International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

# Differential Approaches to Improve Recommendation System: Issues and challenges

**Vikas P. Mapari**[1]

[1]Information Technology Department, D.Y. Patil COE, Ambi, Pune, India

**Abstract:** ***Recommendation system which plays an important role in many applications as WWW, ecommerce etc. The main objective of this paper is to focus on various issues and challenges of recommendation system. Collaborative filtering is one of the techniques in recommender systems, providing personalized recommendations to users based on their previously expressed preferences in the form of ratings and those of other similar users. A recommender system uses Collaborative Filtering or Content-Based methods to predict new items of interest for a user. Although both methods have their own and distinct advantages but individually they fail to provide good recommendations in many situations. Incorporating components from collaborative and content based methods, can overcome these challenges like lack of data, data sparsity, stability, accuracy and correlation of traditional recommender systems. Inadequate ratings lot of time gives poor quality of recommendations in terms of accuracy. By giving the overview of these problems we can improve recommendations by approaching new methods and solutions, which can be used as a highway for research and practice in this area.***

**Keywords:** Collaborative Filtering, Content-Based Recommendation, Recommendation System, Sparsity Problem, Cold Start, over specialization, recommendation diversity, correlation, ranking functions.

## 1. Introduction

Recommender systems [1][2] predict the ratings of unknown items for each and every user, often using other user's ratings, and recommend top N items with the highest predicted ratings. In online applications, items are rated with more or less rating (Rating is scaled in between the range of 1 to 5, from lower to higher order) [14]. Common online applications are online shopping, games, movies, music, videos etc. One of the most promising recommender techniques, Collaborative Filtering (CF) [5][9][11] predicts the potential interests of an active user by considering the opinions of users with same taste. Collaborative Filtering technique provides two main aspects of memory based CF are: i) Simple algorithms ii) accurate recommendation. Memory based [22] CF detect the user's ratings on different items by asking the user or by observing his/her interaction with the systems to store them into a table known as the rating matrix. Then, memory based CF methods use similarity measurement [9] methods to filter users (or items) that are similar to the active user (or the target item) and calculate the prediction from the ratings of these neighbors.

## 2. Recommendation System: General Concepts

### 2.1. What are Recommender / Recommendation System?

Online recommender systems [1] in which they are used to either predict whether a particular user will like a particular item (prediction), or to identify a set of N items that will be of interest to a certain user (Top-N recommendation). Recommender systems (RS) [12] are used in a variety of applications. Examples are web stores, online com-munities, and music players. Currently, people mostly tend to associate recommender systems with e-commerce sites, where recommender systems are extensively used to recommend items / products to the customers and to provide customers with information to help them decide buy which products. Products can be based on the top overall sell on a site, on the demographics of the consumers, or on an analysis of the past buying behavior of the consumers as a prediction for future buying behavior.

### 2.2. Types of recommendation system

Figure 1 shows the approaches to Recommender Systems categorized as follows:

a) Content Based Recommendation: In content based recommendation [5] items those are similar in content of items the user has liked in the past or matched to attributes of the user are recommended.
b) Collaborative Filtering (CF): In Collaborative Filtering systems [11] a user is recommending items based on the previous ratings of all users collectively.
c) Differential approaches (Hybrid Approaches): These methods combine either collaborative and content based approaches or different approaches from CF or CB.

