prediction of links which could disappear in the future. We argue that these kinds of links are important. At least they can do complement for current link prediction processes in order to plan better for the future. In this paper, we propose a link prediction model, which is capable of predicting bi-direction links that might exist and may disappear in the future in dynamic Multi-dimension Networks. Firstly, we present the definition of multi-dimensional networks, reduction dimension networks and dynamic networks. Then we put forward some algorithms which build Multi-dimension Networks, reduction dimension networks and dynamic networks. After that, we give bi-direction link prediction algorithms in dynamic multi-dimension weighted networks. At the end, algorithms above are applied in recommendation networks. Experimental results show that these algorithms can improve the link prediction performance in dynamic multi-dimensional weighted networks.**

***Index Terms*—bi-directional link mining, multi-dimensional networks, reduction-dimensional networks, personalized recommendation, weight similarity**

I. INTRODUCTION

With the rapid development of the network, its scale is getting increasingly large and complex. More and more scholars have been attracted by the research of how to find unknown relationships from the existing relationships in networks, which helps people to understand and recognize something of the future. Link prediction in networks is a kind of method to solve this problem. Most of previous works in this area lay emphasis on the possibility of generating links between nodes based on information of nodes and structure of networks [1]. This method includes both the unknown links prediction and the future links prediction. However, in this paper, we put forward a kind of link prediction method called bi-direction link prediction, which not only predicts the possibility of generating links among nodes and also predicts the possibility of links disappearing among nodes in the future. We hold that this kind of link prediction method is of great universal significance, for the reason that possibilities of these two sorts exist simultaneously in the real world. As we know, an enterprise might grow or might decline. If using bi-directional link prediction method correctly, on one side the enterprise can expand its production gradually, on the other side it can take precautions to reduce losses and avoid failure. So, link prediction of the two kinds are equal important.

A great number of real systems existing in the nature can be described by complex networks [2]. For example, the network formed by the relationship between authors and their published papers is a typical complex network, where there are many characteristics of complex networks such as non-linear and self-organizing phenomena. Methods and ideas of link prediction can be used to study the intricate social relations among authors, to predict the possibilities of publishing papers cooperatively among different authors.

In reality, complex networks are usually multi-dimensional networks in which relationships among objects are diverse. For instance, behaviors of people on the Internet generally are involved in several network systems such as film and television networks as well as reading networks. If each kind of networks is regarded as a sub-network system, the set of these sub-networks will constitute a multi-dimensional network, which is dynamic changed, or grow up or decay. The generation of new edges in the network indicates the emergence of new relations among nodes [3], and the disappearance of existing edges demonstrates that the mutual relations between nodes in networks no longer exist. It is undoubtedly very important for our future planning if we could accurately forecasted when and where a new edge emerges, an existing one disappears and a disappeared edge appears again between two nodes. So, we should

doi:10.4304/jsw.9.8.2223-2231

consider the time effect on link prediction in multi-dimension networks.

This paper presents a kind of link prediction method suitable for dynamic multi-dimension networks. Firstly, we construct a model of multi-dimensional networks and reduction dimensional networks. Then, bi-directional link prediction algorithms in dynamic multi-dimensional weighted networks are put forward, in which factors of time and weight are considered. This paper taking the personalized recommendation network as an example, is to build the vector space model of users and products, and then to reduce dimensionality of the network, in which the time factor is taken into account, and finally do the most suitable recommendation for users.

II.RELATED WORK

In recent years, link prediction problem in complex networks has been paid more attention to by scientists from different areas with different backgrounds. Reference [4] proposed a kind of link prediction method which utilizing the hierarchical structure of networks. Based on the similarity of the network topology, reference [5] analyzed the effects of some indices on link prediction of social cooperation networks. Reference [6] took advantage of the random block model to predict missing edges and error edges. Reference [7] used techniques of matrix and vector to carry out link prediction in static networks. Reference [8] provided description of the implementation process of both static and dynamic networks through similarity. Reference [9] adopted a method of internal links and weighted mapping to predict possible links in bipartite social networks that might appear in the future. Reference [10] conducted link prediction of dynamic social networks by means of methods of social network analysis, proving that link prediction had been a powerful auxiliary tool to accurately analyze the structure of social networks. Reference [11] applied methods of link prediction on the basis of nodes (including gender, age, etc.) similarities to the research of biological agents in the treatment of cancer. Reference [12] put forward the concept of network medicine, which pointed out that the origin and development of various diseases were not isolated, on the contrary they were interrelated, and the theory of complex networks could be used to study networks constituted by the origins of the diseases. Reference [13]and reference[14] presented effective methods of recommending products to users by the construction of user-product relationships based on bi-graph theory, which defining and updating the resource allocation vector model according to the user's vector model and weight information. There are other kinds recommendation methods such as collaborative filtering [15], content-based recommendation [16] and hybrid recommendation[17].

