IMPROVED SPECTRAL CLUSTERING ALGORITHM AND ITS APPLICATION IN RECOMMENDER SYSTEM

Xian Li [1,2] and Zhijia Wang[1*]
[1] College of Computer and Information Technology Engineering, Hohai University, Nanjing, 211100, China
[2] Faculty of Computer and Software, Huaiyin Institute of Technology, Huaian, 223003 China

Abstract

For the problems of collaborative filtering recommender algorithm, such as it is greatly influenced by the sparse rating data, the data clustering pretreatment is easily trapped in the local optimum in the non-convex sample space, etc, we propose an improved spectral clustering algorithm to optimize the recommender system. Firstly, this method improves the standard spectral clustering algorithm based on the feature difference and orthogonal feature vector, and then the clustering number will automatically determine. Secondly, it uses the improved spectral clustering algorithm to cluster the user and item of the original rating matrix. Thirdly, it fills the missing value for the clustered rating matrix. Finally, it recommends new items for users. By the simulation experiment on Epinions and MovieLents data sets, the results show that this method can effectively alleviate the data sparseness, and improve the prediction accuracy and generalization ability of recommender system.

Keywords: Spectral Clustering; Recommender System; Collaborative Filtering; Matrix Decomposition.

1. INTRODUCTION

With the popularization and development of the Internet; mobile Internet; Internet of things and other information technology; the information resources are growing rapidly at an exponential speed. The large growth of information makes people unable to get meaningful information timely and accurately from this information in the face of huge amounts of information; so it reduces the utilization of information; this is information overload problem [1]. In order to solve the information overload problem; the information retrieval system represented by search engines and personalized recommender system stands out [2].

Based on the users' historic behavior data; the personalized recommender system analyzes the users' interest characteristics and builds the personalized model of users; and thus effectively recommends users the information that meets their own needs [3]. Therefore; the recommender system is widely used in electronic commerce; social networks; video on demand; and other fields; such as the Amazon; Facebook; YouTube; etc. From the view of information filtering; the traditional recommender system is mainly divided into the collaborative filtering recommender system; content-based recommender system and mixed recommender system. With the rapid development of mobile terminal equipment; the context-aware recommender system appears [4].

Collaborative filtering recommendation is the most widely used and the most successful personalized recommendation technology at present; which has greatly promoted the research on recommender system [5]. Collaborative filtering recommendation obtains the historic browsing history or purchasing records of users and completes the recommendation by an implicit way; the whole process does not need to explicitly ask users to provide their own interest preferences; such as filling in the questionnaire or checking the preferences category when users register; etc [6]. Another advantage of collaborative filtering could deal with unstructured data;
such as music; movies; products and other complex items; but not restricted to recommended items. However; the collaborative filtering recommendation has some disadvantages; such as sparsity; extensibility; cold start and so on [7].

In order to improve the recommendation quality of collaborative filtering recommender method; the researchers improve it with the aid of dimension reduction; matrix factorization; association rule mining; probability-based analysis and other methods and models. In matrix factorization techniques; the Principal Component Analysis and Singular Value Decomposition methods are representative. These two techniques have been proved to effectively improve the prediction accuracy of recommender system in the Nexflix competition of 2009. Sarwar; et al [8] compared and analyzed the impact of SVD-based dimension reduction on recommendation quality by using SVD dimension reduction techniques and the existing collaborative filtering method in two kinds of database. Research shows that in some settings; the dimension reduction techniques can filter out the noise in some data; which showed a better recommendation quality; however; it was the opposite in some cases. The reason is that. The reason is that the recommendation quality depends on correctly selecting the number of singular values; which are kept in SVD method [9]. Koren; et al [10] proposed a model called Time SVD++ to handle the rating data in high dimension situation; and on the Netflix data set they validated the model and other matrix decomposition models based on root mean square error as the evaluation index. The experimental results show that the model has the optimal recommendation performance. Baltrunas L; et al. put forward three different kinds of context-aware recommendation model [11] based on matrix decomposition technique to cope with different situations. Tensor Factorization as the extension of matrix decomposition on the multidimensional data has been widely applied in the context-aware recommender system (CARS) [12]; Tensor Factorization can not only effectively discover the potential correlation relationship among multidimensional data; but also alleviate the existing sparseness in multidimensional data. Wang Z proposed to solve the sparseness combined with label-based neighborhood algorithm and rating-based neighborhood algorithm [13]. The label in the text is created by online users to express their preferences for products.

In addition; clustering is widely used in the rating matrix of row and column vector classification; but one of the major problems in cluster analysis is the determination of the number of clusters in unlabeled data; which is a basic input for most clustering algorithms. Huang; et al [14] proposed an automatic clustering number determination for the classical FCM (Fuzzy C-Means) algorithm. The proposed automatic clustering number determination is based on the cardinality of clustering fuzzy membership used in the CA (Competitive Agglomeration) algorithm. D. Sharmilarani; et al[15] proposed a new method for automatically estimating the number of clusters in unlabeled data sets; which is based on an existing algorithm for Spectral Visual Assessment of Cluster Tendency (SpecVAT) of a data set; using several common image and signal processing techniques. Its basic steps include: generating a VAT image of an input dissimilarity matrix; constructing Laplacian matrix; normalizing the rows and applying SpecVAT.

