ISSN:0975-9646

Dipika Anil Hole et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2) , 2015, 1213-1216

Using Ranking – Based Techniques Improvement of Aggregate Recommendation Diversity and Accuracy

Dipika Anil Hole#1, Ashvini Sitaram Bangar*2, Priyanka Hanumant Rokade#3

Department of Information Technology

JSPM's ICOER, Wagholi, Pune, India

Abstract— Recommendation systems are becoming necessary for individual user and also for providing recommendations at individual level in various types of businesses. Recommender system is a personalized information filtering technique used to identify desired number of items based on interest of user. The system uses data on past user ratings by applying various techniques. This techniques concentrate to improve accuracy in recommendations, with recommendation accuracy it is also necessary improve aggregate diversity of recommendation. In this paper, we proposed number of item ranking techniques and different ratings prediction algorithm to improve recommendation accuracy and aggregate diversity by using real-world rating dataset.

Keywords: ***Recommender system, Recommendation diversity, collaborative filtering, Ranking function.***

I. INTRODUCTION

Recommender systems express a progressively famous and considerable firm of personification techniques that help people to operate through the huge bulk of information. These systems attempt to find out the ratings of unexplored items or products for particular user, in accordance with the other users ratings and recommended the items with the maximum anticipated ratings. These proposed systems evaluate ratings by using the ratings of items earlier being used. Over the far 10-15 years, recommender systems technologies are helping the peoples to pact with these large bulks of information [1], [3], [5], [9], [10], [11] various researches as well as e-commerce applications, Amazon and Netflix are using recommender systems.

Recommender systems find out ratings of items (or products) that are consumed by users in the past, based on the ratings of items already used by the users. Recommender systems find out the ratings of unexplored items for particular user and suggest candidate items with the maximum anticipated ratings. The Recommender system uses the frequent new algorithms for recommendations. They can develop the divining accuracy of recommendations. But accuracy may not be enough to find relevant items for each user[12].In particular , the importance of diverse recommendation has been insisted in several studies[4],[6],[8],[13],[14],[16].These studies have proposed new recommendation methods to increase the diversity. Several studies explored recommendation diversity by considering individual users perspective (individual diversity) but some studies examined the impact of recommender system on items diversity according to aggregate diversity of recommendations among all users. For example ,if there is recommendation of same five best movies items that are different to each other, therefore the recommendation for each user is diverse (i.e. high individual diversity), but only these five movies are recommended to all users which resulting in low aggregate diversity.

Recent studies have overlooked that the awareness of importance of aggregate diversity in recommended system is growing. But in contrast with this there has been some significant work done on improving individual diversity because of that the aggregate diversity in recommender system has been largely unfocused. Therefore, this paper focuses on improving aggregate diversity of recommendations by developing algorithmic techniques.

II. RELATED WORK

1) *Rating prediction using Recommendation Techniques:*

Mainly there are three approaches that are content-based, collaborative, and hybrid approaches [1],[3].The recommender systems are splat accordingly. The recommender systems which recommends items same as the user selected once in the past those systems are content based systems. Some recommender systems decide the items to recommend based on the same preferences of the users have liked in the past (i.e. "neighbors") these systems are collaborative filtering systems. Content based and collaborative approach can combined in certain different manner and finally that are hybrid approaches of recommendatory systems. Heuristic techniques and model-based techniques are the algorithmic techniques used in recommender systems. Heuristic techniques are memory based techniques. This techniques work out recommendations based on past user activities such that transactional data or rating values. Neighborhood based prediction technique is one of the heuristic techniques that predict the unknown ratings of users by finding their nearest neighbors that have tested similar items as the target users. Recommender systems perform the main two tasks for recommendations of items–

1) The unknown ratings of the user for items are determined by using some prediction techniques in the same manner with neighborhood –based

approach. These techniques use the previous known user ratings or information about the items.

2) After that system finds the items by using ranking based techniques to increase the utility of user by considering the anticipated ratings and suggest them to the users.

The ranking techniques advised in this paper focuses on elaborating the diversity of recommendation to increase user utility in second task. In this paper we preferred the ranking techniques together with most popular and extensively used collaborative filtering mechanism that is heuristic neighborhood- based technique. These proposed techniques provide flexible solution for recommendations.

2) Rating Prediction Technique:

We proposed the Neighborhood based CF technique. By using this technique we find the similarity between the target user and other users. Before we provide an overview of this technique it is necessary to introduce the notations and terminology.

U- Set of users of recommender system,

I- Set of all possible items of recommender system,

i- Targeted user,

$R: U * I \rightarrow$ Rating-Utility function represents the preference of item $i \epsilon I$ by user $u \epsilon U$.

