

# Techniques of Recommender System

Harleen kaur, Gourav Bathla

**Abstract** – The term Recommender system is described as any organization that provides personalized suggestions as a result and it effects the user in the individualized way to favorable items from the large number of opinions. The voluminous inflation of the reachable data online and also the number of users have lead to the information overload problem. To overcome this problem the recommender system came into play as it is able to prioritize and personalize the data. Recommendation systems have developed alongside with the net. Recommender system has mainly three data filtering methods such as content based filtering technique, collaborative based filtering technique and the hybrid approach to manage the data overload problem and to recommends the items to the user the items they are interested in from the dynamically generated data. This paper makes a comprehensive introduction to the recommender system with its types, content based filtering , collaborative filtering and the hybrid recommendation.

**Keywords** – Recommendation system, Collaborative filtering, Content based filtering, Hybrid recommendation

## INTRODUCTION

In this competitive world every product has its reviews and the ratings given by the users on the e-commerce sites they are using. The new users are always willing to go through the product reviews before buying that particular product[1]. Same is in the case of movies people will read the reviews of the movie they want to watch. Meanwhile, the increasing online information leads to the information overload problem. To deal with this problem recommender system automatically suggests the item to the particular user according to the user's profile or the ratings[2]. To determine the interest of the users and to help the users in making the search easier the recommendation system sorts the huge amount of data. It can perform its task only when the user's past information, his or her browsing history, previous purchases and the feedback is available[3].

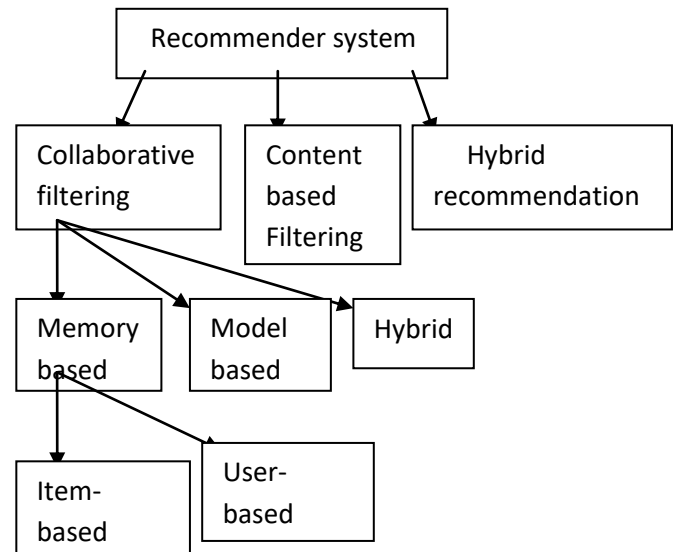
**Manuscript Received on July 02,2019.**

**Harleen Kaur**, CSE, Chandigarh University, Gahrnan, Mohali, Punjab, India

**Gourav Bathla**, CSE, Chandigarh University, Gahrnan, Mohali, Punjab, India

Recommender system that provides the accurate and the fast results will attract the users as per their interest[4].

## A. Types of Recommender systems



**Fig. 1:** Types of recommender system

Data mining has a sub domain known as Recommendation system. Personalized recommendation system have mainly three approaches such as Content based filtering , collaborative filtering and Hybrid approach.

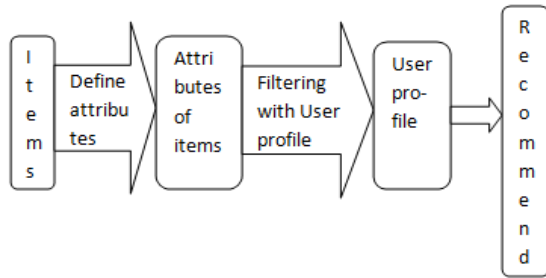
## B. Content based filtering

Content based filtering method is based on the semantic search that is the information retrieval[1]. The other name for content based filtering technique is the cognitive filtering. This technique suggests items on the basis of customer's profile and the customer's item profile. The user's profile is created when the user starts the system[3]. It collects the interest of the users and recommend the items after analyzing the features of the items and the users[4]. The recommended items are identical to the things that were liked by the customer earlier or previously and they also match the attributes of the user. This technique works well only when the attributes are presented in a clear and proper way[5]. In content based filtering technique the item description and user's profile have a lot of importance. It follows the following steps:

- For recommendation attributes of the

items are to be defined.

- The active user's preferences and the attributes of the items are compared.
- Things are recommended according to the interest of the user.



