International Journal of Security and Its Applications
Vol.9, No.7 (2015), pp.99-108
http://dx.doi.org/10.14257/ijsia.2015.9.7.09

Collaborative Filtering Recommendation Algorithm based on Trust Propagation

Miao Duan

Jilin Jianzhu University, Jinlin 130118, china
279332269@qq.com

Abstract

Aiming at the problems that the existing model-based collaborative filtering algorithm has low recommendation accuracy and small recommendation coverage, we propose a collaborative filtering recommendation algorithm based on the trust propagation by introducing the trust information of social network to extend the matrix factorization-based recommendation model. We first design a set of trust propagation rules based on the direct trust relationships of the social network, so as to propagate the trust relationship in the social networks, and get to quantize the new trust relationship. Then we load the quantitative trust relations after the trust propagation as the trust weight into the matrix factorization-based model according to the characteristics that the matrix factorization technique can reduce the dimension of large-scale datasets.

Keywords: *Recommendation, Trust Propagation, Matrix factorization*

1. Introduction

The traditional collaborative filtering (CF) [1-3] recommendation algorithm regards item rated information by system users as the only information source, ignoring the trust relationship among users in the online social network. So the recommendation system can't recommend different merchandise items to target users according to different users having unlike social relationship [4-5]. Likewise in the reality, in the virtue network world, users have different social relationship. Their preferences and demands differ greatly. With popularity of online social network, it becomes a hot topic in the recommendation system field about how to use such trust relationship to solve problems faced by traditional collaborative recommendation methods [6-7].

In recommendation algorithms, matrix decomposition technic is used to deal with massive data set as to reduce the dimension [8-9]. Moreover, the technic has been widely applied in model-based collaborative recommend methods. Koren [10] et al. introduced a recommendation algorithm based on matrix decomposition, not considering the trust relationship among users. Salakhutdinov [11] et al. suggested similar method based on probability matrix decomposition, which neither considers the trust relationship among users. Ma et al. proposed on the basis of social network a collaborative recommendation method based on matrix decomposition. It incorporates the direct trust relations among users, but overlooks the transition of them, leading to narrow coverage of the algorithm [12-13].

To overcome them, we propose a collaborative recommendation algorithm based on trust propagation, TPMR in short. The major points to be concerned include: it brings up a set of trust propagation rules [14-15]; by using the direct trust relationship in social network, the indirect trust relationship among different users is obtained through certain transitive rules; then such relationship is quantified [16-17]; also it raises a collaborative recommendation algorithm of which the trust propagation weight is added in the matrix factoring collaborative recommendation model; the algorithm is evaluated in Epinions data set [18-19].

ISSN: 1738-9976 IJSIA

2. Related Definitions

Figure 1 is a social network, a non-binary trust network. In the paper it studies a binary trust network, that is, the trust value is not 0 but 1. In the network, there are six nodes, each node for a user. Each directed edge stands for one social trust relationship. Question mark in the picture means trust degree after trust transition between user u_1 and u_3.

