

An asymmetric graph recommendation algorithm based on user scores and interests

Yi-Wen ZHANG ^{1,a}, Ran WANG ^{2,b}, Ting-Wei WANG ^{1,c}, Guang-Da HE ^{1,d}, Jun HU ^{1,e}

¹ An Hui Xin Hua University, Hefei, China 230088

² An Hui University, Hefei, China 230601

E-mail: ^ayiwenzh@ustc.edu.cn, ^bransaber@gmail.com, ^c1203901016@qq.com, ^d1639213032@qq.com, ^e417517314@qq.com

Abstract. The proposed algorithm uses the symmetry similarity calculation method based on the common project, which can not reflect the user's interest very well and the recommendation accuracy is low. In order to improve the accuracy of recommendation, this paper proposes a graph recommendation algorithm which integrates user rating and interest. In order to reflect the interests of users, using the principle of diffusion of the substance user rating information asymmetry; Experiments show that the proposed algorithm for user rating and interest fusion improves the prediction accuracy.

1 Introduction

Currently, the amount of Internet information is growing, how to quickly help users find interested content from a lot of information is one of the important challenges of Internet information [1]. In this context, the recommendation system came into being and it is different from the search engine, it does not match the explicit requirements provided by the user, but to explore the user's interest and interest in the content of active recommendation to enhance the user experience [2]. At present, many large e-commerce website using the recommended system to improve the recommended user purchase rate [3] to improve the recommended results.

2 Related work

The current methods of personalized recommendation are as follows: content-based recommendation algorithm based on collaborative filtering [4], pagerank recommendation algorithm [5][6], and the network structure of the recommended algorithm [7][8] and so on.

At present, the recommendation algorithm of PageRank random walk [9][10] received a good recommendation effect, has been widespread concern. the recommendation algorithm of PageRank is based on bipartite graph, which is superior to traditional cooperative filtering algorithm in new user recommendation who without history. In addition, the bipartite graph recommendation algorithm [11][12] is also concerned, which is also called the Bipartite Graph Projection and Ranking.BGPR algorithm. The algorithm mainly uses random travel algorithm based on common project, and statistics the influences between projects and users in a symmetrical way. In order to describe the user preference more accurately, we added the user score in the common project and use the material diffusion principle [13] to secondary allocation of resources to establish a random



travel algorithm of asymmetric bipartite graph random walk algorithm which embodies the user's interest to improve the recommended accuracy rate.

This paper first describes an algorithm which based on user's interest degree asymmetric Bipartite Graph Projection and Ranking (IN-BGRP) [14] [15] , which reflects the asymmetric effect of the project or user, and then adds the user's score and updates the graph weight. Proposed a method of integrating asymmetric graphs with user scoring and interest (SIN-BGRP) . And experiment and verify on the standard data set MovieLens [16] to improve the recommended accuracy.

3 Algorithm introduction

3.1 IN-BGPR algorithm

Based on the user's interest degree in the Bipartite Graph Projection and Ranking method (IN-BGRP) . According to the user - project matrix, calculate the number of users to watch the number of movies, and through the user - project matrix, the secondary allocation of resources, so as to more accurately describe the user preferences.

The IN-BGPR algorithm is described as follows: If there are two items p, q, calculate the degree of influence between p and q. If a few users at the same time on the project p and q played too much, but the items p and q corresponding to the number of scoring records vary, of course, each user has played over the number of items are also different records. If W_p, q represents the degree of influence of the q item on the p item, the value of W_p, q and W_q, p should be unequal due to this asymmetry. For example, the binary diagram between users and projects is shown in Figure 1:

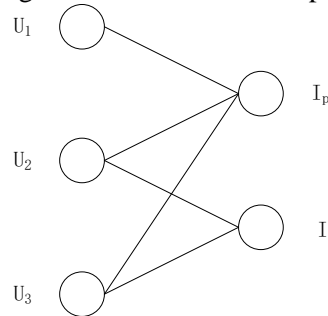


Figure 1: User-item scoring chart

The matrix between users and projects is shown in table 1 as follows:

$$C_{I,u} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

From Figure 1 we can find that U2 and U3 give score to the project Ip and Iq at the same time, but because of the number of scoring is different Ip and Iq so W_p, q and W_q, p value should be different. The formula for calculating W_{pq} and W_{qp} is as follows:

$$W_{pq} = \frac{1}{\deg(I_q)} \sum_{t=1}^m \frac{S_{pt} S_{qt}}{\deg(u_t)} \quad (1)$$

$$W_{qp} = \frac{1}{\deg(I_p)} \sum_{t=1}^m \frac{S_{qt} S_{pt}}{\deg(u_t)} \quad (2)$$

