A Brief Survey On Recommendation System

Anam Khan[1], Anusha Yadav[2], Aman Chand Rahi[3*]

[1,2,3]*Student, Department of Computer Science and Engineering, Galgotias College of Engineering and Technology, Greater Noida, India*

Corresponding author: aman.rahi10@gmail.com

Abstract: Today recommendation system has become very convenient as it is the best technique to make online websites. It has a large number of applications in many fields including economic, education and scientific research. Recommendation systems are the search engines which analyses the user's browsing history and suggest the product according to his taste. Recommendation system applies various approaches to predict users interest on information, products and services among the tremendous amount of available items. In this paper we first describe the recommendation system and then explaining the related research done in this field. The vast growth of information on the internet as well as number of visitors to websites add some key challenges to recommendation system which we discussed lastly in the paper.

Keywords: Recommendation system, Collaborative filtering, Content based filtering, Hybrid approach.

1. Introduction

Web is used by most of the real world applications. We are facing the issue of rapid growth of information on the web which is causing information overloading. It becomes difficult for the users to find the desired information on time among the huge amount of information. By considering this work of searching as an extremely time- consuming task. The enormous data is created by the increasing number of publications, which is growing at a very high rate. In order to solve the problem, recommendation systems can be a game changer for the providers specially e-commerce service providers because in e-commerce small increase in user interaction percentages can generate a huge profit gain for the market [1]. Recommendation system can be seen as a technology of filtering the useful information which helps to make a prediction about any user's preference. Recommendation system takes a user profile and compare it with some specific characteristics, and then predict the chances that a user will like an item or not.

It uses a personalized filtering technology either to predict whether a user will like a particular item or to recognize a set of N items which will of user's interest. There are many areas where recommender system commonly used is Netflix, YouTube, Amazon, Face book and Twitter. Recommendation system used data mining technique to filter out the data which is stored in the user's browsing history. It comprises of three steps: Data preprocessing, Data analysis and Result interpretation.

The various types of recommender systems are:

1. Content based filtering
2. Collaborative filtering
3. Hybrid filtering

Out of all the approaches, collaborative filtering is the most common used as it promotes the evolution of new techniques and methods.
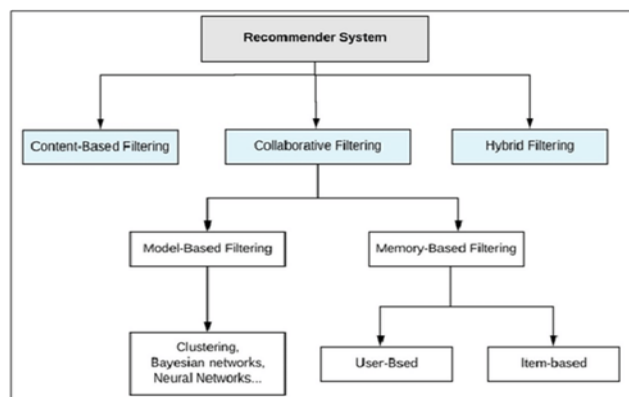

Fig. 1. Recommendation approaches

The power of current search engines lies in the ability to find documents containing particular keywords. Due to synonyms and ambiguous nomenclatures, this approach does not deliver satisfying results.

In this research paper we are presenting recommendation system, which is a combination of both content-based and collaborative based techniques. This approach can reduce the problem of finding appropriate research papers. Recommendation systems are tools and techniques that help users in order to make decision instead of only depending on text mining, this approach combines citation analysis, ratings and source analysis to a recommendation system. It is in under development stage and open for suggestions [2].

First part of the research paper gives an overview of the work done in the field of recommendation system and how ratings help the users to get the desired information in less span of time. Filtering helps to rectify unwanted information available on the web. The last part of the paper depicts the challenges and issues persisting in recommendation systems.