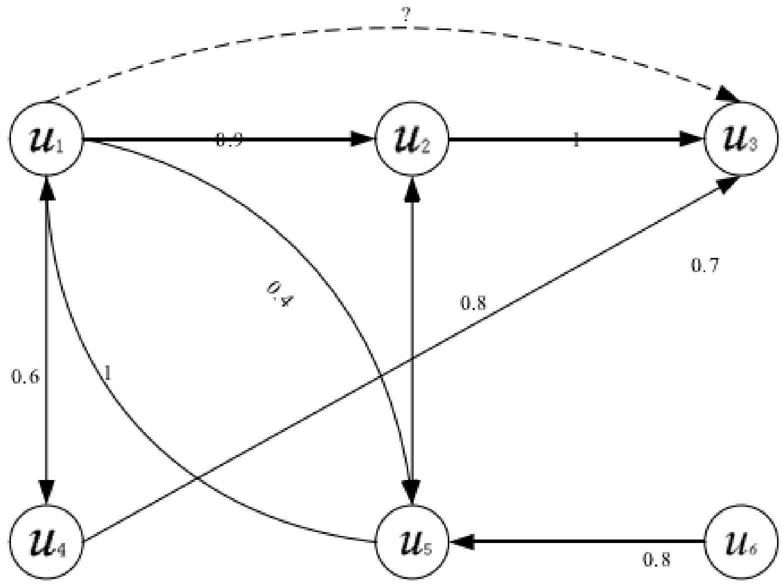


Figure 1. Social Network

Definition 2.1: hop k for trust transition: let $T=\{u_1,u_2,u_3,...,u_m\}$ be a trust network formed by user nodes; from T, any two nodes are selected; u_i and u_j stand respectively for current user node and target user node; then in the trust network T, one hop is from current user node u_i to the next node u_{i+1}; hops of trust propagating from u_i to u_j are marked k.

Definition 2.2: user's parent node u_F: in the trust network $T=\{u_1,u_2,u_3,...,u_m\}$, user node u_i in the previous hop of user u_{i-1} is named father node u_F of user u_i.

In a social network, have a set of direct trust neighbor set N_u for a per user u, Put trust value in trust matrix $T=[t_{u,v}]_{N\times N}$. Where, the non-zero elements $t_{u,v}$, said user u to user v trust value, the value is 0 or 1, where, 0 said no trust, 1 said trust.

3. Collaborative Filtering Recommendation Model and Algorithm based on Trust Propagation

The collaborative filtering recommendation model based on trust propagation involves the process of fusing into the recommendation the new trust relationship which derives through trust transition rules from the social and trust relationship among users. In other words, it utilizes direct trust relationship among users to elicit indirect trust relationship among new users, as to make current users match with more trust users and discover trust-ability among new users for eventually the improved recommendation accuracy of items.

The recommended model is divided into three parts: the original data processing, trust information recommendation, recommendation service.

The original data processing: this part is mainly to get the user-item rating data and user trust of user's rating data. First, for the system obtained data set is preprocessed, non-rating data needs to be quantified, score data needs to appropriate modifications, For example, deleted some aggressive attacker users data, the last obtained item-rating matrix and trust rating matrix.

Trust information recommendation: this part is mainly directed against the traditional collaborative filtering algorithms in the scoring matrix data are very sparse, and it is hard to calculate the similarity between users is proposed under the situation, because the trust degree and similarity are the same, can also be used as a recommendation algorithm right weight, So considered the use of trust to replace the traditional similarity as the weight of the final recommendation, the characteristics are:

One is the effective well integrated in trust relationship in recommended process;

The other is calculating transitive trust rating matrix after the user's trust score than the previous number more, no longer like the previous so sparse rating matrix.

Recommendation service: the main function of the part is the extraction of rating prediction data sets, generated recommendation list for the target user.

In addition, the model also contains one of the most cases, it is the target users in the user item-rating matrix directly score for the goal of the project, generated its recommended list.

3.1. Recommendation Algorithm based on Matrix Factorization

The entity set in the recommendation system includes user set $U=\{u_1,...,u_N\}$ and item set $V=\{i_1,...i_M\}$. Items usually refer to retrieval objects like movies, music, news, books etc. User's rating data of items is stored in rating matrix $R=[r_{ui}]_{M\times N}$, which is generally expressed with r_{ui}, meaning the known score of item i by user u. r_{ui} is an integral figure from 1 to 5, which indicates user's interest in merchandise items. The bigger the figure is, the more one user loves one item. $U\in R^{K\times N}$ is a user-item potential feature matrix. $V\in R^{K\times M}$ is user's potential feature matrix. Column vector u_u and v_i refers to separately K-dimensional user u and item i's potential feature vector. The predicted rating $\hat{r}_{ui}$ can be reached by:

$$\hat{r}_{ui}=u_u^T v_i \tag{1}$$

The calculation of vector u_u and v_i, defines the method of least squares:

$$u^*,v^*=\arg\min\sum_{r_{ui>0}}(r_{ui}-u_u^T v_i)^2+\lambda(\|u_u\|^2+\|v_i\|^2) \tag{2}$$

Equation 2 can be effectively optimized by stochastic gradient descent method,

Constantly updated parameters for a given r_{ui} along the gradient in the opposite direction, until the two time function value is less than ε stop. ε is the stop condition. We got the following formula:

$$\begin{aligned} u_u &\leftarrow u_u+\gamma(v_i e_{u,i}-\lambda u_u) \\ v_i &\leftarrow v_i+\gamma(u_u e_{u,i}-\lambda v_i) \end{aligned} \tag{3}$$