**Fig. 2:** Content based filtering

When the profile of the user and the item's attributes are known then the main role of CBF is to foresee whether the active user would like the recommended item or dislike the item[5].

### ***Advantages of Content based filtering***

1. Content based filtering technique gives the user independency on the basis of the exclusive ratings that are used to build the user profile.
2. In this technique the user gets to know how the recommender system processes that is there is lot of transparency.
3. CBF also recommends those items which have not been used by any person and this is indirectly favorable for the current user.

### ***Limitations of Content based filtering***

1. Creating the attributes of the items are difficult in some particular field.
2. Content based filtering suffers from the overspecialization problem as it recommends items that are all of same type[5]. The user won't get the items beyond the predefined boundary which leads to poor recommendation[1].
3. In this technique we can't check whether the recommended is correct or not because in CBF the user's feedback is not collected as the users don't provide the ratings for the items.
4. CBF only uses the statistic data of the person but the interaction data of the person is not used[4].

### ***C. Collaborative filtering***

In recommender system the most widely implemented and the most popular is the collaborative filtering. In CF recommendation system recommendation is done on the basis of past ratings of the items of the people that are analogous to the target person[3]. The main idea of the collaborative filtering is to gather the information regarding the old user's behavior and opinions so as to correlate with the new user whether they are similar or not[4]. Every user's ratings, viewing history and the purchasing are recorded which grants the recommender system to build the connection among the people that have identical nature and among the

things of interest to the look a like user[7]. CF believes that the users that have analogous preferences previously will surely have analogous taste in the forthcoming too[8]. Implementing collaborative filtering technique is easy when dealt with the low data dependency and further gives accurate suggestions for recommendation[9]. Collaborative filtering can be further branched into three main categories, model based, memory based and hybrid based collaborative filtering.

### ***Model-Based***

Model based method fills matrix by predicting the things that have not been seen by the user previously. To predict the ratings of the unrated items can be done by using the data mining techniques[3]. In case of this technique on the basis of the information collected from the repository is used to generate a model for recommending without making the use of dataset everytime[5]. Speed and scalability are the two factors that are improved when this approach is used. It also enhances the prediction accuracy value for the algorithm[6]. The most commonly used model based technique is Matrix factorization[10].

### ***Memory-Based***

In memory based technique similarity among people or things is calculated for recommending items. The users with similar likings will have higher similarity value[3]. This technique finds the users that resembles the target user's liking and disliking and then predicts the liking of the target user for new items. It is an effective and easy method to implement[5]. Each user having similar

taste are grouped together[6].

### User-based

User based algorithm evaluates the similarity among the users on the basis of the ratings they have provided to a particular item[8].

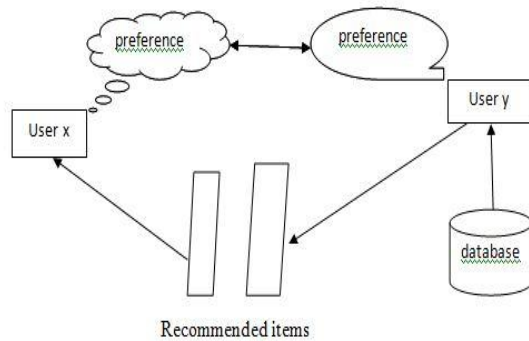


Fig. 3: User based algorithm

### Item-based

Item based algorithm is an memory based method for recommendations[5]. It calculates similarity among items to make predictions[8]. In item based collaborative filtering there are two phases. Foremost is to evaluate the similarity among two things using the similarity measures. Then the similarity values are used to forecast the ratings for unknown things.

### Similarity measures

Euclidian distance

In collaborative filtering approach uses k-nearest neighbor (KNN) classifier the most. KNN depends on the similarity measure or the distance measure. The easiest and most widely used distance measure is the Euclidian distance.

$$d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \quad (1)$$

Where n is denoted by number of attributes and  $x_k$  and  $y_k$  are the  $k^{\text{th}}$  attributes of data objects x and y.

Cosine similarity

It considers items as document vectors of an n-dimensional space and evaluate their similarity as the cosine of the angle that they form:

$$\cos(x, y) = \frac{(x, y)}{\|x\| \|y\|} \quad (2)$$

Where . denotes the vector dot product and  $\|x\|$  is the norm of vector x.

