IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-2, Feb 2019 E-ISSN: 2347-2693

Hybrid Approach for product Recommendations using Collaborative filtering

Chinmay Puranik[1*], Grantej Otari[2]

[1,2]Dept. of Computer Science & Engineering, KIT's College of Engineering, Maharashtra, India

[*]*Corresponding Author: chinmay19puranik@gmail.com Tel.: 9404027798*

DOI: https://doi.org/10.26438/ijcse/v7i2.564568 | Available online at: www.ijcseonline.org

Accepted: 17/Feb/2018, Published: 28/Feb/2019

Abstract— a successful recommendation approach in data mining can be done with the use of Collaborative Filtering (CF). It deals with the information which is recommended by people. People's Choice is one of the better aspects of future recommendations. Typically, CF methods are mostly used for solving the problem of data sparsity and cold-start problem. A novel Domain-sensitive Recommendation (DsRec) is an algorithm used for the rating prediction by exploring the user-item subgroup analysis simultaneously. The Proposed work is an extension to DsRec using Trust-based system that considers the trust of the recommender. This type of recommendation system can help to get the information of user's preferences in different types of domains which make rating predictions trust-worthy and efficient. A trust-based recommendation is complementing the developed algorithm.

Keywords—Collaborative filtering, data mining , recommendation system.

I. INTRODUCTION

With the wide quality of products and services available on the web, it is difficult for different types of users to choose the product or service that most meet their requirements. In order to overcome or even reduce this difficulty, recommender systems have risen. A recommender system is used in several fields to recommend items of interest to users. One of the main areas where this concept is currently used is e-commerce that interacts directly with customers by suggesting products of interest with the aim of improving its sales.

To provide a high-quality personalized recommendation from multiple numbers of choices, the recommendation system has been widely used for all types of users. The supply of online information is increasing throughout the internet. Thus recommender systems are highly supportive to users with different and unique opinions. To witness this situation one of the personalized filtering mechanisms is used that is- Collaborative Filtering (CF) [1]. Collaborative filtering is a process of filtering or evaluating items using the beliefs of other people. While the term collaborative filtering (CF) takes its roots from humans have been doing for centuries - giving ideas to others.

For ages, people have stood over the back or in the office break room and reviewed books they read, restaurants they tried, and movies they have seen – then used these

discussions to form opinions. For example, when enough of Jack's colleagues say they liked the latest release from Bollywood, he might decide that he also should see it.

Similarly, if many of them discovered it a disaster, he might decide to spend his money elsewhere. Better yet, Jack might observe that Matt recommends the types of films that he finds enjoyable, Raj has a history of recommending films that he despises, and Mary just seems to recommend everything. Over time, he learns whose opinions he should listen to and how these opinions can be applied to help him to determine the quality of an item.

II. RELATED WORK

Jing Liu et al. [1] proposed a most successful recommendation approach to make the rating predictions by exploring user-item subgroup analysis in which a user-item subgroup is deemed as a domain consisting of a subset of items with similar characteristics and a subset of users who have interests in these items.

Barley N et al.[2] proposed a method that provides an efficient approach to find the appropriate rating predictions. In other words, it provides better recommendations to the users of the same interest.

SanghviBet al.[3] proposed a method to integrate the involvement of human mood in deciding the user preferences. It is used to increase the accuracy of recommender systems and to generate more reliable results.
Shah Jet al.[4] proposed a system that mainly describes various limitations in the recommendation system. The system mainly consists of hybridization like weighted methods which are used overcome certain limitations.

Schafer Bet al.[5] proposed one of the personalization technologies powering the adaptive web is collaborative filtering. Collaborative filtering is the process of filtering items through the opinions of other people.

Zhang Y et al.[6] proposed an effective recommendation method in which the preference of a user on an item is predicted based on the decisions of other users with related interests.

Zhang X et al.[7] revealed Collaborative Filtering that similar users have similar responses to similar items.
Jiang Y et al.[8] proposed a novel product recommendation method called TCRec, which takes advantage of consumer rating history record. The problem with this algorithm is its performance degradation.

Bin Xuetal.[9] proposed a natural extension of traditional clustering CF models. Experimental outcomes show that using subgroups is a hopeful way to further improve the top-N recommendation performance for many popular CF methods.

Herlocker J et al.[10] presented an algorithmic framework that breaks the collaborative prediction process into components, and provided with pragmatic results regarding variants of each component.

Guo G et al.[11] proposed A simple but effective method to incorporate trusted neighbors in recommender systems.
Hao Ma et al.[12] proposed a factor analysis approach based on probabilistic matrix factorization to solve the data sparsity and poor prediction accuracy problems by employing both users' social network information and rating records.

ManhCuong P et al.[13] suggested that instead of using rating data, use a social relationship between users to identify their neighborhoods. A complex network clustering technique is applied to the social network of users to find groups of similar users. After that, the traditional CF algorithms were used to efficiently generate recommendations.

Si L etal.[14] presented a paper involving flexible mixture model (FMM) for collaborative filtering. FMM extends existing partitioning clustering algorithms for collaborative filtering by clustering both users and items together simultaneously without assuming that each user and item should only belong to a single cluster. Furthermore, with the introduction of 'preference' nodes, the proposed framework was able to explicitly model how users rate items, which can vary dramatically.

Mohsen J et al.[15] described SocialMF to handle the transitivity of trust and trust propagation, which was not captured by the STE model.

Thomas G et al.[16] proposed a weighted co-clustering algorithm that involves instantaneous clustering of users and items. By designing incremental and parallel versions of the co-clustering algorithm it used to build an efficient real-time CF framework.