In a word, studies mentioned above have achieved better predicting results and have been successfully applied to many areas by depicting different characteristics of the network structure from different perspectives. However, there are also some disadvantages: most of the previous works in this area have considered the two-dimensional relation networks or dynamic networks separately, but neglecting the prediction of links in dynamic multi-dimensional networks comprehensively; most of the previous works attach considerable importance to predict the emergence of new links in the future while think little of predicting the disappearance of current links in the future; moreover, application areas of link prediction can be further extended. Therefore, this paper presents a kind of bi-directional link prediction in dynamic multi-dimensional networks, and applies it to relationships networks of personalized recommendation.

III. DYNAMIC MULTI-DIMENSION NETWORKS CONCEPTS

In this paper, we defined networks which consisted of a variety of nodes as Multi-mode network. Then reduction-mode networks are formed when the Multi-mode network projects to one type of node, which are defined as mapping networks. Definitions of some related concepts are firstly given as the foundation for the study of link mining.

Definition 1 (Multi-dimensional Networks): supposing $V = \{V_1, V_2, \ldots, V_n\}$ is a set which is composed by nodes of n categories, in which $V_i (1 \le i \le n)$ is a set of nodes that belong to the same sort. $\forall v_i \in V_i$, $v_j \in V_j$ ($j = 1, 2, \ldots i-1, i+1, \ldots n$), supposing a unordered couples (v_i, v_j) indicates the edge between v_i and v_j, then $W = \{(v_i, v_j) \mid v_i \in V_i, v_j \in V_j, j = 1, 2, \ldots i-1, i+1, \ldots n\}$ is a set of all possible edges between any node v_i in set V_i and any v_j in $V_j (1 \le j \le n, j \ne i)$, we call a network a Multi-dimensional Network (MD), which takes V as the node set and takes a subset of W as the edge set.

Definition 2 (Single-dimensional Networks): in definition 1 of multi-dimensional networks, let V_k represents any particular $V_j (1 \le j \le n, j \ne i)$, supposing $W_k = \{(v_i, v_k) : v_i \in V_i, v_k \in V_k\}$ indicates all possible edges between nodes in set V_i and V_k. Then we called a network a Single-dimensional Network of MD, which takes all nodes in V_i and V_k as the node set and W_k as the set of all connecting edges. The overall number of all single-dimensional networks in MD is called the dimension of MD.

Definition 3 (Reduction-dimensional Networks): in definition 1 of multi-dimensional networks, with regard to $\forall p, q \in V_i$, supposing node sets V_{jp} and V_{jq} ($V_{jp} \subseteq V_j$, $V_{jq} \subseteq V_j$) represent respectively the set of nodes from $V_j (1 \le j \le n, j \ne i)$ that has connecting edges with p and q. The condition for connecting points p and q is defined as if and only if the intersect collection of V_{jp} and V_{jq} is not null. A network is called a reduction dimensional network of MD, which takes V_i as the node set and has connecting edges mentioned above.

Definition 4 (Dynamic Networks): in definition 1 of multi-dimensional networks, we consider the influence of time. Supposing MD_i is a MD at time of T_i, and MD_{i+1} represents MD at time of T_{i+1}, $i=1,2,$ ... and so on, then $S=\{(T_i, MD_i) \mid i=1,2,\ldots\}$ is called Dynamic Networks.

Fig. 1 and 2 shows the state of a MD and its corresponding RD at two different time points, in which oval nodes are defined as object nodes while the other kinds of nodes such as triangle, square and diamond nodes present different kind of subject nodes respectively. There is a connection edge between them when a subject uses an object. We can get the reduction-dimensional network of this multi-dimensional network if relationships among subjects can be obtained according to that between subjects and objects.

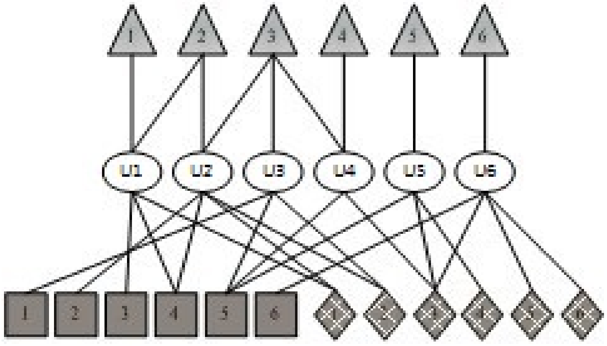

Figure 1. Three-dimensional network

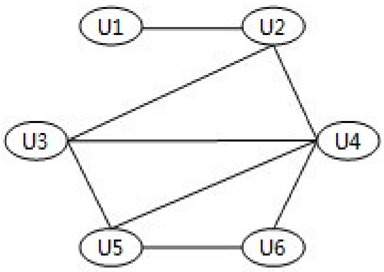

Figure 2. Reduction-dimensional network of figure 1

In figures above, node *U1* to *U6* are subject nodes, while other nodes represent object ones.

