Improving Aggregate Recommendation Diversity Using Top K Queries

S.Abirami[1]
PG Scholar, ME-Software
Engineering

A.Carolin Arockia Mary[2]
PG Scholar, ME-Software
Engineering

V.R.Azhaguramyaa[3]
Assistant Professor, Computer
Science and Engineering

Abstract

Information overload is one of the major problems in today's life. Recommender systems are there to help user and business by giving the personalized information instead of global information. Recommender systems are usually divided into three categories: CF Approach, content related, and hybrid techniques. Content based is one of the type of recommender systems that recommend the items which is similar to the ones the user already preferred in the previous work. Collaborative filtering is the second type of recommender systems that recommend the items that users with similar preferences have liked in the previous. Last one is, hybrid approaches can combination or collection of both content-based and collaborative methods in various different ways. There were many algorithms were proposed to describe the recommender system but all techniques are described only about recommendation accuracy instead of discussing about the recommendation quality, such as diversity of recommendation. Top K Query is a ranking based technique which is used to achieve recommendation quality by achieving the recommendation accuracy. Top K Query is also used to predict ranking and rating. Diversity gain is achieved by using data set and rating prediction algorithm [Top K Query and Item Popularity Based Approach].

Keywords- Recommender system, Recommendation diversity, Recommendation accuracy, Ranking function, Top K query.

1. Introduction

Recommender system is one, which is mainly used to help the user by providing the extract result for their request finding the extract data while searching is one of the toughest works. E-commerce application is the real time example for recommender system. Notion of rating is the main issue in recommender system. Recommender system is mainly used to compute the ranking and rating of prediction. Recommender system is further classified into in to three types matching to the suggestion. Such as,

> - Content Related
> - Collaborative
> - Hybrid Technique

The first one is used to recommend the item compare to the previous item selected by the people collaborative filtering that is CV, which is same as that of content based. Third one is hybrid it is nothing but the combination of both content-based and collaborative in various ways.

Ranking and rating is one, rating is given by the user Recommender system suggest the top N items based on the rating given by the user. Various algorithm re arise to achieve diversity of recommendation. All those things are fail to achieve. Two algorithms are introduced here.

> - Top K query
> - Item popularity based approach.

Top K query is one of the techniques, used to give the rank for the item to achieve the recommendation quality and recommendation diversity. The aim of the recommender system is to provide a user with detailed and extract item. Diversity is of two types. They are,

> - Individual
> - Aggregate

Individual diversity is one which is used to produce the unique item to the user, unique item is accurately relevant to the search for people, but user is not satisfies with the single item even the result is suitable for user. User always goes for comparison so individual diversity is not suitable for recommender system so recommender system go for aggregate diversity. Aggregate diversity is just opposite to the individual diversity, because individual diversity of the recommender system is produced unique results, but

aggregate diversity of the recommender system produces multiple result. The result should be relevant to the search for the people. Item popularity is next approach used for recommender system. Rating is given manually for the item, based on the rating and price of the item ranking is given by using the top K query technique.

Aggregate diversity is not considering Compare to individual diversity. Awareness only there for aggregate diversity top K query is a technique used to improve aggregate diversity of recommendation system. Recommendation is given by measuring the various items which are recommend across many of the people. Increase diversity results that increase in accuracy. Accuracy diversity trade-off is low is previous technique Top K -query & item popularity technique are arise to achieve the accuracy diversity trend off. The aim of the technique is to achieve diversity without affecting the accuracy [1], [2], [3]. They do not produce new technique into recommendation process with available top K query.

Reasonable solutions for developing the recommendation diversity are: they are used any after the rating was given to the new item (i.e.) unrated item and to achieve the diversity gain by comparing the different rating by comparing the previous approach. The remainder of the paper is organized as follows: Section 2 reviews relevant literature on traditional recommendation algorithms and the evaluation of recommendation quality. Section 3 describes problem definition which deals with the existing system. We then propose the system design in Section 4, and the main performance evaluation follow in Section 5. Lastly, Section 6 concludes the paper by summarizing the contributions and future directions.

2. Related Works

2.1. Recommendation Techniques

Recommender system are further divided into three approaches they are, first one is content based, second one is collaborative and finally hybrid technique [1], [3]. Recommender system of the content based deals with the item which is previously used by the user collaborative technique one is the second one which is same s that of content based by considering similar preference have liked previously hybrid approach is combination of both the first and second technique. Recommender system is not only classified into three, it is further classified into two techniques.

- ➢ Heuristic technique
- ➢ Model based technique

Heuristic technique used to analyze the Recommendation which is already given by the user. One of the heuristic technique is

- ➢ Neighborhood technique.