Among them, $\deg(I_p)$ is degree of the project p, which is the number of users scored by project p; $\deg(I_q)$ is degree of the project q, which is the number of users scored by project q; $\deg(u_t)$ is degree of the user t, which is the number of items scored by user t ;

$$S_{qt} S_{pt} = \begin{cases} 0 & \text{User didn't score p, q at the same time} \\ 1 & \end{cases}$$

According to the relationship between each node in Figure 1 and the formulas (1) and (2)

$$W_{pq} = 0.5, \quad W_{qp} = 0.33$$

3.2 SIN-BGPR algorithm

The IN-BGPR algorithm mainly focuses on the use of asymmetric calculations, reflecting the user's interest, but in practice, the user's rating on the project can be prepared to reflect the user's preference for the project, to a large extent reflect the user of the interest in the project. Therefore, this paper also proposes an asymmetric graph recommendation algorithm (SIN-BGPR) which integrates user scoring and interest. The algorithm adds the score on the basis of the original interest based on the IN-BGPR algorithm, and fuses the user interest Weights. as shown in picture 2:

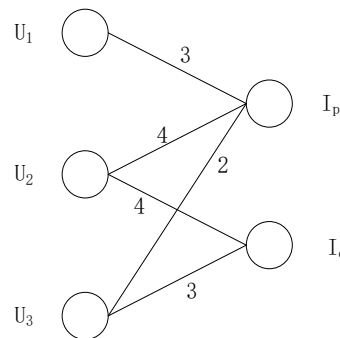


Figure 2: User - Project Score Weighted Chart

Create a weighted user based on Figure 2 - the project score weight matrix is as follows:

$$C_{I,U} = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 4 & 3 \end{bmatrix}$$

It can be seen from Fig. 2 and the matrix that U2 and U3 pay close attention to the items Iq and Ip, and because Iq and Ip are concerned with the number of users and the user's score is different, so $W_{p,q}$ and $W_{q,p}$ Not the same. Therefore, in the SIN-BGPR algorithm, $W_{p,q}$ and $W_{q,p}$ are calculated as follows:

$$W_{pq} = \frac{1}{\text{sum}(I_q)} \sum_{u \in U(p,q)} \frac{R_{up}}{\text{sum}(u)} \quad (3)$$

$$W_{qp} = \frac{1}{\text{sum}(I_p)} \sum_{u \in U(q,p)} \frac{R_{uq}}{\text{sum}(u)} \quad (4)$$

among them, $\text{sum}(I_q)$ is sum weight of the item q, that is, the total score of the users associated with the project q, $\text{sum}(I_p)$ is sum weight of the item p, that is, the total score of the users associated with the item p; $\text{sum}(u)$ is sum weight of the user u, that is, the total score of the users associated with items, R_{up} represents the score of the user u for item p, and R_{uq} represents the score of user u for item q; $U(p,q)$ represents the user set of common scoring items p and q.

4 Experimental results and analysis

Experiment using the MovieLens data set. The data set is composed of 943 users of 1682 movies, a total of about 100,000 ratings. The data set is divided into training set and test set according to the ratio of 8: 2, and is verified by 5-cross verification method.

4.1 Evaluation criteria

In this paper, the use of R [17][18] as the evaluation criteria. R is the average ranking, meaning that, for the target user u_i , the recommended algorithm calculation will produce a sorted length of the project recommendation list I_j . If the item I_j is present in the test set, the relative position of the statistical item I_j in the L queue is shown as equation (5): $r_{i,j}$,

$$R = \frac{\sum_{i=1}^n r_{i,j}}{L} \quad (5)$$

If the item I_j exists in the test set, the recommended item is scored by the user U_i , the recommended length is L , L is the recommended sequence after sorting, the position of item I_j in L is represented by $r_{i,j}$, the smaller the value of $r_{i,j}$ Said the recommended project ranking more forward, the better the accuracy. The value of the evaluation standard R is as small as possible.

