# Critical Investigation on collaborative Recommendation System based on User's credibility

**Sanjeev Dhawan[1], Kulvinder Singh[2]**
Faculty of Computer Science and Engineering[1, 2] , Department of Computer Science & Engineering, University Institute of Engineering and Technology (U.I.E.T), Kurukshetra University, Kurukshetra, Haryana
Email(s): kshanda@rediffmail.com[1], rsdhawan@rediffmail.com[2]
**Monika[3]**
M.Tech.(Computer Engineering)[3], Department of Computer Science & Engineering, University Institute of Engineering and Technology (U.I.E.T), Kurukshetra University, Kurukshetra, Haryana
Email: monikachoudharylather@gmail.com

**Abstract**
**Nowadays, there are many recommendation systems, accessible via internet, which attempt to recommend to users several products such as music, movies, books, etc. In order to understand them first it is necessary to have a description. In a general way, recommendation systems are systems that intend to acquire opinions or preferences about items from a community of users, and use those opinions to present other users with items that are interesting to them and thus increased the ratings for that particular item and increased the revenue of a company. From this general description, one can see that recommendation systems need two basic things to work properly: Information about the preferences of the users', and a way to determine if an item is interesting for a user. Normally, the users' information includes external information, such as user profiles, purchases histories, and product ratings. The way to determine whether an item is interesting to a user or not, depends on the kind of recommendation system. To perform recommendation different techniques have been introduced such as content-based, collaborative and hybrid techniques. Therefore, in this paper an attempt is made to distinguish the recommendation systems which are normally used in various online social websites (like Twitter, Facebook, Instagram etc.) and also in variety of some most popular online shopping websites (like Amazon, e-bay, Flipkart, Myntra, Jabong, Snapdeal etc.) by the ways of recommendations techniques viz. collaborative filtering, Content-Based and Hybrid.**
**Keywords: Collaborative Filtering Technique, Content-Based Technique, Hybrid Technique, Recommendation Systems.**

## 1. Introduction

Recommender systems have changed the view that how people will find products, information, and even other people. They study patterns of behavior to know what someone will prefer from among a collection of things he has never experienced. It will study the most important of those tools, including how they work, how to use them, how to evaluate them, and their advantages and drawbacks in practice. Recommender systems deal with information overload by automatically suggesting to users items that may fit their interests. Accurate recommendations enable users to locate accurate items that are of their interest without overwhelmed by unwanted information. Vendors have to recommend theirs visitor those products that match their interests and hopeful turn them into satisfied and returning customers.

## 2. Types of Recommendation System

### 2.1 Collaborative Filtering

Collaborative Filtering (CF) based recommendation techniques help people to make choices based on the opinions of other people who share similar interests Koren Yehuda and Robert Bell [1].

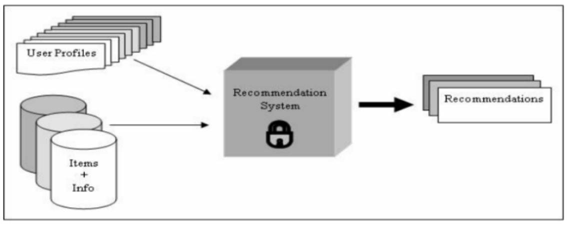


**Fig.1: Recommendation process as a black box**

CF technique has divided into user-based and item-based CF approaches Badrul *et.al* [2].

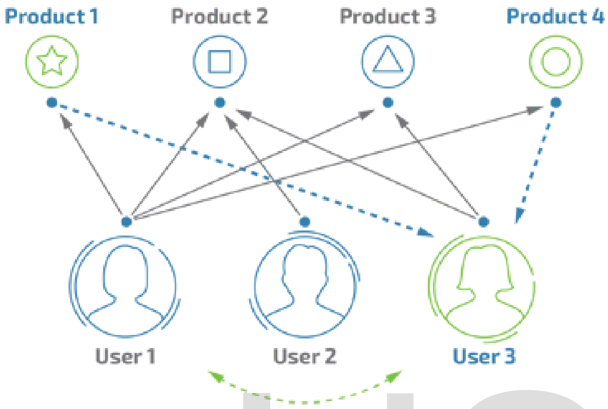


**Fig. 2: General Process of Collaborative filtering**

Collaborative filtering algorithms often require users' active participation, an easy way to represent users' interests to the system, and algorithms that are able to match people with similar interests.