$R_{(u,i)}$ - Known ratings that is user u gave to item i.

$R^*(u,i)$ - Unknown ratings that user u would give to item i.

u - Other users.

We compute the similarity between user u other users that is u by using formula [18].

$$sim(i,j) = \frac{\sum_{u \epsilon U}(R_{u,i} - R_i)(R_{u,j} - R_j)}{\sqrt{\sum_{u \epsilon U}(R_{u,i} - R_i)^2}\sqrt{\sum_{u \epsilon U}(R_{u,j} - R_j)^2}} \quad ....(1)$$

Where, $I(u,u')$ - Set of all items rated by both user u and u'.

Based on the calculation of similarity, $N(u)$ is obtained which is the number of closest neighbors of user u. Set $N(u)$ can run in size, allover between 1 and $|U| - 1$, that is all else users in the dataset. Then, $R^*(u,i)$ is calculated as the adjusted sated sum of all known ratings $R(u',i)$, Where $u' \epsilon N(u)$[18],

$$R^*(u,i) = \overline{R(u)} + \frac{\sum_{u' \epsilon N(u)} sim(u,u').(R(u',i) - \overline{R(u')})}{\sum_{u' \epsilon N(u)} |sim(u,u')|} \quad ....(2)$$

Here, $R(u)$ - Average rating of user u.

A neighborhood-based CF technique can be of two types that are user based and user based. If the alikeness is estimated between users then it is user based. If it estimated between items then it is item based. Formulae (1) and (2) manifest the user-based mechanism, but they can be forthrightly rewritten for the item-based mechanism due to the equality between users and items in entire neighborhood-based CF calculations.

3) Concept of Accuracy and Diversity in Recommendation:

While recommending the items for users the correctness of users choices and interest should be maintained by the system. Set of recommendation techniques such that statistical accuracy metrics and decision-support measure have been employed various metrics for mapping the accuracy of recommendation.

4) Need of Recommendation Re-Ranking:

Recommendation accuracy and aggregate diversity are two inversely proportional things of the recommender system. These are quality measures of Recommender system. During the implementation of system we can achieve higher individual and aggregate diversity at the expense of accuracy. It can be said that there is possible trade off between the diversity and accuracy leading to less personalized recommendations. The table shows the accommodation between accuracy and diversity [18].

Quality Metric: / Top-1 recommendation of:	Accuracy	Diversity
Popular Item (item with the largest number of known ratings)	82%	49 distinct items
"Long-Tail" Item (item with the smallest number of known ratings)	68%	695 distinct items

Table:1

The table summarized that if it is tried to obtain the higher diversity the accuracy will decrease. There is 82% accuracy by considering enormous number of known ratings. If the items are considered with smallest known ratings and enormous number of predicted ratings the accuracy will decrease by 14% that is 68% and system can recommend 695 different items.

III. PROPOSED RANKING TECHNIQUES

1) Standard Approach:

Standard Approach Recommender systems calculate unknown ratings by considering known ratings using prediction approach such as neighborhood–based CF techniques. The proposed system use the anticipated rating value as the ranking criteria [18],

$$rank_{standard}(i) = R^*(u,i)^{-1}$$

Where, the power of -1 in indicates that the items with highest predicted ratings i.e. $R^*(u,i)$ are recommended to user.

2) Proposed approach:

In this paper we proposed item popularity–based ranking approach for recommendations based on item demand, from lowest to highest. Item demand-based ranking function written as follows [18] where popularity is represented by the numbers,

$$rank_{ItemPop}(i) = |U(i)|, where\ U(i) = \{u \in U | \exists R(u,i)\}$$

3) General Steps of Proposed Ranking Technique:

Here we summarized the typical concept at the heels of the suggested ranking methods based on previous discussion illustrated by figure [18].

Step 1: Fig (a) ranks candidate N items according to the rating values found by using standard approach, as long as they are above the highly predicted rating threshold T_H.