In this paper; we proposed the recommender algorithm based on improved spectral clustering. Based on feature difference and orthogonal feature vector; this method improve the criterion clustering algorithm and automatically determine the clustering number. We use the improved spectral clustering algorithm to cluster users and items of the original rating matrix. Finally we fill the missing values and recommend new items to users. By the simulation experiment on Epinions and MovieLens data sets (1M); we prove the effectiveness of this method.

1. BACKGROUND AND RELATED WORK

2.1 Collaborative Filtering Recommendation

The core of the collaborative filtering recommendation is using the specific items' evaluation from similar users of target users to generate the item' evaluation prediction of this user [16]. For example; we assume that the historical behaviors and preferences of user A and user B are very similar; such as watching the same movie. The user A recently saw a movie that B hadn't seen; then the principle based on collaborative filtering is to recommend the user B this movie. The user A and user B have similar historic behaviors and hobbies; so that the system will naturally recommend the similar user B the item that the user A purchased or liked; and when choosing a movie we filter out what we are most likely to be interested in from a large number of movies; and in the process of producing recommendations; the user collaborates with other users implicitly; so this technique is called the collaborative filtering recommendation [17-19]. The idea of collaborative filtering recommendation is easy to understand; in daily life people often refer to the recommendations of others; and the collaborative filtering introduces this idea into the personalized recommendation; that is to recommend the target users by referring to the rating of similar users on a particular item [20-21].

The general process of collaborative filtering recommendation is mainly divided into three steps [22].

(1) Collect the historic behavior or interest preference data of each user.

(2) Find similar users or items by calculating the similarity.

(3) Recommend by using similar users or items.

2.2 Standard Spectral Clustering Algorithm

Spectral clustering algorithm is a high-performance computing method; which has received lots of attention in recent years. Its main idea is derived from the spectrum graph partition theory; and converts the clustering problem to the optimal partitioning problem of graph. It maximizes the similarity among subgraphs and minimizes the similarity between graphs [23; 24]. The dividing criterion maximizes the internal similarity between subgraphs and minimizes the similarity. Considering the continuous relaxation form of a problem; we can convert the graph partition problem into the spectrum decomposition of solving similar matrix or Laplacian. We can think of the spectral clustering as the approximation of the graph partition criterion [25]. Common rules of spectral clustering partition include Minimum Cut; Normalized Cut; Min-max Cut; and Ratio Cut [26].

According to different criterion function and spectrum mapping method. The standard spectral clustering algorithm has many implementation methods; and the implementation process is generally divided into three steps.

(1) Define the similarity measurement between data sample points; and establish the similar matrix between data points.

(2) Build a new data feature space by calculating the first *k* eigenvalues and eigenvectors of similarity matrix.

(3) Use the K-means algorithm or other traditional clustering algorithms for clustering the feature vector in feature space.

2. IMPROVED SPECTRAL CLUSTERING RECOMMENDER ALGORITHM

3.1 Improved Spectral Clustering Algorithm

In the spectral clustering algorithm; the selection of eigenvalue and eigenvector has a great influence on the clustering results; while the selection of the eigenvalue and eigenvector is based on clustering number [27]. Therefore; this article puts forward to improve the standard spectral clustering algorithm by using the improved spectral clustering algorithm based on the feature difference and orthogonal feature vector; and thus automatically determines the clustering number. The basic idea of the improved spectral clustering algorithm is as follows: firstly; to construct similar matrix by using the sample data; make a spectral decomposition for the standardization similarity matrix generated by similarity matrix; and get the corresponding eigenvalue and eigenvector; secondly; to arrange the characteristic value in descending order and describe the difference between adjacent eigenvalues by using a proper clearance; to automatically determine the number of classes through the position where the maximum eigen clearance occurred; Finally; to realize the data classification by combining the class number and the angle between eigenvectors.

Set the input data sets as the user number M and the item number N; the output is the clustering number k and the clustering results.

The specific steps of the improved spectral clustering algorithm are as follows.

eigenvector [28] will be. At this point; each row in the matrix H is used as a point in k dimensional space to form k clusters. They are distributed orthogonally each other on the unit ball in k dimensional space; and these k clusters formed on the unit ball correspond to the formed k clusters of all points in the original space. Thus; the clustering number is determined according to the eigenvalue difference of Laplacian matrix; where the difference is [0.06; 0.1] [29]. Compared with the standard spectral clustering algorithm; the improved spectral clustering algorithm can automatically determine the clustering number; establish a standardized similar matrix and spectral decomposition for sample data. It uses eigengap to automatically determine the clustering number of sample data; and realize the classification of the sample data according to the determined clustering number and the angle between eigenvectors after spectral decomposition.

3.2 Steps of Recommender Algorithm Based on Improved Spectral Clustering

In this paper; the recommender algorithm based on improved spectral clustering is divided into two stages. The first stage is the improved spectral clustering of users; and the second stage is the improved spectral clustering of items.

Input: the original rating matrix

Output: the rating matrix after bi-directionally improved spectral clustering

Step 1 The rating matrix improves the spectral clustering based on user (row data). See the improved spectral clustering algorithm in section 2.3 for detailed steps of improving spectral clustering.

Step 2 The rating matrix improves the spectral clustering based on item (column data). See the improved spectral clustering algorithm in section 2.3 for detailed steps of improving spectral clustering.

Step 3 Fill the missing values for the rating matrix after bi-directionally improved spectral clustering.

Step 4 Output the rating matrix after the completed bi-directionally improved spectral clustering.