3.2. Recommendation Algorithm based on Trust Propagation Mechanism

The idea of the recommendation algorithm proposed in the paper includes: by trust transitivity, we execute trust propagation algorithm for the sparse trust network to get indirect trust relationship among users; then by factoring such trust relationship, we combine it into the collaborative recommendation model based on matrix decomposition in order to optimize target function by means of random gradient descent algorithm. It is illustrated as follows:

3.2.1 Measurement of Trust: The degree of trust between users is trust-ability. Before trust factors introduced to the collaborative filtering recommendation method, the primary task is to manage quantitatively trust relationship as thus to use in the calculation formula.

3.2.2The Rating Prediction Formula of Trust Relationship: b_{ui} is a reference offset, said user project preference related terms, the formula is as follows:

$$b_{ui} = \mu + b_u + b_i \tag{4}$$

b_u , b_i expressed on user u preference and commodity i preference; μ represents the average score of project matrix in the whole scoring matrix.
By the least square method to calculate b_u and b_i formula is as follows:

$$\min_{b*} \sum_{(u,i\in h)} (r_{ui} - \mu - b_u - b_i)^2 + \lambda_1 (\sum_u b_u^2 + \sum_i b_i^2) \tag{5}$$

Where, $b*$ said deviation of all users and projects score value.

3.2.3 The Algorithm Description: In summary, collaborative recommendation algorithm based on trust propagation (TPMR) is given. The algorithm description is as follows:

TPMR recommendation algorithm

Input: $R, T, k, m, n, \gamma, \lambda$
Output: $\hat{r}_{ui}$
(1) For u=1 to m Do
(2) $\text{Trust}_u \leftarrow$ callTPA(u, T, k)
(3) Initialize p_u, z_v with random values
(4) $t_{uv} \leftarrow p_u^T z_v$
(5) Compute $p_u, z_v \in R^f$
(6) For iterations=1 to 15 Do
(7) For i=1 to n Do
(8) $v_i \leftarrow \sum_{v\in R(i)} (r_{vi} - b_{vi}) z_v$
(9) $sum \leftarrow 0$
(10) For each u $\in$ R (i)Do
(11) $\hat{r}_{ui} \leftarrow \mu + b_u + b_i + p_u^T v_i$
(12) $e_{ui} \leftarrow r_{ui} \leftarrow \hat{r}_{ui}$
(13) $sum \leftarrow sum + e_{ui} . p_u$
(14) For each u $\in$ R (i)Do
(15) $z_u \leftarrow z_u + \gamma (r_{ui} - b_{ui}) . sum - \lambda . z_u$
(16) End For
(17) End For
(18) End For
(19) End For
(20) $\hat{r}_{ui} \leftarrow \hat{r}_{ui} + (\mu + b_u + b_i + p_u . \sum (z_u))$
(21) End For
(22)Return

4. Experimental Analysis and Results

4.1. Experiment Data Set

The experiment adopts Epinions data set, collected from a famous e-commerce merchandise rating website. The data set statistics are listed in Table1, inclusive of two data sets frequently used by the recommendation system. Table1 shows the quantity of rated trust relationship among users. It can be deferred that the coverage of trust appraising is below 1%. The numerical score of user's rating about merchandise items is integral figure from 1 to 5. Such scores represent user's different fondness of items. Trust relationship among users can be rated 1 or 0. 1 means belief between two users, while 0 means no trust.

Epinions data set used in the experiment is rather sparse in terms of either user-item rating data or user-user trust relationship rating data. According to statistics, 48.4% of users in the set have less than five rating records, the density of rating matrix below 0.015%; 52.2% of user trust relationship rating records is below 5, the coverage rate of trust rating less than 1%. From Table1 it's noted that Movielens and Eachmovie are two most well-known data sets for the collaborative filtering recommendation system. Also based on scores of items, we can get scoring density of the two sets: 4.25% and 2.29%. But the experiment used Epinions data set instead of Movielens or Eachmoive data set because the latter two don't have user-user trust rating information.

