International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 04 | Apr-2016 www.irjet.net p-ISSN: 2395-0072

A Study Of Recommender Systems With Hybrid Collaborative Filtering

Kaustubh Kulkarni [1], Keshav Wagh[2], Swapnil Badgujar[3], Jijnasa Patil[4]

[1.2.3] *Veermata Jijabai Technological Institute , Mumbai*

[4]*Assistant Professor, VeermataJijabai Technological Institute, Mumbai*

Abstract - *Today people are flooded with many options on internet. Recommender system collect information about the item according to the preferences of the users. Recommender system are successfully implemented in different e-commerce setting. The objective of this paper is to show various techniques being used for recommendation and to discuss various challenges especially for the web media sites.*

Keywords: Collaborative Filtering, Content Based Filtering, Hybrid Filtering, Personalized Recommendation, Recommender System**.**

1. INTRODUCTION

The idea of information reuse and persistent preferences is the origin for the idea of recommender system. The idea of recommender system comes from following in the footstep of others to find what you want.

Recommender systems support users in personalized way for the identification of product based on the history of the user. This can be very useful in case of large product space [1]. Recommender system is an information filtering technology. It is commonly used on e-commerce websites to present items that are likely to be of interest to the customer. The recommender system uses details of the registered users' profile, opinions and habits of their whole community of users and compares the information to reference characteristics and present the recommendations [1]. As a research discipline, recommender systems have shown enormous growth interms of algorithmic developments as well as in terms of deployed applications. Practical experiences from the successful deployment of recommendation technologies in e-commerce contexts (e.g., amazon.com and netflix.com) contributed to the development of recommenders in new application domains.

Recommender Systems are the instances of personalization software. Personalization concerns with the adapting to the individual needs, interest and preferences of each user. It includes:

- Recommending (For example, suggesting list)
- Filtering (For example, Email filters)
- Predicting (For example, product rating)

Basically there are two types of recommendation techniques, personalized recommendation and non-personalized recommendation system. There exist three basic recommendation approaches for personalized recommendation at the algorithmic level.

First is a content-based filtering, Content based Filtering characterizes the affinity of users to certain features (content, metadata) of their preferred items. The filtering is done on the basis of the attributes of the item. For example, in the case of movie, who are the actor, director, genre and so forth? Machine Learning Algorithm is used to induce a profile of the user's preferences from examples based on a feature description of content. There are various other classification technologies under this area to decide whether the item is applicable to the user.

Second is the collaborative filtering. It looks for similar buying pattern among users. It maintains the database of users' ratings of wide range of items. For a given user, it finds other similar users whose ratings strongly correlate with the current user's ratings. It then recommends items rated highly by these similar
users, even though they may not have been rated yet by the current user. Majority of commercial recommender systems use this technique (e.g. Amazon) [1]. There are two variation with this technique, user to user collaborative filtering and item to item collaborative filtering.

Third is the hybrid filtering approach. It is a combination of both content based filtering and collaborative filtering and enjoys the advantages of both content based filtering as well as collaborative filtering.

2. METHODOLOGY

A recommendation system requires a concrete model and for every recommendation system there is a notion of users.

Users have preferences for items and are the source of data in the case of collaborative recommendation system. Items are the entities that a recommender system suggests to the user. Ratings express an opinion of users about items. Each user may have set of attributes for example demographic attributes, cultural attributes, socio-economic attributes etc.

For each user there is a model such as their preference of type of book or type of clothes they like etc. Similarly for each item there are attributes, for example , author , genre, publisher of a book Normally we think of a preference expressed as a user likes a particular item or user purchase that item or user may express an opinion over a things that is not an individual item.

3. NON-PERSONALIZED RECOMMENDATION SYSTEM

It involves summary statistics or product association. It uses external data from community. For example a book that is a best seller or a video that is most popular. It provides summary of community ratings, for example , how much somebody like a restaurant. This summary of community ratings can turn into a list like which are the best hotels in town. The algorithms used to implement non personalized recommender are aggregated opinion recommenders and basic product association recommender. An example of aggregated opinion recommender is imdb.com. Refer framework of recommender system for the understanding of aggregated opinion recommender.

