

A Network-based Multi-dimensional Recommendation Algorithm

Xiangyun Xiong^{1, a}, Yuchen Fu^{2, b} and Zhaoqing Liu^{3, c}

^{1, 3} Mailbox 205, No.1 Shizi Street, Soochow University, Suzhou, Jiangsu, China

² No.1 Shizi Street, Soochow University, Suzhou, Jiangsu, China

^axiangyun_xiong@126.com, ^byuchenfu@suda.edu.cn, ^cliuflysky@yahoo.cn

Keywords: Bipartite network, Projection, Recommendation, Multi-dimensional recommendation.

Abstract. Personalized recommendation based on bipartite network has attracted more and more attention. It's obviously better than CF (Collaborative Filtering). In this paper, we propose a multi-dimensional recommendation algorithm called BNPM. It combines item-based, user-based and category-based recommendation model to improve recommendation quality. The experimental results show that the algorithm can improve the diversity and reduce the popularity on the base of holding the accuracy of the recommendation

Introduction

With the exponential growth of the Internet, the problem of information overload is getting more and more seriously. There is too much data to effectively filter out the pieces that are most appropriate for users. Personalized recommendation [1] is a promising way to filter out them. Up till present, various kinds of algorithms have been proposed. Bipartite-network-based recommendation performance outperforms the traditional collaborative filtering [2, 3] approach.

Inspired by the network-based resource-allocation dynamics, Tao zhou et al. [4] raised a weighting method to compute the resources quota assigned by one item to another and recommend items with high resource. Jie Liu et al. [5] proposed NBIw algorithm based on Tao zhou's NBI algorithm and solved the recommendation problem on weighted bipartite network. Tao zhou et al. [6] and Jianxun Xia et al. [7] proposed initial resource allocation sensitive algorithm and INBI algorithm respectively to increase the recommendation accuracy by decreasing the impact of high-degree item. Current viewpoint is focusing on improving precision of the algorithm. However, improving the algorithmic accuracy is often at the expense of personalized. Therefore, we propose BNPM (bipartite network projection based multi-dimensional recommendation algorithm) to improve diversity and reduce popularity on the base of holding the accuracy of the recommendation.

Basic Recommendation Model

Item-based Recommendation Model

The item-based recommendation model is constructed based on NBIw[5] and initial configuration NBI[6] and needs item-based projection network and the item resource initial configuration.

The user-item bipartite network's edges are weighted by users' ratings on items and r_{ij} is used to donate the ratings. The prediction score vector can be expressed as $O = Wf$. W^0 gets through resource allocation process which consists of two stages: first from items to users, and then back to items. The element of W^0 , w_{ij} , is the resource item i assign to item j . It can be calculated by

$w_{ij} = \frac{1}{k(o_j)} \sum_{l=1}^m \frac{r_{il}r_{jl}}{k(u_l)}$, and $k(o_j) = \sum_{l=1}^m r_{lj}$, $k(u_i) = \sum_{l=1}^n r_{il}$. To user i , the initial resource located

on item j is $f_j^i = r_{ji}k^\beta(o_j)$, and $\beta = -0.8$. Finally, we gain the element of the predictive score matrix O , which means the prediction scores of user i to the unrated item j . It can be calculate by Eq.1.

$$\Pr(o_{ij}) = \sum_{t=1}^n w_{jt} f_t^i = \sum_{t=1}^n \left(\left(\frac{1}{k(o_t)} \sum_{l=1}^m \frac{r_{jl} r_{tl}}{k(u_l)} \right) r_{ji} k^{-0.8}(o_j) \right) \quad (1)$$

User-based Recommendation Model

The weighted bipartite network projection contains two one-mode networks: user projection network G^u and item projection network G^o . In order to make full use of original bipartite network information, we build user-based recommendation model based on G^u .

Similarly, user-based recommendation model is based on NBIw[5] and initial configuration NBI[6]. W^u is the weight of user projection network G^u 's edges. The element w_{ij} means the influence

of user i to user j . It's expressed as $w_{ij} = \frac{1}{k(u_j)} \sum_{l=1}^n \frac{r_{il} r_{jl}}{k(o_l)}$. To item j , the initial resource located on user i is $f_i^j = r_{ij} k^{(-0.8)}(u_i)$. Finally, we gain the prediction scores of user i to the unrated item j :

$$\Pr(u_{ij}) = \sum_{t=1}^m w_{it} f_t^j = \sum_{t=1}^m \left(\left(\frac{1}{k(u_t)} \sum_{l=1}^n \frac{r_{il} r_{tl}}{k(o_l)} \right) r_{ij} k^{(-0.8)}(u_i) \right) \quad (2)$$

The prediction scores based on user-based recommendation model are expressed as predictive score matrix U .