IV. Constructing Multi-dimensional Networks

Personalized recommendation systems are usually multi-dimensional networks in which relationships among objects are diverse. For instance, behaviors of people on Internet are generally involved in several network systems such as film networks, television networks as well as reading networks. If each kind of network is regarded as a sub-network system, the set of these sub-networks constitute a multi-dimensional network, which is dynamic changing, growing up or decay. The generation of new edges in the network indicates the emergence of new relations among nodes[18], and the disappearance of existing edges demonstrates that the mutual relations between nodes in networks may no longer exist in future.

A. Construction of Multi-dimensional Networks

We express a single-dimensional network as a bipartite graph $G=(U, V_i, E_i)$, $i=1,2,3\ldots n$, in which n is the dimension of multi-dimensional networks, U is the set of subject nodes in multi-dimensional networks, V_i represents the collection of object nodes and E_i is its set of edges in the i^{th} single-dimensional network, the weight on an edge indicates the strength of the relationship. For convenience, we normalize the edge weight, which makes the maximum value of weight as 1.A single-dimensional network is constructed by adding an edge between a subject node and an object node if the subject node uses the object node and the weight will be marked on the edge. Therefore, the construction algorithm of multi-dimensional networks can be presented as follows.

Algorithm 1: Construction algorithm of multi-dimensional networks.

Input: Data set.
Output: multi-dimensional weighted networks.

```
MultiDimentionModel ()
{
    forEach (node1 in U)
    {
        for(int i=1; i<=n; i++)
        {
            forEach (node2 in Vi)
            {
                if(node1 uses node2)
                {
                    Add an edge between
                    node1 and node2 labeled
                    with a weight value;
                  Normalize the weight value;
                }
            }
        }
    }
}
```

B. Dimension Reduction of Multi-dimension Networks

As described above, Multi-dimensional Networks represent multiple relationships that the subject nodes participate in. In order to find more important implicit relationships between subject nodes, multi-dimensional networks need to be projected to get their reduction-dimensional networks of subject nodes. It is noticed that simply overlay of these single-dimensional networks could not be done in the process of dimension reduction, which certainly results in the loss of some important information. This is because the importance of every single-dimensional networks are different, further more they are not independent but interact with each other another, which is of great significance for link prediction in dynamic multi-dimensional networks. So the dimension reduction of multi-dimensional networks is divided into three steps.

- **Step 1:** Each Single-dimension Network needs to do projection to the subject node set, to obtain n (n is dimensions of Multi-dimension Networks) projection networks.

- **Step 2:** A correlation analysis is done to these n projection networks, in order to remove redundant information from original relationships and to select the most effective relationships for analysis.
- **Step 3:** The Multi-dimension Networks is represented by the first L relationships which have the minimum correlation value got from step 2. In this way, we can obtain the Reduction-dimension networks of subject nodes.

In the first step, each single-dimensional network does projection to the subject node set, to obtain n (n is dimensions of multi-dimensional networks) projection networks, which is expressed as follows.

Algorithm 2: Network projection of multi-dimensional networks.

Input: multi-dimensional networks $G=\{G_1,G_2,\ldots G_n\}$ (n is the dimension of Multi-dimension Networks), in which $G_i=(U, V_i, E_i)$ is a single-dimensional network ($i = 1,2,3 \ldots n$) .

Output: the collection of projection networks

```
NetProjection()
{
    forEach (node1 in U)
    {
        forEach (node2 in U)
        {
            forEach (edge in Ei)
            {
                if(there are N1 neighbor nodes between node1and node 2)
                Add an edge between node1 and node2 labeled with N1;
            }
        }
    }
}
```

Easy to description, the collection of n projection networks obtained from Algorithm 2 are expressed as $P=\{P_1,P_2,\ldots, P_n\}$, in which $P_i=(U, E_i)$ is a projection network, ($i=1,2,3\ldots n$). Let C_i ,a square matrix of $k\times k$, be the corresponding weight matrix of P_i ,in which k is the number of nodes contained by subject node sets U. We do normalization to C_i, thus making the value of each element of matrix C_i between $[0,1]$. As C_i is symmetric, it can be represented by vector V_i with $k(k-l)/2$ dimensions.

After the first step, we begin the second step. A correlation analysis is done to these projection networks, in order to remove redundant information from original relationships and to select the most effective relationships, which is expressed as follows.

Algorithm 3: Relationship selection based on the correlation analysis.

Input: The relationship feature vectors $\{V_1, V_2, \ldots V_n\}$ corresponding to the weight matrixes of projector networks.

