

A Study on Different Phases and Various Recommendation System Techniques

Mallari Vijay Kumar, P.N.V.S. Pavan Kumar

Abstract--- Now-a-days, recommender systems(RS)are playing a crucial role in human tasks. The online websites like movies, restaurants, education and much more, uses the recommender systems to suggest the customers for E-commerce. Recommender systems make money by attracting users with recommendations. Each system may use different datasets from various sources to analyze the user behavior and to find interesting patterns that predict the user's future purchase or taste. Most of the times lack of data results to the inappropriate recommendations (means bad recommendations). This paper reviews different phases involved in implementing RS and various recommender methods including the study of those methods that are used in several papers of various authors. We also included the advantages and disadvantages of each method. Finally, this paper also gives analyses of various challenges and issues (problems) faced in the implementation of RS algorithms.

Index terms-phases, collaborative filtering, content-based filtering, hybrid filtering, problems.

I. INTRODUCTION

In the world of E-commerce people have various choices in finding the items of their interests where recommender systems help people in finding and evaluating the items of interest. These RS tools uses various parameters like purchase, view, listen to, etc., to interact with the users who like to consume the items, by associating the opinions of other people with the consumers opinions. Recommender systems[1] filters out the vital information from huge amount of overloaded information which is generated dynamically according to the user's preferences and interests. At the beginning so many approaches were applied to the basic problem for making effective and responsive recommender systems. The earliest system was content filtering system which is designed to remove the overloaded information in domain specific textual corpus. Recommendation systems which uses information retrieval IR methods that incorporate information retrieval methods are often used to satisfy short term needs which involve static databases.

Without computers, a person often receives recommendations by listening to what people around him have to say. If several people mention that they have liked a particular movie then we can take that as a recommendation. In general, Collaborative filtering (CF) is the processing of words that came from the mouth of other people. The CF models gives the results which are evaluated from a large set of words or descriptions [1].

The techniques, user-user correlations and nearest-neighbor algorithms, which are proposed by the Resnick and

Varian in 1997 [1], were employed by most of the personalized systems. In recommender systems user-to-user correlations are used to identify the ratings given by the users which contains a higher predictable value for the items. Users who have only one or two items which are rated in common should not be treated as strongly correlated. Nearest-neighbor algorithm evaluates the gap among the users based on their purchase or interests or preference history. This distance may vary depending on the amount of users, domains, number of items recommended and the degree of inter-related items between users. Both correlated and nearest-neighbor techniques show that those are effective and efficient in their recommendations.

The term data mining, in RS domain, can be described as the collection of techniques which are used to produce recommendation rules from the huge data sets. One can use the recommender systems and may incorporate data mining techniques to make their recommendations from the knowledge obtained or gathered from the actions and attributes of users. These techniques are clustering, classification, association, and Mining which is used to generate similarity graphs, in which it contains nodes and edges where nodes are users and edges contains a weight for similarity.

II. PHASES IN RECOMMENDATION SYSTEM

There are three main phases in the recommendation process; they are Information collection phase, Learning phase, Prediction/Recommendation phase [3]. Fig.1 shows the phases of recommendation system.

A. Information collection(IC) phase

This phase collects user's details to produce a model for prediction tasks which includes user's attributes, behaviors or content of the resources. The system has to gather much information about the user in order to provide best recommendations. Recommendation systems depends upon different types of inputs, for example a high quality direct feedback which contains interests in items as its input with respect to the user and indirect feedback by concluding user preferences indirectly by analyzing user's behavior. From the combination of indirect and direct feedbacks we can also obtain Hybrid feedback [3]. A recommendation system will not produce correct and accurate output until the user model is well constructed.

1. Implicit feedback

The system automates the user's preferences by analyzing the user's history, purchases, links clicked by the user and

Revised Manuscript Received on February 22, 2019.

Mallari Vijay Kumar, Assistant Professor Department of CSE-GPREC, AP, India (vijay930.vk@gmail.com)

Sri P.N.V.S. Pavan Kumar, Assistant Professor Department of CSE-GPREC, AP, India (pavan.cse@gprec.ac.in)

time spent on number of web pages, contents of emails and so on. The need for user's action is not required for this kind of feedback instead it automatically provides recommendations by analyzing the above mentioned contents.

2. Explicit feedback

This system asks the users for their feedback to producerecommendations. The quality and efficiency of the RS relies on the user ratings. Despite of implicit feedback this system expects more work from the user.

3. Hybrid feedback

The weaknesses of both indirect and direct feedbacks are removed and strengths of them are combined to form this hybrid feedback. This feedback can be attained by using the indirect data as an attribute for recommendation while allowing users to provide direct feedbacks and ratings.

B. Learning Phase

This phase applies learning algorithms on the user's data which are obtained from the feedbacks in IC phase. The learning algorithms are the methods which are helpful in drawing out the patterns appropriate for application in certain situations.

C. Prediction/recommendation phase

By analyzing the patterns which are obtained from learning phase, this phase provides recommendation or predictions for the given data. The trained data in learning phase provides certain patterns and then which are subjected to envisionthe user's course of action or future interests.