The problem with this approach is that the average may be misleading. Context in a recommender system can be quite important. It is important to know a particular recommendation is being used based on what situation. This leads to the concept of product association recommenders. In this case, the recommender system suggests the item that frequently brought together. For this, the concept of association rule mining is used. This non personalized recommendation still faces challenges in a clustered diverse population. Interests are not uniformly distributed among everyone.

4. PERSONALIZED RECOMMENDATION SYSTEM

1. Content based filtering

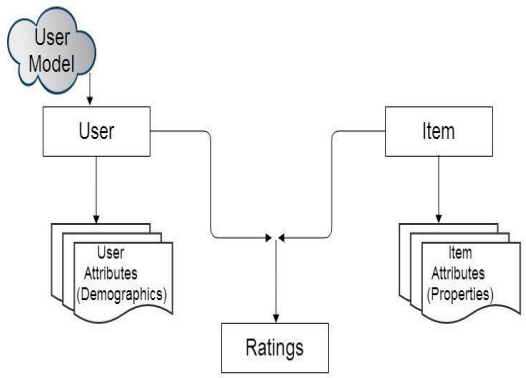

Figure 1: Framework of recommendation system

In content based filtering user rate items and from that we build a model of user preferences. For example movies will have some important properties like genre, actor/actress director and when user submits a rating for a particular movie that updates a model. The model is about which properties of movie user likes and inside that model there is going to be a table, frequently referred to as keyword vector or a test vector. Every time user rates a movie , the model is updated.

An example where the content based filtering is be applicable is the *personalized news feeds*, because news release quickly and the system cannot wait for the other people's opinion.

2. Collaborative Filtering system

It uses opinions of other users to recommend rather that attribute of data. In collaborative filtering , the items versus the users matrix is filled in very sparsely. If the matrix is dense, a recommender system is not needed as we know everybody's opinion.

There are again two types of collaborative filtering. In user to user collaborative filtering the system selects a neighborhood of similar-taste and use their opinion. In item to item collaborative filtering , the system establish relationship among items via ratings Issue with User-User collaborative filtering is *sparsity*, with large item set, small number of ratings too often there are points where no recommendation can be made for a user.

Solution to this issue can be found using filter bot. Filter bot is nothing but a combination of content based technique with collaborative by creating agent that rate everything based on person's profile.

Another issue with user-user collaborative filtering is computational performance, with millions of users computing all pair's correlation is expensive. Even incremental approach is expensive. User's profiles would change very quickly and require updating in real time. The idea behind item to item similarity is predicting user's item preferences using correlation between rating vectors. For example look at the following table

Table 1 - Item to item similarity prediction

	Item 1	Item 2	Item 3	Item 4
Amar		5	3	1
Bahadur	4	5	3	
Chand	2	1		5
Dinesh		2	3	
Jack	5	1	2	4

We can see from the table that item2 and item4 is having negative correlation while item2 and item3 are strongly correlated The system wants to predict the missing value for Chand for item3.There is a strong correlation between item2 and item3. Since Chand didn't like item2 so the system can predict that Chand won't like item3. If we have more than one column with correlation than it takes the average of all[4].

5. PROPOSED METHODOLOGY

The system we have created is a music recommender system.

The proposed recommendation system is based on hybrid collaborative filtering . It is a combination of collaborative filtering along with pattern finding algorithms.

It has the following four steps:

(i) Data Preparation / preprocessing .

(ii) Clustering the web log files by user

(iii) Finding associative rules.

(iv) Web page recommendation.

The user interaction details are recorded in the web server in the form of web log . Web log files are maintained in the the form of plain text files. Preprocessing techniques are necessary for the web logs to discover the knowledge from them. Web logs are also maintained in Proxy Servers and Browser. The data preprocessing is considered as the important activity in web usage mining technique and treated as a key to success.

