A Comparison of Collaborative Filtering-based Recommender Systems

Jyoti Shokeen
Research Scholar
Department of Computer Science and Engineering,
UIET, Maharshi Dayanand University, Rohtak, India

Abstract : ***The proliferation of Internet has made people to rely on virtual recommendations. Recommender systems help out in giving important recommendations. Collaborative filtering is the most successful and widely used approach in designing recommender systems since the introduction of the concept of recommender systems. This approach uses the known tastes and preferences of a set of users to make predictions or generate recommendations about the unknown tastes and preferences of the target user. This paper discusses various works which use collaborative filtering approach to design recommender systems. The paper also gives a comparison of these approaches.***

IndexTerms – Recommender Systems, Collaborative filtering, Social networks.

I. INTRODUCTION

Technology continues to evolve at an unimaginable pace. With the emergence of internet connectivity and data analytic technologies on the web, information is easily accessible to people. The enormous amount of available information is called Big data. According to a survey result of year 2017 discussed in [1], half of the population of the world is connected through Internet. But the selection of useful information from the Big data is the main challenge. The use of internet has changed the user behavior and shifted their interests towards digital content. Recommender systems (RSs) are becoming viral to resolve the issues of information overload and customer satisfaction [2]. The main purpose of a RS is to suggest meaningful and relevant information to users. A recommender system is an information retrieval system that uses the interests, profile and context of users to give recommendations.

II. TYPES OF RECOMMENDER SYSTEM

A number of recommendation system algorithms have been proposed in past years. However, traditional recommendation algorithms were centered on three types: collaborative filtering, content-based filtering and hybrid algorithms. This section discusses these three major algorithms along with other recent algorithms.

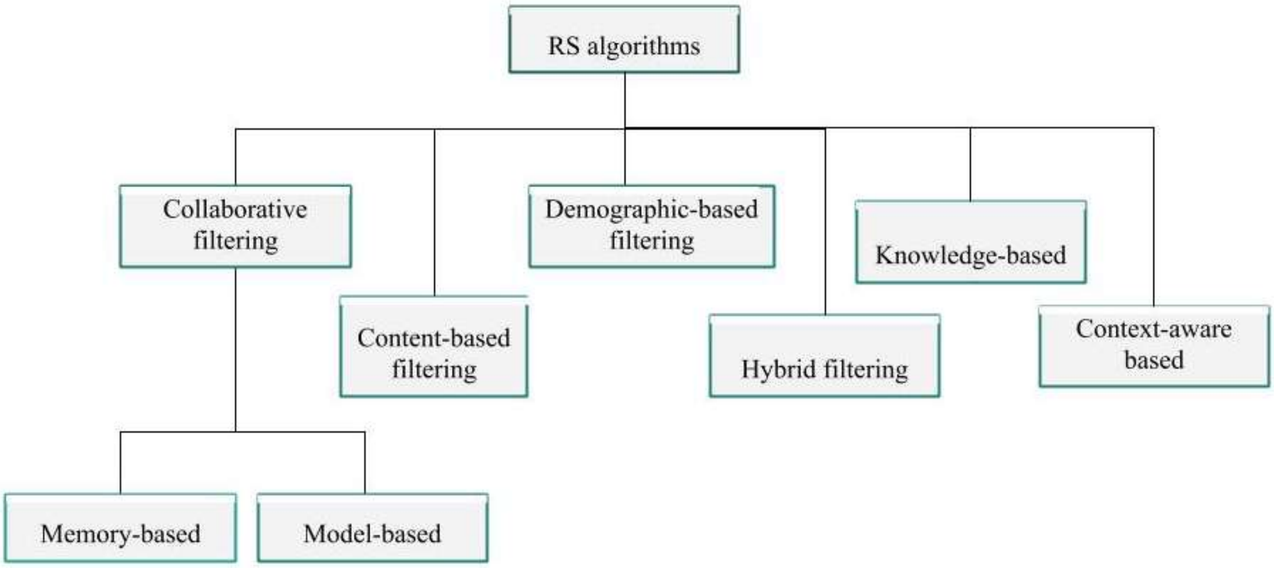

Fig.1 Types of Recommender System Algorithms

2.1 Collaborative filtering:

Collaborative filtering technique recommends the items to the users which are already liked by similar users [3]. Collaborative filtering techniques can be further split into memory-based CF techniques and model-based CF techniques.

2.2 Content-based filtering:

Content-based filtering RSs [4] attempt to find the similar items to the items which have been already liked, rated or purchased by the users. However, the performance of these systems is solely dependent on user feedback.

2.3 Hybrid filtering:

Hybrid filtering techniques basically use the mixture of collaborative filtering and content-based filtering techniques. The combination of the techniques assists in alleviating the drawbacks faced by the techniques when implemented separately. Hybrid RS [5] leverages the benefits of the techniques to build a better RS.

