REVIEW PAPER ON COLLABORATIVE FILTERING

Mayuri Dalvi [1], Prof. S.V. Gumaste [2]

[1]Student ME Computer Engineering, Gokhale Education Society's, R. H. Sapat College of Engg, Nashik, Maharashtra, India-422003.

[2]Assistant Professor ME Computer Engineering Gokhale Education Society's, R. H. Sapat College of Engg, Nashik, Maharashtra, India-422003.

---***---

Abstract - *As more and more information became available electronically, the need for effective information retrieval and implementation of filtering tools have became essential for easy access of relevant information. This paper presents a literature review of the field of recommender system and the recommendation methods that are usually classified into three categories. Content based Filtering, Demographic and Hybrid systems. The proposed system incorporate users' response models into the probabilistic matrix factorization (PMF), a popular matrix factorization CF model, to establish the response aware probabilistic matrix factorization (RAPMF) framework*

Key Words: PMF, RAPMF

1. INTRODUCTION:

Recommendation system has become an important research field. The recommendation system is defined as the supporting system which is used to help users to find information services, or products (such as Books, Music, Movie, Digital Products, Web sites & TV Programs) by analyzing the suggestions from other users, that reviews from other authorities and user attributes. It provides the personalized recommendation services and contents to the different users. Recommendation system is an information filtering system, it is also called as recommendation engine, used to recommend informational items.

In everyday life, people rely on recommendation from other people by spoken words, news reports from news media, reference letters, general survey, travel guides etc. Recommender system assist & augment this natural social process to help people sift through available books, articles, web pages, movies, music, restaurants, jokes, grocery products & so forth to find the most interesting & valuable information for users. The recommendation system can be distinguished between 1) Recommendation class 2) Recommendation approach 3) Recommendation algorithm & 4) Recommendation implementation.

The "recommendation class" is broad concept that describes how recommendations might be given. The recommendation concepts i.e:. Collaborative filtering & content based filtering fundamentally differ in their underlying ideas. The idea of content based filtering is that users are interested in items that are similar to item the users previously liked. On the other hand the idea of collaborative filtering is that users like items that the users peers liked.

A "Recommendation Approach" is a model of how to bring a recommendation class into practice. The idea behind collaborative filtering, content based on collaborative filtering [1][2]. This approach are quite different but are each consistent with the central idea of collaborative filtering.

A" Recommendation Algorithm" precisely specifies a recommendation approach. An algorithm of a content based filtering approach would specify whether terms were extracted from the title of the document or from the body of text, & how terms are processed(e.g stop word removal or stemming) & weighted (e.g TF-IDF), pseudo-

code might contain only the most important information & ignore basic, such as weighting schemes.
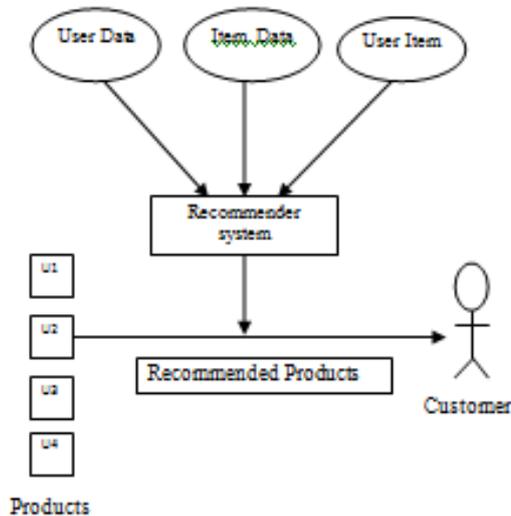

Fig 1: Recommendation Process

The "Implementation" is the actual source code of an algorithm that can be compiled and applied in a recommender system.

2. RELATED WORK

Mainly the Recommendation Methods are classified into three categories:

A. Content-Based Filtering:

This system recommends items based on product description or content of items rather than other users ratings of the system. This system uses the item-to-item correlation rather than user-to-item correlation for generating recommendation. In this system the recommendation process first starts by gathering data or information about the items. E.g author, title, cost etc. Most of this type of system use feature extraction techniques and information indexing to extract the content data [3].

In content based filtering, this system processes information and data from various sources and try to extract useful features and element about the contents of the items. In this system the constraint based filtering uses features of items to determine their relevance. This approach does not require data of other users and it has capability of recommending item to users with unique taste and does not suffer from problems. The disadvantage of this system is that the feature extraction and representation can be achieved automatically i.e:. papers or news but human editors which have to manually insert features from items, i.e movies and songs.

B. Demographic filtering :

Demographic filtering recommender system uses prior knowledge on demographic information about users and opinions of users for the recommendations. This system states the description of people to learn the relationship between a single item and the class or type of people who liked it [3]. This system is stereotypical because this is depends on assumptions that all users are belonging to a certain demographic group have similar taste or preference. In the user model the representation of demographic system can be very grate.

The advantage of demographic system is that this system does not require history of user ratings. This approach is quick, straight forward and easy for making results based on few observations. The disadvantage is that this system mainly based on users interest which are general, and which leads this system recommend the same items to users of same demographic profile and this gives the result of recommendation which is too general.

C. Hybrid Recommender System:

The another category of recommender system is hybrid recommender system. This system tries to overcome the limitations of the other approaches. This technique combines two or more recommendation techniques to gain better system optimization and fewer of the weaknesses of any individual ones. The content

based collaborative filtering is the most popular hybrid approach. The hybrid algorithm use both items attributes and the ratings of all users [3].