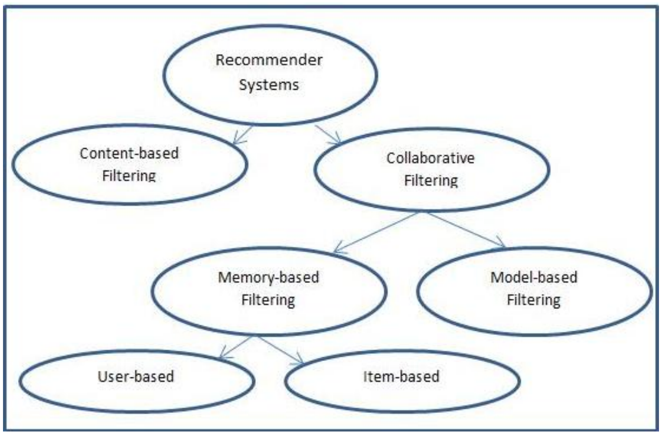


**Figure 1:** An overview of recommender techniques

#### 2.2.1. Content-based filtering

In content-based methods [3], items with similar content features comparing to a user's past favorite items will be recommended to the user. The weakness of this kind of methods is that it depends on the features, and effective features are di cult to find in some recommendation applications. For example, in the Amazon, many users have very incomplete pro ling information, and the items in their history have a quantity of diversity. Thus there are not enough features for accurate predictions. This is the reason that collaborative filtering comes up.

### 2.2.2. Collaborative filtering

Collaborative filtering (CF) is different from other filtering technologies in that information is filtered by using evaluation instead of analysis, thus categorizing in-formation based on the user's opinion of the information instead of the information itself. In addition, CF stresses the concept of community by letting recommendations be a result of the opinions of the current user and other similar users. As figure 1 shows, all users contribute with ratings based on their preferences. Recommendations for the current user are produced by matching the user's ratings with ratings given by other users. In this way, similar users are linked together to form a community. The prediction is based on the common behavior patterns analyzed from the large real dataset. The key point is that CF finds similar users for each user, according to the similarity of their rating history. Then the prediction is made by the ratings of his/her similar users.

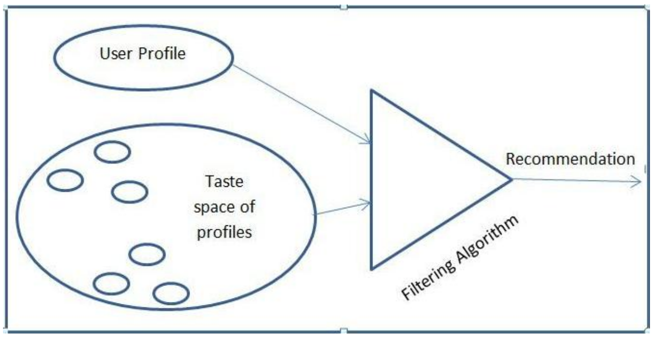


**Figure 2:** Collaborative Filtering

### 2.3. Applications of Recommendation Systems

### 2.3.1. Product Recommendations

Perhaps the most important use of recommendation systems is at on-line retailers. We have noted how Amazon or similar on-line vendors strive to present each returning user with some suggestions of products that they might like to buy. These suggestions are not random, but are based on the purchasing decisions made by similar customers or on other techniques.

### 2.3.2. Movie Recommendations

Netfiix offers its customers recommendations of movies they might like. These recommendations are based on ratings provided by users. The importance of predicting ratings accurately is so high, that Netflix offered a prize of one million dollars for the first algorithm that could beat its own recommendation system.

### 2.3.3. News Articles

News services have attempted to identify articles of interest to readers, based on the articles that they have read in the past. The similarity might be based on the similarity of important words in the documents, or on the articles that are read by people with similar reading tastes.

### 2.3.4. E-commerce

Recommendations for consumers of products to buy such as books, cameras, PCs etc.

### 2.3.5. Services

Recommendations of travel services, recommendation of experts for consultation, recommendation of houses to rent, or matchmaking service.

## 3. Research Challenges of Recommendation System

Different experiments and technics are used to improve recommender system for accurate and dynamic ranking and recommendation, but still there are some of the hurdles that may be described in terms of basic challenges as:

### 3.1. Sparsity Problem

From [11] [14] [16], it focused to alleviate the sparsity problem and improve the recommendation accuracy and precision in collaborative filtering systems. Use the different association retrieval technology to alleviate the sparsity problem and find a new collaborative filtering algorithm to increase the recommendation precision. The benefits of the approach were evaluated using data from the movie lens data set. It was indicating the approach alleviated the sparsity problem and improved recommendation quality than the standard collaborative filtering approaches. But, still there is a problem for the refined system is the volume of data these systems utilize will continue increasing over time, so the system causes the data overload problem. So the system still not scalable and need focus on scalability problem.