Pearson correlation

Correlation between the items calculates the linear relationship among the objects. There are many correlation coefficients that can be implemented, but Pearson correlation is widely implemented.

$$Pearson(x, y) = \frac{\Sigma(x, y)}{\sigma_x \times \sigma_y} \quad (3)$$

$\Sigma$  symbolizes covariance and  $\sigma$  as standard deviation[12].

### Hybrid collaborative filtering

Hybrid approach enhances the all over quality of recommendation system as it combines the other collaborative filtering techniques among themselves to make the recommendations. They are used to overcome the drawbacks of CF such as loss of information and sparsity. But these methods are quite expensive and complex[3].

### Advantages of collaborative filtering

1. Implementing memory based CF makes the recommendation process easier.
2. Using modeling based CF enhances the performance of the prediction.
3. Addition of the new data in an incremental manner is made easy in memory based technique.

### Disadvantages of collaborative filtering

1. Cold start: Cold start complication emerges when a new user signs in and there are no past records of the user[7]. In CF approach enormous amount information is required to provide the correct recommendations[6].
2. Scalability: Recommendations are made when there are billions of people and things in the organization. These systems have to recommend things to the users without the past history or the ratings to the online requirements[5]. Therefore, computation power should be high to suggest items[6].
3. Sparsity: The items are sold on the e-commerce sites in large amount. But every item is not rated by every

user. Hence, small part of the database is rated[6].

4. Synonymy: When a same item has different names then this problem arises. In recommender system it is hard to discover the latent association so the system considers these items as different items. For example, ‘adult movie’ and ‘adult film’ are same items but they are considered different by the system. In case of memory based CF they will not calculate the similarity between these two terms as they find it different. Therefore, synonymism abrupts the performance of the recommendation process.

5. Gray sheep: People who don’t like or dislike the items in the group of users they don’t get benefit from the recommender system[5].

7. Shilling attack: It is refered as ‘‘profile injection’’ attack, as fake profiles are made to influence the behavior of the recommendation process[13]. Some users would give positive comments for their own product and negative comments for their rivalries[5].

### C. Hybrid recommendation

For retrieving the valuable information content and collaborative based filtering techniques are implemented. Both the techniques have their pros and cons. Hybrid recommendation aggregates both the methods to eradicate the drawbacks of both filtering techniques. Hybrid filtering approach helps in enhancing the accuracy and efficiency of the recommender process . Hybrid recommendation approaches can be used in different ways:

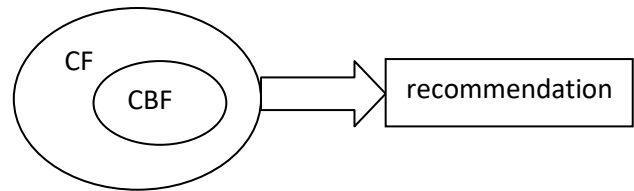
- 1) Make use of content based and collaborative filtering algorithms individually and then put together their suggestions.



**Fig. 4:** Recommendation 1 of CB and CBF

Figure 4 depicts that combining the content based and collaborative filtering techniques would give actual recommendations to the users.

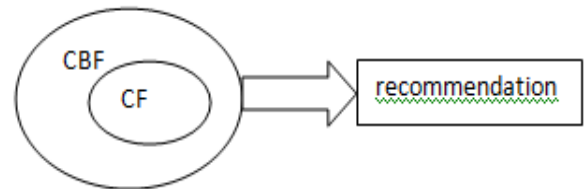
- 2) Using some of the features of content based method in the collaborative filtering technique.



**Fig. 5:** Recommendation 2 of CB and CBF

Figure 5 shows adding some features of CBF in CF helps in resolving two major drawbacks of each method. When a new user sighs in the system, he or she does not have any past history regarding his or her likes and dislikes. So the system is not able to recommend. This complication is known as cold start issue in collaborative filtering. And overspecialization in CBF.

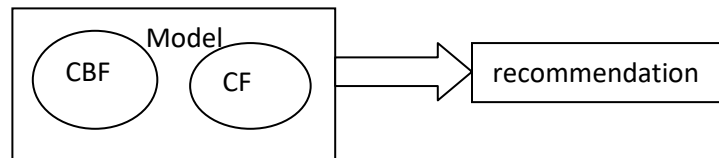
- 3) Adding some collaborative features in to content based filtering method.



**Fig. 6:** Recommendation 3 of CB and CBF

Figure 6 illustrates that adding some features of CF in CBF helps in resolving drawbacks of each method.

- 4) Constructing a model including content based and collaborative based filtering features.