4.2 Experimental results and analysis

The BGPR algorithm and the IN-BGPR algorithm are compared and compared with the R value of the result after 5-cross validation. The comparison result is shown in Fig. 3, where the abscissa is 5 random data sets and the ordinate is for each data Concentrated R mean.

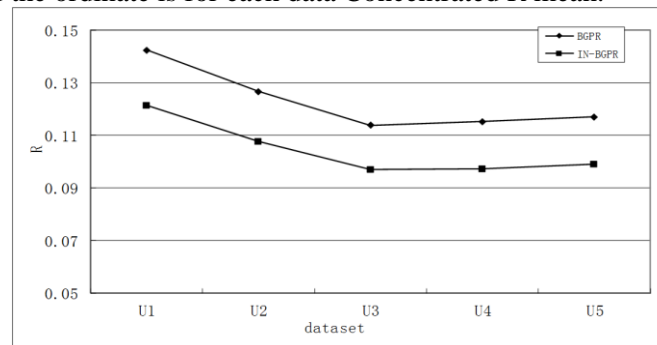


Figure 3 BGPR and IN-BGPR algorithm average ranking comparison

As can be seen from Figure 3, the R -value of the IN-BGPR algorithm used in this paper is significantly lower than that of the traditional BGPR algorithm and cooperative filtering algorithm. It is shown that the asymmetric bipartite graph recommendation algorithm can improve the average ranking and can reflect to some extent User's interest.

The BGPR algorithm is selected and the SIN-BGPR algorithm is compared with the R value of the result after 5-cross validation. The results are shown in Fig. 4, where the abscissa data set represents five random data sets and the ordinate R represents each data concentrated R mean.

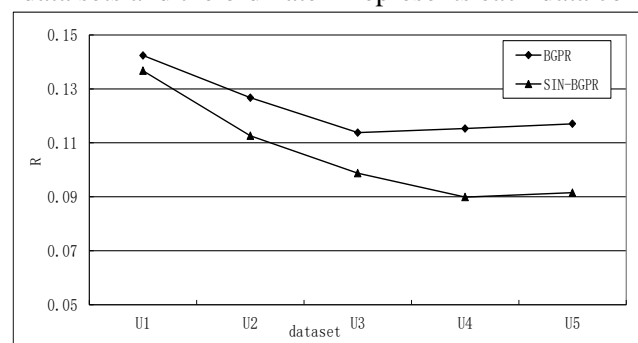


Figure 4 Average ranking of IN-BGPR and SIN-BGPR algorithms

It can be seen from Figure 4 that the SIN value of the SIN-BGPR algorithm used in this paper is lower than that of the common project BGPR algorithm and the cooperative filtering algorithm. The asymmetric bipartite graph recommendation algorithm which does not integrate the score can also increase the average ranking, To the extent that the user's interest.

The BGPR algorithm, IN-BGPR algorithm and SIN-BGPR algorithm are compared. The comparison result is shown in Figure 5.

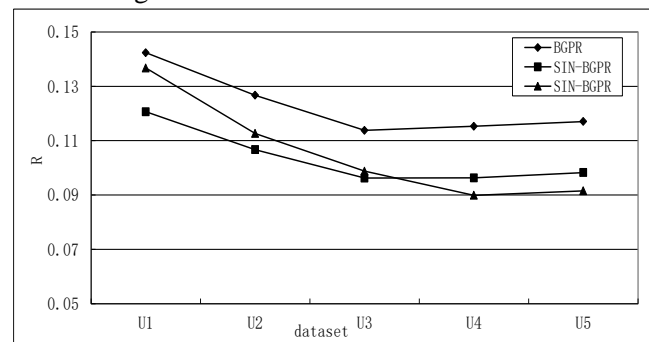


Figure 5 BGPR, IN-BGPR and SIN-BGPR algorithm average ranking comparison

It can be seen from Figure 5 that the IN-BGPR algorithm and the SIN-BGPR algorithm have lower R value than the BGPR algorithm on the five data sets, but the SIN-BGPR algorithm has less on the first three data sets. The fourth and fifth data sets are significantly lower. From the experimental results analysis, the optimization of BGPR algorithm with IN-BGPR algorithm and SIN-BGPR algorithm has a great relationship with the contents of the initial data set.

