

Survey on Content Based Recommendation System

Nymphia Pereira^{#1}, SatishKumar Varma^{*2}

^{#1,2} Department of Information Technology
University of Mumbai
PIIT, New Panvel, India

Abstract— Today, as the Web is rapidly growing at a faster rate, finding relevant information is becoming extremely difficult. Information or content can be in any form such as music, video, images or text which is of interest to the users. Therefore Recommendation systems come into picture. Recommendation systems are a sub-category of information filtering system that help people find products, correct information and even other people as well. This paper represents Content-Based Recommendation Techniques that will help personalize the search and provide only relevant information to the user based on previous ratings and predictions along with a comparative study.

Keywords— recommendation techniques, data sparsity, cold start problem.

I. INTRODUCTION

Every day an individual, spends about half of their time searching for relevant information. Recommendation systems plays a major role in the information filtering systems dealing with how best it can recommend items or information relevant to the user. Recommendation system or recommender system can be applied to a variety of applications. E-Commerce site being widely dependent on it. Its task is to automatically personalize the efforts of the user by recommending items from a large collection of items to what the user needs.

However, current recommendation techniques commonly suffer from data sparsity and cold start problem, which lacks the ability to recommend items to user until sufficient amount of user ratings are not available. Therefore combining two techniques i.e hybrid systems overcomes these drawbacks and is much better in terms of accuracy and efficiency. A comparative study of the various techniques mentioned in the paper will help us find out which technique is better suited for a specific domain.

This paper is organized into 5 sections. Section 1 gives an overview of the paper. It describes the use of recommendation system in today's world. Section 2 describes the related work done by different authors. Section 3 describes relevant literature that uses various techniques for different recommendation systems. Section 4 gives a comparative analysis of the different recommendation techniques that are widely being applied to different applications. The last section of this paper summarizes and gives an outline of the different recommendation techniques. It reveals that a lot of improvement can still be done in the field of recommendation systems.

II. RELATED WORK

There is a lot of improvement being done in the information filtering systems using recommendation techniques. But still personalizing the search is a major issue. Thus in this paper we cite previous relevant literature that can be used to recommend relevant information to the user. Many proposals are made to combine different recommendation techniques to overcome the drawbacks of information filtering systems.

III. LITERATURE SURVEY

In this paper we describe the relevant literature survey that uses various techniques for different recommendation systems.

TABLE 1. Literature Survey on various techniques

SN	Techniques	Author & Year of Publication	Advantages and Disadvantages
1.	Content-based Recommender Systems	Pasquale Lops, Marco de Gemmis and Giovanni Semeraro. 2011	Advantages: Learning of profile is made easy. Quality improves over time. Considers implicit feedback. Disadvantages: Does not completely overcome the problem of over-specialization and serendipity.
2.	Hybrid Recommender Systems:	Robin Burke 2010	Advantages: The survey shows combine techniques for improved performance using hybrid methods. It improves the user preferences for suggesting items to users.
3.	Association rule Mining for recommendation system on book sale	Luo Zhenghua. 2012	Advantages: The web site based on this has shown great performance. Disadvantages: It does not recommend quality content to the users. Suffers new user problem. Not very efficient in terms of performance

4.	Collaborative filtering for recommender systems: User-based and Item-based CF	Gilbert Badaro, Hazem Hajj, Wassim El-Hajj and Lama Nachman. 2013	<p>Advantages: solves the problem of finding the ratings of unrated items in a user-item ranking matrix. It improves data sparsity problem.</p> <p>Disadvantage: It does not consider the demographic features which would give better results and solve the user cold start problem.</p>
5.	Content Based Filtering, Collaborative Filtering and Association Rule Mining	Anand Shanker Tewari, Abhay Kumar, and Asim Gopal Barman. 2014	<p>Advantages: It considers various parameters like content & quality of the book by doing collaborative filtering of rating of other buyer. It does not have performance problems. It builds the recommendation offline.</p> <p>Disadvantage: It still lacks the new user cold start problem.</p>
6.	Non-Personalized Recommender Systems and User-based Collaborative Recommender Systems	Anil Poriya, Neev Patel, Tanvi Bhagat, and Rekha Sharma. 2014.	<p>Advantages: System helps users find items they want to buy from a business. It overcomes the lack of personalization involved with non-personalized recommender systems. It is domain independent.</p> <p>Disadvantages: The recommendations are not very specific. It still lacks personalization. The computational time is low.</p>

IV. RECOMMENDATION TECHNIQUES

Recommendation systems are a sub category of information filtering system which recommends items to users based on previous ratings and predictions done by other users. The different types of recommendation techniques are content based recommendation system, collaborative filtering, demographic methods, utility based filtering, knowledge based filtering and hybrid systems. Each of these techniques is explained in brief below.