### 3.2. Cold Start problem

From [20], it focused to a method combining social sub community division and ontology decision model to solve the new user cold-start problem in collaborative filtering algorithm, which builds relationships between user static information and dynamic preferences by learning. But, still there are several points where need a focus such as, the system needs to be improved to make it apply to both new users and ordinary users and determine the optimal solution, also user privacy and security also need to be research to get better solution.

### 3.3. Scalability

From [12] [15], the recommender system supports with the enormous information, products and services evolution, and becomes more and more challenging to create robust, and scalable recommender systems that are able to perform in real time. The approaches are there to provide increased scalability and decreasing the time complexity of recommender systems, involves user clustering, based on their profiles and similarities. Clustering approach provides recommendations for the other cluster members; but, here complexity of recommendation depends only on cluster size. In other approach clustering methods have been often used, the requirements of user clustering in recommender systems, are quite different from the typical ones. So, there is no reason to create disjoint clusters or to enforce the partitioning of all the data. To overcome such issues system focuses on data clustering method that is based on genetic algorithms, which shows this method is faster and more accurate than classic clustering schemes for significantly better recommendation quality.

### 3.4. Over Specialization Problem

From [22], the recommender system focusing on the problems of overspecialization problem, here it provides an approach for recommending items in the neighborhood based collaborative filtering. In first phase it uses probabilistic neighborhood selection, where it uses an efficient method for weighted sampling of k neighbors that takes into consideration the similarity levels between the target user (or item) and the candidate neighbors. In other phase it focuses on the system increases the coverage, dispersion, and diversity reinforcement of recommendations by selecting diverse sets of representative neighbors. The system also focuses on item prediction accuracy, utility based ranking, and other popular measures, across various experimental settings.

### 3.5. Shilling Attacks

From [8], system focused on use of statistical metrics for detecting patterns of shilling attackers in a recommender system. The system evaluates and shows that attackers must indeed special, noticeable rating patterns. The system proposes an additional metric, Rating Deviation from Mean Agreement (RDMA), to measure a user's disagreement with the other users in the database, weighted by the inverse rating frequency of her rated items. Based on these investigations the system has developed an algorithm that computes the probability of a user to be a shilling attacker by studying the rating patterns within the system.

### 3.6. Privacy

From [21] [27], the system addressed on enhancement of privacy in OSNs by aggregating contacts' information and their live streams from various social online services in order to provide valuable privacy recommendations. The system furtherly focuses on semantic core and the trust engine, which are central components of the system userware, have been enhanced by adding intelligent information extraction techniques and utilized in order to unintended information disclosure, possibly compromising the privacy of the user or his/her contacts. There are three step evaluation in system. The first step is to define ontology based access rights, followed by the semantic equivalence detection, in second step which is used detect contacts possibly being different accounts of the same person and to suggest to merge those into one contact, on the other hand, to give a mechanism that helps to prevent contacts to detect that two (or more) partial identities belong to the user and enable them to link them. In the third step is for privacy recommendations when disclosing information through microblogging and related communication channels. The advanced NLP mechanisms are being utilized in order to process live text inputs and to detect sensitive information before posting to other services and provides advanced privacy recommendation.

### 3.7. Latency Problem

From [26] [28], the recommender system faces latency problem when new items are added more frequently to the database, where the recommender suggests only the already rated items as the newly added items are not yet rated. system proposed an approach to Recommender systems for application domains where items are frequently added. Also provides that sufficient categorization is possible, and shows that category based filtering enables handling the latency problem. In system users are represented partly by individual user models, and when the knowledge about an individual user is too limited to draw the needed conclusions for recommending items, offline clustering, are used. The system will automatically attempt classification of new users by comparing the user's behavior with the user stereotype cases, selecting the most similar one. Personalized information is divided into two categories: appreciation-known and appreciation-assumed. While the former represents item selections based on a user's known previous behavior, appreciation-assumed items are chosen because of high appreciation probabilities among other users belonging to the same user stereotype case as the current user.