Category-based Recommendation Model

The greater part of previous recommendation techniques are one-dimensional. They can't reflect the real preferences of the individual users. Therefore, we design the third recommendation model. Donate the genres [8] vectors of item j as $O_j^C = (C_1, C_2, \dots, C_k)$, where k represent the genres item j may belong to. $C_k = 1$ if item j belongs to the genre k and $C_k = 0$ otherwise. We get the genres vectors of the user preferences by transferring the genres vectors of items in the bipartite network. Donate the genres vectors of the user i preferences as $U_i^C = (C_1', C_2', \dots, C_k')$. Considering that one user may give different rates to different items, the user preference genres vectors should be decreased by $\sum_{j=1}^n r_{ij}$, and U_i^C can be calculated by Eq.3.

$$U_i^C = \left(\sum_{j=1}^n r_{ij} \right)^{-1} \cdot \sum_{j=1}^n r_{ij} O_j^C \quad (3)$$

The score between the given user i and an item j which user i has not rated is defined as the similarity of the genres vector of item j and user i preferences vector. We use cosine of vectorial angle [9] to calculate the similarity. The equation is

$$\text{sim}(i, j) = (U_i^C \cdot O_j^C) / \left(\|U_i^C\| \|O_j^C\| \right) \quad (4)$$

The element of the predictive score matrix L , which means the prediction scores of user i to the unrated item j , can be calculate by Eq.5.

$$\Pr(c_{ij}) = r_{ij} \text{sim}(i, j) \quad (5)$$

BNPM Recommendation Algorithm

The prediction scores getting from the three recommendation models above are O , U and C respectively. Considering the different effects of the three models on recommendation result, we use dynamic adaptive method to balance the influence of them.

The weight vector of U , O and C is donated as $W_a = (\alpha_1, \alpha_2, \alpha_3)^T$, and $\sum_{i=1}^3 \alpha_i = 1$. The last predictive score matrix is defined as L , so the prediction score of item j which user i has not rated is:

$$\Pr(l_{ij}) = (\Pr(u_{ij}), \Pr(o_{ij}), \Pr(c_{ij})) * Wa \quad (6)$$

BNPM algorithm's steps are as follow:

1) calculate two predictive score matrix O, U through item-based and user-based recommendation model; 2) gain item's category and calculate the third predictive score matrix C through category-based recommendation model; 3) adaptive calculate the weight vector and the last predictive score matrix L; 4) order each rows of matrix L descending and recommend front items to target user.

Weight vector adaptive allocation

Accuracy and personalization are the two types of recommendation algorithm evaluation criterion [3,10]. Ranking score and hitting rate are used to evaluate accuracy. Hamming distance can measure the strength of personalization and the popularity can be measured directly by averaging the degree <k> over all the recommended items.

We choose hitting rate to adaptive calculate weight vector and all four evaluations to evaluate BNPM. Hitting rate's value is determined by the last predictive score matrix L and the adjacency matrix A of user-item pair in testing set. If the items with high prediction scores in L are exactly the items which were rated by user in the testing set, the L has high hitting rate. Define the hitting rate approximation function as

$$Hi = \sum(L * A) \quad (7)$$

$$Hi = \alpha_1(m \cdot n \cdot \sum((U-C) * A)) + \alpha_2(m \cdot n \cdot \sum((O-C) * A)) + \sum(C * A)$$

The m and n are the number of users and items respectively.

Adaptive calculating weight vector is the process of finding the maximum of Hi. The initial value of (α_1, α_2) is (α_1^0, α_2^0) , and calculate the partial derivative of this point follow Eq.2.

$$\begin{cases} Hi_{\alpha_1^0}(\alpha_1^0, \alpha_2^0) = m \cdot n \cdot \sum((U-C) * A) + \alpha_2^0(m \cdot n \cdot \sum((O-C) * A)) \\ Hi_{\alpha_2^0}(\alpha_1^0, \alpha_2^0) = \alpha_1^0(m \cdot n \cdot \sum((U-C) * A)) + m \cdot n \cdot \sum((O-C) * A) \end{cases} \quad (8)$$

The gradient vectors of the point $(\alpha_1, \alpha_2) = (\alpha_1^0, \alpha_2^0)$ is

$$\text{grandHi}(\alpha_1^0, \alpha_2^0) = Hi_{\alpha_1^0}(\alpha_1^0, \alpha_2^0)\vec{i} + Hi_{\alpha_2^0}(\alpha_1^0, \alpha_2^0)\vec{j} \quad (9)$$

The speed of the point grow is η , and $\Delta Wa_{i+1} = \Delta Wa_i + \eta \text{grandHi}(\alpha_1^i, \alpha_2^i)$. The initial value of ΔWa_i is (α_1^0, α_2^0) . Iterate the weight vector with ΔWa_{i+1} until the function Hi is maximized.

Experiment

We use a benchmark data-set, MovieLens, to judge the performance of described algorithms. The data contains 1682 movies and 943 users. Each user rates movies in five discrete ratings 1-5. Here we define that a movie has been rated by a user if the giving rating is at least 3. The original data contains 105 ratings, 85.25% of which are ≥ 3 , thus the user-movie bipartite network contains 85,250 edges. The data set is randomly divided into two parts: The training set contains 90% of the data, and the remaining 10% of data constitutes the probe.

Numerical results

The function of hitting rate is look like fig.1. In the experiment, we choose $(\alpha_1^0, \alpha_2^0) = (0.5, 0.4)$, and the value of hitting rate is small. The weight vector changing along the surface like fig.1 until the function Hi is maximized.