2.4 Demographic-based:

Demographic-based RSs [6] exploit the demographic data of users such as gender, age, education, occupation and nationality to make predictions. However, demographic-based RSs are implemented along with other RSs to achieve better performance and generate meaningful recommendations.

2.5 Knowledge-based:

A knowledge-based RS [7] models the profile of users to determine the correlation between the existing products and the preferences of users. Ontology-based RS are a type of knowledge-based recommender systems [8].

2.6 Context aware-based:

A context aware-based RS [9] exploits the contextual information like time, location, activity and individual context. Contextual information can be both static and dynamic. Villegas et al. [10] give a systematic literature survey of context aware recommender systems. They perform a study on 87 papers of context-aware based RS and classify these papers on the basis of collaborative filtering, content-based and hybrid techniques.

III. Collaborative Filtering

Collaborative filtering (CF) technique finds similar users on the basis of the past interests of target user. For example, if two users are doing research in the field of recommender systems and the other most similar user is using deep learning to build a recommender system. Then, the system will recommend deep learning technique to the target user. This is the reason that this technique is often called the word-of-mouth technique. Collaborative filtering techniques can be further split into memory-based CF techniques and model-based CF techniques. Memory-based CF technique exploits the whole database of user to give recommendations. On the other hand, model-based technique exploits the pre-defined models and selects the most suitable model for the database and then generates recommendations. Further, CF-based techniques can be classified into user-based CF and item-based CF approaches. The former approach searches for the nearest-neighbors of the target user to whom the system will generate recommendations. The system then uses the ratings of nearest neighbors to predict the ratings for unrated item of the target user. When a user *u* rates an item *i*, then the corresponding element in the rating matrix can be denoted by $R_{u,i}$. If the user does not give rating for an item, then the corresponding element is substituted with value 0. Generally, user-based CF approaches use cosine similarity to search top-k nearest neighbors. The similarity between the users is used to find how similar the tastes of two users are? The similarity between two users A and B is defined as follows:

$$Sim_{A,i}(A,B) = \frac{R_A \cdot R_B}{\|R_A\|\|R_B\|} \quad (1)$$

where R_A and R_B denote two rating vectors for users A and B respectively . If there are n items in the rating matrix, then the similarity formula can be expressed as:

$$Sim_{A,i}(A,B) = \frac{\sum_{i=1}^{n} R_{A,i} R_{B,i}}{\sqrt{\sum_{i=1}^{n} R_{A,i}^2}\sqrt{\sum_{i=1}^{n} R_{B,i}^2}} \quad (2)$$

Users	**Average ratings**	**I**	**j**	**k**	**l**	**m**
A	3.75		4	6	3	2
B	3		4	5	2	1
C	2.75	4	3		2	2
D	3.2	4	4	5	2	1
E	4.25	5	4	3		5

In item-based CF approaches, the ratings of users who have rated both the items are used to make predictions. If i and j are the two items for which similarity is to be calculated, then the item-based similarity is expressed as follows:

$$Sim(i,j) = \frac{\sum_{u\in U}(R_{u,i}-\bar{R}_u)(R_{u,j}-\bar{R}_u)}{\sqrt{\sum_{u\in U}(R_{u,i}-\bar{R}_u)^2}\sqrt{\sum_{u\in U}(R_{u,j}-\bar{R}_u)^2}} \quad (3)$$

where $\bar{R}_u$ is the average rating of user *u*. For example, there are 5 users and 5 items and the rating matrix is given as follows:
We calculate the similarity between items *i* and *j*. *T*he ratings of users C, D and E are used to find the similarity between items *i* and *j*.

$$Sim(i,j) = \frac{\sum_{u\in C,D,E}(R_{u,i}-\bar{R}_u)(R_{u,j}-\bar{R}_u)}{\sqrt{\sum_{u\in C,D,E,}(R_{u,i}-\bar{R}_u)^2}\sqrt{\sum_{u\in C,D,E}(R_{u,j}-\bar{R}_u)^2}}$$

$$Sim(i,j) = \frac{(4-2.75)(3-2.75)+(4-3.2)(4-3.2)+(5-4.25)(4-4.25)}{\sqrt{(4-2.75)^2+(4-3.2)^2+(5-4.25)^2}\sqrt{(3-2.75)^2+(4-3.2)^2+(4-4.25)^2}}$$

$$Sim(i,j) = 0.5260$$

Schafer et al. [11] discuss the tasks where people prefer to use this technique. The first task where CF technique is useful is when the user wishes the recommendations of new items that he might like. Secondly, when the user wishes to take advice on a specific item from a group of users, CF comes to the play. Thirdly, CF techniques are effective in finding similar users who might be interesting people for the target user. Other tasks where this technique is beneficial are the tasks to search the domain-specific areas, tasks searching a combination of new and old items, etc.

