

A Survey on Recommendation System

Tejal Arekar
IT Dept./Pune University

Mrs. R.S. Sonar
IT Dept./Pune University

Dr. N. J. Uke
IT Dept./Pune University

Abstract—Recommendation systems are defined as the techniques used to predict the rating one individual will give to an item or social entity. These items can be music, videos, books, movies, restaurants and things on which individuals have different choices. These choices are being predicted using two basic approaches first content-based approach which involves characteristics of an item and second collaborative filtering approaches which takes into account user's past behaviour to make choices. The issues involved in all the techniques of the recommendation system.

Keywords—Recommendation systems, Personalized Recommendation, Non- Personalized Recommendation, Amazon, YouTube.

I. INTRODUCTION

In the recent years, the Web has undergone a tremendous growth of Internet access, various types of multimedia (e.g., text, images, audios, and videos) have become ubiquitously available anytime [1]. But the amount of information available on the Internet has become immense and is still growing at an unbelievably fast rate. The emergence of social networks (e.g., Facebook and Twitter) and Internet-enabled mobile devices (e.g., smart phones and tablets) has further boosted the volume of online information resources, since these technologies enable online users to freely create, upload and share information involves media items, such as texts, images, videos. On one hand, the abundance of online information may virtually guarantee that users are able to find what they are looking for. On the other hand, this same abundance also makes the useful information difficult to find, a problem referred to as “information overload”. Two major Internet techniques, to wit, information search and recommendation, have been developed to help online users handle the information overload problem.

In the search case, illustrated in Fig 1(a), users actively express their information needs by submitting queries to the search system (engine), and then the system tries to find the items (e.g., texts, images, videos, music) in the collection that best match the queries.

In the recommendation case, illustrated in Fig. 1(b), recommenders intend to provide people with suggestions of products they will appreciate, based upon their past preferences, history of purchase, or demographic information.

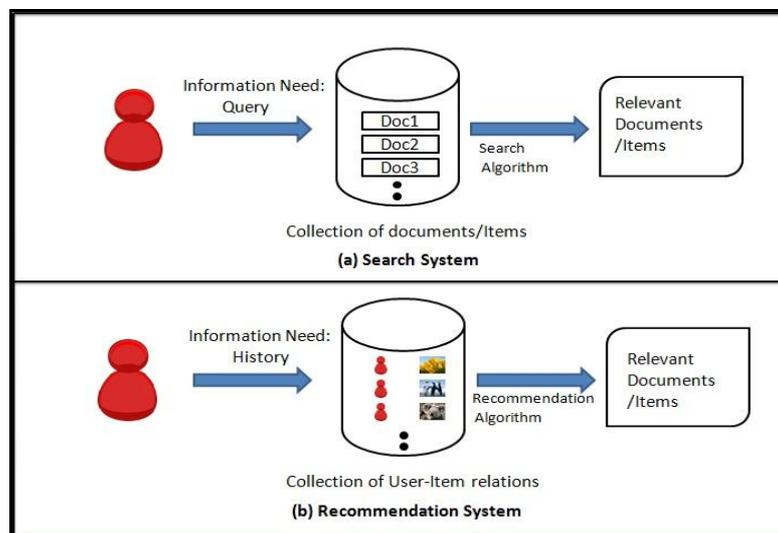


Fig. 1 Types of Information Retrieval System

II. RECOMMENDATION SYSTEM AND TECHNIQUES

Recommendation systems apply data mining techniques and prediction algorithms to predict users interest on information and products among the tremendous amount of available items. They have been widely advocated as a way of coping with the problem of information overload for knowledge workers. Recommendation systems are software agents that elicit the interests and preferences of individual consumers and make recommendations accordingly. They have the potential to help and improve the quality of the decisions consumers make while searching for and selecting products online.

Due to the tremendous growth of e-commerce introduced information overload problem where users are not able to effectively search items on the web. Now a day's electronic world has introduced the need for information filtering techniques that are use to help users by filter out information in which they are interested in.

Recommendation systems are one approach to this issue, based on providing possible items of interest to a user instead of the user to go searching for them. Recommendation systems changed the way as the websites communicate with their users. Instead of providing a static feel in which users search for and potentially purchase products, recommendation systems increase communication to provide a higher experience. Recommendation systems recognize recommendations autonomously for individual users based on past browsing history, rating given to item and other users behaviour.