3.3 Application Examples of Recommender *Table 1 are the ratings of 8 users for 8 movies; 8 movies are respectively "Pearl Harbor"; "Dolphins"; "Saving Private Ryan"; "Warcraft"; "The Alps Climb of Your Life"; "Bean: The Ultimate Disaster Movie"; "Jurassic World"; "Ted". They are respectively expressed with the letter a-h; where a and c are war movies; b and e are documentary films; d and g are science fictions; f and h are comedies.*

Table 1 Rating of User for Movies

Users	a	b	c	d	e	f	g	h
User 1	?	5	3	2	5	1	2	?
User 2	5	3	5	1	3	2	?	2
User 3	3	5	?	?	?	?	?	1
User 4	2	1	?	5	1	3	5	3
User 5	4	3	4	2	?	?	?	5
User 6	2	?	2	?	1	?	5	3
User 7	?	3	?	2	3	5	2	5
User 8	?	?	5	?	3	2	?	?

Fig. 1 is the improved user spectral clustering; the improved item spectral clustering; and the improved user - item bi-directional spectral clustering and schematic diagram. Each row of the matrix represents a user; and each column represents an item. The improved bi-directional spectral clustering in the process of clustering and column repeats iteration until convergence; which can assign each row and each column of the matrix to different rows and columns for clustering. data subset. Fig. 2 is the schematic diagram of the results that the movie rating matrix in Table 1 uses the improved user spectral clustering; improved item spectral clustering and improved user-item bi-directional spectral clustering proposed in this paper.

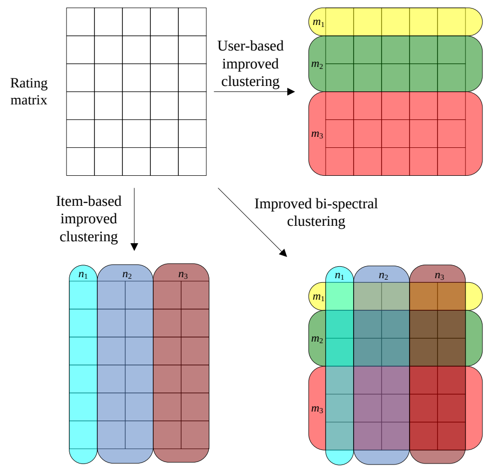

Fig. 1. Schematic Diagram of Improved Bi-Spectral Clustering; User-Based Improved Clustering; and Item-Based Improved Clustering

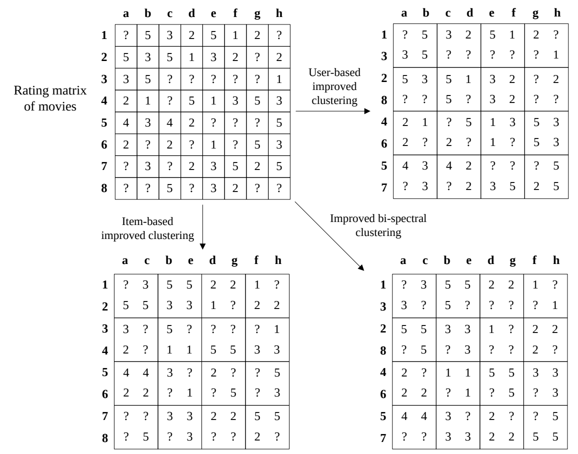

Fig. 2. Improved Spectral Clustering of Movie Rating

From Fig. 2; we can see the improved user spectral clustering and the improved item spectral clustering have achieved a certain effect of clustering based on the original rating matrix; which have made respectively the users and items gathered correspondingly; while the improved user–item bi-directional spectral clustering has the best effect; making the users and items gathered.

3. EXPERIMENT AND RESULT ANALYSIS

4.1 Datasets and Metrics

The experiment selected two real data sets of Epinions and MovieLens (1M) for simulation experiment. Epinions dataset (http://www.epinions.com) contains 49290 users; 139783 items; 664824 ratings and 487181 friendship data on the online service website *epinions.com*. MovieLens dataset (http://movielens.umn.edu) was collected by the computer science and engineering GroupLens item team from the University of Minnesota college by ratings of a large number of users on the MovieLens website; the rating grade is 1-5; where 5 expresses the most favorite and 1 expresses the least favorite; the users express their hobbies by rating values. MovieLens (1M) contains 1 million ratings from 6000 users on 4000 movies.

In this work; we select the Normalized Discounted Cumulative Gain (NDCG) [29] and Expected Reciprocal Rank (ERR)[30] as evaluation indexes. NDCG measures the performance of a recommendation system based on the graded relevance of the recommended entities. It varies from 0.0 to 1.0; with 1.0 representing the ideal ranking of the entities.

$$DCG_k = \sum_{i=1}^{k} \frac{2^{rel_i} - 1}{\log_2(i+1)}$$

Where k is the maximum number of entities that can be recommended and rel_i is the graded relevance of the result at position i .

$$NDCG_k = \frac{DCG_k}{IDCG_k}$$

Where $IDCG_k$ is the maximum possible DCG for a given set of queries; documents; and relevances.

The ERR is defined as a cascade based metric that using function $\varphi(r) = 1/r$; where φ is a utility function; a cascade based metric is the expectation of $\varphi(r)$.

$$ERR = \sum_{r=1}^{n} \frac{1}{r} P_r = \sum_{r=1}^{n} \frac{1}{r} \prod_{i=1}^{r-1} (1 - R_i) R_r$$

Where P_r stands for the probability that the user stops at position r ; and n denotes the number of documents in the ranked list; R_i denotes the document i satisfies the user with probability R_i .