Table 1. Data Set Description

Data Sets	Users	Items	ItemsRatings	TrustRatings
Epinions	40100	149856	663478	377191
Movielens	5060	3800	1111309	-
Eachmovie	68834	1578	3811789	-

4.2 Test Environment Configuration

This experiment mainly using Java language in this paper. In Eclipse platform to achieve. The experimental configuration specific includes two parts of hardware configuration and software configuration.

(1) Hardware configuration: Intel Core Duo processor, 4G memory, 500G hard disk.
(2) Software configuration: the development tools of Eclipse7.0, the compiler environment is Jdk1.6.0, the operating system is win7.

4.3. Evaluation Index in the Experiment

To assess the quality of the recommendation algorithm, we introduce two key indicators:

4.3.1 Accuracy Rate: Root mean square error (RMSE) is the most common indictor for assessing accuracy rate of the recommendation method. It's used to measure the square deviations between the assumed rating value and user's actual rating value. If one user's comments are known about some items, the standard is adopted. Here we use it to evaluate the precision rate of the recommendation method. By calculating the difference between user's predicted score and the actual one, RMSE examines the accuracy rate of such prediction. The smaller the RMSE is obtained, the higher the accuracy rate implies and the better quality the recommended service proves.

RMSE Formula is defined:

$$RMSE = \sqrt{\frac{\sum_{(u,i)|R_{test}} (r_{ui} - \hat{r}_{ui})^2}{|R_{test}|}} \tag{6}$$

4.3.2 Coverage Rate: The coverage rate is the percentage of the number of successful predicted scores by the recommendation system against that of all ratings in the whole test set. If the coverage rate is lower, the quantity of items recommended by the system is fewer and that the user gets worse recommendation results. Otherwise, the user gets better recommendation results because of higher coverage rate and more items recommended by the system. The formula is as follows

$$COV_R = \frac{N_d}{N} \tag{7}$$

4.4. Validation of the Proposed Algorithm

4.4.1. Results and Analysis: To validate the performance of rating algorithm, we make comparison between the proposed algorithm and the following ones:

(1) BaseMF: Recommendation method based on matrix decomposition, ignorance of trust relationship among users;
(2) RSTE: Probability matrix decomposition model fused with trust relationship, without account of trust transition;
(3) TidalTrust: Recommendation method applying TidalTrust model;
(4) MoleTrust: Recommendation method applying MoleTrust model

Epinions data set used in the experiment is randomly and averagely divided into 80% training set and 90% training set. Considering the time consumed in each iteration and recommendation result, we set 15 iterative times. Step length factor γ and regular parameter λ are acquired through cross validation. For all tests in this part, make $\gamma = 0.005, \lambda = 0.003, f = 100$, and the propagation step length $k = 4$ as experiment basic. All test results we reached are under the basic: $k = 4, f = 100$. Repeat the test for 5 times.

Results of root mean square error of the above algorithms are seen in Figure 2.

In the experiment were calculated coverage, used the Epinions data set of 80% training set segmentation, the experimental results are shown in Table 2.

It can be seen from Table 2, the TPMR method proposed has higher score coverage, the ratio of BaseMF, RSTE, TidalTrust and MoleTrust methods respectively increased by about 2.4 times, 0.7 times, 0.06 times and 0.09 times. Recommendation system coverage rate is high, it means that user selects the quantity of the goods more, the user's satisfaction is higher, in recommender systems of trust propagation maked the system recommended coverage is wide, so compared with other methods, the method presented in this paper has higher practical value.

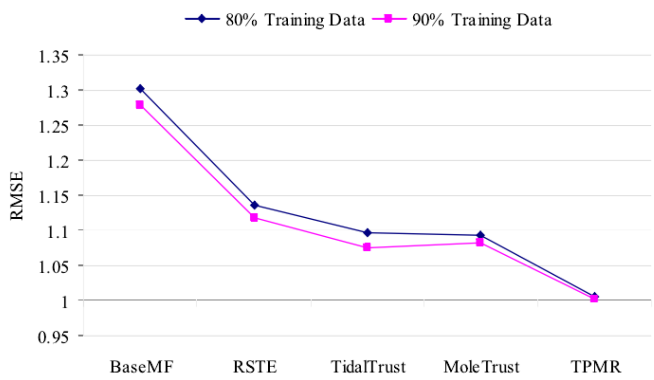


Figure 2. RMSE Value of Different Recommendation Approaches in Different Training Set