Certain strategies are given by which hybridization can be achieved [10].

1) Weight: In this method ratings of several recommendation technique are combined together to produce a single recommendation.

2) Switching: Depending on the current situation the system switches between recommendation techniques.

3) Mixed: In this method the recommendation from several different recommenders are presented at the same time.

4) Feature Combination: The several features from different recommendation data source are thrown together into a single recommendation algorithm.

5) Cascade: In this method one recommender refines the recommendations given by another.

6) Feature Augmentation: In this method the output of one technique is given an input to another technique.

3.LITERATURE SURVEY

The concept of recommender systems introduced in mid-1990s. In past 10 years there has been a tremendous growth in the development of recommender sites. The people using the recommender systems is increasing exponentially which makes it very important for these systems to generate recommendations that are close to the items of users interest.

Jia Zhou and Tiejan Luo [4], it has published a paper on Collaborative Filtering applications. The paper describes about the collaborative filtering techniques which were currently in used in this generation. This paper states that the Collaborative Filtering techniques used in this generation that could be divided into heuristic-based method and model-based method. The paper discusses about the limitations of the Collaborative

Filtering techniques in that generation and suggests some improvements to increase the recommendation capabilities of the systems.

SongJie Gong and Zhejiang [10], proposes a 'personalized recommendation systems' is widely utilized in e-commerce websites to provide recommendations to its users. The paper states that the recommendation systems use Collaborative Filtering technique which has been successful in providing recommendations. A technique to solve the common problems that are encountered in recommender systems namely, scarcity and scalability is suggested in this paper. This paper suggests the recommender system which combines both user clustering and item clustering can be used to provide recommendations. This approach is employed to provide recommendations in this project which makes the prediction smoother. In this approach, item clustering is done using the two techniques Pearson correlation technique and Adjusted cosine similarity technique to find the similarity between the items. Then, users are clustered depending on alikeness between the user targeted and cluster center. Users are grouped into clusters based on their likes and dislikes for an item and every cluster has a center. The authors state that the proposed method is more accurate than the traditional method in generating recommendations.

Robert M Bell and Yehuda Koren [11], state that recommender systems provide recommendations to the users based on past user-item relationship. Based on past user-item relationship the neighbors are computed which makes the prediction easy. The weights of all the neighbors are calculated separately and are interpolated concurrently for many interactions to provide optimized solution to the problem. The proposed method is stated to provide recommendation in 0.2 milliseconds. The training also takes less time unlike very lengthy time in large scale applications. The proposed method was tested on Netflix

data which consisted of 2.8 million queries which was processed in 10 minutes.

Micheal Pazzani [12], discusses about recommending data sources for news articles or web sites after learning the taste of the user by learning his profile. This paper mentions various types of information that can be considered to learn the profile of a user. Based on ratings given by a user to different sites, ratings that other users have given to those sites and demographic information about users the recommendations can be made. This paper describes how the above information can be combined to provide recommendations to the users.

Lee W. S [13], proposed a method in which he assumes that each user is likely to belong to any one of the 'm' clusters and the rating of each user depends upon one of the items that belong to the n cluster of items. Bayesian sequential probability is used to calculate the performance of this method. Heuristic approximations are proposed to Bayesian sequential probability for making experiments on the data set comprising of the ratings of movies. The method suggested is believed to have good performance and tested results are observed to be near to the actual values.

4. CONCLUSION

Different approaches of recommender systems have been discussed in detail. Due to the overload of information on the World Wide Web, the necessity of recommender systems to generate efficient solutions have evolved. In the present scenario, finding the right recommender for evaluating the Credibility of recommender systems is an essential feature. Retrieval of information from huge volumes of data in diversified areas results in a tedious process. Hence, Collaborative filtering recommender systems have evolved to make the recommendation process trivial.

REFERENCES

[1] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," in Proceedings of the 1994 ACM conference on Computer supported cooperative work, 1994, pp. 175–186.

[2] P. Melville, R. J. Mooney, and R. Nagarajan, "Content-boosted collaborative filtering for improved recommendations," in Proceedings of the National Conference on Artificial Intelligence,2002,pp.187–192.

[3] " A categorical review of recommender Systems" International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.5, September 2012

[4] Jia Zhou, Tiejian Luo, "Towards an Introduction to Collaborative Filtering". 2009

[5] SongJie Gong, Zhejiang "A Collaborative Filtering Recommendation Algorithm based on User Clustering and Item Clustering". Business Technology Institute, 2010

[6] Robert M. Bell and Yehuda Koren ,"Improved Neighborhood-based Collaborative Filtering"., 2007

[7] Micheal Pazzani, "A Framework for Collaborative, Content-Based and Demographic Filtering". University of California, 2004

[8] Lee W.S ,"Online Clustering for collaborative filtering.", School of Computing Technical Report, 2009.

[9] "Boosting response aware model-based Collaborative filtering", ieee transactions on   knowledge and data engineering, vol. 27, no. 8, august 2015

[10] Z. Liu,W. QU, H. Li, C.Xie, "A hybrid collaborative filtering recommendation mechanism for P2P networks". 2010

[11] Y. Shi, M. Larson, and A. Hanjalic, "Collaborative Filtering Beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges," ACM Comput.Surv., vol. 47, no. 1, pp. 3:1–3:45,2014.