Neighborhood technique is one of the heuristic technique used to the nearby items [9], [6], [7], [8], [11].Opposite to the heuristic the model based. Model based approach considers the previous activities of the user using statistics and machine learning method.

Content based, collaborative and hybrid approaches are the three main types of recommender system. First one (i.e.) content based recommender system consider the previous usage of the people CF technique of the recommender system is same as that of content based recommender system by using the previous recommender user . Last one is hybrid approaches hybrid recommendation system and collaborative filtering Recommender system are not only divided in to three types it is further classified into two technique

- ➢ Heuristic approach.
- ➢ Model related technique.
- ➢ Model based technique [1], [9]

In general application recommendation to every people first analyze the rating for the new item depends on the information available to the item (new item means the item is not yet getting rating from the user) rating is calculated manually by using the feedback and comments of the user. Second function is to provide ranking. Top K query is a technique which is used to provide ranking for the item. Once the racking is given to the item the next work is to list the top N items by using a technique item popularity based approach. Top K query and item popularity provide a suitable solution for the user. They are not mentioning any new regulations in the recommendation process.

U- Uses of the recommender system.

I-Set of all available items suitable to the user.

Depends upon the preference of the user and the quality of the product. Rating is calculated depends on the rating of the item and price of the item.

2.2. Recommendation Accuracy

Several technique are try to measure the recommendation accuracy both statistics and decision both MAE and RMSE techniques product the

performance of the system and rating of the specified item for example the item should be good for all users. The rating must be used displayed as star values. This star value denotes the item is good one and liked by the user. The goal of the recommended system is to produce best N items liked by the user.

Precision-in-top-N= $\sum_{u \in U} |$ **correct(L$_N$(u))** $|/ \sum_{u \in U} |$ **L$_N$(u)** $|$

Accuracy is not enough for the item each and every items should have both accuracy and diversity (quality) [10], [11].

U- Uses of the recommender system.

I-Set of all available items.

L- List of Items.

Recommendation system is fulfill only after achieving the recommendation accuracy and recommendation diversity in equal way.

2.3. Diversity of Recommendation

Recommendation diversity is calculating in dual way.

- Individual
- Aggregate

Individual diversity is one which is used to produce the unique item to the user, unique item is accurately relevant to the search for people, but user is not satisfies with the single item even the result is suitable for user. User always goes for comparison so individual diversity is not suitable for recommender system so recommender system go for aggregate diversity. Aggregate diversity is just opposite to the individual diversity, because individual diversity of the recommender system is produced unique results, but aggregate diversity of the recommender system produces multiple results. The result should be relevant to the search for the people. The goal the technique is to provide multiple results for the same user ass of accuracy is a major failure. Our goal is to provide item with recommendation diversity with our affecting accuracy.

Various metrics are the two measures diversity group (Aggregate) considering the percentage of the item the performance of the recommender system based on top N product in the list. Diversity-in-top-N defines as follows:

Diversity-in-top-N = $| U_{u \in U} $ **L$_N$(u)** $|$.

Diversity in top N metric is able to also serve up as a pointer of the stage of personalization provided by a recommender system.

3. Problem Statement

Re-ranking of the candidate items whose predictions are above threshold can be affected by the accuracy and diversity trade off and various item ranking factors, popularity based approach. The general discussion about personalizing information listing is not good; e.g., its preference has been discussed in information extraction literature [15], [18], by combining both attempts for decreasing the redundancy and improving the diversity of retrieved pattern by ranking them again [12], [19], [21].

Recommender systems not focused on improving recommendation diversity, other important aspects of recommendation accuracy is only consider, such as the quality of recommendations, have not been discussed.

3.1. Standard Ranking Approach

Distinctive recommender systems predict ratings for the new item, ratings based on known ratings, using any conventional proposal technique such as neighbourhood based or matrix factorization CF techniques that predicted ratings are used to help the user's view of making. In particular, each user u gets recommended a list of top-N items, L$_N$ ðuÞ, selected according to some ranking criterion. More properly, item i$_x$ is ranked in advance of item i$_y$ [i.e., i$_x$ i$_y$] if rankði$_x$Þ < rankði$_y$Þ, where rank: I! IR is a function representing the ranking criterion. The vast greater parts of in progress recommender systems use the predicted mark value as the ranking criterion:

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

The power of 1 in the above expression indicates that the items with highest predicted [as opposed to lowest predicted] ratings R ðu; iÞ are the ones being recommended to user. The standard ranking approach and it shares the motivation with the widely used probability ranking principle in information retrieval literature that ranks the documents in order of decreasing probability of relevance [20]. Recommending the most highly predicted items selected by the standard ranking approach is designed to help get better advice accuracy, but not proposal diversity. Therefore, new statuses criterions are considered necessary in organize to accomplish diversity improvement. Since recommending preeminent advertising bits and pieces to each user normally leads to diversity diminution, recommending

less trendy matter intuitively should have a consequence toward increasing recommendation diversity.