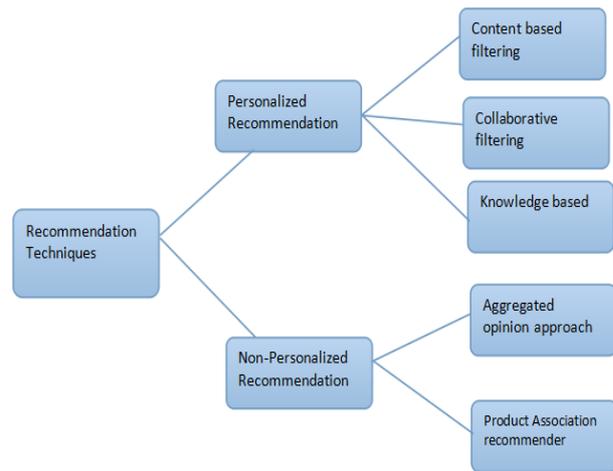


Figure 1: Recommendation Techniques

1] Personalized Recommendation

Personalization deals with adapting to the individual needs, interests and preferences of each user. Detail explanation of different personalized techniques are described below.

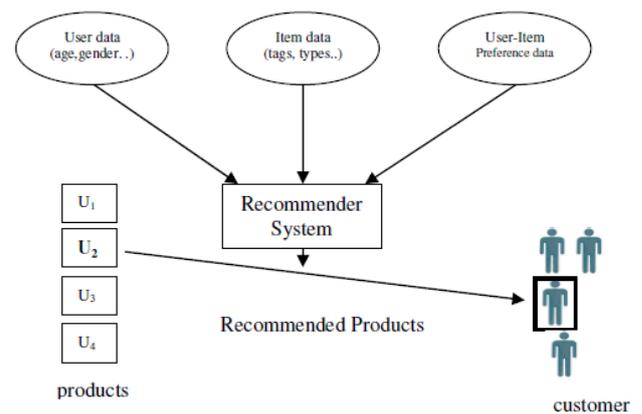


Figure 2: Recommendation Process[11]

A. Content based Recommendation system

Content based recommendation system(CBRS) recommends items to the user based on the users previous buying history. In order to provide relevant information to the user or buyer, a user profile has to be created using web usage mining or information retrieval methods with attributes and features of the items. Content recommendation system filters items based on the similarity of the contents the user is interested in. It recommends or analyzes items based on the items which are highly rated and similar to the users preferences. Applications of content based recommendation are documents with textual information, websites etc.

B. Collaborative Filtering

Collaborative filtering recommends items to the user based on the ratings of other users. The basic idea is automating the word of mouth. The users first give ratings to the items and the system matches these ratings with other users and based on similar taste items are recommended to the user. The drawback of collaborative filtering is that not

enough is known about the user to recommend items and it suffers from issues such as data sparsity and cold start problem. This technique is used mostly by e-commerce websites where in they recommend items based on other users ratings. Therefore quality items are recommended which is lacked by automatically system generated results.

C. Demographic

Demographic techniques as the name suggests recommends items based on the personal characteristics given by the user such as age, religion, community, designation etc. Demographic groups or classes are formed by taking information from marketing research and models are formed. These models are then computed and items are recommended. It does not use explicit ratings. They form people to people correlations. The advantage of using this technique is that it does not suffer from cold start problem. Drawback is it cannot track changes in population as well as rules have to be given by experts by which items can be inferred. This technique is used in tourism to recommend attractions.

D. Utility based Recommendation System

Utility based system is a function that describes the degree of happiness. It determines the users satisfaction and makes computation based on the utility of each attribute of the user to recommend items. The user profile [2] is the utility function that the system has derived for the user, and the system employs constraint satisfaction techniques to locate the best match. The benefit of utility-based recommendation is that it can factor non-product attributes, such as vendor reliability and product availability, into the utility computation, making it possible for example to trade off price against delivery schedule for a user who has an immediate need. Utility based methods improves the drawbacks of new user problem, data sparsity and new items which collaborative filtering cannot overcome. This technique can be applied to applications where short term goals need to be achieved.

E. Knowledge Based Recommendation System

Knowledge based recommendation system recommends items by drawing inferences from users needs and preferences. It creates a user model. The user model draws inferences which can be in the form a query, a case, an adaptive based similarity metrics or a part of an ontology. Explicit knowledge is inferred by rules given by experts as well as implicit knowledge by the users. This techniques is used only when content based and collaborative filtering fails. Many e-commerce websites use this technique to make decisions based on personal interest of the users.

F. Hybrid Systems

Hybrid systems are a combination of recommendation techniques which inherit the advantages of the techniques and eliminates the drawbacks. Many techniques use collaborative filtering as a combination with other techniques. Many methods have been computed using hybrid methods such as weighted, switching, mixed, feature combination, cascade, feature augmentation and meta level.

The table below shows each of the methods and their description.

Table 2 Hybridization Methods[2]

Hybridization Method	Description
Weighted	The scores(or votes) of several recommendation techniques are combined together to produce a single recommendation,
Switching	The system switches between recommendation techniques depending on the current situation
Mixed	Recommendations from several different recommenders are presented at the same time
Feature combination	Features from different recommendation data sources are thrown together into a single recommendation algorithm
Cascade	One recommendation refines the recommendations given by another
Feature augmentation	Output from one technique is used as an input feature to another
Meta-level	The model learned by one recommender is used as input to another.