5 Conclusion and future work

This paper put forward an algorithm which based on user's interest degree asymmetric Bipartite Graph Projection and Ranking (IN-BGRP). Firstly, the asymmetric graph model is used to establish the recommendation algorithm IN-BGPR which based on the user's interest. The validity of the algorithm is verified. Based on the IN - BGPR algorithm, the user 's interest is further extracted, and the user' s score information is added. The IN - BGPR algorithm is further optimized by using the graph model of the fusion user 's score. The experimental results show that the SIN-BGPR algorithm is superior to the traditional BGPR algorithm in the recommended ranking, and it is superior to the asymmetric IN-BGPR algorithm based on user interest in some data sets.

The experimental process shows that user interest is an effective means to improve the accuracy of recommendation. At the same time the user rating on the evaluation model also has a certain impact and significance, the user rating can also reflect the user's interest. But the model can be further optimized, the research process found that the user's interest is to change over time, in order to be able to more accurately represent the user's preferences, the next step will be based on the establishment of a time series of user interest preference model, Which is used to excavate and express the change law of user's interest, and to describe the change of interest preference of users in different time periods according to the advance of time, so as to further improve the accuracy of the recommended algorithm.

References

- [1] A. B. Author, C. D. Author. "Title of the article", The Journal, volume, pp. 110-120, (2000).
- [1] XU Hai Ling WU Xiao LI Xiao Dong YAN Bao Ping, Cnnic L O, Sciences T C A O, et al. Comparison Study of Internet Recommendation System[J]. Journal of Software, 2009, 20(2):350-362.
- [2] Liu Q, Chen E, Xiong H, et al. A Cocktail Approach for Travel Package Recommendation[J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(2):278-293.
- [3] Ricci F, Rokach L, Shapira B. Introduction to Recommender Systems Handbook[J]. Recommender Systems Handbook, 2010, 1-35(3):1-35.
- [4] Breese J S, Heckerman D, Kadie C. Empirical Analysis of Predictive Algorithms for Collaborative Filtering[J]. New Page, 2013, 7(7):43--52.

- [5] Mitliagkas I, Borokhovich M, Dimakis A G, et al. FrogWild! -- Fast PageRank Approximations on Graph Engines[J]. Computer Science, 2015, 8(8):874-885.
- [6] Hotho A, Jäschke R, Schmitz C, et al. Information Retrieval in Folksonomies: Search and Ranking.[J]. Lecture Notes in Computer Science, 2006, 4011:411-426.
- [7] Liu R R, Liu J G, Jia C X, et al. Personal recommendation via unequal resource allocation on bipartite networks[J]. Physica A Statistical Mechanics & Its Applications, 2010, 389(16):3282-3289.
- [8] Zhou T, Ren J, Medo M, et al. Bipartite network projection and personal recommendation[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2007, 76(2):046115.
- [9] Li X, Chen H. Recommendation as link prediction in bipartite graphs: A graph kernel-based machine learning approach[J]. Decision Support Systems, 2013, 54(2):880-890.
- [10] Zhou T, Jiang L L, Su R Q, et al. Effect of initial configuration on network-based recommendation[J]. Physics, 2008, 81(5):58004.
- [11] Liu Q, Chen E, Xiong H, et al. Enhancing collaborative filtering by user interest expansion via personalized ranking.[J]. IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society, 2012, 42(1):218-233.
- [12] Fouss F, Pirote A, Renders J M, et al. Random-Walk Computation of Similarities between Nodes of a Graph with Application to Collaborative Recommendation[J]. IEEE Transactions on Knowledge & Data Engineering, 2007, 19(3):355-369.
- [13] Newman M E J, Watts D J. Renormalization group analysis of the small-world network model[J]. Physics Letters A, 1999, 263(4-6):341-346.
- [14] Wang G X, Liu H P, Li Q. Bipartite network projection and its application in recommendation systems[C]// Control and Decision Conference. 2013:5052-5057.
- [15] Liu J G, Wang B H, Guo Q. Improved collaborative filtering algorithm via information transformation[J]. International Journal of Modern Physics C, 2009, 20(02): 285-293.
- [16] Movielens datasets [EB/OL] .[2013-02-13]<http://www.grouplens.org/node/12>,2010.
- [17] Fouss, Francois, et al. "Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation." Knowledge and Data Engineering, IEEE Transactions on 2007,19(3): 355-369.
- [18] Wang G X, Liu H, Li Q. Bipartite network projection and its application in recommendation systems[M]. 2013.