Collaborative Filtering System work as follows:

(a) A user expresses his or her interest by rating items (e.g. books, movies or CDs). These ratings can be viewed as an approximate representation of the user's interest in the corresponding domain.

(b) Given user's ratings are matched against other users' and finds the people with most "similar" tastes.

(c) With similar users, the system recommends items that the similar users have rated highly.

The main problem of collaborative filtering is how to combine and weight the preferences of user neighbors. Sometimes, users can immediately rate the recommended items. As a result, the system gains an increasingly accurate representation of user preferences over time.

Collaborative Filtering Technique is categories into three main categories, each of them use some algorithms for prediction (Su and Taghi, 2009).

**2.1.1 Memory-based Collaborative Filtering Technique**

Memory-based collaborative filtering uses many algorithms to utilize the user-item database in order to generate a prediction. It uses some statistical techniques to find a set of users called neighbors who share a similar history. On other hand, it finds the users who rated similar items or purchased similar sets and by employing algorithms to combine the preferences of the neighbors; it produces the recommendation for the active user. The nearest-neighbor algorithm is a well-known algorithm that is used by memory-based CF to generate a predication

**2.1.2 Model-based Collaborative Filtering Technique**

This algorithm provide user recommendations based on learned models. This technique depends on learning concept, that is, the system that can analyze the training data, summarize the complicated patterns into the learned models, and then make predications based on the learned models. The model building process can be processed by different machine learning algorithms such as Bayesian network, clustering and rule-based approaches.

**2.1.3 Hybrid Collaborative Filtering Technique**

Hybrid CF systems technique combines CF with other recommendation techniques, such as content-based recommended systems, and demographic-based recommender systems in order to make recommendations.

**2.2 Content Based Recommendation**

Content-based (CB) recommendation techniques recommend articles or commodities that are similar to items previously preferred by a specific user. The basic workflow of CB recommender systems are:

(a) To emphasis, the description of the items preferred by a particular user the common attributes can be used to distinguish these items.

(b) To compare each item's attributes with the user profile so that only items that have a high degree of similarity with the user profile will be provided by recommendation. In CB Recommendation Systems, two techniques have been used to generate recommendations. First approach generates recommendations using traditional information retrieval methods.The other technique generates recommendations using statistical learning and machine learning methods, largely building models that are capable of learning users' interests from the historical data (training data) of users.

**2.2.1 Advantages**

Idea: a user is likely to have similar level of interest for similar items

System learns importance of item features, and builds a model of what user likes

**2.2.2 Limitations**

Idea: a user is likely to have similar level of interest for similar items

- Need to know about item content
  - requires manual or automatic indexing
  - Item features do not capture everything
- "User cold-start" problem
  - Needs to learn what content features are important for the user, so takes time

**2.3 Hybrid Recommendation**

To achieve higher performance and eliminate the drawbacks of collaborative and content based recommendation techniques, a hybrid recommendation technique that combines the best features of two or more recommendation techniques into one hybrid technique. The existing hybrid recommendation techniques is to combine the CF recommendation techniques with the other recommendation techniques in an attempt to avoid cold-start, sparseness and/or scalability problem Switching.

**2.3.1 Advantages**

- There is no need to have an electronic representation of the items so the computer can parse them.
- Furthermore, the recommendations produced might be very different in content from the original preferences of the user. This means that there is more variety in the kind of music recommended.
- Besides, Recommendations based more on the quality of items, rather than in objective properties such as content.
- Finally, collaborative-filtering techniques are domain independent and can work perfectly in domains where it is hard to extract content information from the items, or where there is not content at all associated with the items.

**2.3.2 Limitations**

- The other issue explained here is relate with popularity of certain items. Recommendations using collaborative techniques may not correspond to actual musical preferences but could be biased on the popularity of a certain item.
- According to collaborative systems produce interesting recommendations only for naïve profiles, they get stuck when using bigger profiles, and they cannot handle correctly heterogeneous profiles.