2. Background

Recommendation system uses a number of different

International Journal of Research in Engineering, Science and Management
Volume-3, Issue-6, June-2020
www.ijresm.com | ISSN (Online): 2581-5792

322

technologies to predict the ranking. The outcome of recommendation system is mainly two out of many approaches. These are collaborative based filtering and content-based filtering. Content based filtering [3] approaches are based on the user's profile of the preference to indicate the types of item user like. In this all the information about the user's choice in stored in his profile and from there his interest is analyzed and depending on his taste suggestions are made to the user. It also includes opinion-based recommender systems.

The most widely used algorithm is the collaborative filtering which is based on the assumptions that people agreed in past will agree in future. It requires the user's rating [4] to preference the item. The more users rate items, the more accurate recommendation becomes. It focuses on social community that helps in deciding the user having the same likes and if the users having the same choices and taste then they will fall in same category. However, collaborative filtering algorithms have come to be the best as it does not depend on the machine analyzing content.

Various approaches used in collaborative filtering are:

1. User-based approach: It is based mainly on the user. The item is suggested by finding the same users. If majority of the user has the same likes or preference they tend to fall in the same category. The computation is difficult due to its dynamic nature [5].
2. Item-based approach: It is based on the notion of items. The filtering is done by finding the similarity between the different items. Recommendation is made on similar item user has likes. As the items do not change rapidly so computation can be done offline [5].

Hybrid filtering [6] consists of both content-based filtering and collaborative filtering for the better output. The combination of approaches can be done in many ways:

1. Implementation of separate algorithms and combining the results.
2. Utilizing some methods of collaborative filtering in content based approach
3. Utilizing some methods of content based filtering in collaborative approach
4. Creating a mixed recommendation system which brings both the approaches

Some of the modern recommendation approaches are-semantic based approach, context based technique, peer to peer approach and cross lingual approach.

3. Related Research

Research paper recommendation systems do not exist practically. But concepts have been published that could be used for their awareness. Considerable amount of work has been done in the field of academic research. Mostly work is done on the profiling of researchers, which intent at ranking the authors.

Collaborative filtering and ratings are suggested by some authors. The main focus is to rank the research papers.

For conferences and authors, the rankings are obtained from the ranks of research paper. Ratings could be obtained straightway by contemplating citations as ratings [7] or completely generated by monitoring actions of readers such as bookmarking or downloading [8], [9]. Mostly work on the research make use of number of citations in the form of metric. But this metric only considers the quantity of citations not the quality of citations.

Databases of citations apply citation analysis such as bibliographic coupling co-citation analysis [10] for the identification of papers that resemble with input papers [11]. Impact factor depends on the number of citations, and not always gives correct results.

We can use a modified version of the PageRank algorithm. Search engines such as Google Scholar mainly focuses on text mining and counting of citations. Every concept have some disadvantages, which restrict its propriety to generate recommendations.

For example, homographs cannot be identified by citation analysis [12]. Similarly, lists can hold inappropriate entries.

Limitations occur with text-based analysis, which has to handle ambiguous nomenclatures. Therefore, recommendation systems that are text-based are not able to identify related papers if distinct terms are used.

To improve recommendation system performance when only ratings are not adequate to calculate the resemblances for user. Filtering's in the field of recommendation is condemned for many reasons. Some claims that this filtering would be not effective in domains where more items than exist [13]. Some believes that users would not be convinced to spend time for rating research papers directly.

These filtering has to deal with the feasibility of manipulation [14]. Other drawback is that the ratings and the users need to get useful recommendations.

4. Challenges and Issues

A. Cold-start [15]

In this the new user does not have any data stored in his profile so it is difficult to made recommendation as his taste is almost unknown. This is the cold start problem. This problem can be solved by using hybrid approach.

B. Scalability [15]

There are millions of users and products for which the recommendations are made in different environments. Thus, a large number of resources are required to get perfect and efficient recommendations.

C. Sparsity [15]

In online shops there are large number of items and most of the active users have rated small subset of database which can result in the lowest rating of the popular items. Sparsity arises due to lack of information.