In the experiment; we forecast users' rating for item in the test dataset; and compare by calculating the values of NDCG@5; NDCG@10; NDCG@20 and ERR@5; ERR@10; ERR@20.

4.2 Contrast experiment and result analysis

Different selection of experimental parameters will affect the result of the experiment. In order to make the experiment has a better representative; the training data set randomly selects 60% and 80% two ratios. The spectral clustering partitioning criteria selects the Normalized Cut and Ratio Cut methods. The potential feature dimensions of matrix decomposition *D* take respectively 8 and 16 two dimensions. In order to compare the performance of the BSCRM method in this paper; we select User-based Nearest Neighbor Recommendation (UBNNR) [31]; Item-based Nearest Neighbor Recommendation (IBNNR) [32]; Weighted IBNNR (WIBNNR) [33]; Soft Margin Ranking MF (SMRMF) [26]; Quadratic Matrix Factorization (QMF) [34] for comparison and Matrix Factorization(MF)[35]; Biased Matrix Factorization (Biased MF)[36] as reference lines. All these above methods are based on the matrix decomposition; and ignore the conditions of preference.

User-based on his Neighbor Recommendation (UBNNR) is given a user-item rating matrix to find out other users; who had similar preferences with current users in the past. That is the process of looking for neighbors. For the item that the current users have not seen; we use the historic rating of user's neighbors for the item to calculate the predictive value of users for the preference degree of items.

Item-based Nearest Neighbor Recommendation (IBNNR) is to calculate the similarity of items based on users' historical data. Using the similarity between items replaces the similarity between users; and then recommend users the items that are very similar with the preferential items of users.

Matrix Factorization (MF) is to estimate rating by using the inner product of potential users' eigenvectors and potential item eigenvectors.

Biased Matrix Factorization (Biased MF) is adding the offset value into the decomposition of the extended matrix.

Quadratic Matrix Factorization (QMF) uses the quadratic polynomial approximation condition; literature [37] prove that the expressive ability of quadratic polynomial is stronger than that of linear function for conditional preference and propose to use quadratic polynomial to approximate conditional preference. Plackett–Luce model based permutation probability and cosine based permutation probability are used; which are denoted as

QMF-PL and QMF-COS in the comparison; respectively.

Soft Margin Ranking MF (SMRMF) changes the objective function of MMMF [38] to the ordinal regression score.

In the experiment; in addition to the proposed BSCRM and WIBNNR; QMF methods; other methods participating in comparison are implemented in the reference [39] and the default parameters of methods participating in comparison is set as the optimal value described in of the original documents.

Weighted IBNNR (WIBNNR) extends IBNNR by giving weight for each pair of comparison in training data.

The experiments composed of different parameters are repeated five times; and the experimental result is the average result of five times. In the experiment; the mean and standard deviation of NDCG and ERR are shown in Fig. 3 and Fig. 4. The values of NDCG @20 and ERR @ 20 are shown in Table 2 and Table 3.