It consists of the following steps:

- Data Collection:
  Data can be collected from web server, proxy server or browser
- Data Cleaning :
  This involves removal of irrelevant and noisy data from the log files.
- Clustering the web log file using K-means Algorithm.
- Finding association rules using the CHARM algorithm
  It involves two steps.

  i. Finding the set of all frequent item sets.

  ii. Generating rules among item sets.

  In the second step the set of *closed* frequent item sets is mined, which is mostly used than the set of all frequent item sets. It is the main advantage of CHARM algorithm that redundant rules can be eliminated, as a result prediction time and quality are improved.

- Web page recommendation:
  To recommend web pages a model is used to identify the next pages based on the sequence of previously visited pages. When a new user enters the website, the sequence path of that user is matched with the association rules and it would recommend the web page using the support and confidence values found using probability definition.

6. CONCLUSION

Through Recommender System, we can achieve personalization into E-Commerce web sites. In this paper, we studied recommender system that used different techniques like content based filtering, collaborative filtering and hybrid filtering. Recommender systems, providing personalized favorite recommendations, have been prevalently adopted in these services to boost the sales of retailers and trigger the growth of business. Due to the prominence of the commercial value and technical challenges, recommender techniques have attracted the interests of researchers from academia and practitioners from industry. Collaborative filtering technologies, aiming to automatically predict consumers' preferences by analyzing their previous behaviors, e.g., the transaction history or product ratings, become mainstream techniques for recommender systems.

7. REFERENCES

1. Alexander Felfernig, Michael Jeran, Gerald Ninaus, Florian Reinfrank, and Stefan Reiterer, "Toward the Next Generation of Recommender Systems: Applications and Research Challenges," Springer Link, vol. 24, pp. 81-98, 2013.
2. Gideon Dror, Noam Koenigstein, and Yehuda Koren, "Web-scale media recommendation system,"Procedding of IEEE, vol. 100, Issue:9 no. 5, pp. 2722-2736, May 2012.
3. Lalita Sharma, Anju Gera, "A Survey of Recommendation System: Research Challenges," International Journal of Engineering Trends and Technology, vol. 4 Issue 5, 2013.
4. Zhongyun Ying, Zhurong Zhou Fengjiao Han and Guofeng Zhu "Research on personalized web page recommendation algorithm based on user context and collaborative filtering," IEEE Trans. Software Engineering and Service Science., pp. 220-224, May 2013.
5. Haiqin Yang, Member, IEEE, Guang Ling, Yuxin Su, Michael R. Lyu, Fellow, IEEE, and Irwin King, Senior Member, IEEE, Boosting Response Aware Model-Based Collaborative Filtering, 2064 IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 27, NO. 8, AUGUST 2015

6. Greg Linden, Brent Smith, and Jeremy York , Amazon.com,Amazon.com Recommendations Item-to-Item Collaborative Filtering, Industry Report
7. Haifeng Liu, Xiangjie Kong, XiaomeiBai, Wei Wang,TeshomeMegersaBekele, Feng Xia, Context-Based Collaborative Filteringfor Citation Recommendation ,IEEE Access, September 28, 2015.
8. Mohamed Amine Chatti, SimonaDakova, Hendrik Thus, and Ulrik Schroeder,Tag-Based Collaborative FilteringRecommendation in Personal Learning Environments, IEEE Transactions On Learning Technologies, Vol. 6, No. 4, October-December 2013
9. Mohamed Sarwat, James L. Avery, Mohamed F. Mokbel, RECATHON: A Middleware for Context-Aware Recommendation in Database Systems, 2015 16th IEEE International Conference on Mobile Data Management
10. L. C. Wang, X. Y. Zeng, Senior Member, IEEE, L. Koehl, and Y. Chen, Intelligent Fashion Recommender System: Fuzzy Logic in Personalized Garment Design, IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS, VOL. 45, NO. 1, FEBRUARY 2015 95