4. System Design

In actual globe settings, recommender systems in the main perform the following two tasks in order to provide recommendations to each user. The ratings of unrated items are estimated based on the available information using top k queries. The system finds items that maximize the user's utility based on the predicted ratings, and recommends them to the user. [Item Popularity]. Ranking approaches [Top k Queries] designed to improve the recommendation diversity in the second task of finding the best items for each user.
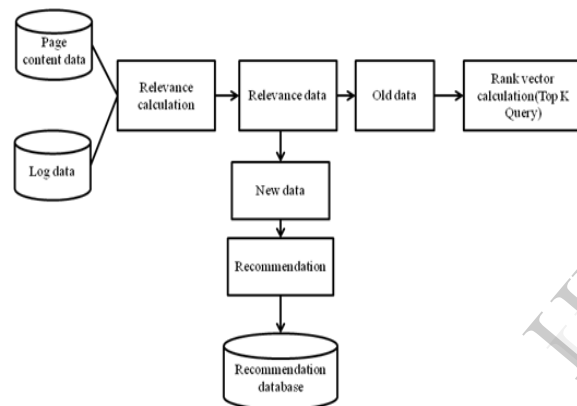

Figure 1. Data flow diagram

Recommendation diversity is achieved by using recommendation algorithms such as Item Popularity based technique and Top k queries. Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques Architecture follows page content data and log data can be calculated by relevance data where rank vector calculation can be done by rank vector data. By using the previous systems found projects at finding relevant data. The relevance data can be stored in recommendation system. User posting the feedback, comment, ratings and admin calculate ranking are stored in database. Finally, overall result display in the recommender system.

This function can be split up into four modules they are Posting the opinion, Recommendation Technique, Rating Prediction and Ranking Approach.

In posting the opinion, the opinions from various people about business, internet commerce and products through internet. The opinions consist of two types. They are direct opinion and comparative opinion. Straight view is to place a comment about the components and attributes of products frankly. Comparative opinion is to post a feedback about the products based on comparison of two or many products. The comments may be good or bad depends on the product.

Moreover, the product quality of recommendations product can be evaluated in many number of dimensions, and presence on the accuracy of recommendations only not be enough for the item find the most relevant items for each.

One of the main goals of recommender systems is to provide a user with highly personalized items, and more diverse recommendations result in more opportunities for users to get recommender such items. With this motivation, new recommendation methods that can increase the diversity of recommendation set for a given individual user. They can give the feedback of such items. The ratings of unrated items are estimated based on the available information [typically using known user ratings and possibly also information about item content] using Top K algorithm Item popularity algorithm. Heuristic techniques are used to analyze the recommendations based on the feedback of the previous user activities. For each user, ranks will the analyze items quality according to the noted rating value ranking the candidate [maximum predicted] items based on their predict rating value, from minimum to maximum.

Ranking items according to the rating variance of neighbours of a particular user for a particular item. There exist a number of various ranking techniques are there to improve recommendation quality by recommending the items other than the ones with top most predicted items suitable to a user. A wide-ranging set of experiments was performed by means of every rating prediction technique in concurrence with every suggestion ranking function on each dataset for poles apart number of top *N* recommendations.

5. Performance Evaluation

Recommender system focused on improving recommendation accuracy, additional significant aspect of proposal quality, such as the variety of recommendation, has been unnoticed. This drawback is overcome by replacing the standard ranking approach by Top k queries and Item popularity based technique.

Top K Query are extremely proficient, because they are based on scalable sorting related heuristics that make decisions depends only on the "general" data

[i.e., only on the candidate stuff of each character user] without having to keep track of the "worldwide" information, such as which items have been recommender crossways all users and how many times. Recommender system focuses on improving recommendation accuracy and also recommendation Quality.

5.1. Top k query

The top k algorithm is a simple iterative method to partition a given dataset into a user specified number of clusters, k. Information systems of different types use various techniques to rank query answers. In many application domains, end-users are more interested in the most important top-k query answers in the potentially huge answer space. Different emerging applications warrant efficient support for top k queries. For illustration, in the background of the Web, the success and good organization of Meta investigate engines, which combine rankings from different search engines, are greatly associated to resourceful rank aggregation ideas. Comparable applications live in the framework of in sequence retrieval and data mining. Most of these applications work out queries that involve joining and aggregating numerous inputs to make available users with the top-k results.