**3. Application of Recommendation System**

- E-government recommender systems- It refers to the use of the Internet and other information and communication technologies to support governments in providing improved information and services to citizens and businesses.
- G2C service recommendation (Government to Customer)- To support citizens in their access to personalized and adapted services supplied by public administration offices. The proposed system identifies and suggests the most interesting services for a user by considering both the user's profile and the profile of the device being used.
- G2B service recommendation (Government to Business)- In G2B services, many items from a business perspective are one-time items, such as events, which typically receive ratings only after they have ended. Traditional CF techniques cannot recommend these kinds of items due to the sparse rating data.
- E-commerce/E-shopping recommender systems- In the last few years, a number of unique E-shopping recommender systems have been developed to provide guidelines to online individual customers. E Shopping is a

specialized and highly popular field of E-commerce.

- E-library recommender systems: Recommender systems have been used in digital library applications to help users locate and select information and knowledge sources.
- E-learning recommender systems- E-learning recommender systems have become increasingly popular in educational institutions since the early 2000s based on the development of earlier e-learning systems. This type of recommender system usually aims to assist learners to choose the courses, subjects and learning materials that interest them, as well as their learning activities (such as in-class lecture or online study group discussion). In more than ten years' accumulated study on this topic, many practicable e-learning recommender systems are discovered.
- Tag recommendation- Tags are arbitrary words specified by users to label and manage the resources that are uploaded to the Internet. Users want tags to be personalized and convenient to enable the easy sharing of resources, but it is often difficult for users to select appropriate tags from the wide range of possibilities. Tag recommender systems thus become increasingly important for making tag selection easy and personalized.

