Challenges & Limitation in Recommender Systems

Mani Madhukar
Technical Lead
IBM India Pvt. Ltd., Noida, India

Abstract- Recommender systems have made a wide inpact in online marketing. The recommender systems have been instrumental in forging a mental alliance with the buyer and hence influencing the decision of the buyer. However the recommender systems are highly successful and advisable for people with strong online presence, but for new users the recommendation algorithms need various parameters to propose suggestions to the user and hence the challenges and limitations of recommender systems come to the fore.

Keywords – Recommender system, Content Filtering, Collaboration Filtering, Cold start, sparsity, privacy

I. INTRODUCTION

Recommender systems or recommendation systems are a subclass of information filtering system that seek to predict ‘rating’ or ‘preference’ that a user would give to an item (such as music, books or movies) or social element (e.g. people or group) they had not yet considered, using a model built from the characteristics of an item (content based approaches) or the user’s social environment (collaborative filtering approaches).A recommender system is a system which provides recommendations to a user. The Recommender System is to generate meaningful recommendations to a collection of users for items or products that might interest them. 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 books, movies, restaurants and things on which individuals have different preferences. These preferences are being predicted using two 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. In collaborative filtering, partners are chosen who will make recommendations because they share similar ratings history with the target user. One partner who have similar ratings to the target user may not be a reliable predictor for a particular item. So the past record of the partner of making a reliable recommendation also needs to be take into consideration which is dictated by trustworthiness of a partner. In order to keep track of past records of a recommender reputation systems comes into the picture those who actually assign reputation ratings to the partners.

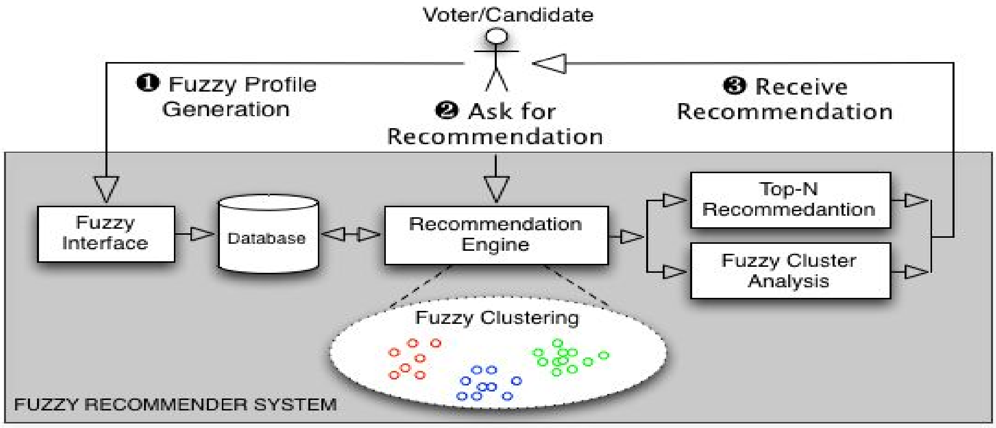

Figure 1. Fuzzy Recommender System

Applications: Books, music CDs, movies. Even documents, services and other products such as software games.

II. TECHNIQUES USED IN RECOMMENDER SYSTEM

We will discuss the primary approaches used in analyzation of Recommender system. Broadly the Recommender systems are categorised as of three types namely -

1. Content Based Recommender System
2. Collaboration Based Recommener System
3. Hybrid Recommender System

A. Content based Recommender System approach -

Content based recommendation systems recommend an item to a user based upon a description of the item and a profile of the user's interests. Such systems are used in recommending web pages, TV programs and news articles etc.

