Asian Journal of Science and Applied Technology
ISSN 2249-0698 Vol.10 No.2, 2021, pp.22-25

DOI: https://doi.org/10.51983/ajsat-2021.10.2.3059

Filtering Techniques in Recommendation Systems: A Review

S. Shargunam and G. Rajakumar
Department of Electronics and Communication Engineering,
Francis Xavier Engineering College, Tirunelveli, Tamil Nadu, India
E-mail: shargunamguna@gmail.com, gmanly12@gmail.com

Abstract **- Recommendation systems are not new to the world, they have rapidly become prevalent, appearing in almost every type of technology on a daily basis. As a result, recommendation systems were necessary to reduce the amount of time spent looking for the best and most essential items. Information filtering, user personalization, collaborative filtering, and hybrid filtering are just some of the ways used by recommendation systems in diversion, streaming, software, and other areas to present users and customers with customized content and products. The various filtering methods are compared and analyzed in order to improve the accuracy and quality of the recommendation system.**
Keywords: **Filtering, Streaming, Personalization, Content Gathering**

I. INTRODUCTION

In the Proposal frameworks, partner degree endeavor attempts to deliver forecasts on 'rating' or 'inclination' that client would propose to a thing or social part, even the slightest bit times are regularly arranged into 3 principle classes: content-based, mixture proposal approaches agreeable sifting ways, inside which the ideas for clients are based for the most part with pertinence elective clients that have the near preferences and inclinations, which might be extra coordinated into thing based frameworks and client based frameworks. Here the advice algorithms which are the core of this technique. Ancient recommendation algorithm are content based mostly recommendation, collaboration based mostly recommendation algorithms, and so on. At the large information times, the challenges have tendency to meet are information scale, performance of computing, and alternative aspects. The proposed work successfully provides consumers with tailored and personalized item recommendations and prescribes the most suited items. Key phrases are frequently used to express user preferences, and a user-based cooperative filtering algorithmic method has been used to create relevant recommendations.

II. TYPES OF FILTERING

A. Content-Based Filtering

Another option for a recommender system is content-based filtering. These methods are backed up by item descriptions and user preferences. Keywords are used to describe items in order to highlight features that can be utilised to generate suggestions. To put it another way, these algorithms strive to propose items that are similar to persons who a user has liked in the past (or is examining in the present). Specifically, a large number of potential goods are compared to items previously assessed by the user, and the best-matching items are then recommended. The following steps must be completed in order to establish a content-based filtering system.

1. Terms Allocation
2. Terms Representation
3. Learning Algorithm Selection
4. Provide Recommendations.

B. Collaborative Filtering

Collaborative filtering methods create a model based on a user's previous behaviour (items previously purchased or elected, and/or numerical ratings given to those things), as well as similar selections made by other users, and then use that model to predict items (or ratings for items) that the user is also interested in. Content-based filtering methods use a set of distinct features of a single item to identify other items with comparable qualities.

1. Simple Support Vector Machine Algorithm

It's a looping algorithm in which the basic Support Vector Machine algorithm's steps are as follows.

a. Configuring the Training Data
b. Configuring the SVM`s Parameters and Training it
c. Obtaining Information about the support Vectors.

2. Adjusted K-Means Algorithm

The number of clusters and item attribute features are given as input to this method. A collection of clusters has been obtained as output which reduces the squared error. The probability of each movie belonging to each cluster centre is also represented. Then, at random, some items are chosen as cluster centres. Each object is reallocated to the cluster with the best match based on its mean value. The cluster means are then updated, and the likelihood between objects and each cluster centre has been calculated. This technique is repeated until no more centroids change its way.

C. Recommendation System with User-Interested Object

First, the contents are divided into genres. As a result, phase information are derived from the entire contents. The preference score of user-interested objects has been calculated once the users viewed the contents. Finally, the advice system displays the top-Nth contents based on the user's preference score for victimization.

D. Cosine Similarity Measure

For two vectors, the cosine similarity is a metric that calculates the cosine of the angle between them. This metric assesses the direction rather than the magnitude. The cosine similarity formula is given as:

$$\vec{a} \cdot \vec{b} = \|\vec{a}\|\|\vec{b}\| \cos\theta$$

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|\|\vec{b}\|}$$

E. Hybrid Filtering

There are two approaches to use a hybrid filtering strategy. Both content-based and collaborative filtering are used individually, with the results combined as needed. Second, collaborative filtering has been used initially, followed by content-based filtering on the results.