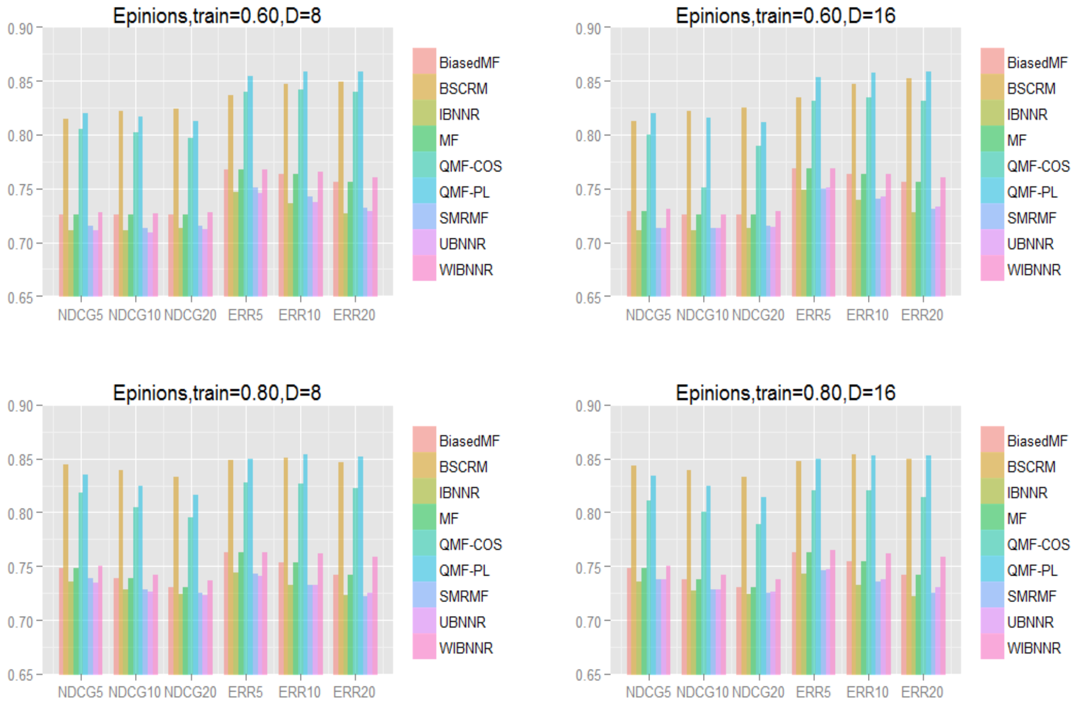

Fig. 3.Comparison Results on Epinions Dataset

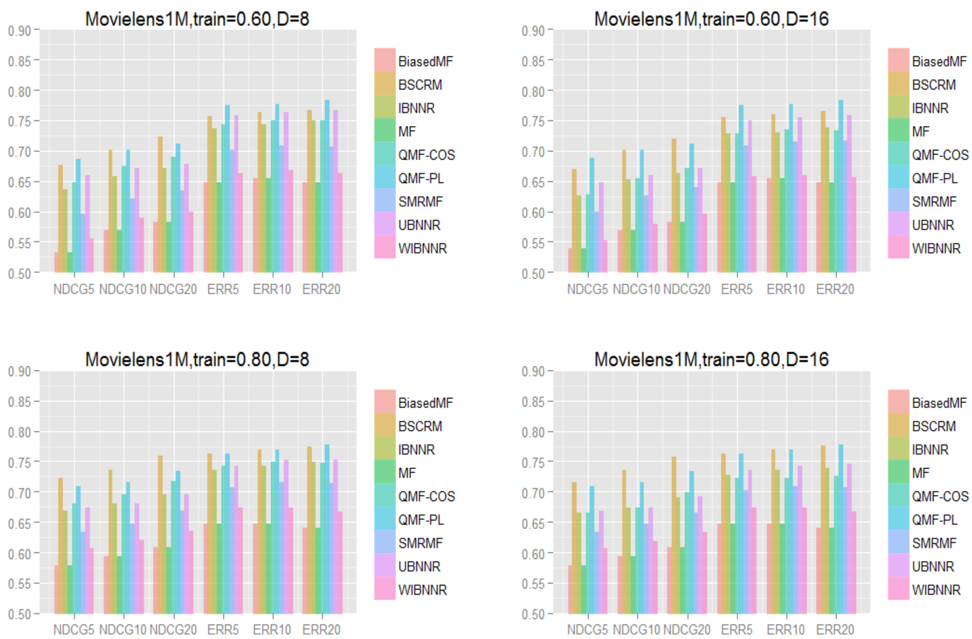

Fig. 4.Comparison Results on Movielens (1m) Dataset

From Fig. 3; Fig. 4 and Table 2; Table 3; we can see the proposed recommender algorithm based on improved spectral clustering has a higher forecasting accuracy; more stable forecasting and stronger generalization ability. In the case of ignoring condition preferences; the method proposed in this paper is superior to MF; BiasedMF; UBNNR; IBNNR; WIBNNR and SMRMF in all selected data sets of experiments. The condition preferences have widely existed in real life [40]; so that if we consider the setting of condition preferences; this method can obtain better recommendation results.

Fig. 5 is the data set iteration process of Epinions and Fig. 6 is the data set iteration process of Movielens1M.

Table 2 Comparison Results on Epinions Dataset

Method	Train=0.60				Train=0.80			
	D=8		D=16		D=8		D=16	
	NDCG@20	ERR@20	NDCG@20	ERR@20	NDCG@20	ERR@20	NDCG@20	ERR@20
MF	0.7254	0.7557	0.7254	0.7557	0.7302	0.7423	0.7302	0.7423
BiasedMF	0.7254	0.7557	0.7254	0.7557	0.7302	0.7423	0.7302	0.7423
UBNNR	0.7124	0.7283	0.7144	0.7332	0.7230	0.7253	0.7259	0.7306
IBNNR	0.7127	0.7269	0.7127	0.7276	0.7241	0.7230	0.7236	0.7215
WIBNNR	0.7275	0.7604	0.7283	0.7604	0.7367	0.7581	0.7375	0.7585
SMRMF	0.7150	0.7317	0.7148	0.7305	0.7249	0.7215	0.7253	0.7250
QMF-PL	0.8129	0.8580	0.8116	0.8581	0.8161	0.8514	0.8144	0.8524
QMF-COS	0.7969	0.8392	0.7896	0.8312	0.7957	0.8226	0.7894	0.8142
BSCRM	0.8243	0.8495	0.8248	0.8519	0.8333	0.8467	0.8326	0.8497

Table 3 Comparison Results on Movielens (1m) Dataset

Method	Train=0.60				Train=0.80			
	D=8		D=16		D=8		D=16	
	NDCG@20	ERR@20	NDCG@20	ERR@20	NDCG@20	ERR@20	NDCG@20	ERR@20
MF	0.5821	0.6476	0.5821	0.6476	0.6088	0.6404	0.6088	0.6404
BiasedMF	0.5821	0.6476	0.5821	0.6476	0.6088	0.6404	0.6088	0.6404
UBNNR	0.6780	0.7665	0.6712	0.7589	0.6954	0.7517	0.6918	0.7460
IBNNR	0.6706	0.7494	0.6634	0.7377	0.6954	0.7492	0.6898	0.7397
WIBNNR	0.5991	0.6624	0.5954	0.6564	0.6345	0.6675	0.6343	0.6665
SMRMF	0.6349	0.7069	0.6390	0.7164	0.6691	0.7134	0.6653	0.7072
QMF-PL	0.7112	0.7836	0.7109	0.784	0.7340	0.7781	0.7339	0.7781
QMF-COS	0.6892	0.7507	0.6714	0.7325	0.7165	0.7473	0.6983	0.7254
BSCRM	0.7230	0.7672	0.7200	0.7647	0.7596	0.7744	0.7582	0.7752