## 4. Related Work

There is an incessant growth in the Information Technology and web services on the internet. The growth of the internet in this era has made it more difficult to the user to extract information from online sources, user have to spend a lot of time on the corresponding sources to extract useful information. The Grundy System presented by E.Rich [3] it was an early step towards automatic recommender systems. It was primitive, grouping users into "stereotypes" based on a short interview and using hard-coded information about various stereotypes' book preferences to generate recommendations. Later on Goldberg *et al.* [4] discovered a Tapestry system, which was a manual Collaborative Filtering (CF) system.Marlin [5] defined that Collaborative Filtering was work for filtering data established on the preferences of users. Furthermore examines continuing methods for the task of locale forecast from a contraption discovering perspective, display that countless continuing methods counseled for this task. Jacob Abernethy *et al.* [6] they defined a finished way for cooperative filtering (CoF) employing spectral regularization to discover linear operators from "users" to a set of perhaps wanted "objects". Karen HL [7] discovered a Tag-aware recommender systems by fusion of Collaborative Filtering(CoF) algorithms which defines that Recommender Arrangements (RA) target at forecasting items or ratings of items that the user are interested in. To enhance recommendation quality, metadata such as content data of items have utilized as supplementary knowledge. With the rising popularity of the cooperative tagging arrangements, tags might be interesting and functional data to enhance RS algorithms. Collaborative filtering for implicit feedback datasets was considered by Yehuda Koren [8] in which they described that a public task of recommender arrangements is to enhance client experience across personalized recommendations established on prior inherent feedback. They inherent user observations ought to transform into two paired magnitudes: preferences and assurance levels. In supplementary words, for every single user-item pair, they derive from the input data a guesstimate to whether the user should like or disgust the item ("preference") and couple this guesstimate alongside an assurance level. Mobasher Bamshad et al. [9] they scrutinize discovered model-based collaborative filtering as a defense against profile injection attacks. In particular, they ponder two recommendation algorithms, one established on k-means clustering and the supplementary established on Probabilistic Latent Semantic Analysis (PLSA). Momentous improvements in stability and robustness above the average k-nearest acquaintance way after attacked. They have clarified the comparative robustness and stability of model-based algorithms above the memory-based way. Robert Bell M. *and* Yehuda Koren [10] they have Concluded that cooperative filtering is an area established ("k-nearest neighbors"), whereas a user-item preference locale is interpolated from ratings of comparable items and/or users in past. First, they remove precise so-called "global effects" from the data to make the disparate ratings extra comparable, thereby enhancing interpolation accuracy. Second, they display how to derive simultaneously interpolation weights for all nearest neighbors. In the same year, Robert M. Bell and Yehuda Koren [11] discovered neighborhood-based collaborative filtering in which they display how to derive simultaneously interpolation weights for all nearest acquaintances, unlike preceding ways whereas a single heaviness was computed separately. Single value decomposition (SVD) at preprocessing stage produce less Root Mean Squared Error (RMSE) and increase in prediction accuracy and normalization allowed to KNN method also improved estimation accuracy. Ruslan Salakhutdinov *et al.* [12] discovered a Restricted Boltzmann machines in which they find out that how a class of two-layer undirected graphical models, shouted Restricted Boltzmann Mechanisms (RBM's), can be used to ideal tabular data, such as user's ratings of movies. They present effectual discovering and inference procedures for this class of models and clarify that RBM's had been prosperously requested to the Netflix data set, encompassing above 100 million user/movie ratings. Additionally display that RBM's somewhat outperform carefully tuned SVD models. Ahn and Hyung Jun [13] launch a new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, In which they concluded that Collaborative filtering is one of the most prosperous and extensively utilized methods of automated product recommendation in online stores. They both gave a new heuristic similarity compute shouted PIP for cooperative filtering that is extensively utilized for automated product recommendation in Internet stores. The PIP compute was industrialized by employing area specific clarification of user ratings on produce in order to vanquish the flaw of established similarity and distance measures in new user cold-start conditions. PIP was tested employing three openly obtainable datasets for completeness, whereas it displayed superior presentation for new user cold-start conditions. A hybrid CF way was additionally counseled that can join the strengths of PIP and supplementary similarity measures, displaying extremely prosperous aftermath. Heung-Nam Kim *et al.* [14] they concluded that counsel a cooperative filtering method to furnish an increased recommendation quality derived from user-created tags. Experimental aftermath display that the counseled algorithm proposals momentous gains both in words of enhancing the recommendation quality for sparse data and in dealing alongside cold-start users as contrasted to continuing work. Concluded they analyzed the possible of cooperative tagging arrangements, encompassing personalized and biased user preference scrutiny, and specific and vibrant association of content for requesting the recommendations. Additionally they gave a new and exceptional recommendation algorithm via cooperative tags of users to furnish enhanced recommendation quality and to vanquish a little of the limitations in CF systems. Furthermore, they noted that their method could furnish extra suitable items for user preference even nevertheless; the number of suggested items is tiny. In addition, Yehuda Koren [15] discovered the Scalable and accurate collaborative filtering that define recommender arrangements furnish users alongside personalized suggestions for servicing. These arrangements frequently rely on CF whereas past deals were analyzed in order to institute connections amid users and products. The most public way to CF to established on area models that initiate from similarities amid produce or users. Furthermore, they familiarize a new area ideal alongside enhanced forecast accuracy. Unlike preceding ways that are established on heuristic similarities, author ideal area relations by minimizing a globe price function. They counseled a new area established ideal that unlike preceding area methods to established on properly optimizing a globe price function. This leads to enhanced forecast accuracy, as maintaining merits of the area way such as clarify skill of forecasts and skill to grasp new ratings (or new users) lacking retraining the model. Including accuracy, diversity, skill to surprise alongside unexpected recommendations, clarify skill, appropriate top-K recommendations, and computational efficiency. A little of those criteria are moderately facile to compute, such as accuracy and efficiency. Further model based Collaborative Filtering Recommendation was developed by Zhao Xu *et al.* [16] they particular Aim at the long reply period, inaccurate recommendation and cold-start setbacks that confronted by present recommendation algorithm, seizing movie recommendation arrangement as an example , proposes a cooperative filtering recommendation ideal established on user's credibility clustering.

It is clear that although there are various techniques are available to improve the recommendations in order to attract a number of users. However, in most of the recommendation systems (like collaborative filtering, Content-Based and Hybrid), it is difficult to maintain the memory when network size becomes large. In order to avoid this problem, an methodology will be made to design the model-based collaborative filtering to define the recommendation based on User's Credibility.