IV. Literature Review

Koren [12] highlighted the importance of temporal dynamic in building the RS. Koren emphasize the need of exploiting time as a context in RS. The preferences of people change with time and thus it is an important parameter in enhancing the potentiality of a RS.

Konstas et al. [13] used social information from Last.fm to improve the collaborative filtering approach. The social metadata like tags and friendships of users is used to enhance the capabilities of this approach.

Ma et al. [14] propose another collaborative filtering RS called *SoRec* which is based on trusted social relations. SoRec helps in alleviating the data sparsity problem.

Yang et al. [15] emphasized that a user trusts his friends differently based on the context of items. They proposed a circle-based RS in online social networks by using the matrix factorization approach of collaborative filtering. Using this approach, they discover the expert level of their friends for different set of items. The drawback of this approach is that it uses only item category as the context.

Ma et al. [16] emphasized that there is difference between trusted relationships and social relations. They highlighted that social fiends may have different tastes. A user having multiple friends on social networks do not mean that he trusts all his social friends equally. They incorporated the social information into RS. They define two social regularization models where the first model is:

$$\frac{\alpha}{2}\sum_{i=1}^{m}||U_i - \frac{\sum_{j\epsilon F^+(i)} Sim(i,j)\times U_j}{\sum_{j\epsilon F^+(i)} Sim(i,j)}||_F^2 \qquad (4)$$

where U_i and U_j denote the tastes of users u_i and u_j respectively, $\alpha > 0, F^+(i)$ is the user u_i's friends list. $Sim(i,j) \in [0,1]$ is the similarity function that indicates the similarity between user u_i and u_j and $||.||_F^2$ expresses the Frobenius norm. The second regularization term defined by them is:

$$\frac{\beta}{2}\sum_{i=1}^{m}\sum_{j\epsilon F^+(i)} Sim(i,j)||U_i - j||_F^2, \text{ where } \beta > 0. \qquad (5)$$

Recently, Lian et al. [17] proposed an integrated framework that combines CF into content-based filtering technique. The framework uses factorization approach to integrate the above techniques and propose a cross-domain RS called CCCFNET.

Li et al. [18] have integrated deep learning approach into collaborative filtering to learn latent representation of data. They have integrated deep feature learning into the matrix factorization approach of CF. They have proposed mDA-CF and mSDA-CF as two models and implemented these models on Movielens and Book-Crossing dataset to measure the efficiency of these models.

Rand and Kumar [19] propose a temporal data based evolutionary clustering approach *EVAR* that can be applied in collaborative filtering recommender system to predict the dynamic preferences of users.

V. Comparison

In this section, a two-dimensional scheme is introduced that classifies the existing collaborative filtering based RS and perform a comparative analysis on the listed works.

For a better bird-view of existing studies, the classification scheme is proposed that categorizes the existing works of collaborative filtering recommender systems into two perspectives: recommender systems which rely solely on collaborative filtering approach and recommender systems which rely on integrated collaborative filtering with other approaches.

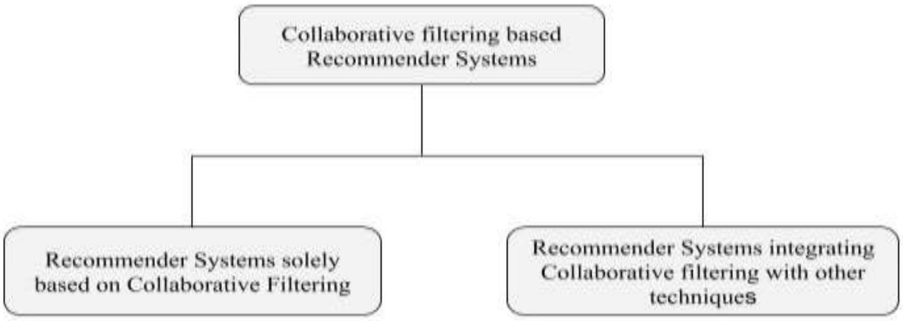

Fig 2. Classification of Collaborative Filtering-based recommender Systems

Table 1 lists the shortlisted papers on collaborative filtering based recommender systems. The techniques are either based solely on CF approach or use CF-approach in combination with approached like deep learning, content-based filtering, clustering . Some algorithms use information from online social networks to enhance the recommendations.