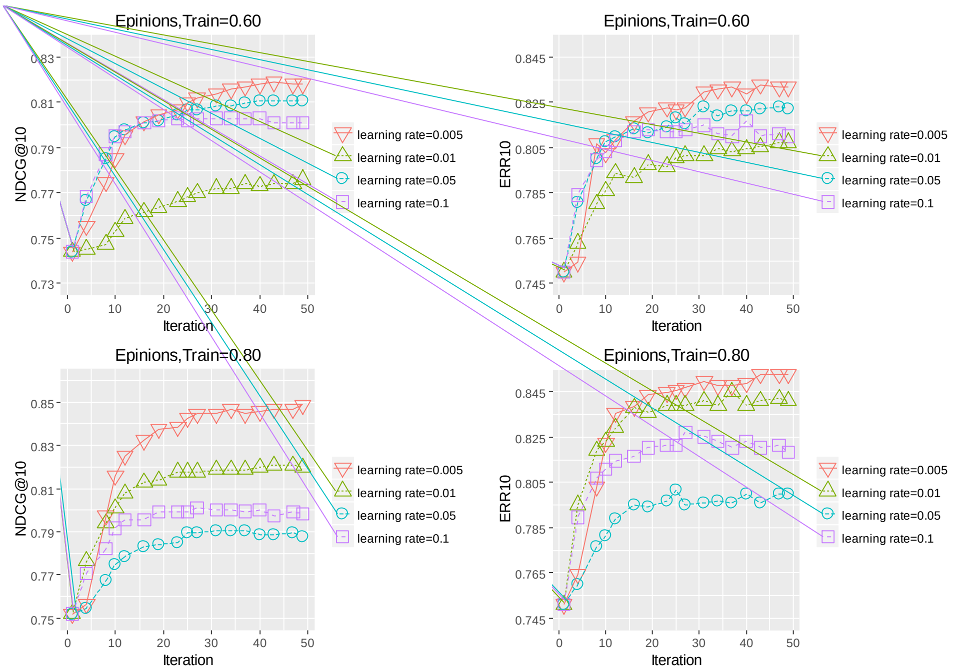

Fig. 5. NDCG@10 and Err@10 at Each Iteration during Learning Process on Epinions Dataset

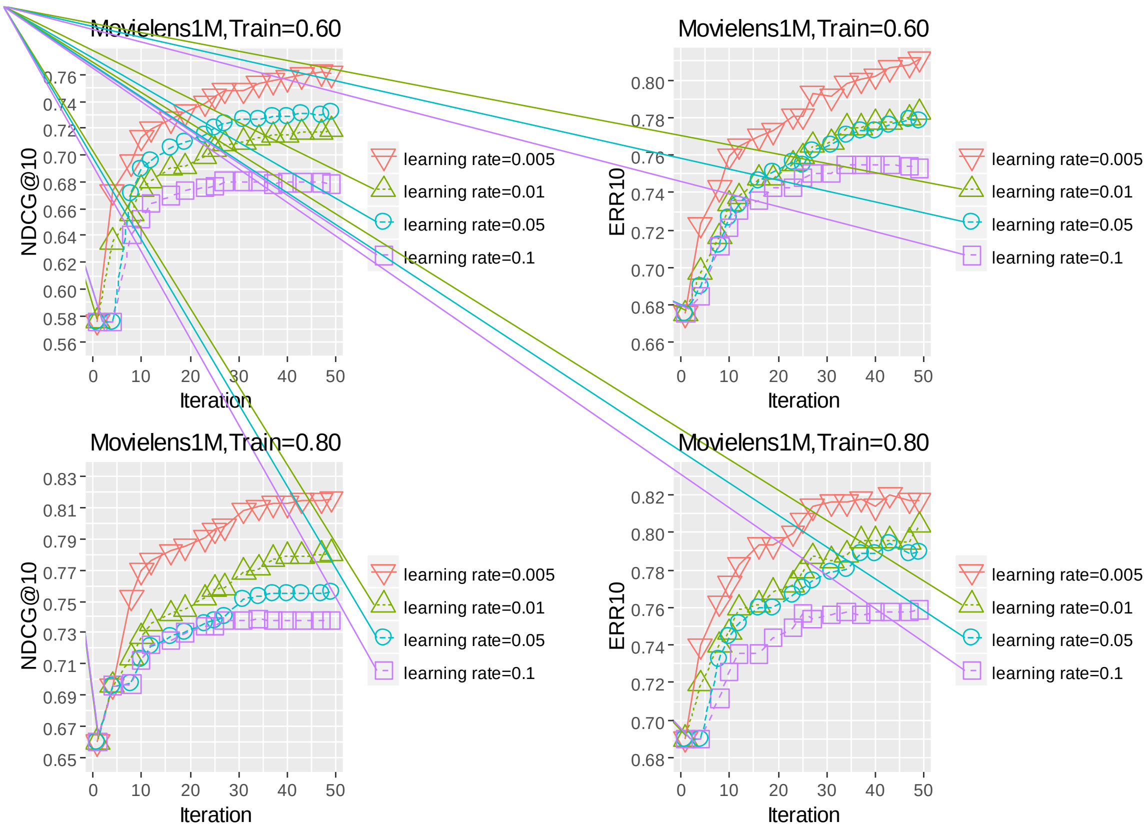

Fig. 6. NDCG@10 and Err@10 at Each Iteration during Learning Process on Movielens1m Dataset

From Fig.5 and Fig.6; we can see that NDCG and ERR converge fast at the beginning of the training process and after several iterations the convergence slows down. The training processes with higher learning rate converge more quickly than those with lower learning rate. This shows that the model proposed in this paper has a better generalization ability under low learning rate.