Recommendation systems are a subdivision of information filtering systems that recommend items available in sites, amusement items (books, music, videos, news, images, events etc.) or people (e.g. on dating sites) that are likely the interest of the user. Recommendation systems generally generate a list of recommendations in one of given ways (Fig. 2)

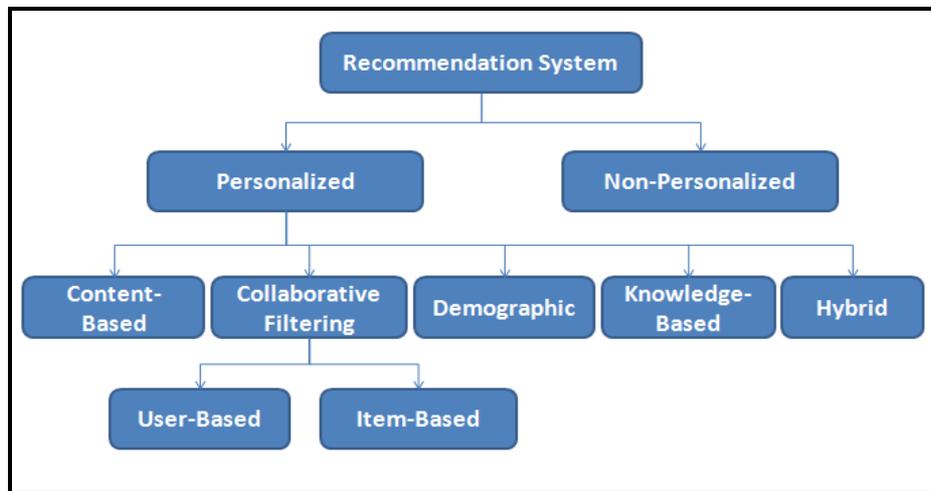


Fig. 2 Types of Recommendation System

PERSONALIZED RECOMMENDATION

Personalized recommendation enables the online introduction insertion, customization or suggestion of data in any format that is relevant to each and every user, based on the users implicit behaviour and tastes and explicitly provided details [2].

Personalized recommendation engines are classified into five types depends on their approach to recommendation [3]:

1. Content-Based Filtering

Content-based recommendation method is based on the information about item content and ratings a user has given to items. This technique combines these ratings to profile of the user's interests based on the features of the rated items. The recommendation engine then can find items with the preferred in the past as illustrated in Fig 3. The recommendations of a content-based system are based on individual information and ignore contributions from other users.

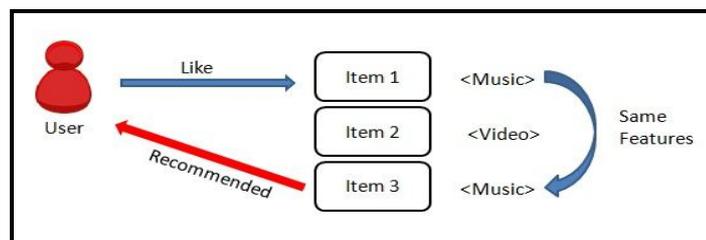


Fig. 3 Content-Based Recommendation

2. Collaborative Filtering

Collaborative filtering technique based on users history in the form of rating given by the user to an item as their information source [5]. It can be accomplished by making relation between the users or between items. Collaborative filtering is categorized into three types: user-based, item-based, model-based.

User-based-

User-based Approach makes recommendation based on the interest of the user having the similar taste. It correlates user as per the rating given to the items.

From the Fig. 4, first user related to third user instead of second because the rating given by third user is quite similar to the first one. That's why item 3 is recommended to the user as it's the only remained item.

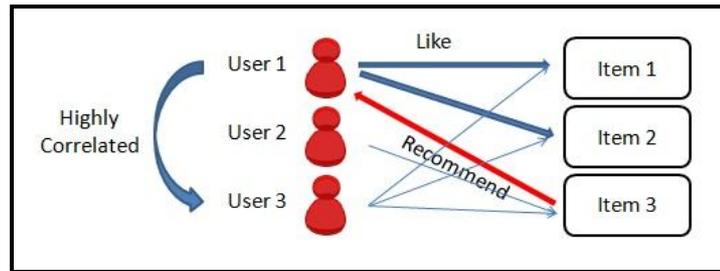


Fig. 4 User-Based Collaborative Filtering

Item-based-

Item-based Approach is based on the items as the user rated items similarly are probably similar. From Fig. 5, 2nd and 3rd user rated item 1 and 3 so it assumes that item 1 and 3 are become similar. As 1st user like item 1, item 3 is recommended.