Output: A collection of L eigenvectors with the smallest correlation coefficient.

```
SelectRelation()
{
    for(int i=1;i<=n;i++ )
    { Calculate mean characteristics ui of Vi};
    Select k relationship eigenvectors from the n relationship eigenvectors to get the set with n!/k!(n-k)! vectors;
    For any two vectors in each vector collection calculating their correlation coefficient in accordance with Equation 1;
    for(int x=1;x<=n;x++ )
    { Calculate Σ_{y=1}^{n} |ρ_xy| };
    Sort Σ_{y=1}^{n} |ρ_xy| in ascending order (x=1,2,...n);
    Get the first L elements from the sorting above to form a collection of the corresponding relationship vectors;
}
```

In the algorithm above, the correlation coefficient formula of any two vectors among the N vectors is presented as follows[19].

$$\rho_{xy} = \frac{\frac{1}{N}\sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_y)}{\sqrt{\frac{1}{N}\sum_{i=1}^{n}(x_i-\mu_x)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{n}(y_i-\mu_y)^2}} \tag{1}$$

In formula (1), x_i and y_i are the i[th] characteristic components of vector x and y respectively; and u_x and u_y are the mean characteristics of vector x and y respectively. Correlated characteristics being added or removed from the original vector sets, the classifications of the original vector sets would not be affected, therefore we chose vectors with smaller correlation coefficients to represent the original networks.

The multi-dimensional network is represented by the first L relationships in algorithm 3.

Now we start the third step. As we know, in real networks, different relationships usually impose different effects on multi-dimensional networks, so weight factors should be taken into account. The dimension reduction process of multi-dimension networks is as follows.

Algorithm 4: Construction the dimension-reduction networks.

Input: A collection of the L relationship eigenvectors with the smallest correlation coefficient.

Output: Dimension-reduction network.

```
DimensionReduction()
{
    Obtain the corresponding projection network set P= {P1, P2,...PL} according to the collection of the L relationship eigenvectors with the smallest correlation coefficient;
    Determine the weights collection Λ= {α1, α2, α3,...αL} which is made up of L numbers between 0 and 1;
    forEach(x,y in U)
    {
        initialize the weight of the edge between node x and node y called Wxy;
        for(int i=1;i<=L;i++ )
        { Wxy= Wxy + W[Pi]xy×αi};
        Get a weighted network with nodes in U;
```

}
}

C. Dynamic Model of Multi-dimension Networks

In multi-dimensional networks, relationships among the subject nodes are not only pluralistic but also changing dynamically. The changes, such as adding new nodes, removing older ones and changing relationships between nodes, will exert new impacts to the link prediction of the networks, which makes us give sufficient consideration to the dynamic changing factors during the construction process of the multi-dimensional networks. With notation S_i representing the network diagram of the multi-dimensional networks in time slice t_i, the status of the multi-dimensional networks in time slice t_i can be expressed as a 2-tuples $(t_i,S_i(V_i,E_i,W_i))$, in which V_i is nodes set, E_i is edges set and W_i is edge weight. So the dynamic multi-dimension networks can consequently be represented as a sequence of the two-tuples above, $G=\{(t_1,S_1(V_1,E_1,W_1)),(t_2,S_2(V_2,E_2,W_2)),\ldots\ldots(t_i,S_i(V_i,E_i,W_i)),\ldots\ldots\}$. In this way we can effectively describe the changes of the multi-dimensional networks to show their dynamic characteristics. The algorithm of building dynamic networks is as follows.

Algorithm 5: Construction algorithm of dynamic networks.

Input: The multi-dimension networks and the time sequence $T = \{t_1, t_2, \ldots t_i, \ldots t_n\}$, wherein $t_1<t_2<\ldots t_i<\ldots<t_n$.

Output: Dynamic networks.

```
DynimicReductionNetwork()
{
    forEach (x in T)
    {
        MultiDimentionModel();
        NetProjection();
        SelectRelation();
        DimensionReduction();}
        Get
        G={(t1,S1(V1,E1,W1)),(t2,S2(V2,E2,W2)),…
        …(ti,Si(Vi,Ei,Wi)),……
    }
}
```

It is noticed that the characteristics of two sorts should be reflected in the link analysis of dynamic multi-dimension Networks.

- At the same moment, each single-dimensional network plays a different role in the link prediction of the entire networks. For example, the effect of joining friends' reception is less important than that of treating patients in order to reflect the doctors' working relations.
- With networks dynamic changing as time going on, the role of the older status of networks gives less contribution to reflecting the current networks state or predicting the future possible networks state than that of recent one. For instance, if two authors have early co-operations to publish papers together, but they have no recent co-operations, the possibility of their future co-operations will reduce. On the contrary, the possibility will increase.

The first characteristic has been embodied according to the edge weights in the process of dimension reduction of the multi-dimensional relationship networks, in which different edges has been set different weights, to reflect the relationship familiarity degrees among nodes in dimension reduction networks.

In order to reflect the second characteristic, the weights of edges need to be re-corrected in the dynamic model, that is, W_i in algorithm 5 should be a combination of W_{xy} in different periods of algorithm 4. As a result, node weights should be endowed with different factors in different moments, the earlier with lower factors while the recent with higher ones. For example, if a dynamic network is divided into three periods of time in chronological order: $T_1, T_2, T_3 (T_1<T_2<T_3)$, let W_1, W_2, W_3 be three weights of some pair of nodes (x, y) during the three periods, and let different factors $\alpha_1, \alpha_2, \alpha_3$, ($\alpha_1<\alpha_2<\alpha_3$) be assigned to these three weights, then the final weight of the pair (x, y) is $W= W_1\times\alpha_1+ W_2\times\alpha_2+ W_3\times\alpha_3$.

