

# A Survey on Recommendation System in E-Commerce

M.B.Nirmala

Associate Professor

Department of Computer Science and Engineering  
Siddaganga Institute of Technology, Tumkur  
nirmalamb@gmail.com

Manushree H K

Mtech Student

Computer Network Engineering  
Siddaganga Institute of Technology, Tumkur  
manuhkswamy@gmail.com

**Abstract**—Today in internet, there is number of information need to be filter to reduce the information overload, which created problem to many users in internet. Recommendation system will overcome from this problem by filtering the information in the internet and provide the useful information to the user. The Recommendation system will search the information in the large database which is generated dynamically to provide users with useful content and services. In this paper it gives the different characteristics of various prediction techniques in order to serve as a compass for research and practice in recommendation system field.

**Keywords**— Collaborative filtering; Content based filtering; Hybrid filtering

\*\*\*\*\*

## I. INTRODUCTION

In the rapid development of internet today, the number of visitors to the internet is increasing day by day, so there is a biggest challenge of information overload. Several data retrieval systems such as Google, DevilFinder and Altvista are solved partially this problem but priority and personal information is absent, so to overcome from this, recommendation system is introduced to solve the problem. Recommendation system is a filtering technique to deal with information overload problem. It will recommend an item based on the user preference and interest. The recommendation system will have an ability to predict an item if the user is interested in this or not.

Recommendation system is very useful for service providers and the users. It reduce the time and cost of transactions to finding and selecting items in an e-commerce environment. Recommendation system will also demonstrate to improve the quality and decision making process. In the present e-commerce world the recommendation system will enhance the revenue and improve the selling more products effectively. In libraries the recommendation system will allow users to move beyond catalogue searches so need to use better recommendation techniques with system that will provide relevant recommendation system for users. The immense growth of the e-commerce there is a poor decision making by the user to search for an item from the vast database so recommendation system will help the user by suggesting the better items which they are interested.

## II. RELATED WORK

In the past few years the cloud computing has dragged the attention of everyone towards it. A cloud market is a place where different cloud providers deliver their service to the cloud users. For the user to select the best service provided by the service provider a recommendation system is proposed where the user by making use of this system select the best desired service provider. Based on the factors like network quality, virtual machine platform factors and some other factors the system will provide the recommendations to the user for the service. In traditional recommendation system one of the widely used technologies is Collaborative filtering. This collaborative filtering aims to figure out the similarity between the users and the existing data of interaction in order to find out the matches and suggest recommendations further. In e-commerce there are several methodologies to achieve the recommendation in real time they are mentioned below.

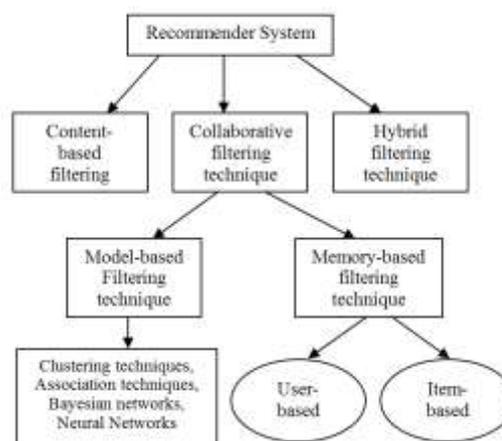


Figure 1: Different recommendation techniques

**A. Content-based filtering:**

Content based filtering is a domain dependent algorithm. In this the system recommends the item which is based on the similarity of the item which is purchased by other person and also it recommends the item based on the positive ratings to the particular item which is purchased. In this system mainly it records the user’s history and it computes the similarities of user’s using statistical methods and machine learning, by using this it will recommend to the other user of same similarities.

Advantages: In this system it will recommend the items if there is any information which is available in the history it will recommend as soon as it get.

Disadvantages: In this the items and the attributes must be recognized by the machine and it cannot filter the items based on the quality, style and the view point of the user. Content based filtering won’t support the personal recommendation it just suggest the similar item to the user.

**B. Collaborative filtering technique**

Collaborative filtering technique is a domain independent algorithm [1]. It recommends the items based on the similar taste, it collects the information of an item who as experienced an item which is to be recommend to the current user. In CF it requires the user’s to give opinion on each items which they purchase so it collects all the information regarding an item and recommends the same to current user’s who as same taste.