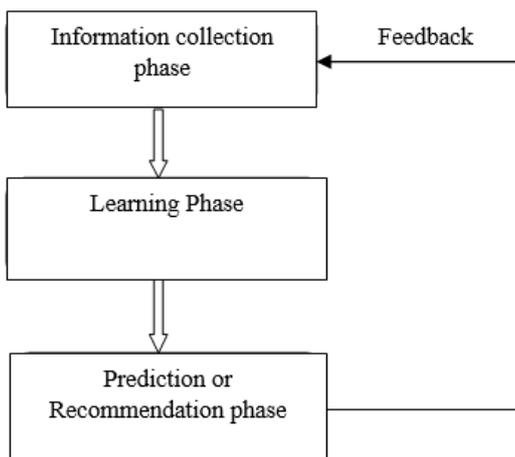


Fig.1 Phases in Recommendation system

III. METHODS OF RECOMMENDER SYSTEMS

This section gives the details of the three main recommender system methods as follows: content-based filtering, collaborative filtering and hybrid filtering technique. Fig.2 shows these three recommender methods.

A. Content-based filtering

This technique suggests some identical objectsusing the interests shown on previously selecteditem(s). This technique usesobject's metadata such as titles, writers, number of pages, publisher names etc., for a book

recommendation. The general idea behind this system is that if he/she likes a particular item, then he/she may also like a similar item [4].

To infuse the metadata of the items into the system an algorithm called Item-Presentation can be utilized. A conventional algorithm is the TF-IDF representation, which is also called as vector space representation. The system mainly focuses on two things: 1. A plan which is designed from the user's choices and 2. The past data which was obtained from the user's communication with the RS. In this method, some weights are credited to the items. These weights indicate the user's top visited or rated items and by adopting various kinds of techniques these weights are estimated. Some simple machine learning algorithms such as Bayesian Classifiers, cluster analysis, decision trees, and artificial neural networks are used to estimate the probability that user is likely going to like or show interest an item [5]. This content-based technique may won't work properly when there is a single content source.

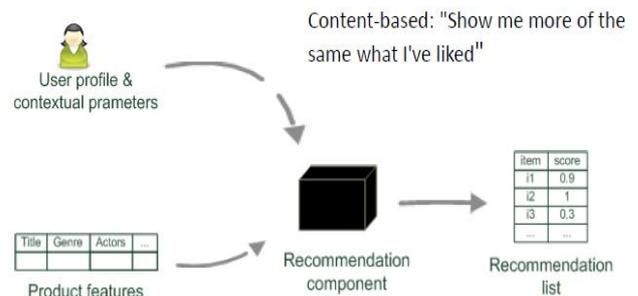


Fig.3 Content-based filtering technique

B. collaborative filtering(CF)

CF is a prediction technique which is domain independent. Recommendations will be shown to the user according to the other user's behavior and set of commonalities between those users. Some datasets might be produced depending on the user desires and according to that recommendations will be provided to the users [4].

The CF does not rely on the content which is an advantage and thus it is proficient enough to recommend complex items such as movies, where the analysis of its metadata is not needed [5].

Several approaches have been proposed to estimateuser and item similarity in recommendation systems, for example, k-nearest neighbor(k-NN) approach and the Pearson correlations proposed by Allen.

According to Saha [6], two major kinds of CF are there; they are "user-user" and "item-item".

1. User-user

The main concept of user-user collaborative filtering recommender is to select a neighborhood of users who are similar to the user and analyze their tastes and then providing suggestions of the items to the user which are liked by the neighborhood users.

2. Item-item

Here the similarity is calculated from the item ratings. If a person gives rating to an item, then similar items are recommended which have same ratings. Suppose, for example, In Amazon when a user rate an item very high, then other items is suggested which have higher ratings.

In [4], CFclassified into three types: memory-based, model-based, and hybrid methods.

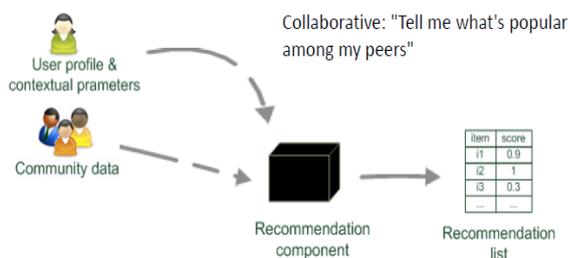


Fig.4 Collaborative filtering technique

C. Hybrid filtering

Hybrid filtering is the combination of different filtering approaches; one can make use of the advantages of both the above filtering methods by avoiding the disadvantages in them and can combine to form a hybrid filtering technique. One of the good examples for hybrid filtering is the Netflix, it compares the user’s and similar users history (i.e., Collaborative filtering) and offer movies based on the user ratings which contains common aspects(Content-based filtering).

There are some other techniques that can be combined to form various hybrid filtering techniques; they are demographic and knowledge-based.