V. Bi-directional Link Mining in Personalized Recommendation Networks

The model of dynamic multi-dimension weighted network can be obtained according to the previous section. In this section, the bi-direction link mining of this modal will be carried, which means that when predicting the possibility of future links, the possibility of existing links that would disappear in the future are also predicted. This section presents a link mining algorithm based on similarity indicators.

A. Correction of Similarity Indicators

Similarity metric is an important indicator in the link mining algorithm. We use calculation method of the adjusted Cosine similarity[20] when computing the similarities between the subject nodes. As to the set of subject nodes $U = \{u_1, u_2, \ldots u_n\}$, the formula of similarity between nodes is as following.

$$sim(i,j)=\left[\sum_{k\in I_{ij}}(R_{i,k}-\overline{R}_i)(R_{j,k}-\overline{R}_j)\right]\Big/\left[\sqrt{\sum_{k\in I_i}(R_{i,k}-\overline{R}_i)^2}\sqrt{\sum_{k\in I_j}(R_{j,k}-\overline{R}_j)^2}\right] \quad (2)$$

In formula (2), I_{ij} indicates the common set of neighbor nodes between subject i and j, I_i and I_j represent the sets of neighbor nodes of subject i and j respectively, wherein neighbors of one node are defined as those directly connected with it. R_{ij} is the weight between node i and j.

Indexes of common neighbors similarity used in this paper are simple, easy to understand and easy to be processed. However, similarity indicators need to be adjusted to fit for model of dynamic Multi-dimension Networks.

Similarity indicators are the main innovations of this article. Node similarities should be endowed with different weights in different moments, the earlier with Lower weights while the recent with higher values. For example, the dynamic network was divided into three periods of time in chronological order: T_1, T_2, $T_3(T_1<T_2<T_3)$, let S_1, S_2, S_3 be three similarities of some

node during the three periods, and let the three different weights $\alpha_1, \alpha_2, \alpha_3$, ($\alpha_1 < \alpha_2 < \alpha_3$) be assigned to the three similarities, thus the final similarity of the node would be $S= S_1\times\alpha_1+ S_2\times\alpha_2+ S_3\times\alpha_3$. Hence the formula of node similarity will be adjusted to.

$$sim(i,j) = \sum_{k=1}^{n} \alpha_k \times sim(i,j)_k, (i,j) \in E_k \quad (3)$$

A. Forward-directional Link Mining

The forward-directional link mining algorithm for personalized recommendation networks is shown below, in which similarity indicators is used as evaluation criteria.

Algorithm 6: Forward-directional link mining for personalized recommendation networks.

Input: multi-dimensional networks.

Output: The possible link sequences of subject nodes in the future.

```
LinkPrediction()
{
    MultiDimentionModel();
    NetProjection();
    SelectRelation();
    DimensionReduction();
    forEach (int x=1; x<n; x++)
    {
        Compute similarity for each pair of nodes
        according to formula 3;
    }
    Put these similarity numbers in a list and sort them
    descending;
    Get the first L … in the list;
}
```

B. Backward-directional Link Mining

The purpose of backward-directional link mining is to gain those current links that might disappear in future. The main thought is to change weights during the dimension reduction of the network, which means changing the original bigger weight into a smaller one, or vice versa. Then the forward-direction link mining is conducted, and the outcome is the links that might disappear most likely in future. The following is the algorithm.

Algorithm 7: Algorithm of the backward-directional link mining in dynamic multi-dimension networks.

Input: Multi-dimensional networks and time sequence $T= \{t_1,t_2,\dots t_n,\dots\}$, $t_1<t_2<\dots t_n<\dots$.

Output: Link sequences of subject nodes that might disappear in the future.

```
ReverseLinkPrediction()
{
    MultiDimentionModel();
    NetProjection();
    SelectRelation();
    DimensionReduction();
    Get reduction dimensional networks called
    RD=(U,E,W);
    forEach (w_xy in W) w_xy=1- w_xy;
    LinkPrediction();
    Put these similarity numbers in a list and sort
    them descending;
    Get the first L ⋯ in the list;
}
```

VI. Experiments and Analysis

The personalized recommendation system based on dynamic multi-dimensional network combines multi-dimensional social network and dynamic factor in practice to model user. This system constructs multi-dimensional network according to user data and local world involving theory, makes use of algorithms in previous section to find neighbor users, does recommendation by figuring out the predict ratings finally.

A. Personalized Recommendation System

Fig. 3 is the structure of personalized recommendation system of dynamic multidimensional network.