5.2. Item popularity based approach

Item popularity related ranking approach position items in a straight line based on their status, from buck to peak, where popularity is represented by the number of predictable ratings that each piece have. More legitimately, item popularity related position function can be written as follows:

$$\text{Rank}_{\text{item popularity}}(i) = \lvert U(i) \rvert$$

The performance of the item-popularity-based ranking approach with the standard ranking approach using data set and item-based CF, present this comparison using the accuracy-diversity. In particular, the results demonstrate that, as compared to the normal ranking approach, the item popularity related ranking approach amplified proposal diversity; however, recommendation accuracy drop from 89 to 59 percent. Here, regardless of the important diversity expand; such a noteworthy accuracy defeat [30 percent] would not be good enough in most general life personalization applications. Therefore, a universal method to parameterize suggestion status

approaches, which allows accomplishing noteworthy diversity gains while scheming accuracy wounded [e.g., according to how much hammering is unobjectionable in a specified purpose].

6. Conclusion

Recommender systems have made significant progress in recent years and many techniques have been proposed to improve the recommendation quality. However, in most cases, many techniques are calculated to pick up the accuracy of recommendations, whereas the proposal diversity has often been ignored. They are diversity preserving algorithm, content based, item re ranking, neighbourhood Collaborative Filtering techniques, all these techniques are try to improve the recommendation excellence in different way but fail to attain the recommendation excellence devoid of moving recommendation accurateness. Top k query technique is planned to carry out score and grade predictions so that it is easy to improve both the accuracy of advice and quality of counsel. In particular, additional important item ranking criterion should be explored for likely diversity improvements. This may include consumer dependent or producer oriented status mechanisms, depends on the given submission domain, as well as peripheral factors, such as communal networks.

7. Future Enhancement

Travelling around to the recommendation diversity while recommending thing bundles or sequences as an alternative of cluster of items also constitutes fascinating topics for upcoming research.

REFERENCES

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

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

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

[4] 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. Knowledge and Data Eng., vol. 17, no. 6, pp. 734-749, June 2005.

[5] M. Balabanovic and Y. Shoham, "Fab: Content-Based, Collaborative Recommendation," Comm. ACM, vol. 40, no. 3, pp.66-72, 1997.

[6] S. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis ofPredictive Algorithms for Collaborative Filtering," Proc. 14th Conf. Uncertainty in Artificial Intelligence, 1998.

[7] J. Delgado and N. Ishii, "Memory-Based Weighted-Majority Prediction for Recommender Systems," Proc. ACM SIGIR Workshop Recommender Systems: Algorithms and Evaluation, 1999.

[8] A.Nakamura and N.Abe,"Collaborative Filtering Using Weighted Majority Prediction Algorithms," Proc. 15th Int'l Conf. Machine Learning, 1998.

[9] S.T. Park and D.M. Pennock, "Applying Collaborative Filtering Techniques to Movie Search for Better Ranking and Browsing," Proc. 13th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 550-559, 2007.

[10] B. Smyth and P. McClave, "Similarity vs. Diversity," Proc. Fourth Int'l Conf. Case-Based Reasoning: Case-Based Reasoning Research and Development, 2008.

[11] Z. Huang, "Selectively Acquiring Ratings for Product Recommendation," Proc. Int'l Conf. Electronic Commerce, 2007.

[12] 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.

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

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

[15] 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.

[16] B. Smyth and K. Bradley, "Personalized Information Ordering: A Case-Study in Online Recruitment," J. Knowledge Based Systems, vol. 16, nos. 5/6, pp. 269-275, 2003.

[17] J. Carbonell and J. Goldstein, "The User of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries," Proc. ACM Conf. Research and Development in Information Retrieval [SIGIR], pp. 335-336, 1998.

[18] M. Sanderson, J. Tang, T. Arni, and P. Clough, "What Else Is There? Search Diversity Examined," Proc. European Conf. Information Retrieval, pp. 562-569, 2009.

[19] C. Zhai, W.W. Cohen and J. Lafferty, "Beyond Independent Relevance: Methods and Evaluation Metrics for Subtopic Retrieval," Proc. ACM Conf. Research and Development in Information Retrieval [SIGIR], 2003.

[20] S.E. Robertson, "The Probability Ranking Principles in IR,"Readings in Information Retrieval, pp. 281-286, Morgan Kaufmann Publishers, 1997.Subtopic Retrieval," Proc. ACM Conf. Research and Development in Information Retrieval [SIGIR], 2003.

[21] A. Umyarov and A. Tuzhilin, "Using External Aggregate Ratings for Improving Individual Recommendations," ACM Trans. Web,vol. 5, p. 3, 2011.