In this paper; the method is also better than QMF-COS method. In Epinions and MovieLen (1M) data sets; we select 80% data as the training data; and our method has a higher NDCG value than QMF-PL. In other cases; our method is not better than QMF-PL. QMF is a list-wise method; while BSCRM is point- wise. The literature [40] mentioned the list-wise method can obtain more accurate results than the point-wise. This may be one of the reasons that the proposed method is not better than QMF method in some cases.

4. CONCLUSION AND FUTURE WORK

Collaborative filtering algorithm has been widely used in the field of recommender systems; and has got better recommendation results; but it is still affected greatly by the sparsity of rating matrix. In this paper; the proposed recommender algorithm has combined the advantages of spectral clustering and collaborative filtering; so that it can produce better recommendation results; providing reference for the application of collaborative filtering recommendation.

In the future; we will try to use other clustering algorithm instead of K-means algorithm in the improved spectral clustering; to further improve the clustering effect of bi-directional spectral clustering. The other possible research point is to decompose the rating matrix in this model and get the rating matrix of sharing group level; and then use the sharing group level rating matrix and migration learning method for rating prediction and recommendation to improve the cross-domain knowledge migration learning ability of the recommender algorithm.

5. ACKNOWLEDGMENTS

This work is supported by University Science Research Project of Jiangsu Province (15KJB520004); Science and Technology Projects of Huaian (HAG2015060) and Science and Technology Project of Jiangsu Province (BE2015127).

6. REFERENCES

[1] G. Xiong; F.H. Zhu and X.S Dong; A Kind of Novel ITS Based on Space-Air-Ground Big-Data. IEEE Intelligent Transportation Systems Magazine. 8(1) (2016); pp.10-22.

[2] J. P. Shim; J. Koh and S. Fister; Phonetic Analytics Technology and Big Data: Real-World Cases. Communications of the ACM. 59(2) (2016); pp.84-90.

[3] R. Canas; G. Santos and M. Juan; A Recommender System for Non-traditional Educational Resources: A Semantic Approach. Journal of Universal Computer Science. 21(2) (2015) pp. 306-325.