Table 2. Score Coverage

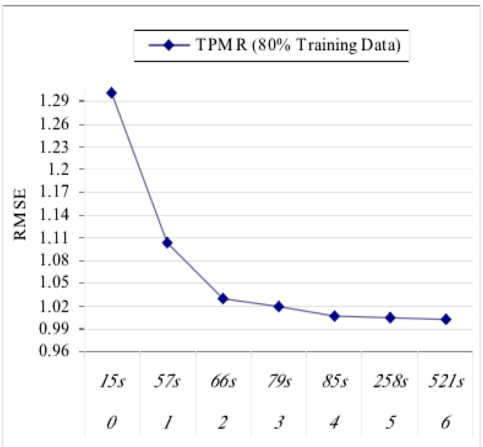
Method	COV_R
BaseMF	20.1%
RSTE	49.7%
TidalTrust	83.6%
MoleTrust	81.2%
TPMR	90.5%

From validation results, it's obvious that in Figure 2 and Table 2, the proposed TPMR has no better effect on RMSE than others. Put it in another way, the proposed algorithm didn't achieve higher accuracy rate of recommendation and the recommendation service quality was not improved effectively. However, the new method is superior to others as far as the coverage rate of recommendation is concerned. For consideration of the two aspects, TPMR is much improved. On the regard, when one recommendation algorithm is evaluated, it's advisable to measure by focusing on one or more assessment indexes.

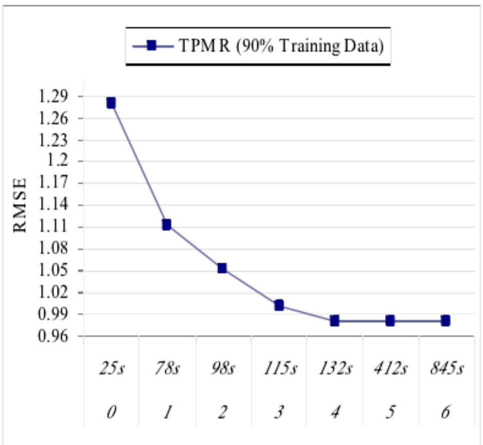
4.4. Validation of Parameter k's Influence on Experiment Results

For the goal, we define parameter k mentioned in Definition 2.1 as hop of trust transition. Parameter k is used to control the size of neighboring collections. It is an important value of trust transition to the validation here. Figure 3(a) and (b) display iteration time and precision achieved by the algorithm for different values of k in different data sets. We describe it here: when k=0, the method in the part approximates to BaseMF in the above experiment, which considers no trust relationship but employs only matrix factorization recommendation technics. With k becoming bigger and bigger, it means user's neighbor of bigger and bigger scale. When k is maximal, according to Six Degrees of Separation, the entire social network is user u's neighbor and the recommendation precision reaches the best. But too bigger size of the neighbor will affect the model's computational time. So in the experiment, k is selected dependently on the recommendation efficiency and iteration time.

Figure 3 shows the TPMR method proposed, we choose 80% training set and 90% for the training set with different values of hop k. We use recommendation accuracy and iteration time in the algorithm, it is shown in Figure 3.



(a)



(b)

Figure 3. k Takes Accuracy of Different Values and the Iteration Time

It can be seen from Figure 3, whether in the training set of 80% or 90% in the training, the experimental results are obtained. With the increase of the k value, RMSE value reducing continuously, it said that recommendation accuracy increases constantly. When k is in the interval [0,4], RMSE decreases greatly. But the magnitude of iterative time is not large. When the interval k=4, the magnitude of RMSE tends to be flat, or iterative time variation amplitude is big. RMSE values and the iterative time are obtained in weigh experiment, it can be obtained when k=4 proposed to obtain the best performance of the recommendation algorithm.