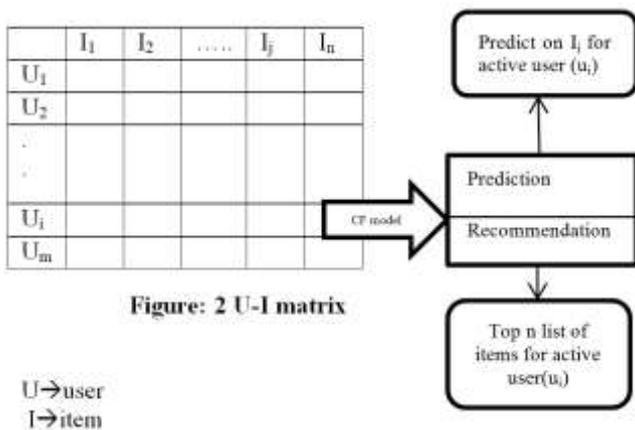


Figure: 2 U-I matrix

Prediction will give the numerical value, it gives the score of prediction for the item j for the user i. In the recommendation model it gives the top n items which user like is as shown in the figure 2.

In collaborative filtering technique there are two types

1. Model-based filtering technique
2. Memory-based filtering technique

Model-based filtering technique: In model based it will give the previous rating for the user to improve the performance of the recommendation. This technique will use the machine learning and data mining technologies for the quick recommendation to the user. In model based technique it will evaluate the user and the item matrix to find out the relation between the items. so this technique will compare the most top

n list of the items. This technique will come up with solution for sparsity problem in recommendation.

There are several learning algorithm in the model based filtering technique to change the recommendation manner. They are mentioned below:

Clustering technique: This technique will be used in various domains like image processing, Pattern recognition and Data analysis. Clustering algorithm will divide the huge information into a set of sub cluster to give a meaningful group exists within the group. After creation of cluster it will take the average opinions of each user in the cluster and recommend for the individual users. If the intra-cluster similarity is high it will give a high quality cluster else if the intra cluster similarity is low it will give a poor quality of recommendation. If the user as partial participation in the cluster it will take the average of all the users and recommend.

The clustering technique has most commonly used methods they are k-means and self -organizing map (SOM). K-means takes an input parameter and then partitions a set of n items into K clusters. The SOM is a method for an unsupervised learning, based on artificial neurons clustering technique. In collaborative based algorithm the candidate set will reduce using clustering techniques.

Association technique: In a transaction the prediction of the occurrence of an item will be calculated based on the presence of another item and it will be extracted by the association rule algorithm. Association rules will give an awfully compact illustration of preference information which will improve potency of storage moreover as performance. For an example take a set of items in a transaction where each transaction is a set of items so an association rule applies the form  $I_1 \rightarrow I_2$ , where  $I_1$  and  $I_2$  are the two sets of an item.

Bayesian networks: The Bayesian network (BN) is represented by an acyclic graph which shows the dependencies between random variables and it gives a graphical representation of probabilistic model [7]. The below figure 3 shows the example for Bayesian network.

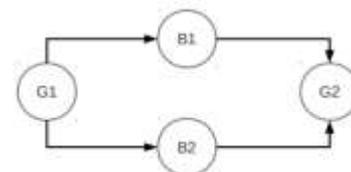


Figure 3: Bayesian network

The advantage of Bayesian network is that the output of a recommendation is a set of probabilities, which allows for ranking from the most to the least probable. For example as shown above to calculate the probability value of the random variable B1 the below equation can be used. The G1 and G2 can be represented as BPMN gateways, B1 and B2 as other

blocks, e.g. Events or tasks, thus by using given any of these blocks, we can calculate a particular block probability.

$$P(B1) = \sum_{G1} \sum_{G2} \sum_{B2} P(G1)P(B1|G1)P(B2|G1)P(G2|B1, B2)$$

Memory based algorithm: To make the prediction the memory based CF algorithms can utilize the entire or portion of the user-item database. People with same kind of interests are grouped together and just by knowing the adjacent user of the active or new user, a prediction can be made on new topics.

The neighborhood-based CF algorithm, a pervasive memory-based CF algorithm takes after the accompanying strides: by calculating the likeness or weight, which is similar to the distance, correlation, or weight between two clients or two things, and create an prediction for the active client by taking the weighted normal of the considerable number of appraisals of the client or thing on a specific thing or client, or utilizing a basic weighted normal. By discovering the most comparable clients or things (closest neighbors) it is possible to generate a top-recommendation, we have to discover most comparable clients or things (closest neighbors) the neighbors are aggregated once the similarities are computed, to get the top-most frequent items as the recommendation.