International Journal of Research in Engineering, Science and Management
Volume-3, Issue-6, June-2020
www.ijresm.com | ISSN (Online): 2581-5792

323

D. Trust [15]

It is irrelevant to compare the user who rarely used his profile and rate the item with the most active user having great history. This arises the issues of trust while evaluating. These problems can be solved by giving preference to the user.

E. Privacy [15]

Privacy is the most arising problem. To give user the most precise recommendation, the system uses the data about the user like demographic data and location. This arises the eye on the reliability, security and confidentiality of the user. Many online shops use certain algorithm to protect the privacy of the user.

F. Loss of neighborhood transitivity [15]

When user 1 is connected with user 2 and user 2 is related with user 3 then possibly user 1 would connect with user 3. This type of relation is not understood by recommender system but this can be possibly understood with knowledge from ontology.

G. Synonymy [16]

There are some words which have the same meaning causing the problem of synonymy. In this case the recommender system is not able to predict whether the terms represent same item or different item. In order to deal with this problem different techniques like ontology, SVD and LSI is used.

H. Latency problem [16]

This problem arises in collaborative filtering where the new items are added continuously to the database but the system recommends only the old items as the new items are must be reviewed before recommending. Mixture of category based filtering and user stereotype must be used to reduce this problem.

5. Conclusion

This survey paper has represented a brief review on recommendation system. Recommendation system helps to find the user's interest by analysing his history and suggest the product according to his likes. Recommender systems are rapidly becoming an important tool in E-commerce on the Web. Different techniques have been incorporated in recommended systems. A large number of computations is required to give better results. Recommender system are being stressed by huge volume of data in database. New technologies are needed that can dramatically improve the scalability and efficiency of recommender systems.

References

[1] Witten I. H. and Frank I. Data Mining, Morgan Kaufman Publishers, San Francisco, 2000.

[2] Karypis, George. "Evaluation of item-based top-n recommendation algorithms." Proceedings of the tenth international conference on Information and knowledge management. ACM, 2001.

[3] Das, A. S., Datar, M., Garg, A., & Rajaram, S. (2007, May). Google news personalization: scalable online collaborative filtering. In Proceedings of the 16th international conference on World Wide Web (pp. 271-280).ACM.

[4] K. O. et. al. Context-aware svm for context-dependent information recommendation. In International Conference On Mobile Data Management, 2006.

[5] Jun Wang, Arjen P. de Vries, Marcel J. T. Reinders, "Unifying User-based and Item-based Collaborative Filtering Approaches by Similarity Fusion" (2006).

[6] David H. Stern, Ralf Herbrich, and Thore Graepe, "Matchbox: large scale online bayesian recommendations," Proceedings of the 18th international conference on World Wide Web, page 111--120. (2009)

[7] Middleton, S.E. Shadbolt, N. R. and De Roure, D. C. 2004. Ontological User Profiling in Recommender Systems

[8] S. Brin, and L. Page, "The Anatomy of a Large Scale Hyper Textual Web Search Engine"

[9] Ley, Michael. (2002). The DBLP Computer Science Bibliography: Evolution, Research Issues, Perspectives. SPIRE. 1-10.

[10] Y. Ding, E. Yan, A. Frazho, J. Caverlee, "PageRank for Ranking Authors in Co-Citation Networks,"

[11] J. E. Hirsch, "An Index to Quantify an Individual's Scientific Research Output," Proceedings of National Academy of Sciences, 2005.

[12] Journal of the American Society for Information Science and Technology, 2009.

[13] J. Klienberg, "Authoritative sources in a hyperlinked environment"

[14] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank Citation Ranking: Bringing Order".

[15] Cover, T., and Hart, P., Nearest neighbor pattern classification. Information Theory, IEEE Transactions on, 13(1):21–27, 1967

[16] Sollenborn, M., Funk, P, Category-based filtering and user stereotype cases to reduce the latency problem in recommender systems, in Advances in Case-Based Reasoning, pp. 395–405. Springer (2002).