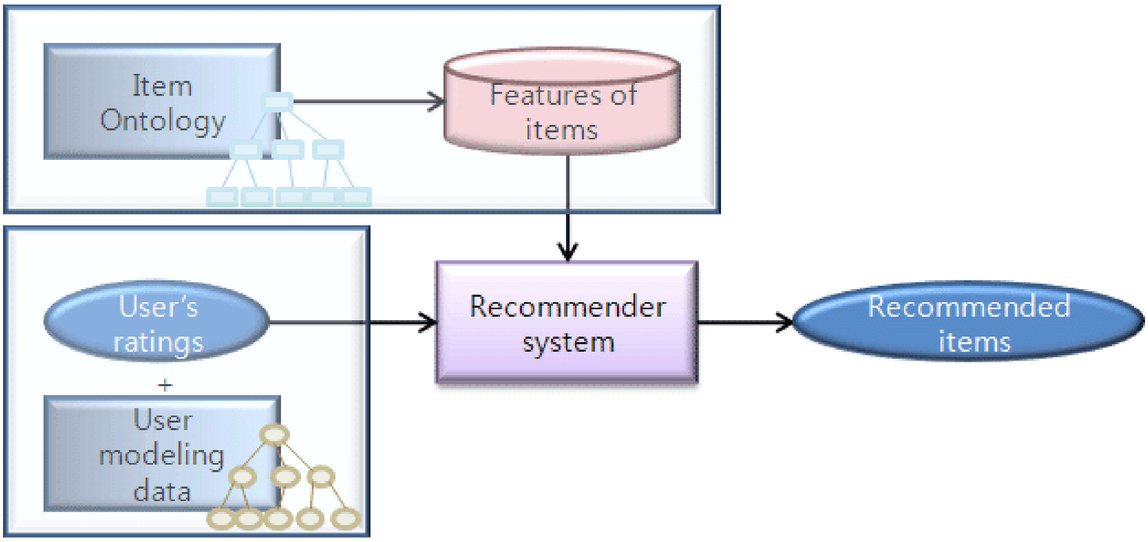

Figure 2: Content based approach

All content based recommender systems has few things in common like means for description of items, user profiles and techniques to compare profile to items to identify what is the most suitable recommendation for a particular user. Content-based recommendation systems analyze item descriptions to identify items that are of particular interest to the user. Because the details of recommendation systems differ based on the representation of items.

B. Collabarotive Filtering based Recommender system Approach -

Collaborative Filtering (CF) systems work by collecting user feedback in the form of ratings for items in a given domain and exploiting similarities in rating behaviour amongst several users in determining how to recommend an item.

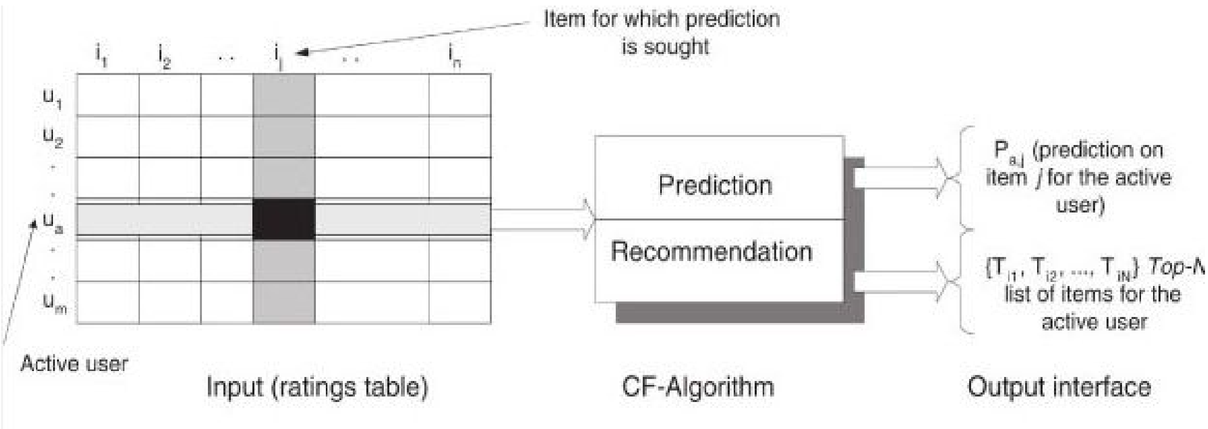

Figure 3. Collaborating filtering

The term "collaborative filtering" was introduced in the context of the first commercial recommender system which was designed to recommend documents drawn from newsgroups to a collection of users. The motivation was to leverage social collaboration in order to prevent users from getting inundated by a large volume of streaming documents. Collaborative filtering, which analyzes usage data across users to find well matched user-item pairs, has since been juxtaposed against the older methodology of content filtering which had its original roots in information retrieval. Collaborative Filtering systems analyze historical interactions alone, while Content-based Filtering systems are based on profile attributes; and Hybrid techniques attempt to combine both of these designs. The architecture of recommender systems and their evaluation on real-world problems is an active area of research. Web sites that uses a collaborative filtering to present information on items and products that are likely to be of interest to the reader. In presenting the recommendations, the recommender system will use details of the registered user's profile and opinions and habits of their whole community of users and compare the information to reference characteristics to present the recommendations.

Collaborative Filtering, accumulates customer product ratings, identifies customers with common ratings and offers recommendations based on inter-customer comparison. In other words, recommendations for a given c1ustomer are based on the behaviour and the evaluations of the other customers. Suggestions for books on Amazon, or movies on Netflix, are real world examples of the operation of industry-strength recommender systems. For example, movie watchers on Netflix frequently provide ratings on a scale of 1 (disliked) to 5 (liked). Such a data source records the quality of interactions between users and items. Additionally, the system may have access to user-specific and item-specific profile attributes such as demographics and product descriptions respectively. Recommender systems differ in the way they analyze these data sources to develop notions of affinity between users and items which can be used to identify well-matched pairs.

The motivation for collaborative filtering comes from the idea that people often get the best recommendations from someone with similar tastes to themselves. Collaborative filtering explores techniques for matching people with similar interests and making recommendations on this basis. Collaborative filtering algorithms often require (1) users' active participation, (2) an easy way to represent users' interests to the system, and (3) algorithms that are able to match people with similar interests.