## 4. Research Issues In Recommendation System

### 4.1. Lack of data [10]

Perhaps the biggest issue facing recommender systems is that they need a lot of data to effectively make recommendations. It's no coincidence that the companies most identified with having excellent recommendations are those with a lot of consumer user data: Google, Amazon, Netflix, Last.fm.

The more item and user data a recommender system has to work with, the stronger the chances of getting good recommendations. But it can be a chicken and egg problem to get good recommendations, you need a lot of users, so you can get a lot of data for the recommendations.

### 4.2. Changing Data [10]

Systems are usually "biased towards the old and have difficulty showing new". Due to trendy users its critical to have data recommendation, e.g. 1. as per fashion changes users expects to have new recommendation according to their choice. 2. As per discount on various website.

### 4.3. Changing User Preferences [25]

The issue here is that while today user have a particular intention when browsing e.g. Amazon – tomorrow I might have a different intention. A classic example is that one-day user will be browsing Amazon for new books for himself, but the next day he will be on Amazon searching for a birthday present for my sister.

### 4.4. Evaluation and the Availability of Online Datasets [29]

In recommendation system mostly offline data sets are used in research or to write research paper. With research and discussion on offline and online dataset it gives contradiction in available offline and online data set. Offline data sets are available in huge amount but still it is not up to date. It is available with some reference (study carried out during that period) and it could not be used for all research. e.g. movie lens dataset is available with 9,000 movies by 700 users with 100000 ratings lastly updated on 10, 2016. But still it will not be used for any resent application as each application is distinct and may require less or more data for analysis. Another important issue with this data set is that, it not has any evaluation how many users are satisfied and how system improves the accuracy or scalability or diversity.

Most common factor in offline data set is that they are always incomplete because insufficient user knowledge of the literature, or biases arising in the citation behavior of some researchers, such datasets may have the same negative effects on different algorithms.

In other situations, they have different effects on different algorithms, which is why offline evaluations could only sometimes predict results of online evaluations. Since we see no way of knowing when negative effects of incomplete datasets would be the same for two algorithms, we concluded that user-offline-datasets are not suitable for predicting the performance of recommender systems in practice.

### 4.5. Limited Content Analysis and Overspecialization [30]

Limited content analysis is the big issue in recommender system that arises due to difficulty in extracting reliable automated information from various content (e.g., images, video, audio and text), which can greatly reduce the quality of recommendations.

Another issue of overspecialization occurs due to the phenomenon in which users only receive recommendations for items that are very similar to items they liked or preferred; therefore, the users are not receiving recommendations for items that they might like but are unknown (e.g., when a user only receives recommendations about fiction films). Recommendations can be evaluated for novelty.

### 4.6. Loss of neighbor transitivity

Assume that user A is highly correlated with user B, user B is highly correlated with user C. Possibly, user C is also highly correlated with user A. Such relationships are not captured by recommender systems, but can be captured with knowledge of users from, for instance, ontology. For example, people aged 22-70 are correlated as adults, when people aged 3-5 as children. A recommendation quality plays one of the central roles in recommender systems. Among other details, the user is sensible for false negatives (incorrect recommendations, which the user does not like). Assume the user likes genre Sci-Fi and highly rated many other Sci-Fi movies. If the recommender system will rate “The Matrix” as bad one, but the user likes it, the prediction will be a false negative. In such cases users lose trust in the system and stop using it. Therefore, it is important to keep recommendation quality at the highest possible level.

### 4.7. Recommenders in Mobile Devices

Location based services are becoming more popular these days with the swift development of wireless networks and mobile devices; hence the geographical information is having a vital role to play here. The users may seek different recommendations especially when they are to move across cities, hotels, restaurants, shops etc. Such type of scenario demands for the possible computational solutions along with mobile user interfaces that can effectively and efficiently utilize the available limited resources such as the screen size and computing power of the mobile devices.

## 5. Conclusion

Recommender system plays vital role in our todays day to day life. Human beings are blindly dependent on recommender system, as most of tech-savvy population is search online for various items as per availability of different recommender systems. The recommender systems are providing results as per available technics (algorithms) or data sets, but still they are not complete or not providing accurate results. e.g. discount on items are differs on various recommender system websites, rating of items are differing because most of users are not interested to rate item.