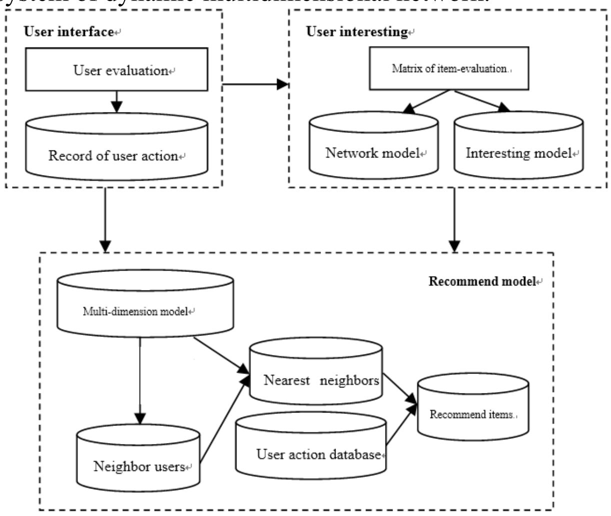

Figure 3. Structure of Recommendation System

The system includes three models, which are user interface model, user interesting description model and personalized recommendation model.

- The user interface model mainly is used to collect user information, which includes three kinds such as user login information, user action information and item-evaluation matrixes. The item-evaluation matrixes are inputs for user interesting description model.
- The user interesting description model is used to construct dynamic and multi-dimensional network model of user. Firstly, dynamic and multi-dimensional networks are constructed according to item-evaluation matrixes. Then reduction-dimensional networks of users are formed based on improved similarity of users. All of these results above will be input to the personalized recommendation model.
- The personalized recommendation model is the most important part in the system, which mainly responsible for finding the nearest neighbor users and doing recommendation. Firstly, neighbor users

are found according to the improved clustering algorithm. Then the nearest neighbor uses sets are obtained. Finally do personalized recommendation.

B. Experimental Data Set

The experimental data in this paper are mainly obtained from the back-end access logging of a comprehensive video website in the last three months. The website offers a lot of resources such as movies, TV show and music. We implement the recommendation algorithm in a three-dimension network based on these three kinds of data. There are about 1496, 1253 and 1800 pieces of data in evaluation data sets of movie, TV series and songs respectively. Each dimension of the experimental data turned into a two-dimensional table after data processing, with the following structure as shown in table 1.

TABLE I.

EVALUATION TABLE STRUCTURE OF USERS

Identifier	Meaning
userId	User ID
itemNum	Project category ID
itemId	Project ID
rating	rating
timestamp	Timestamp

After data processing, each dimension of the experimental data is described in a two-dimensional table, and each dimension corresponds to a type of project network. As the consistency of the users, the evaluation tables are joined according to userID, a foreign key, to generate the three-dimensional data table for each user.

Examples of users' score for items is shown in table 2.

TABLE II.

EVALUATION EXAMPLES OF USERS

userId	306	429	278	107	335
itemNum	1	1	1	1	1
itemId	394	232	508	455	112
rating	3	2	4	1	1
timestamp	92337 6508	89902 5782	90087 9351	91266 5873	90061 2468

C. Evaluating Criterions

There are different evaluating criterions for different experiments.

- **Accuracy:** Accuracy is defined as the percentage of number of correct predictions to all predictions generated by the system, which is shown in formula (4), in which n_cp is the number of correct predictions and n_tp is the number of all predictions. Accuracy reflects the degree of users' interests to recommended items.

$$Accuracy = n\_cp / n\_tp * 100\% \quad (4)$$

- **Recall:** Recall is defined as percentage of items which both are liked by users and predicted by the system to all items liked by users, shown in formula (5), in which P_{fr} is the number items which both are liked by users and predicted by the system and P_{ff} is the number of items disliked by users. Recall describe the percentage of items liked by users can be recommended.

$$Recall = \frac{P_{fr}}{P_{fr} + P_{ff}} \quad (5)$$

- **MAE:** MAE is Mean Absolute Error，which is used to measure the mean absolute error between the predicting scores by the recommendation system and the actual scores by users. It is shown in formula (6), in which p_i presents the predicting scores, q_i presents the actual scores by users and n is users' number.

$$MAE = \sum_{i=1}^{n} |p_i - q_i| / n \quad (6)$$

Generally speaking, the smaller MAE value is, the more accurate the prediction is. As it is easier for computing MAE, MAE is often used to compare the accuracy of different prediction systems.

D. Experimental Results

Data sets of users and resources include 328 users, 1496 comments on films, 1253 comments on teleplays and 643 comments of songs. In our experiments they are randomly divided into a training set and a test set, with a proportion of 80% and 20% separately. The training set is used to generate results of predictions and the test set is used to evaluate superior and inferior of recommendation results.

In order to test the performance of our bi-directional link mining algorithms in personalized recommendation networks abbreviated as DMRS, we compare DMRS with classic methods such as collaborative filtering recommendation based on project's scoring and content-based recommendation based on comparability. With test users divided into ten groups, we carries out three kinds of recommendation methods above separately in each group and get their Mean Absolute Error(MAE), Accuracy rate, Recall rate and Precision rate of each group. The experimental results are shown in Fig. 4, 5 and 6.