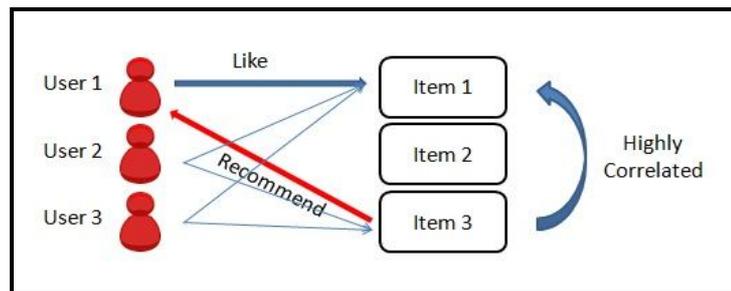


Fig. 5 Item-Based Collaborative Filtering

3. *Demographic*

Demographic recommendation technique uses information about user only. The demographic types of users include gender, age, knowledge of languages, disabilities, ethnicity, mobility, employment status, home ownership and even location. The system recommends items according to the demographic similarities of the users.

4. *Knowledge-Based Filtering*

Knowledge based recommendation system is based on the explicit knowledge about item classification, user interest and recommendation standard (which item should be recommend in which feature) [6]. It is an alternative approach to the collaborative filtering and content-based filtering.

5. *Hybrid approach*

Hybrid approach is a combination of all above types [8].

NON-PERSONALIZED RECOMMENDATION

Non-personalized recommendation system recommend items to consumers based on what other consumers have said about the product in an average. That is, the recommendations are independent of the customer, so all customers gets the same recommendation.

III. ISSUES IN RECOMMENDATION SYSTEM

1. *Data Collection*

The data used by recommendation engines can be categorized into explicit and implicit data. Explicit is all data that user themselves feed into the system. The collection of explicit data must not be intrusive or time consuming. Implicit data source in e-commerce is the transaction data including the purchase information. Implicit data needs to be analysed first before it can be used to describe user features or user-item ratings.

2. *Cold Start*

The cold start problem occurs when too little rating data is available in the initial state. The recommendation system then lacks data to produce appropriate recommendations. Two cold start problems are new user problem and new item problem.

3. *Stability vs. Plasticity*

The converse of the cold start problem is the stability vs. plasticity problem. When consumers have rated so many items, their preferences in the established user profiles are difficult to change.

4. *Sparsity*

In most use cases for recommendation systems, due to the catalogue sizes of e-business vendors, the count of ratings already obtained is very small related to the count of ratings that need to be predicted. But collaborative filtering techniques focuses on an overlap in ratings across users and have difficulties when the space of ratings is sparse (few users have rated the similar items). Sparsity in the user-item rating matrix degrades the quality of the recommendations.

5. *Performance & Scalability*

Performance and scalability are important issues for recommendation systems as e-commerce websites must be able to determine recommendations in real-time and often deal with huge data sets of millions of users and items. The big growth rates of e-businesses are making the sets even larger in the user dimension.

6. *User Input Consistency*

Recommendation techniques that work with user-to-user correlations, like collaborative filtering or demographic, depend on more correlation coefficients between the users in a data set. Users can be categorized into three classes based on their correlation coefficients with other users. The majority of users fall into the class of “white sheep”, where there is a high rating correlation with many users. Resulted in engines can easily find recommendations for these users. The opposite type is the “black sheep” where there are only few or no correlating users. This makes it quite difficult to find recommendations for them. The bigger problem is the “gray sheep” problem where users have different opinions or an unusual taste that results in low correlation coefficients with many users. They fall on a border between user tastes. Recommendations for them are very difficult to find and they also cause different recommendations for their correlated users.

7. *Privacy*

Privacy is an important issue in recommendation systems. To provide personalized recommendations, recommendation systems must know something regarding to users. In fact, the more the systems know, the more precise the recommendations can get. Users are concerned about what information is gathered, how it is used, and if it is stored.

These privacy affect both, the collection of explicit and implicit data. Regarding explicit data, users are not interested to disclose information about themselves and their interests. If questionnaires get too personal, users may give false information in order to protect their privacy.