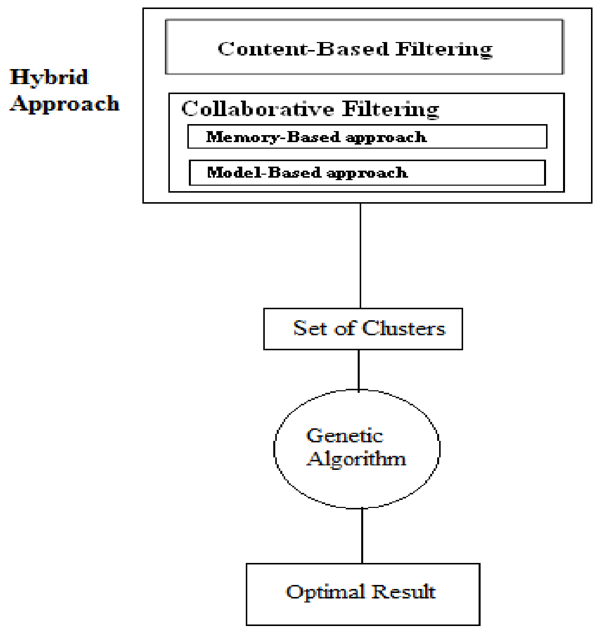

Fig. 1 Hybrid Approach with Genetic Algorithm

III. LITERATURE SURVEY

Hirdesh Shivhare *et al.*, [1] discussed a movie recommendation system, researchers presented an integrative method that combines the fuzzy c-means clustering method with a genetic algorithm-based weighted similarity measure. The suggested movie recommendation system provides finer similarity metrics and quality than the existing movie recommendation system, but the proposed recommendation system takes longer to compute than the existing recommendation system. The clustered data points can be used as an input dataset to solve this problem.[1].

Gaurangi Tilak *et al.*, [2] introduced Movie GEN, an expert system for movie recommendation, was introduced. On the basis of a hybrid recommendation method, they constructed the system utilizing machine learning and cluster analysis. It chooses movies from the collection, groups them, and produces questions for the users based on the support Vector Machine prediction.

An article in Cybernetics on collaborative recommender systems works with various sorts of matrix factorizations, including Regularized Matrix Factorization (RMF), Maximum Margin Matrix Factorization (MMMF), and Negative Matrix Factorization (NMF) [3].

Takacs G *et al.*, [4] described Positive MF technique, momentum-based MF approach, and hybrid MF neighbour based methods which are some of the proposed matrix factorization methodologies.

Xin Luo *et al.*, [5] worked on a regularized matrix factorization approach and simplified the matrices' training algorithm. They stated that this produces positive results, as evidenced by their work published.

Xin Luo *et al.*, [6] worked on a project in which they used adaptive learning rate approaches to construct a recommender system.

Bu Sung Kim *et al.*, [7] approached on the sparcity problem of collaborative filtering and employed a user tag matrix to tackle the problem.

Mohammed Wasid *et al.*, [8] worked on accuracy of memory-based collaborative filtering and the scalability of model-based collaborative filtering.Both the models are improved and also employed fuzzy sets and the Particle Swarm Optimization (PSO) technique to achieve.

Dheeraj *et al.*, [9] used Singular Value Decomposition (SVD), Principal Component Analysis (PCA), and Probabilistic Matrix Factorization (PMF) was among the matrix factorization models used. They also highlighted how to use these matrix Factorizations to overcome the problems of CF.

IV. ISSUES AND CHALLENGES IN RSS

A. Cold-Start Problem

To RSs24, the cold-start problem is a combination of a new item and a new user. When a new item is introduced to a CF system with no ratings, it cannot be recommended at first. For example, Movie Lens (movielens.org) cannot recommend new films unless they have received some early feedback. The problem of new users is difficult to solve because it is impossible to discover comparable users or develop a CB profile without knowing a user's previous preferences.

B. Sparse RSs

In general, the majority of users do not evaluate the majority of the objects, resulting in a highly sparse ratings matrix. As a result, the data sparcity problem occurs, reducing the likelihood of finding a group of users with comparable ratings. This is the most serious flaw in the CF method.

C. Scalability Problem

The scalability of algorithms with big real-world datasets is a critical and pressing topic for RSs today. Dealing with large and dynamic data sets generated by item-user interactions such as preferences, ratings, and reviews are becoming more difficult. When some recommendation algorithms are applied to relatively small data sets, it produces the best results, but when applied to very big datasets, it may produce inefficient or worst outcomes. As a result, advanced large-scale assessment approaches are needed to address this problem.