2] Non-Personalized Recommendation System:

Non personalized recommendation system are the most simple type of recommendation system. They do not take into account the personal preferences of the user. Similar recommendations are provided to all users. There are two types of algorithms used aggregated opinion approach and basic product association recommender.

A. Aggregated Opinion Approach

This algorithm uses scores or ratings given by different customers to recommend items to all the users. This technique is most widely used by E-commerce websites and restaurants. Items with the highest ratings are displayed to all the customers. These scores are calculated by removing the average mean ratings given by different customers and displays the items to the customers. The drawback of the techniques is it does not consider the context during recommendation. Also if one bad rating is given it can bring down the overall average score. Therefore instead of considering the rating overall percentage of the people who rated the items would serve best.

B. Basic Product Association Recommender

This algorithm considers the overall percentage of the people who rated the items. It is the most useful type of non personalized recommendation. Most of the online shopping websites make use of this technique such as by recommending people who brought product 1 also brought product 2 feature. It uses the market basket analysis technique over here. The recommendations done by this system is by what the user is actually doing at present either viewing or buying or what is currently present in his cart. It uses the percentage of product who bought x also bought y. It uses association rule mining which does not look for individuals but for the basket of products to recommend.

V. COMPARATIVE STUDY

A comparative analysis of different applications against recommender techniques are given below. Instead of comparing individual techniques a combination of two techniques are used in every application such as content based filtering(CN) along with collaborative filtering. Collaborative filtering(CF) is used with every other technique as it improves the quality of recommending. For example [7] using demographic technique(DM) along with collaborative filtering in online websites provides better recommendation as it combines the predictive ratings by using an adaptive weighting scheme.

Table 3 Comparative study of different techniques and applications

Techniques	Applications					
	Online Websites (B2B,C2 C,B2C)	Travel Packages	Health Insurance	Movies	Online Book Recommendation System	Google Search Engine
CF/CN	✓		✓		✓	✓
CF/DM	✓	✓		✓		
CF/Utility		✓				
CF/Knowledge	✓					✓

VI. CONCLUSION

A complete study of different recommendation techniques is presented in this paper. A comparative study of the different techniques we can conclude that using hybrid approach would give better accuracy and performance results. The survey reveals that a lot of improvement can still be done in content based and collaborative filtering recommendation techniques for obtaining better results. Many applications can be developed such as a financial planning adviser portal which can recommend how to save money and manage your savings for the future. It can be used in trekking recommendation system wherein by taking personal details trekking attractions around the world can be recommended along with feasible costing to the user .

REFERENCES

- [1] Luo Zhenghua, "Realization of Individualized Recommendation System on Books Sale", IEEE 2012 International Conference on Management of e-Commerce and e-Government. pp.10-13.
- [2] Tewari, A.S. Kumar, A and Barman, A.G, "Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining", International Advance Computing Conference (IACC), 2014 IEEE, pp 500 – 503.
- [3] Gilbert Badaro, Hazem Hajj, Wassim El-Hajj and Lama Nachman, "A Hybrid approach with collaborative filtering for recommender systems",9th International Wireless Communications and Mobile Computing Conference (IWCMC), 2013 .
- [4] Pasquale Lops, Marco de Gemmis and Giovanni Semeraro, "Content-based Recommender Systems: State of the Art and Trends" Springer Science+Business Media, LLC 2011.
- [5] Robin Burke, "Hybrid Recommender Systems: Survey and Experiments", California State University, Department of Information Systems and Decision Sciences. Volume 12, Issue 4, pp 331-370.
- [6] Anil Poriya, Neev Patel, Tanvi Bhagat and Rekha Sharma, Ph. D, "Non-Personalized Recommender Systems and User-based Collaborative Recommender Systems",International Journal of Applied Information Systems (IJ AIS) – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA Volume 6– No. 9, March 2014.
- [7] Jyoti Gupta and Jayant Gadge, "A Framework for a Recommendation System Based On Collaborative Filtering and Demographics", 2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA), pp 300 – 304.
- [8] Jyoti Gupta and Jayant Gadge, "Performance Analysis of Recommendation System Based On Collaborative Filtering and Demographics", 2015 International Conference on Communication, Information & Computing Technology (ICCICT), Jan. 16-17, Mumbai, India.
- [9] Shenshen Liang, Ying Liu, Liheng Jian, Yang Gao and Zhu Lin, "A Utility-Based Recommendation Approach for Academic Literatures", Web Intelligence and Intelligent Agent Technology (WI-IAT), 2011 IEEE/WIC/ACM International Conference on (Volume:3).
- [10] Bahram Amini, Roliana Ibrahim and Mohd Shahizan Othman, "Discovering the impact of knowledge in recommender systems: A comparative study", International Journal of Computer Science & Engineering Survey (IJCSES) Vol.2, No.3, August 2011.
- [11] RVVSV Prasad and V Valli Kumari, "A categorical review of recommender systems", International Journal of Distributed and Parallel Systems (IJDPSS) Vol.3, No.5, September 2012.