Table 1 Classification of shortlisted papers on Collaborative filtering based RS

References	Recommender System	Technique	Parameter	Dataset
Koren [12]	timeSVD++	Singular Value Decomposition (CF)	Time	Movie (NetFlix)
Konstas et al. [13]	-	User-based CF	Tag, social networks	Last.fm
Ma et al. [14]	SoRec	Probabilistic matrix factorization (CF)	Social network data	Epinion
Yang et al. [15]	CircleCon	Matrix factorization + Social networks	Trusted relations, Social relations	Epinion
Ma et al. [16]	Average-based regularization and Individual-based regularization	Collaborative filtering + Trust-based method	Social regularization	Douban, Epinion
Lian et al. [17]	CCCFNET	Collaborative filtering + Content-based filtering + Neural Networks	Cross-domain knowledge	Douban, MovieLens
Li et al. [18]	mDA-CF, mSDA-CF	Deep learning + Collaborative filtering	-	MovieLens, Book-Crossing
Rana and Kumar [19]	EVAR	Collaborative filtering+ Clustering	Time	MovieLens

VI. Conclusion

Collaborative filtering has been the most successful and used technique in RSs. This papers discusses different recommender systems which are based either solely on collaborative filtering technique or integrate CF technique with other techniques to implement RSs. Use of information and relations from social networks is playing an increasing role in improving the performance of recommender systems. In the future work, I will focus on collaborative filtering techniques that are only based on social networks.

References

[1] S. Papneja, K. Sharma, and N. Khilwani, “Context aware personalized content recommendation using ontology based spreading activation,” *Int. J. Inf. Technol.*, vol. 10, no. 1, 2018.

[2] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, “Recommendation systems: Principles, methods and evaluation,” *Egypt. Informatics J.*, vol. 16, no. 3, pp. 261–273, 2015.

[3] M. D. Ekstrand, J. T. Riedl, and J. Konstan, “Collaborative filtering recommender systems,” *Found. Trends Hum. Comput. Interact.*, vol. 4, no. 2, pp. 81–173, 2011.

[4] P. Lops, M. de Gemmis, and G. Semeraro, *Recommender Systems Handbook*, vol. 532. Springer US, 2009.

[5] R. Burke, “Hybrid Web Recommender Systems,” in *The Adaptive Web*, 2007, pp. 377–408.

[6] M. Y. H. Al-Shamri, “User profiling approaches for demographic recommender systems,” *Knowledge-Based Syst.*, vol. 100, pp. 175–187, 2016.

[7] C. C. Aggarwal, “Knowledge-Based Recommender Systems,” in *Recommender Systems*, Springer, Cham, 2016, pp. 167–197.

[8] J. K. Tarus, Z. Niu, and A. Yousif, “A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining,” *Futur. Gener. Comput. Syst.*, vol. 72, pp. 37–48, 2017.

[9] I. Ben Sassi, S. Mellouli, and S. Ben Yahia, “Context-aware recommender systems in mobile environment: On the road of future research,” *Inf. Syst.*, vol. 72, pp. 27–61, 2017.

[10] N. M. Villegas, C. Sánchez, J. Díaz-Cely, and G. Tamura, “Characterizing context-aware recommender systems: A systematic literature review,” *Knowledge-Based Syst.*, vol. 140, pp. 173–200, 2018.

[11] J. Schafer, D. Frankowski, J. Herlocker, and S. Sen, “Collaborative Filtering Recommender Systems,” *Adapt. Web*, pp. 291–324, 2007.

[12] Y. Koren, “Collaborative Filtering with Temporal Dynamics,” in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge discovery and data mining (KDD ’09)*, 2009, vol. 53, no. 4, pp. 89–97.

[13] I. Konstas, V. Stathopoulos, and J. M. Jose, “On Social Networks and Collaborative Recommendation,” in *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2009, pp. 195–202.

[14] H. Ma, H. Yang, M. R. Lyu, and I. King, “SoRec : Social Recommendation Using Probabilistic Matrix Factorization,” in *Proceeding of the 17th ACM conference on Information and knowledge management*, 2008, pp. 931–940.

[15] X. Yang, H. Steck, and M. Hill, “Circle-based Recommendation in Online Social Networks,” in *Proceedings of the 18th ACM International Conference on Knowledge discovery and data mining*, 2012, pp. 1267–1275.

[16] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, “Recommender Systems with Social Regularization,” in *Proceedings of the fourth ACM international conference on Web search and data mining*, 2011, pp. 287–296.

[17] J. Lian, F. Zhang, X. Xie, and G. Sun, “CCCFNet: A Content-Boosted Collaborative Filtering Neural Network for Cross Domain Recommender Systems,” in *Proceedings of the 26th International Conference on World Wide Web Companion*, 2017, pp. 817–818.

[18] C. Rana and S. Kumar, “An evolutionary clustering algorithm based on temporal features for dynamic recommender systems,” *Swarm Evol. Comput.*, vol. 14, no. Elsevier, pp. 21–30, 2014.