D. Privacy Issue

To provide high-quality personalized suggestions, RSs must collect as much user data as possible and use it to its greatest potential. However, because the system knows far too much about the users, this may leave a negative impression in their minds about their privacy. As a result, procedures must be developed that may use user data rationally, painstakingly, and properly, guaranteeing that information about the users' genuine preferences is not widely accessible to malicious users.

E. Robustness of RSs

Another important issue with RSs is their resistance to attacks. RSs' robustness is a performance metric. An attacker may create phone user accounts based on attack models such as Push/Nuke Attacks to make particular target items more/less popular in order to achieve specific revenues. Shilling attacks, also known as profile injection attacks, are a type of assault.

F. Proactive RSs

Even if it is not explicitly requested, RSs are responsible for offering recommendations. The majority of the RSs that have been created so far use a “pull approach,” in which users initiate recommendation requests. In today's environment, where consumers are constantly connected to the world of computing and the internet, systems that can forecast what, when, and how to “push” recommendations on implicit requests are desirable. As a result, recommender systems can take a proactive approach when recommending goods in the users best interests.

V. COMPARATIVE ANALYSIS

TABLE I COMPARISON OF FILTERING TECHNIQUES

S. No.	Techniques	Merits	Demerits
1	Content based filtering	User independence, Transparency, No cold start	Limited content Analysis, overspecialization
2	Collaborative filtering	Captures the change in user interest over time and inherent subtle characteristics	Data sparcity, scalability, shilling Attacks
3	Cosine similarity measure	Low complexity for sparse vectors and Only non-zero dimensions are needed	High correlation vectors are expected
4	Hybrid filtering	Higher performance, accuracy than filter	Several combinations may lead to cold-start problem.

VI. CONCLUSION

The number of studies on cooperative filtering recommendation systems is growing in order to provide more exact recommendations. However, measurability and cold-start difficulties plague these studies. To unravel these issues, this paper proposes the advice system with a Hybrid approach by unifying content based mostly filtering and cooperative filtering. The usually recommendation algorithms are suited to ancient datasets, like the content-based recommendation, collaboration recommendation. However, with the event of huge information, the advice algorithms ought to have the power to cater to the large information.

REFERENCES

[1] Anshul Gupta, Hirdesh Shivhare and Shalki Sharma, “Recommender system using fuzzy c-means clustering and genetic algorithm based weighted similarity measure,” In *2015 International Conference on Computer, Communication and Control (IC4), IEEE*, pp. 1-8, 2015.

[2] A. Eyrun, T. Gaurangi and L. Nan, “Moviegen: A movie recommendation system,” *Hewlett-Packard, August*, 2008.

[3] Mingrui Wu, “Collaborative filtering via ensembles of matrix factorizations,” In *KDD Cup and Workshop 2007*, pp. 43-47, 2007.

[4] Takacs Takács, Gábor, István Pilászy, Bottyán Németh and Domonkos Tikk, “Investigation of various matrix factorization methods for large recommender systems,” In *2008 IEEE International Conference on Data Mining Workshops, IEEE*, pp. 553-562, 2008.

[5] Xin Luo, Yunni Xia and Qingsheng Zhu, “Incremental collaborative filtering recommender based on regularized matrix factorization,” *Knowledge-Based Systems*, Vol. 27, pp. 271-280.

[6] Xin Luo, Yunni Xia, and Qingsheng Zhu, “Applying the learning rate adaptation to the matrix factorization based collaborative filtering,” *Knowledge-Based Systems*, Vol. 37, pp. 154-164, 2013.

[7] Bu Sung Kim, Heera Kim, Jaedong Lee and Jee-Hyong Lee, “Improving a recommender system by collective matrix factorization with tag information,” In *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS), IEEE*, pp. 980-984, 2014.

[8] Mohammed Wasid and Vibhor Kant, “A particle swarm approach to collaborative filtering based recommender systems through fuzzy features,” *Procedia Computer Science*, Vol. 54, pp. 440-448, 2015.

[9] Dheeraj Bokde, Sheetal Girase and Debajyoti Mukhopadhyay. “Matrix factorization model in collaborative filtering algorithms: A survey,” *Procedia Computer Science*, Vol. 49, pp. 136-146, 2015.

[10] Schafer, J. Ben, Dan Frankowski, Jon Herlocker and Shilad Sen, “Collaborative filtering recommender systems,” In *The adaptive web, Springer*, pp. 291-324, Berlin, Heidelberg, 2007.

[11] Xiaohua Sun, Fansheng Kong and Song Ye, “A comparison of several algorithms for collaborative filtering in startup stage,” In *Proceedings. 2005 IEEE Networking, Sensing and Control, IEEE*, pp. 25-28, 2005.