Recommender systems are still in adaption mode and are get improved by service provider due to high competition or to attract customers by providing accuracy in item recommendation. The paper concludes that if recommender systems are overcoming challenges and various issues by updating or research the better quality and accuracy will be for their customers and in results for researcher.

## References

[1] Paul Resnick and Hal R. Varian. Recommender systems, volume 40. ACM Press, http://doi.acm.org/10.1145/245108.245121, 1997. [cited at p. 1, 8]

[2] R. Mooney, P. Bennett, and L. Roy. Book recommending using text categorization with extracted information. In Recommender Systems. Papers from 1998 Work-shop. Technical Report WS-98, volume 8, 1998.

[3] J. Carbonell and J. Goldstein, The user of MMR, diversity-based reranking for reordering documents and producing summaries, Proc. of the ACM Conf. on Re-search and Development in Information Retrieval (SIGIR), pp. 335-336, 1998.

[4] K. Bradley and B. Smyth, Improving Recommendation Diversity, Proc. of the 12th Irish Conf. on Artificial Intelligence and Cognitive Science, 2001.

[5] B. Smyth and P. McClave, Similarity vs. Diversity, Proc. of the 4th Intl. Conf. on Case-Based Reasoning: Case-Based Reasoning Research and Development, 2001.

[6] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J. Riedl, Evaluating Collaborative Filtering Recommender Systems, ACM Transactions on Information Systems, 22(1), pp. 5-53, 2004.

[7] Gediminas Adomavicius and Alexander Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions", IEEETKDE: IEEE Transactions on Knowledge and Data Engineering, 17, 2005.

[8] Paul Alexandru Chirita, Wolfgang Nejdl, Cristian Zamfir. Preventing Shilling Attacks in Online Recommender Systems, WIDM'05, November 5, 2005, Bremen, Germany.

[9] S.M.. McNee, J. Riedl, J. A. Konstan, Being Accurate is Not Enough: How Ac-curacy Metrics have hurt Recommender Systems, Conf. on Human Factors in Computing Systems, pp. 1097-1101, 2006.

[10] S.T. Park and D.M. Pennock, Applying collaborative filtering techniques to movie search for better ranking and browsing, Proc. of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 550-559, 2007.

[11] Anne Yun-An Chen and Dennis McLeod, "Collaborative Filtering for Information Recommendation Systems, Semantic Information Research Laboratory, Oct 2006

[12] Olga Georgiou and Nicolas Tsapatsoulis. Improving the Scalability of Recommender Systems by Clustering Using Genetic Algorithms, © Springer-Verlag Berlin Heidelberg 2010

[13] Ma, H., Zhou, T.C., Lyu, M.R., and King, I. 2011. Improving recommender systems by incorporating social contextual information. ACM Trans. On Information Systems 29, 2 (April 2011), 9:1-9:23.

[14] YiBo Chen, ChanLe Wu,Ming Xie and Xiaojun Guo. Solving the Sparsity Problem in Recommender Systems Using Association Retrieval, JOURNAL OF COMPUTERS, VOL. 6, NO. 9, SEPTEMBER 2011

[15] Cacheda, F., Carneiro, V., Fernndez, D., and Formoso, V. 2011. Comparison of collaborative filtering algorithms: limitations of current techniques and proposals for scalable, high-performance recommender systems. ACM Trans. on Web 5, 1 (Feb. 2011), 2:1-2:33.

[16] Feng Xie, Zhen Chen, Jiaxing Shang, Geo rey C. Fox, Grey Forecast Model for Accurate Recommendation in Presence of Data Sparsity and Correlation, ACM 1-58113-000-0/00/0010, RSWeb 12, September 913, 2012, Dublin, Ireland.

[17] Yong Zheng, Robin Burke, Bamshad Mobasher. Optimal Feature Selection for Context-Aware Recommendation using Differential Relaxation, CARS-2012, September 9, 2012, Dublin, Ireland.