**Fig. 7:** Recommendation 4 of CB and CBF

In figure 7 it is shown that aggregating the features of CBF and CF for building a model will help in enhancing the accuracy and efficiency of the recommendation[6].

### I. Making good recommendation system

- 1) Relevant recommendations should be suggested to the user.
- 2) Cold start issue should be solved for current user and even for current items.
- 3) Recommendation system should suggest personalized

recommendations according to the likes and the dislikes of the user.

4) Handling data sparsity in the recommendation process is very important.

5) In the recommendation system the users expects the system to have high accuracy and efficiency.

6) Shilling attack is also a major issue in recommendation process which should be eradicated[1].

**Table I:** Comparison Of different filtering approaches

	Collaborative filtering	Content based filtering	Hybrid recommendation
Advantages	1) CF techniques use ratings only instead of the profile of the user and items. 2) CF methods rely on the experience of the other people. 3) As CF methods consider experience of the other people so it produces personalized recommendations.	1) CBF will analyze the profile of the user and the item individually for recommending. 2) CBF will suggest on the basis of the item's features.	1)Hybrid recommendation aggregates the content based and collaborative filtering approach to overcome the drawbacks of these method.
Disadvantages	1) Cold start complication emerges when a new user signs in and there are no past records of the user. 2) In CF approach enormous amount information is required to provide the correct recommendations. 3) CF techniques slow down when dealing with enormous data.	1) In CBF if the profile of the item and the user does not contain the enough information the recommendation will not be precise. 2) CBF suffers from synonymity if two words have different spelling they would be treated as independent words.	1)It is complex because two different techniques are joined to work as one system.

## **II. Related work**

This paper[1] reviews the past, present and the future of the recommendation system by analyzing the different techniques used in the recommendation process. It has also compared the past research work done on the recommendation system using any of the three methods. In this paper[2] the author have explored the Collaborative filtering based social recommender system. The author describes the tasks that are associated with the recommender system and the traditional approaches that have not been used in the social network recommendation process. The prime focus of this[3] paper is to cover all kinds of approaches involved in recommendation system. Mainly three methods are widely used they are content based and

collaborative based and hybrid recommendation. They have also mentioned the drawbacks which cause hindrance in the recommendation process. In[4] recent research work is focused on addressing the cold start complication by giving suggestions to the new users who don't have any past history related to their preferences. They recommended on the basis of top-N recommendations. The prime focus of this paper[5] is to introduce the methods, concepts and applications of collaborative filtering technique. It has also described the different types of algorithms that com under CF like memory based, model based and hybrid collaborative filtering. In this[6] paper an overview of the recommendation system is

presented, It explains the content based filtering, collaborative filtering and the hybrid recommendation. It also elaborates the techniques, advantages and disadvantages of each method. The prime focus of this paper[7] is to handle the cold start user problem by using the group recommender system. The groups are detected automatically by two different methods proposed, one is the analogous feature wise and other is the social performance wise. Social performance wise group detection method helps in solving the cold start user problem. In this paper[8] Collaborative filtering has been introduced. It has also described the applications of CF associated in different kind of domains. In paper[10] a similarity measure is proposed to suggest things to group of people. The author have used collaborative filtering approach tries to improve accuracy and diversity of results. Challenges like data sparsity and data scalability have been solved by proposing a method in[11] paper. They have used collaborative filtering algorithm to propose their own method to overcome the drawbacks. They have incorporated the users likings on the non functional features as the additional information which will lead to the increased accuracy in recommendation system. Group recommendation has become the profound concern of researchers and is now a days an effective research area. In this paper[12] data mining techniques are used in recommendation system. They have also explained the sampling and the dimensionality reduction techniques. The author has describes the techniques that are included in classification and association. This paper[13] has proposed a method to improve the accuracy and efficiency in the context of data sparsity, cold start that emerges when a new user signs in and there are no past records of the user and shilling attack that is when some users would give positive comments for their own product and negative comments for their rivalries. Collaborative filtering is widely used in recommender system is deeply explained in [14]. They have demonstrated the core concepts of CF, its algorithms and its need for users. They have also mentioned the different kinds of rating acquisitions.

### III. Conclusion

In this paper recommendation system has been reviewed. Recommendation systems are useful tools for the users as they provide the actual suggestions

according to their likes and dislikes. Content based filtering and collaborative filtering methods provide good accuracy and efficiency but they even get better when the hybrid recommendation is used.