User based algorithm: By making use of the Pearson correlation or vector-space model, User-based top-recommendation algorithms initially find out the most similar users (nearest neighbors) to the active user, where in every user is taken as a vector in the -dimensional item space and by finding the likeliness between the vectors we can find out the similarities between the active user and other users. Once the most likely users have been identified next step is that their related rows in user-item matrix are combined to get the set of items. With the set, user-based CF strategies then suggest the top-most incessant things in that the dynamic client not acquired. Client-based top-proposal algorithms have impediments identified and real-time performance.

Item based algorithm: The scalability issues of the user-based top-recommendation algorithms are addressed by the item-based top-recommendation algorithms. Based on the similarity the algorithm initially process the most similar items for each item; then a set is formed by identifying the and uniting of the most similar items of recommended items and removing each of the items in the set. The likeness between the each item of the set and set itself is computed by the user who has previously brought. The items which have been listed according to the decreasing order of the similarity will become the final and recommended resulting set of items. One issue of this strategy is, the point at which the joint dispersion of an arrangement of items in the set, the above schemes can conceivably create suboptimal recommendations. Deshpande

and Karypis gave a solution for the above issue by creating higher-order item-based top-recommendation algorithms. These algorithms by utilizing all blends of items up to a specific size, determine the itemsets as a recommendation to a user.

Advantages: CF systems do not require content information about neither users nor items to be machine recognizable. In pure CF system it utilize only the ratings and it do not require any extra information regarding users or items. CF systems can make an assessment of quality of an item and also style and viewpoint by consideration of other user experience.

Disadvantages: The main disadvantage in CF system is that it cannot recommend an item if there is no ratings available for an item. When there is less data or information about the user's ratings it will give a poor accuracy. This system will not consider the information about the items when they recommend to the user because it's not a content based filtering.

### C. Hybrid filtering technique

The hybrid filtering is the combination of collaborative filtering and the content filtering. The implementation of hybrid filtering will be done in different ways. The prediction of content based and collaborative based filtering will be done separately and then adding the content-based capabilities to a collaborative – based filtering technique or vice-versa, even it can be done by unifying the techniques into single model. The comparison of performance of the hybrid filtering with other filtering like collaborative and content-based technique and proved that the hybrid technique will provide more accurate than other recommendation techniques. These techniques will solve the problem in recommendation system such as sparsity and cold start problem.

Summarizing the comparison between the recommendation system techniques is given below:

Comparison of different recommendation technique

Factor		Collaborative filtering	Content-Based filtering	Hybrid Filtering
Heterogeneous	Low	✓	✓	✓
	High	+	-	-
Risk	Low	✓	✓	✓
	High	-	-	+
Churn	Low	✓	✓	✓
	High	-	+	+
Interaction style	Implicit	+	+	-
	Explicit	✓	✓	+
Preferences	Stable	✓	✓	✓
	Unstable	-	-	+
Scrutability	Required	-	-	+
	Not Required	✓	✓	✓

Table 1: ✓ -> Character which is present, +/- -> Amount of character

The above table will give you the overall idea about the recommendation system and the different approaches which can be implemented in the recent years. It's very essential for any user who can quickly find the item in the e-commerce world

### III. CONCLUSION AND FUTURE WORK

Recommendation system is a new powerful technology for suggesting an item from the huge database to the user. It plays an important role in the e-commerce world because it's very difficult to search for an item to the user from the huge volume which is dynamically generated, so it saves the user time by wasting a lot on searching things which they are interested to buy.

In this paper we discussed about the different approaches in the recommendation system and further we can improve this recommendation system by adding new features like merging the social networks into e-commerce world so the users can personally get the reviews about the particular item which is needed by the current user, it is more trusting way of recommend than by predicting on the user ratings.

### REFERENCES

- [1] Xiaoyuan su and Taghi M. Khoshgoftaar "A survey of Collaborative filtering techniques".
- [2] Dehua kong, Yuqing zhai "Trust Based Recommendation system in service-oriented cloud computing".
- [3] Fuxi Tang, Kejian Liu, Ling Feng "Research on the integration strategy of group recommendation based on user's interactive behaviors".
- [4] Zaint T., Aslam M., Imran M R., Martinez-Enriquez A.M "Cloud Service Recommender System using clustering".
- [5] J.A. konstan, j. riedl recommender systems: from algorithms to user experience user model user-adapt interact, 22 (2012), pp. 101–123
- [6] c. pan, w. li research paper recommendation with topic analysis in computer design and applications IEEE, 4 (2010) pp. v4-264.
- [7] Krzysztof Kluza, Mateusz Baran, Szymon Bobek, Grzegorz J.Nalepa. "Overview of recommendation techniques in business process modeling?"