[4] M.T. Ribeiro; Ziviani Nivio and Edleno Silva; Multiobjective Pareto-Efficient Approaches for Recommender Systems. ACM Transactions on Intelligent Systems and Technology. 5(4) (2015); pp. 38-52.
[5] F. Lin; X. Zhou and W.H. Zeng; Online Learning for Collaborative Filtering. International Journal of Computers Communications & Control. 11(2) (2016); pp. 248-258.
[6] W.K. Pan and M. Zhong; Interaction-Rich Transfer Learning for Collaborative Filtering with Heterogeneous User Feedback. IEEE Intelligent Systems. 29(6) (2014); pp. 48-54.
[7] M.Y. Komkhao; J. Lu and Z. Li; Incremental Collaborative Filtering Based on Mahalanobis Distance and Fuzzy Membership for Recommender Systems. International Journal of General Systems. 42(1) (2013); pp. 41-46.
[8] B. Sarwar; G. Karypis and J.A. Konstan; Application of dimensionality reduction in recommender system-a case study . Minneapolis: Minnesota Univ Minneapolis Dept of Computer Science;2000.
[9] H. Ji; J. Li and C. Ren; Hybrid collaborative filtering model for improved recommendation. 2013 IEEE International Conference on Service Operations and Logistics; and Informatics (SOLI). Dongguan: IEEE; 2013: 142-145.
[10] Y. Koren; Collaborative filtering with temporal dynamics.In: Proc.of the ACM SIGKDD Conf.on Knowledge Discovery and Data Mining. ACM Press;2009.89-97.
[11] L. Baltrunas; B. Ludwig and F. Ricci; Matrix factorization techniques for context-aware recommendation. Proceedings of the fifth ACM conference on Recommender systems;2011:301-304.
[12] B.Y. Zou; C.P. Li and L.W. Tan; Social Recommendations Based on User Trust and Tensor Factorization. Journal of Software;25(12) (2014):2852-2864.
[13] Z. Wang; Y. Wang and H. Wu; Tags meet ratings:Improving collaborative filtering with tag-based neighborhood method. Proceedings of the Workshop on Social Recommender Systems;2010:1-9.
[14] C.Q. Huang; A Modified Fuzzy C-Mean Algorithm for Automatic Clustering Number; Applied Mechanics and Materials; 333(2013):1418-1421.
[15] D. Sharmilarani; N . Kousika; G. Komarasamy. Modified K-Means Algorithm for Automatic Stimation of Number of Clusters Using Advanced Visual Assessment of Cluster Tendency. 8th IEEE International Conference on Intelligent Systems and Control (ISCO);2014:236-239.
[16] S.Y. Lin; C.Y. Lai and C.H. Wu; A Trustworthy QoS-based Collaborative Filtering Approach for Web Service Discovery. Journal of Systems and Software. 93(2014); pp. 217-228.
[17] C. Andrew; L. Zhang and D. Agarwal; Collaborative Filtering for Massive Multinomial Data. Journal of Applied Statistics. 41(4) (2014); pp. 701-715.
[18] M.B. Xiao and X.W. Zheng. Collaborative Filtering Algorithm with Stepwise Prediction. Application Research of Computers; 32(11) (2015); pp. 3256-3272.
[19] H. Wu; Y.J. Wang and Z. Wang; Two-Phase Collaborative Filtering Algorithm Based on Co-Clustering. Journal of Software; 21(5) (2010); pp. 1042-1054.
[20] H.C. Zhao; L.L. Zhai and S.C. Zhang; A Collaborative Filtering Recommendation Method Based on Clustering of Gray Association Degree and Factors of Tag Overlap. Computer Engineering & Science; 38(1) (2016); pp. 171-175.
[21] H.X. Zhao; X.H. Wang and J.P. Yang; Mixed Collaborative Recommendation Algorithm Based on Factor Analysis of User and Item. Journal of Computer Applications; 31(5) (2011); pp. 1382-1386.
[22] Z. Huang; D. Zeng and H.C. Chen; A Comparison of Collaborative-filtering Recommendation Algorithms for E-commerce. IEEE Intelligent Systems; 22(5) (2007); pp. 68-78.
[23] Y. Yang; Z.G. Ma and Y. Yang; Multitask Spectral Clustering by Exploring Intertask Correlation. IEEE Transactions on Cybernetics; 45(5) (2015); pp. 1083-1094.
[24] M.A. Oikawa; Z. Diasm and G. Siome; Manifold Learning and Spectral Clustering for Image Phylogeny Forests. IEEE Transactions on Information Forensics and Security; 11(1) (2016); pp. 5-18.
[25] S.B. Fatemi; M.R. Mobasheriv and A.A. Abkar; Clustering Multispectral Images Using Spatial–Spectral Information. IEEE Geoscience and Remote Sensing Letters; 12(7) (2015); pp. 1521-1525.
[26] B. Tandianus; H. Johan and H.S. Seah; Spectral Caustic Rendering of A Homogeneous Caustic Object Based on Wavelength Clustering and Eye Sensitivity. The Visual Computer; 31(3) (2015); pp. 365-370.
[27] S. Mehrkanoon; C. Alzate and R. Mall; Multiclass Semisupervised Learning Based Upon Kernel Spectral Clustering. IEEE Transactions on Neural Networks and Learning Systems; 26(4) (2015); pp. 720-733.
[28] J. Liu; C. Wu; and Y. Xiong; List-Wise Probabilistic Matrix Factorization for Recommendation. Information Sciences; 278(2014); pp. 434–447.
[29] K. Jarvelin; J. Kekalainen. Ir evaluation methods for retrieving highly relevant documents; in: Proceedings of 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval; in SIGIR ’00; ACM; New York; NY; USA; 2000; pp.41-48.
[30] O. Chapelle; D. Metlzer; Y. Zhang. Expected Reciprocal Rank for Graded Relevance. Proceedings of the 18th ACM conference on Information and knowledge management; in: CIKM ’09; ACM New York; NY; USA; 2009; pp.621-630.
[31] X.K. Zhou; S. Wu and Chen; kNN Processing with Co-space Distance in SoLoMo Systems. Expert Systems with Applications; 41(16) (2014); pp. 6967-6982.
[32] A. Saleh; E. Desoulw and H. Shereen; Promoting the Performance of Vertical Recommendation Systems by Applying New Classification Techniques.

Knowledge Based Systems; 75(20156); pp. 192-223.

[33] J.J. Huang; X. Yuan and N. Zhong; Modeling Tag-Aware Recommendations Based on User Preferences. International Journal of Information Technology & Decision Making; 14(5) (2015); pp. 947-970.

[34] W. Liu; C. Wu and B. Feng; Conditional Preference in Recommender Systems. Expert Systems with Applications; 42 (2) (2015); pp. 774–788.

[35] J. Liu; Y. Xiong and C. Wu; Learning Conditional Preference Networks from Inconsistent Examples; IEEE Transactions on Knowledge & Data Engineering; 26 (2) (2014); pp. 376–390.

[36] J. Liu; Z. Yao and Y. Xiong; W. Liu; Learning Conditional Preference Network from Noisy Samples Using Hypothesis Testing; Knowledge-Based Systems; 40 (1) (2013); pp. 7-16.

[37] H.S. Le and T.M. Tuan; General Factorization Framework for Context-aware Recommendations; Data Mining and Knowledge Discovery; 30 (2) (2016); pp. 342-371.

[38] J. Liu; C.H. Sui and D. Deng; Representing Conditional Preference by Boosted Regression Trees for Recommendation. Information Sciences; 327(2016); pp. 1-20.

[39] Z. Gantner; S. Rendle and C. Freudenthaler; et al. Mymedialite: A Free Recommender System Library; In: Proceedings of the 15th ACM Conference on Recommender Systems; in: RecSys'11; ACM; New York; NY; USA; 2011;pp. 305–308.

[40] Z. Cao; T. Qin and T.Y. Liu; Learning to Rank: From Pairwise Approach to Listwise Approach. Proceedings of the 24th International Conference on Machine Learning; in: ICML '07; ACM; New York; NY; USA;2007; pp. 129–136