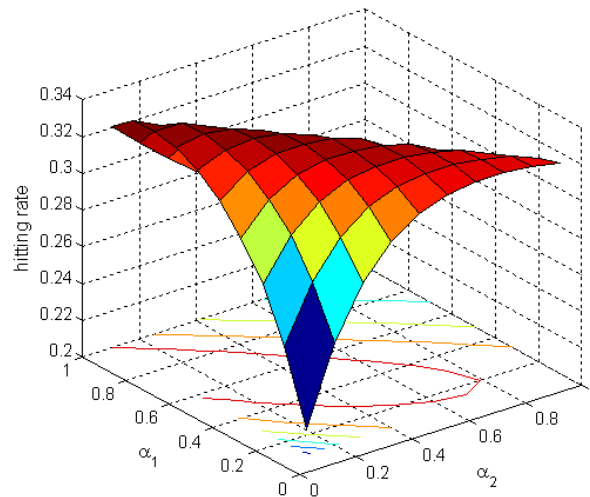


Fig.1 the impact of α_1, α_2 on hitting rate

By iterating W_a , when H_i get the maximum, $\Delta W_{a_i} = (0.792, 0.098, 0.112)$. The weights represent how important user-based, item-based and category-based recommendation are to BNPM.

To confirm the algorithm's effectiveness, comparison experiment on evaluation criterion is done using the other four recommendation algorithms and the new one. The evaluation criterions are ranking score, hitting rate, mean hamming distance S and average degree $\langle k \rangle$, and the other four recommendation algorithms are GRM, CF, NBI, CNBI [6]. The results are shown in Table 1.

Table 1 The four evaluation criterions of different recommendation algorithms

	GRM	CF	NBI	CNBI	BNPM
ranking score	15.23%	13.14%	9.74%	10.08%	9.48%
hitting rate	17.58%	23.60%	31.61%	30.31%	33.12%
S	53.27%	65.57%	71.89%	65.19%	77.13%
degree $\langle k \rangle$	347.58	261.87	285.82	277.05	66.275

Table1 reports the ranking score, hitting rate, meaning hamming distance S and degree $\langle k \rangle$ of different recommendation algorithms. Compared with GRM, CF, the ranking score of BNPM decreased 5.39%, 3.16% respectively. The hitting rate of BNPM, compared with all the other four algorithms, is improved 15.54%, 9.52%, 1.51% and 2.81%. The hamming distance S has raised 77.13%, 11.56%, 5.24% and 11.94%. At last, the average degree $\langle k \rangle$ has dropped 281.3050, 195.5949, 219.5122 and 210.7600. Comparative analysis of the above, the accuracy of BNPM is effectively.

Summary

In this paper, we propose a new personalized recommendation algorithm called BNPM to improve the accuracy and the diversity of the recommendation. BNPM combines three recommendation models: user-based, item-based and category-based recommendation model. User projection network and items' genre are taken into account in BNPM, so more information could be used for recommending. At last, we get the final prediction result by adaptive calculating the weight of the three dimensions. The experimental results show that the algorithm can improve diversity and reduce popularity on the base of holding the accuracy of the recommendation.

References

- [1] D.Jannach, M.Zanker, A.Felfernig, G.Friedrich, Recommender system: an introduction, Cambridge University, 2011

- [2] G.Linden, B.Smith, J.York, Amazon.com recommendations: item-to-item collaborative filtering. Internet Computing, IEEE, (2) 2003, Vol.7, Iss.1, pp76- 80
- [3] J.L.Herlocker, J.A.Konstan, J.Riedl, Explaining collaborative filtering recommendations, n Proceeding of the 2000 ACM conference on Computer supported cooperative work, 2000, pp. 241-250.
- [4] T.Zhou, J.Ren, M.Medo, and Y.C Zhang, Bipartite network projection and personal recommendation, Phys Review E, (12)2007, Vol.76, No.4:046115.
- [5] J.Liu, M.S.Shang, D.B. Chen, Personal recommendation based on weighted bipartite networks, In 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery: IEEE, 2009, pp134-137.
- [6] T.Zhou, L.L.Jiang, R.Q.Su, and Y.C.Zhang, Effect of initial configuration on network-based recommendation, Europhysics Letters, (3)2008, Vol.81, No.5:58004.
- [7] J.X.Xia, F.Wu, C.S.Xie and J.W.Tu, INBI: An improved network-based inference recommendation algorithm, In 2012 IEEE Seventh International Conference on Networking, Architecture and Storage, 2011.
- [8] Y.L Liu, The research of personalized recommendation system based on network structure, Dong Hua University, (2)2012.
- [9] Y.Zhang, Y.D.Liu, J.Zhao, Vector similarity measurement method, Journal of Technical Acoustics, (8)2009, Vol.28, No.4, pp532-536.
- [10] J.G.Liu, T.Zhou, Q.Guo and B.H.Wang, Overview of the evaluated algorithms for the personal recommendation systems, Journal of Complex systems and complexity science, (9)2009, Vol.6, No.3, pp1-12.