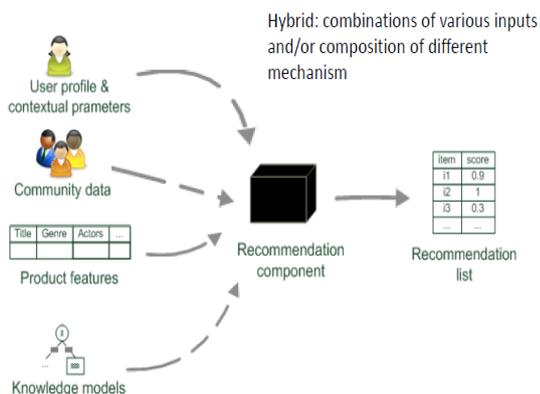


Fig.5 Hybrid filtering

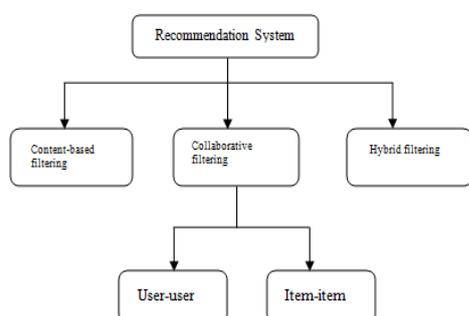


Fig.2 methods of Recommender systems

IV. ADVANTAGES AND DISADVANTAGES OF RECOMMENDER TECHNIQUES

Recommender System methods	Content-based filtering	Collaborative filtering	Hybrid filtering
Advantages	User independence, Transparency	Improve recommendation performance	Overcome cold start, sparsity problem
Disadvantages	Limited content analysis	Data sparsity, Scalability, Synonymy	Increased complexity, expensive in implementation

V. PROBLEMS OF RECOMMENDER SYSTEMS

To build an efficient recommender system, the companies or developers need to overcome some problems; they are, as shown below.

A. Lack of data

The sizable issue encountered by the RS may reasonably is the lack of data but where RS needs lots of data to make an efficient model of recommendation system to make recommendations. Of course, the big companies like Google, Facebook has lots of customer data. As explained in the phases of recommender system, a good RS is a one which mainly needs data about items and users first, then analyze user data, then the algorithms do its work to give recommends as a result. It’s like a hen and egg problem- to get accurate recommendations, one needs a lots of user data, so the user can get a lot of objects as the recommendation.

B. Changing data

Any algorithm will find its difficulty when it comes to the fashion trends. Most fashion- challenged people rely on friends and family who are aware of latest things in fashion to recommend new clothes forthem. In this case, the item recommendation doesn’t work.

C. Changing user preferences

A user always doesn’t want to buy or show interest on a single item every time. For example, if a person browses for a book in Amazon on one day and the other day the user may want to search for an entire different category, even though the recommendation shows familiar objectsbased on his search history.

D. Unpredictable Items

Even though the recommendation system suggests similar items, some suggestions may tend the user to diverse or lose interest. Like, for example, suggesting a user a smart phone brand which the user does like it.



E. Complexity

A recommendation system has to lots of variable which are connected to one another which is complex to analyze and to generate a model based on those variables.

VI. CONCLUSION

In this paper, we have made literature survey of different phases and various techniques in the recommender systems. It has been observed from the study that collaborative filtering has a better advantage over the other techniques and the user-user collaborative filtering gives efficient and correct results than the item-item filtering. Every technique has its advantages and disadvantages. To increase the quality of the results hybrid filtering is used but it has its own complexity and expense for developing hybridized filtering. The phases in recommender system explain the cycle of the data obtained from the user, more the number of cycles more will be the accuracy.

VII. ACKNOWLEDGMENT

I would like to express my gratitude to my guide Sir P.N.V.S.Pavan Kumar for their even willingness to give me valuable and direction. Finally, I would like to thank department of computer science engineering for their support.

REFERENCES

1. Schafer, J. Ben. "Application of Data Mining to Recommender Systems." *Encyclopedia of Data Warehousing and Mining*, 2015, pp. 345–353., doi:10.4018/9781591405573.ch009.
2. Francesco Rikki, et al. "Recommender Systems Handbook." *Recommender Systems Handbook*, 2011, pp. 1–845., doi:10.1007/978-0-387-85820-3.
3. Isinkaye, F.o., et al. "Recommendation Systems: Principles, Methods and Evaluation." *Egyptian Informatics Journal*, vol. 16, no. 3, 2015, pp. 261–273., doi:10.1016/j.eij.2015.06.005.
4. Patel, Bansari, et al. "Methods of Recommender System: A Review." 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017, doi:10.1109/iciiecs.2017.8275856.
5. https://en.wikipedia.org/w/index.php?title=Recommender_system&oldid=86592141
6. Saha, Tanwistha, et al. "Journey to planet Datum and Beyond: Coursera-Introduction to recommender systems." *Encyclopedia of Social Network Analysis and Mining*, 2017, pp. 1–12., doi:10.1007/978-1-4614-7163-9_110164-1.
7. Chu, Pang-Ming, and Shie-Jue Lee. "A Novel Recommender System for E-Commerce." *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2017, doi:10.1109/cisp-bmei.2017.8302310.