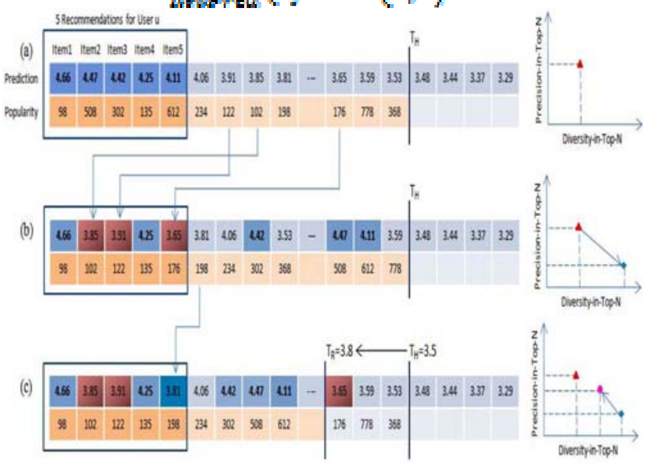
Step2: Fig (b) manifests the recommendations provided by applying average predicted rating values, but still on high of the T_H .Here the top N items are re-ranked and Recommendation diversity is significantly improved, this can be achieved at some cost of recommendation accuracy.

Step3: Fig (c) imported ranking threshold T_R (e.g.3.8 out of accuracy loss of the Fig (b) is significantly minimized by confining candidate items above new ranking threshold T_R.

4) Average predicted rating value:

Now we introduce item ranking function i.e. Average anticipated rating value that supports diversity improvement. This ranking function rank the highly anticipated item on the basis of their anticipated rating value from least to most by choosing fewer famous items, according to Fig 1(a).[18].

$$rank_{RevPred}(i) = R^*(u,i)$$



(a) Recommending top-*N* highly predicted items for user *u*, according to standard ranking approach
(b) Recommending top-*N* items, according to some other ranking approach for better diversity
(c) Confining re-ranked recommendations to the items above new ranking threshold T_R (e.g., ≥ 3.8) for better accuracy

Figure 1.General Overview of Ranking dependent mechanism for elaborating recommendation diversity.

IV. RECOMMENDER SYSTEM DESIGN

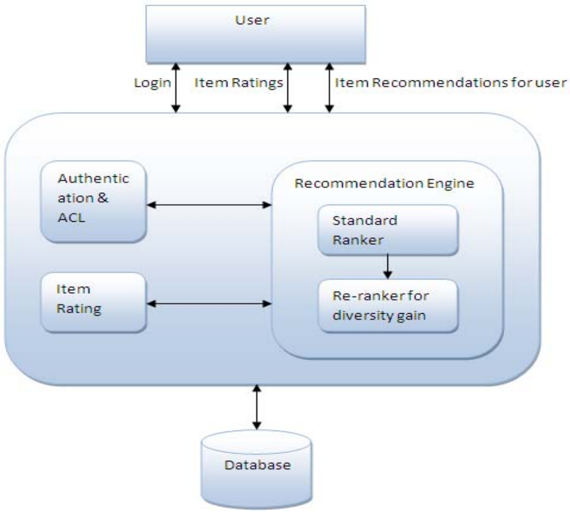


Fig.2 System Design

1) Authentication & Access Control logic:

This module will handle the authentication and authorization task. It will authenticate user with valid username and password. Authentication error will be thrown if username or password does not match with database details. User will be assigned admin/user role based on details available in database. It will allow user to register and create new account. ACL will authorize user before allowing access to admin features. If user has admin role then only if will allowed accessing those features.

2) Item rating:

This module will grant user to add/update rating to items available in application. It will validate user rating that it should greater than minimum possible rating value & shouldn't be greater than maximum possible rating value.
Minimum rating: 1
Maximum rating: 5.
User defined item ratings in database for future references.

3) Recommendation Engine:

Recommendation engine will determine the item recommendations for user depend on user concern and the diversity parameter selected by the user. Recommendation engine ranks all available items with some prediction value. Top N items from the prediction list are suggested to user.
It has two parts:

1. Standard Ranker
2. Re-ranking for diversity gain

4) Standard Ranker:

Standard ranker ranks all items with prediction values calculate with standard ranking technique Neighborhood-based collaborative filtering. Neighborhood-based CF assigns high prediction values to items that users with similar choices (i.e., "neighbors") have liked in the past. In this way all items are ranked with prediction values.

5) Re-Ranking for Diversity gain:

In second step we apply "Item Absolute Likeability" ranking functions, rank $X(i)$ on certain different items (that are not necessarily among N most greatly anticipated, but are yet above (T_H) to re-rank items above the T_H. This manner, a user can acquire recommended more idiosyncratic, long-tail, less frequently recommended items that may not be as widely popular, but can still be very relevant to this user. Therefore, re-ranking of the successor items can extremely raise the recommendation diversity, as discussed, but this generally achieved at some reduced recommendation accuracy. Next we extremely minimize loss of accuracy by inclosing the re-ranked suggestions to the items above recently imported ranking threshold T_R (e.g., 3.8 out of 5). This T_R (Ranking threshold), is provided by user to raise or deplete the diversity in recommendation. The above given parameterization helps to fairly minimize the accuracy loss with still a extreme achievement of diversity (as examined to the standard ranking method).