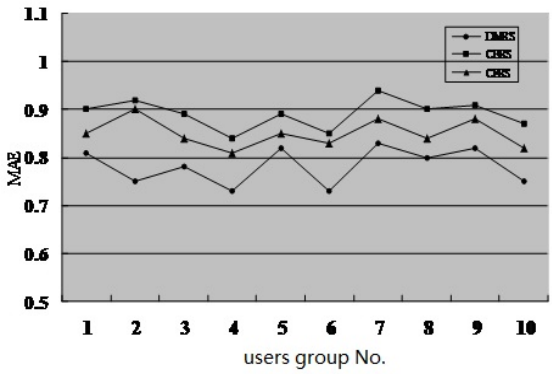

Figure 4. MAE comparisons

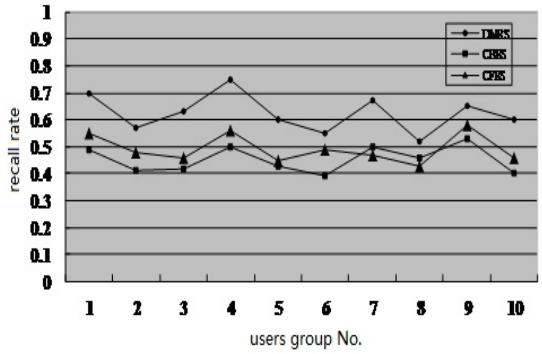

Figure 5. Recall rate comparisons

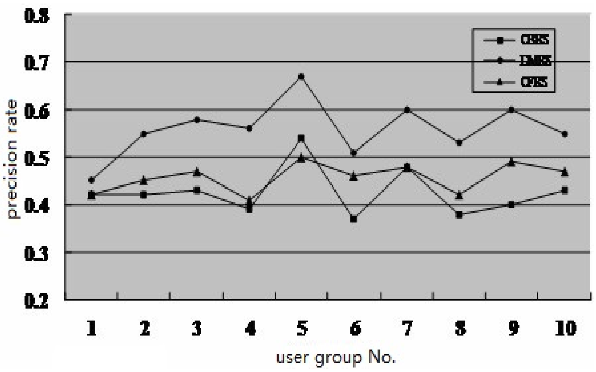

Figure 6. Precision rate comparisons

From the results, we can see that some evaluating indicator value such as MAE, Recall rate and Precision rate of DMRS are higher than other classic methods. Note that in order to avoid the effect of scoring standard, the scoring mechanism use in DMRS is the same as that of classic methods.

Next, we test the effect of dynamic factor. We do compare experiments between algorithms of DDMLP_dyc (Bi-direction Link Prediction in Dynamic Multi-dimension Networks) and DDMLP (without dynamic evolution). We also randomly divide the users into 10 groups and calculate the average results. The mean of accuracy and time-consumption are shown in Fig. 7 and 8.

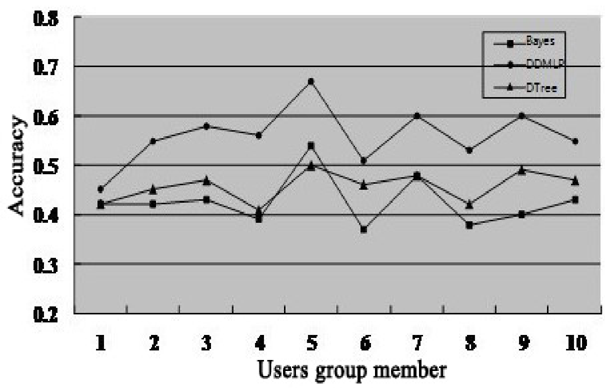

Figure 7. Accuracy comparison

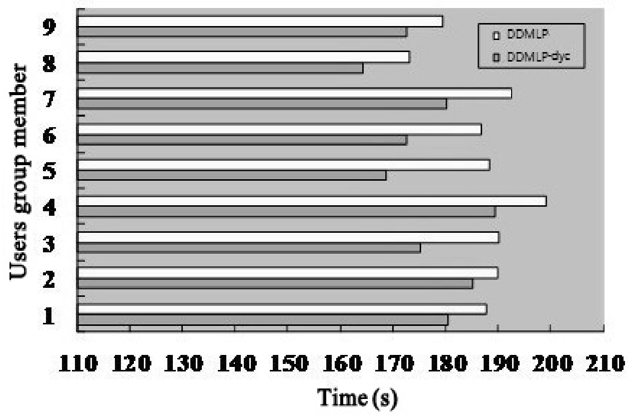

Figure 8. Time comparison

Experimental results show that the accuracy of DDMLP_dyc is higher than that of DDMLP. In addition, the time consumption of DDMLP_dyc is generally less than that of DDMLP, especially in some of the users groups, for the reason that the reconstruction of network when updating users after the formation of dimension reduction network increases data redundancy in DDMLP. From the two aspects of the experimental data, we find that the bi-directional link prediction algorithm in multi-dimensional networks has acquired better recommendation quality and efficiency.