Typically, the workflow of a collaborative filtering system is:

1. A user expresses his or her preferences by rating items (e.g. books, movies or CDs) of the system. These ratings can be viewed as an approximate representation of the user's interest in the corresponding domain.
2. The system matches this user's ratings against other users' and finds the people with most "similar" tastes.
3. With similar users, the system recommends items that the similar users have rated highly but not yet being rated by this user (presumably the absence of rating is often considered as the unfamiliarity of an item)

A key problem of collaborative filtering is how to combine and weight the preferences of user neighbours. Sometimes, users can immediately rate the recommended items. As a result, the system gains an increasingly accurate representation of user preferences over time.

C. Hybrid based recommender system approach -

Hybrid recommender systems are systems that combine multiple recommendations techniques together to achieve a synergy between them. Several researchers have attempted to combine collaborative filtering and content based approaches in order to smoothen their disadvantages and gain better performance while recommendations. Depending on domain and data characteristics, several hybridization techniques are possible to combine CF and CB techniques which may generate different outputs.

Different ways of hybridization are:
• Implementing CF and CB separately and combine their predictions.
• Incorporating some content based characteristics into collaborative approach.
• Incorporating some collaborative characteristics into content based approach.
• Constructing a general unifying model that incorporates both content-based and collaborative.

Characteristics: Many hybrid approaches are based on CF but CB methods are used to maintain the user profiles and such profiles are used to find similar users. Depending on domain and data characteristics, several hybridization techniques are possible to combine CF and CB techniques which may generate different outputs.

III. Challenges And Limitation

A. Data Sparsity

In practice, many commercial recommender systems are based on large datasets. As a result, the user-item matrix used for collaborative filtering could be extremely large and sparse, which brings about the challenges in the performances of the recommendation. One typical problem caused by the data sparsity is the cold start problem. As collaborative filtering methods recommend items based on users' past preferences, new users will need to rate sufficient number of items to enable the system to capture their preferences accurately and thus provides reliable recommendations. Similarly, new items also have the same problem. When new items are added to system, they need to be rated by substantial number of users before they could be recommended to users who have similar tastes with the ones rated them. The new item problem does not limit the content-based recommendation, because the recommendation of an item is based on its discrete set of descriptive qualities rather than its ratings.

B. Scalability

As the numbers of users and items grow, traditional CF algorithms will suffer serious scalability problems. For example, with tens of millions of customers $O(M)$ and millions of items $O(N)$, a CF algorithm with the complexity of n is already too large. As well, many systems need to react immediately to online requirements and make recommendations for all users regardless of their purchases and ratings history, which demands a higher scalability of a CF system. Large web companies such as Twitter use clusters of machines to scale recommendations for their millions of users, with most computations happening in very large memory machines.

C. Cold Start Problem

The cold start problem is a typical problem in recommendation systems. The "cold start" problem happens in recommendation systems due to the lack of information, on users or items. The Cold-Start problem is a well-known issue in recommendation systems: there is relatively little information about each user, which results in an inability to draw inferences to recommend items to users. The Cold start problem refers to the situation when a new user or item just enters the system. Three kinds of cold start problems are: new user problem, new item problem and new system problem. In such cases, it is really very difficult to provide recommendation as in case of new user, there is very less information about user that is available and also for a new item, no ratings are usually available and thus collaborative filtering cannot make useful recommendations in case of new item as well as new user. However, content based methods can provide recommendation in case of new item as they do not depends on any previous rating information of other users to recommend the item.

IV. Conclusion

The Several recommendation systems have been proposed that are based on collaborative filtering, content and hybrid recommendation methods but these have some problems which are the challenges for research work. It is required to work on this research area to explore and provide new methods that can reduce the challenges and provide recommendation in collaborating filtering a wide range of applications while considering the quality and privacy aspects. Thus, the current recommendation system needs improvement for present and future requirements of better recommendation qualities.

V. REFERENCES

[1] Jun Zou, Arash Einolghozati, and Faramarz Fekri-". Privacy-Preserving Collaborative Filtering using Randomized Perturbation Techniques".

[2] Meenakshi Sharma, Sandeep Mann-"**A** Survey of Recommender Systems: Approaches and Limitations" International Journal of Innovations in Engineering and Technology ICAECE-2013, ISSN : 2319-1058.

[3] Lalita Sharma, Anju Gera-"A Survey of Recommendation System: Research Challenges" International Journal of Engineering Trends and Technology (IJETT) - Volume4Issue5- May 2013.

[4] Ofer Arazy,Nanda Kumar, Bracha Shapira-"Improving Social Recommender Systems" IT Pro May/June 2009 Published by the IEEE Computer Society 1520-9202/09/ 2009 IEEE.

[5] Esma A¨ımeur _ Gilles Brassard _ Jos´e M. Fernandez _ Flavien Serge Mani Onana-"ALAMBIC: A Privacy-Preserving Recommender System for Electronic Commerce"

[6] Hao Ma,Dengyong Zhou,Chao Liu, Michael R. Lyu,Irwin King-"Recommender Systems with Social Regularization" February 9–12, 2011, Hong Kong, China. Copyright 2011 ACM 978-1-4503-0493-1/11/02