[18] Yong Zheng, Robin Burke, and Bamshad Mobasher. Recommendation with Differential Context Weighting, S. Carberry et al. (Eds.): UMAP 2013, LNCS 7899, pp. 152–164, 2013. © Springer-Verlag Berlin Heidelberg 2013

[19] Lalita Sharma, Anju Gera. A Survey of Recommendation System: Research Challenges, International Journal of Engineering Trends and Technology (IJETT) - Volume4, May 2013

[20] Meng Chen, Cheng Yang, Jiechao Chen, Peng Yi. A Method to Solve Cold-Start Problem in Recommendation System based on Social Network Sub-community and Ontology Decision Model, © 2013. The authors - Published by Atlantis Press.

[21] Arjan J. P. Jeckmans, Michael Beye, Zekeriya Erkin, Pieter Hartel, Reginald L. agendijk, Qiang Tang. Privacy in Recommender Systems, © Springer-verlag London 2013

[22] Panagiotis Adamopoulos, Alexander Tuzhilin. On over-specialization and concentration bias of recommendations: probabilistic neighborhood selection in collaborative filtering systems, RecSys '14 Proceedings of the 8th ACM Conference on Recommender systems.

[23] Nadia A. Najjar and David C.Wilson. Differential Neighborhood Selection in Memory-Based Group Recommender Systems, Proceedings of the Twenty-Seventh International Florida Artificial Intelligence Research Society Conference, 2014

[24] Mr. Mapari Vikas Prakash, Prof. Patil Pramod. Improving Recommendation using Forecast based approach and Re-ranking Approache, IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 16, Issue 5, Ver. IV (Sep – Oct. 2014), PP 145-153

[25] Soanpet .S. Lakshmi , Dr. T.A. Lakshmi. Recommendation Systems: Issues and challenges, Soanpet .Sree Lakshmi et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4) , 2014, 5771-5772.

[26] Negar Hariri, Bamshad Mobasher, Robin Burke. Adapting to User Preference Changes in Interactive Recommendation, Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

[27] Marcel Heupe, Lars Fischer, Mohamed Bourimi, and Simon Scerri. Ontology-Enabled Access Control and Privacy Recommendations, Research gate, December 2014

[28] F.O. Isinkaye, Y.O. Folajimi, B.A. Ojokoh. Recommendation systems: Principles, methods and evaluation, Egyptian Informatics Journal (2015) 16, 261–273.

[29] Joeran Beel, Stefan Langer, Bela Gipp, Marcel Genzmehr, Andreas Nürnberger. A Comparative Analysis of Offline and Online Evaluations and Discussion of Research Paper Recommender System Evaluation, RepSys '13, October 12 2013, Hong Kong, China. Copyright is held by the owner/author(s). Publication rights licensed to ACM.

[30] Gediminas Adomavicius, and Alexander Tuzhilin. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions On Knowledge And Data Engineering, Vol. 17, No. 6, June 2005.

[31] Shah Khusro, Zafar Ali and Irfan Ullah. Recommender Systems: Issues, Challenges, and Research Opportunities, © Springer Science+Business Media Singapore 2016, Information Science and Applications (ICISA) 2016, Lecture Notes in Electrical Engineering 376, DOI: 10.1007/978-981-10-0557-2_112

[32] Balraj Kumar and Neeraj Sharma. Approaches, Issues and Challenges in Recommender Systems: A Systematic Review, Indian Journal of Science and Technology, Vol 9(47), DOI: 10.17485/ijst/2016/v9i47/94892, December 2016

[33] Shah Khusro, Zafar Ali and Irfan Ullah. Recommender Systems: Issues, Challenges, and Research Opportunities. Available from: https://www.researchgate.net/publication/294860665_Recommender_Systems_Issues_Challenges_and_Research_Opportunities [accessed May 20, 2017].

[34] Robert Jaschke, Leandro Marinho, Andreas Hotho, Lars Schmidt-Thieme, Gerd Stumme. Tag Recommendations in Social Bookmarking Systems, ISSN 0921-7126, IOS Press.

## Author Profile

**Mapari Vikas P.** is working in D.Y. Patil COE, Ambi, Pune, Maharashtra, INDIA as Assistant Professor. Major Field of interest of Data Mining, Wireless Sensor Networks.