IV. MERITS/DEMERITS

TABLE I
MERITS AND DEMERITS OF RECOMMENDATION TECHNIQUES

Sr. No.	Techniques	Merits	Demerits
1	Content-Based Filtering	- No need for data of other users.	- Content analysis is necessary to determine the item features.
		- No cold start and sparsity.	
		- Able to recommend users with unique taste.	- The technique depends not only on the quality of the item metadata but also on the homogeneity of the stock, so items can be categorized.
		- Able to recommend new and unpopular items.	- The quality of items cannot be evaluated. The similarity computation is limited to the item features.
2	Collaborative Filtering	- Like for demographic recommendations no knowledge about the item features is needed. Collaborative filtering works completely independent of machine-readable item representations. It is therefore domain independent.	- The quality of the recommendations depends on the size of the historical rating data set.
		- The quality (not just the relevancy) of items can be evaluated, as it is also expressed through user-item ratings.	- The technique suffers from the cold start problem for new users and new items.
		- Collaborative filtering techniques are able to make recommendations “outside the box” because they look outside the preferences of the individual user.	- Gray sheep problem.
3	Demographic	- It is not based on user-item ratings, it gives recommendation before user rated any item.	- Gathering of demographic data leads to privacy issues.
		- Item feature is not needed, therefore the technique is domain independent.	- Gray sheep problem.
			- Stability vs. plasticity problem.

V. APPLICATIONS

Sr. No.	Name	Year	Recommendation Technique Used	Multimedia Used For Recommendation
1	IMDb	1990	Collaborative Filtering	Video
2	Amazon [12]	1994	Item-to-Item Collaborative Filtering	Product
3	News Dude	1999	Content-Based Filtering	News
4	Pandora Radio	2000	Content-Based Filtering	Music
5	Last.fm	2000	Collaborative Filtering	Music
6	YouTube [11]	2005	Hybrid Approach	Video
7	Netflix [10]	2006	Hybrid Approach	Video (Films)

VI. CONCLUSION

Recommendation systems are used to overcome information overload problem as e-commerce continuously growing. Recommendation systems can benefit both customers and provider. Customers use it by finding new interesting products and provider can enhance their sales.

There are various techniques used for recommendations and their related issues are discussed. Two recommendation systems are widely used that are YouTube.com and Amazon.com for videos and product respectively.

REFERENCES

- [1] K. Palanivel, R. Sivakumar, *A study on Implicit Feedback in Multicriteria E-Commerce Recommender System*, Journal of Electronic Commerce Research, Vol. 11 (2010).
- [2] Lalita Sharma, Anju Gera, *A Survey of Recommendation System: Research Challenges*, International Journal of Engineering Trends and Technology (IJETT) - Volume4Issue5- May 2013.
- [3] Nana Yaw Asabere, *Review of Recommender Systems for Learners in Mobile Social/Collaborative Learning*, International Journal of Information and Communication Technology Research, ISSN 2223-4985, Volume 2 No. 5, May 2012.
- [4] G. Adomavicius, A. Tuzhilin, *Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art*, IEEE transactions on knowledge and data engineering 2005, vol. 17, p. 734-749.
- [5] M. Pazzani, D. Billsus, *Content-Based Recommendation Systems*, The Adaptive Web. Springer Berlin / Heidelberg 2007, p.325-341.
- [6] R. Burke, *Knowledge-based recommender systems*, Encyclopedia of Library and Information Systems 2000, vol. 69, pp. 175-186.
- [7] J. B. Schafer, J. A. Konstan, J. Riedl, *E-commerce recommendation applications*, Journal of Data Mining and Knowledge Discovery 5 (1/2), 115-152, 2001.
- [8] R. Burke, *Hybrid Recommender Systems: Survey and Experiments*, User Modeling and User-Adapted Interaction, Volume 12, Issue 4 (November 2002), Kluwer Academic Publishers, pp. 331-370.
- [9] Daniel Billsus, Michael J. Pazzani, *A Personal News Agent that Talks, Learns and Explains*, Proceeding AGENTS '99 Proceedings of the third annual conference of ACM on Autonomous Agents, Pages 268-275 New York, NY, USA ©1999.
- [10] J. Bennett, S. Lanning, *The Netflix Prize*, Proceedings of KDD Cup and Workshop 2007, Aug. 12, 2007.
- [11] James Davidson, Benjamin Liebald, Junning Liu, Palash Nandy, Taylor Van Vleet, *The YouTube Video Recommendation System*, Proceeding RecSys '10 Proceedings of the fourth ACM conference on Recommender systems Pages 293-296.
- [12] G. Linden, B. Smith, J. York, *Amazon.com recommendations: item-to-item collaborative filtering*, Internet Computing, IEEE (Volume:7 , Issue: 1) Pages:76-80 ISSN:1089-7801, 2003.