5. Conclusion

This paper proposes a collaborative filtering recommendation algorithm based on trust propagation. Firstly, according to the propagation characteristics trust in social networks, presents a set of trust propagation rule, obtains the user between the indirect trust degree by direct trust degree calculation in the network, and to quantify the indirect trust. Secondly, simply introduce the matrix decomposition algorithm. Finally, proposed a kind of the trust propagation of the collaborative filtering recommendation model based on matrix decomposition. By the stochastic gradient descent method to optimize the objective function, it effectively improve the recommendation quality and recommend coverage.

References

[1] J. Liu, T. Zhou and B. Wang, “Research progress in natural science”, personalized recommendation system, vol. 19, no. 1, (**2009**), pp. 1-15.
[2] H. Ma, I. King and M R. Lyu, “Learning to recommend with social trust ensemble”, InSIGIR, (**2009),** pp. 203–210.
[3] H. Yuan, T. Zhou, Y. Han and Y. Chen, “The collaborative recommendation algorithm based on MI clustering”, Journal of Wuhan University (Information Science Edition), vol. 2, (**2015**), pp. 253-257.
[4] Z. Cai and L. C. Chung, “Trust oriented social commerce collaborative filtering recommendation method”, Computer applications, vol. 1, (**2015**), pp. 167-171.
[5] J. Zhongtao, “Design and implementation of collaborative filtering algorithm based on the film of personalized recommendation”, Software guide, vol. 1, (**2015**), pp. 86-88.
[6] Q. Fang and J. Du, “Research on collaborative filtering recommendation trust model considering P2P environment”, Tech Information Development economy, vol. 1, (**2015**), pp. 122-124.
[7] X. Bi, “Project similarity fusion collaborative filtering recommendation algorithm”, The application of computer system, vol. 1, (**2015**), pp. 147-150.
[8] L. Li, H. Zhang, Z. Li and J. Chen, “Application of collaborative filtering recommendation algorithm fusion user attributes in government procurement”, Journal of Chongqing University of Technology (NATURAL SCIENCE), vol. 1, (**2015**), pp. 76-81.
[9] L. Ke and J. Wang, “A collaborative filtering recommendation based on user migration characteristics”, Computer Engineering, vol. 1, (**2015**), pp. 37-43.
[10] Y. Koren, R. Bell and C. Volinsky, “Matrix factorization techniques for recommender systems”, IEEE Computer, vol. 42, no. 8, (**2009**), pp. 30–37.
[11] R. Salakhutdinov and A. Mnih, “Probabilistic matrix factorization”, InNIPS, vol. 20, (**2008**), pp. 1257-1264.
[12] Y. Xu and C. Liu, “The user's interests and recommend trust domain based recommendation”, Micro-blog Telecom Science, vol. 1, (**2015**), pp. 13-20.
[13] H. Rong, H. Xu Chunhua and J. Mo, “Collaborative filtering recommendation algorithm based on user similarity”, Journal of China Institute of communications, vol. 2, (**2014**), pp. 16-24.
[14] Z. Wang and M. Wang, “Collaborative filtering recommendation algorithm based on user behavior model and based on ant colony clustering”, Micro computer application, vol. 3, (**2014**), pp. 5-8.
[15] Y. Du, L. Huang and M. Ho, “A collaborative filtering recommendation method fusion trust computing”, Pattern recognition and artificial intelligence, vol. 5, (**2014**), pp. 417-425.
[16] Z. Li and G. Xu, “A collaborative filtering recommendation algorithm, user based on spectral clustering”, Computer technology and development, vol. 9, (**2014**), pp. 59-62+67.
[17] H. Li, Q. Xie and Q. Ding, “A collaborative filtering recommendation algorithm based on scenarios”, Computer technology and development, vol. 10, (**2014**), pp. 42-46.
[18] X. Bi, “Collaborative filtering recommendation algorithm based on user interest density distribution oriented”, The application of computer system, vol. 12, (**2014**), pp. 115-119.

[19] T. Liang, Y. Liang, J. Fan and J. Zhao, "A collaborative filtering recommendation algorithm based on user interest model", Computer applications and software, vol. 11, (**2014**), pp. 260-263.

Author

Miao Duan, She received her B.S degree from Jilin University, and received her M.S degree from Jilin University. She is a lecturer in Jilin Jianzhu University. Her research interests include Network intelligent information processing.

.