V. Conclusions

Recommender systems have made significant progress in recent years and many methods have been proposed to make the recommendation quality better. In most studies, various new mechanisms are developed for the improvement of the recommendations accuracy, whereas the recommendation diversity has often been overlooked. In particular, while making recommendations to users based on anticipated rating value that give better accuracy in predictions. This implies the low performance of overall diversity of recommendations. Therefore, in this paper, it is proposed that a different number of ranking approaches for recommendation that can provide expressive improvements in diversity of recommendation with little expense of accuracy. With the addition of improvement of diversity, these ranking methods offer extensibility in designing system to the designers. These approaches can be used together with different algorithms used for rating prediction.

Acknowledgment

This is a great pleasure and immense satisfaction to express our deepest sense of gratitude and thanks to everyone who has directly or indirectly helped us in completing our paper work. We express our gratitude towards project guide Prof. R. N. Phursule and Prof. S. S. Fatanagre, Department of Information Technology, JSPM's ICOER, Wagholi who guided an encouraged us in completing paper work in scheduled time.

References

[1] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Trans. On Knowledge and Data Engineering, 17(6), pp. 734-749, 2005.

[2] J. Bennett, and S. Lanning, "The Netflix Prize," Proc. of KDD-Cup and Workshop at the 13th ACM SIGKDD Int'l Conf. on Knowledge and Data Mining, 2007.

[3] Marina Drosou Computer Science Department University of Ioannina, Greece mdrosou@cs.uoi.gr,Evaggelia Pitoura Computer Science Department University of Ioannina "Diversity over Continuous Data" Greece pitoura@cs.uoi.gr,

[4] K. Bradley and B. Smyth, "Improving Recommendation Diversity," Proc. of the 12th Irish Conf. on Artificial Intelligence and Cognitive Science, 2001.

[5] S. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. of the 14th Conf. on Uncertainty in Artificial Intelligence, 1998.

[6] E. Brynjolfsson, Y J. Hu, and, D. Simester, "Goodbye Pareto Principle, HelloLong Tail: The Effect of Search Costs on the Concentration of Product Sales", NET Institute Working Paper, 2007.

[7] Sofiane Abbar QCRI, Doha, Qatar sabbar@qf.org.qa, Sihem Amer-Yahia CNRS, LIG," Real-Time Recommendation of Diverse Related Articles" France Sihem.Amer-Yahia@imag.fr

[8] D. Fleder and K. Hosanagar, "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity", Management Science, 55(5), pp. 697-712, 2009. [15] S. Funk, "Netflix Update: Try this At Home", http://sifter.org/˜simon/ journal/ 20061211.html, 2006.

[9] Sangeetha G M Mr. Prasanna Kumar M CS&E , "A Survey on Approaches for More Accurate and Diverse Recommendations" EWIT & VTU CS&E,EWIT & VTU India,

[10] P. Resnick, N. Iakovou, M. Sushak, P. Bergstrom, and J. Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," Proc. Computer Supported Cooperative Work Conf., 1994

[11] B.M. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of Recommender Algorithms for E-Commerce," Proc. ACM Conf. Electronic Commerce, pp. 158-167, 2000

[12] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J. Riedl, "Evaluating Collaborative Filtering Recommender Systems," ACM Trans. Information Systems, vol. 22, no. 1, pp. 5-53, 2004

[13] S.M. McNee, J. Riedl, and J.A. Konstan, "Being Accurate Is Not Enough: How Accuracy Metrics Have Hurt Recommender Systems," Proc. Conf. Human Factors in Computing Systems, pp. 1097-1101, 2006

[14] D. McSherry, "Diversity-Conscious Retrieval," Proc. Sixth European Conf. Advances in Case-Based Reasoning, pp. 219-233, 2002

[15] B. Smyth and P. McClave, "Similarity vs. Diversity," Proc. Fourth Int'l Conf. Case-Based Reasoning: Case-Based Reasoning Research and Development,2001.

[16] M. Zhang and N. Hurley, "Avoiding Monotony: Improving the Diversity of Recommendation Lists," Proc. ACM Conf. Recommender Systems, pp. 123-130, 2008

[17] C-N. Ziegler, S.M. McNee, J.A. Konstan, and G. Lausen, "Improving Recommendation Lists through Topic Diversification,"Proc.14th Int'l World Wide Web Conf., pp. 22-32, 2005.

[18] Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques Gediminas Adomavicius, Member, IEEE, and YoungOk Kwon