CONCLUSIONS

In this paper we use bi-directional link mining method in personalized recommendation networks, which not only predicts future possible links but also predicts existing links that might disappear in future. We can draw a conclusion that the most of the performance values in DMRS are higher than those of classic recommendation methods. However it should also be noted that execution time of DMRS is longer than the other two algorithms, owing to time spent on iterating user data from the three kinds of networks to constructing multi-dimensional weighted networks. In future our works mainly include some meticulous work such as adjusting and refining the parameters, and optimizing the algorithms and improving their efficiency.

ACKNOWLEDGMENT

The research work was supported by Technology Program of Shandong Province under Grant No. 2012GGB01058.

REFERENCES

[1] GETOORL, DIEHLCP, Link mining: a survey, ACM SIGKDD Explorations Newsletter, 7(2), pp.3–12, 2005.
[2] Stefan Thurner, Christoly Biely, Two statistical mechanics aspects of complex networks, Physica A, 74 (2), pp. 346–353, 2002.
[3] Yoshii K, Goto M, Komatani K, et al., An efficient hybrid music recommender system using an incrementally trainable probabilistic generative model, IEEE

Transactions on Audio Speech and Language Processing,16(2), pp. 435–447, 2008.
[4] Nowell, Kleinberg, The link prediction problem for social networks, Proceedings of the twelfth international conference on Information and knowledge management, ACM, pp.556–559, 2003.
[5] CLAUSET A, MOORE C, NEWMAN, Hierarchical structure and the prediction of missing links in networks, Nature, 45(3), pp.98–101, 2008.
[6] Dunlavy DM, Kolda TG, Temporal link prediction using matrix and tensor factorizations, ACM Trans Knowl Discov Data (TKDD), 5(2), pp.10–18, 2011.
[7] Wang Lin, Shang Chao, Research on Link Prediction in No-scale networks, Computer Engineering, 38(3), pp.67–70, 2012.
[8] Allali O, Magnien C, Latapy M, Link mining in bipartite graphs using internal links and weighted projection, IEEE conference on computer communications workshops, Shanghai, pp. 936–941, 2011.
[9] Tylenda T, Angelova R, Bedathur S, Towards time-aware link prediction in evolving social networks, Proceedings of the 3rd workshop on social network mining and analysis, SNA-KDD ACM, pp.1–10, 2009.
[10] Mooney R J, Bennett P N, Roy L. Book recommending using text categorization with extracted information, Proceedings of AAAI-98/ICML-98 Workshops, Madison, pp. 49–54, 1998.
[11] Zhou T, Jiang L, Su R, et al., Effect of initial configuration on network-based recommendation, Europhys Lett, 81(5), pp.58004p1–p5, 2008.
[12] Liu R R, Jia C X, Zhou T, et al., Personal recommendation via modified collaborative filtering, Physica A, 388(4), pp.462–468, 2009.
[13] Naji G, Nagi M, Elsheikh AM, Gao S, Effectiveness of social networks for studying biological agents and identifying cancer biomarkers, Counterterrorism and open source intelligence, 2, pp. 285–313, 2011.
[14] Sun Jijia etc., Theory of Complex networks with its application in research of traditional Chinese Medicine, Complex Systems and Complex Science, 5(3), pp. 55–61, 2008.
[15] Zeng C, Xing C X, Zhou L Z. A survey of personalization technology, Journal of Software, 13(10), pp.1952–1961, 2002.
[16] Wu, Y.H., Chen, Y.C., Chen, A.L.P., Enabling personalized recommendation on the web based on user interests and behaviors, Proceedings of the 11th International Workshop on Research Issues in Data Engineering, IEEE CS Press, pp.17–24, 2001.
[17] Zhao P, Research on complex networks & personalized information service on Internet, University of Science and Technology of China, 2006.
[18] GALLAGHERB, TONG H, ELIASSI-RAD T, Using ghost edges for classification in sparsely labelled networks, ACM SIGKDD, LASVEGAS, pp. 256–264, 2008.
[19] Ordonez C, Omiecinski E, Efficient disk-based K-means clustering for relational databases, IEEE Trans. Knowledge and Data Engineering, 16(8), pp. 909–921, 2004.
[20] Deng Ailin, Research on Key Technologies of recommendation system in E-commerce, ShangHai, FuDan University, 2003.

Hong Wang (1966-): female, from Tianjin, professor, doctoral supervisor, postdoctoral.

She gained Master's and Bachelor's degree in computer software of Tianjin University in 1988, gained PHD of the Chinese academy of sciences in 2002. She has been working in Shandong Normal University since 1991. Her main research direction includes complex networks, workflow, mobile agent, social software and big data cloud. She writes and publishes over 60 academic articles.

Address: Institute of Information Science and Engineering, Shandong Normal University, East Wenhua road no. 88, Lixia district, Jinan, Shandong, China
E-mail:wanghong106@163.com
Tel: 0531-86180244

Yanshen Sun (1993-): female, from Jinan, undergraduate student.

She is interested in picture processing, big data and cloud.

Address: Department of Earth Sciences, Zhejiang University, Hangzhou, Zhejiang, China
E-mail:sunyanshen@163.com