### References

1. Nehete, S.P. and Devane, S.R., 2018, August. Recommendation Systems: past, present and future. In *2018 Eleventh International Conference on Contemporary Computing (IC3)* (pp. 1-7). IEEE.
2. Yang, X., Guo, Y., Liu, Y. and Steck, H., 2014. A survey of collaborative filtering based social recommender systems. *Computer Communications*, 41, pp.1-10.
3. Patel, A., Thakkar, A., Bhatt, N. and Prajapati, P., 2019. Survey and Evolution Study Focusing Comparative Analysis and Future Research Direction in the Field of Recommendation System Specific to Collaborative Filtering Approach. In *Information and Communication Technology for Intelligent Systems* (pp. 155-163). Springer, Singapore.
4. Yanxiang, L., Deke, G., Fei, C. and Honghui, C., 2013, January. User-based clustering with top-n recommendation on cold-start problem. In *2013 Third International Conference on Intelligent System Design and Engineering Applications* (pp. 1585-1589). IEEE.
5. Mustafa, N., Ibrahim, A.O., Ahmed, A. and Abdullah, A., 2017, January. Collaborative filtering: Techniques and applications. In *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)* (pp. 1-6). IEEE.
6. Thorat, P.B., Goudar, R.M. and Barve, S., 2015. Survey on collaborative filtering, content-based filtering and hybrid recommendation system. *International Journal of Computer Applications*, 110(4), pp.31-36.
7. Ghodsad, P.R. and Chatur, P.N., 2018, August. Handling User Cold-Start Problem for Group Recommender System Using Social Behaviour Wise Group Detection Method. In *2018 International Conference on Research in Intelligent and Computing in Engineering (RICE)* (pp. 1-5). IEEE.
8. Ambulgekar, H.P., Pathak, M.K. and Kokare, M.B., 2019. A Survey on Collaborative Filtering: Tasks, Approaches and Applications. In *Proceedings of International Ethical Hacking Conference 2018* (pp. 289-300). Springer, Singapore.
9. Li, J., Zhang, K., Yang, X., Wei, P., Wang, J., Mitra, K. and Ranjan, R., 2019. Category Preferred Canopy-K-means based Collaborative Filtering algorithm. *Future Generation Computer Systems*, 93, pp.1046-1054.
10. Ortega, F., Hurtado, R., Bobadilla, J. and Bojorque, R., 2018. Recommendation to groups of users using the singularities concept. *IEEE Access*, 6, pp.39745-39761.
11. Fletcher, K.K., 2017, June. A Method for Dealing with Data Sparsity and Cold-Start Limitations in Service Recommendation Using Personalized sPreferences. In *2017 IEEE International Conference on Cognitive Computing (ICCC)* (pp. 72-79). IEEE.
12. Amatriain, X., Jaimes, A., Oliver, N. and Pujol, J.M., 2011. Data mining methods for recommender systems. In *Recommender systems handbook* (pp. 39-71). Springer, Boston, MA.
13. Guo, L., Liang, J., Zhu, Y., Luo, Y., Sun, L. and Zheng, X., 2018. Collaborative filtering recommendation based on trust and emotion. *Journal of Intelligent Information Systems*, pp.1-23.
14. Schafer, J.B., Frankowski, D., Herlocker, J. and Sen, S., 2007. Collaborative filtering recommender systems. In *The adaptive web* (pp. 291-324). Springer, Berlin, Heidelberg.
15. Katarya, R. and Verma, N., 2017, December. Automatically detection and recommendation in collaborative groups. In *2017 International Conference on Intelligent Sustainable Systems (ICISS)* (pp. 218-222). IEEE.

#### AUTHORS PROFILE



**Harleen Kaur**, is persuing M.E in computer science in Chandigarh university,Punjab. She has complemented her B.Tech in information technology from Gurur gobind singh indraprastha university, New Delhi. She is a research scholar and have published 1 paper in a reputed journal. Her area of interest are data minin, Machine learning and data analytics. She can be contacted at [harleenkaur3663@gmail.com](mailto:harleenkaur3663@gmail.com).



**Gaurav Bathla**, is working as Assistant Professor at Chandigarh University, Punjab, India. He has 12 years of teaching experince. He has completed M.E from Delhi College of Engineering. He is GATE qualified with All India Rank 59. He is an active researcher and published 10 research papers in reputed journals and 5 research papers in International conferences. His areas of interest are Big Data, Machine Learning, Deep Learning, Data analytics and Programming Langauges. He may be contacted at [gouravbathla@gmail.com](mailto:gouravbathla@gmail.com).