## 5. Conclusion

When Internet surfers search for information, they rely on recommendations from other people, customer reviews, or recommender systems. Recently, various kinds of recommender systems have attempted to reduce information overload and retain customers by providing personalized recommendations based on preferences. These recommender systems use diversified algorithms to filter data and generate recommendations about items such as books, news, music, Web pages, and even virtual items. Among those various recommender systems, collaborative filtering (CF) has become the most popular recommendation algorithm. This system predicts user preferences based on the opinions of other similar users who have rated the items according to preference. By calculating the level of similarity between users in a rating data set, it becomes possible to find the nearest neighbors with the highest similarity values among all users. Once a user's neighborhood has identified, particular items can be evaluated by forming a weighted summation of neighbors' opinions.

## References

[1] Koren Yehuda and Robert Bell, "Advances in Collaborative Filtering - Recommender Systems Handbook", pp. 145-185. Springer US, 2015.

[2] Badrul Sarwar, George Karypis, Joseph Konstan and John Riedl, "Item-based collaborative filtering recommendation algorithms" In the Proceedings of 10th international conference on World Wide Web, pp. 285-295, ACM, 2001

[3] Rich E. , "User Modeling via Stereotypes", Cognitive Science, Vol. 3(4), pp. 329-354, 1979.

[4] Goldberg D. , D. Nichols, B. M. Oki and D. Terry, "Using collaborative filtering to weave an information tapestry," Communications of the ACM, vol. 35(12), pp. 61–70, 1992.

[5] Marlin Benjamin, "Collaborative filtering: A machine learning perspective" PhD diss., University of Toronto, 2004.

[6] Abernethy Jacob, Francis Bach, Jean-Philippe Vert, and Theodoros Evgeniou, "A New Approach to Collaborative Filtering: Operator Estimation with Spectral Regularization." 2008.

[7] Karen HL, Tso Sutter, Leandro BalbyMarinho and Lars Schmidt-Thieme, "Tag-aware recommender systems by fusion of collaborative filtering algorithms" In Proceedings of the ACM symposium on Applied computing, pp. 1995-1999, 2008.

[8] Koren Yehuda, "Factor in the neighbors: Scalable and accurate collaborative filtering", ACM Transactions on Knowledge Discovery from Data (TKDD) 4, no. 1, 2010.

[9] Mobasher Bamshad, Robin Burke and Jeff J. Sandvig, "Model-based collaborative filtering as a defense against profile injection attacks", In AAAI, vol. 6, pp. 1388. 2006.

[10] Bell Robert M. and Yehuda Koren, "Scalable collaborative filtering with jointly derived neighborhood interpolation weights." In Data Mining, Seventh IEEE International Conference on, pp. 43-52. IEEE, 2007.

[11] Bell Robert M. and Yehuda Koren, "Improved neighborhood-based collaborative filtering", In KDD cup and workshop at the 13th ACM SIGKDD international conference on knowledge discovery and data mining, pp. 7-14, 2007.

[12] Salakhutdinov Ruslan, AndriyMnih and Geoffrey Hinton, "Restricted Boltzmann machines for collaborative filtering", In Proceedings of the 24th international conference on Machine learning, pp. 791-798, ACM, 2007.

[13] Ahn and Hyung Jun, "A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem." Information Sciences 178, no. 1 , pp. 37-51, 2008

[14] Kim Heung-Nam, Ae-Ttie Ji, Inay Ha, and Geun-Sik Jo, "Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation" Electronic Commerce Research and Applications 9, no. 1 , pp.73-83, 2010.

[15] Hu Yifan, Yehuda Koren and Chris Volinsky, "Collaborative filtering for implicit feedback datasets", In the proceedings of 8th IEEE International Conference on Data Mining, ICDM'08, pp. 263-272, 2008.

[16] Xu Zhao and QiaoFuqiang, "Collaborative Filtering Recommendation Model Based on

User's Credibility Clustering", In Distributed Computing and Applications to Business, Engineering and Science (DCABES),13th International Symposium, pp. 234-238, IEEE, 2014.