Slobodan V et al.[17] proposed a regression-based approach to collaborative filtering that searches for similarities between items, building a collection of experts, and combines them in an appropriate way to give predictions for a new user.

Guibing G et al.[18] observed, when the trust is propagated along the web of trust, the proposed method achieved better performance.

III. PROBLEM STATEMENT

To design an effective recommendation system considering User Trust and Item ratings based on the DsRec algorithm. This system can be applied to e-commerce websites for getting user preferences and product recommendation.

IV. OBJECTIVES

Objectives of the dissertation are:

1. Designing an efficient and reliable recommendation system for knowing user preferences and item recommendation based on trust of the recommender.
2. The design recommendation system will be applied for recommending the items of e-commerce websites.
3. Predicting the ratings for unknown items that help in a cold-start problem.

V. METHODOLOGY

The proposed work is planned in the following phases:

A. Phase I: Dataset Collection

The phase I of the proposed work will consist of first collecting the datasets from online shopping systems or movie review rating system. To examine system product review datasets: like Epinions and Ciao or synthetic data can be generated and used. The system will remove the users who rate less than n items and then remove the items which have less than n ratings by the users.

B. Phase II: User Item Subgroup Analysis

In phase II, DsRec algorithm is applied to the collected dataset. The User profile value is generated with the collected ratings for every domain. In short, product base clustering will be done on the basis of type assigned by the system.

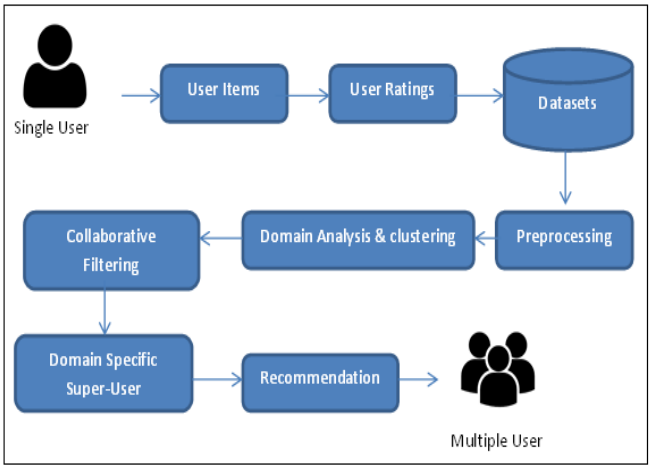

Figure 1. Proposed System Architecture

C. Phase III: cold-start and Trust-based system creation

In phase III, whenever the cold-start condition occurs, the system will assign ratings based on system checks. The new user checks for recommendations for different products with trust based review analysis. A novel "Trust-Based Algorithm" will be developed for the creation of Super-user. The super-user is created by the system depending upon purchases and ratings of product for specific domains. Recommender system will use the ratings of the Super-user and trust value for promoting cold start product for any user. As per the domain preferences, the recommender system will provide recommendations to different users.

VI. METHODOLOGY

Relevant details should be given including experimental design and the technique(s) used along with appropriate statistical methods used clearly along with the year of experimentation (field and laboratory).

VII. RESULTS AND DISCUSSION

The well-known Movielens dataset to conduct the experiments, which is available online, including 100,000 ratings by 943 users on 1682 movies, and assigned to a discrete scale of 1–5. Each user rated at least 20 movies. We use φ to describe the sparsity level of dataset: φml=1–100,000/943 * 1682=0.9369. Then the dataset randomly split into training and test data respectively with a ratio of 8/2. We utilized training data to build the offline model, and the remaining data were used to make a prediction. To verify the quality of recommendation, calculate the mean absolute error (MAE), precision, recall as evaluation measures which have been widely used to compare and measure the performance of recommendation systems. The MAE is a statistical accuracy metric which measures the average absolute difference between the predicted ratings and actual ratings on test users as shown in Eq. (1). A lower MAE value corresponds to more accurate predictions.

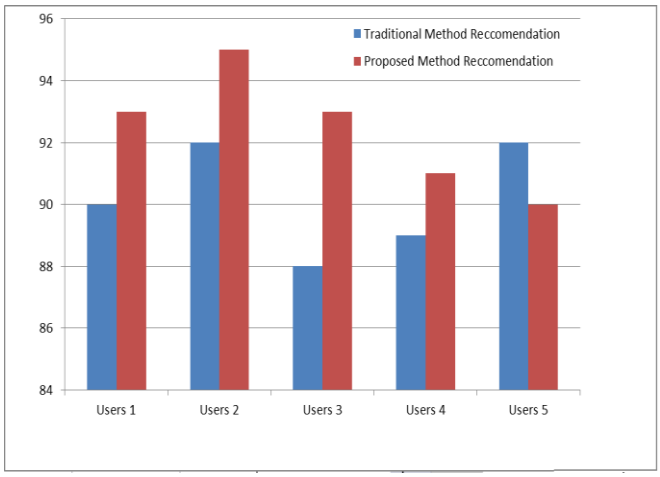

Figure 2. Rating Prediction of Collaborative Filter Recommendation

$$MAE \Rightarrow \sum\left(\left|Pij - rij/M\right|\right) \text{----------(1)}$$

M is the total number of predicted movies, Pij, represents the predicted value for user i on item j, and rij is the true rating. To understand whether users are interested with the recommendation movies, we employ the precision and recall metrics which are widely used in movie recommender systems to evaluate the intelligence level of recommendations. Precision is the ratio of interesting movies retrieved by a recommended method to the number of recommendations. Recall gives the ratio of interesting movies retrieved that is considered interesting. These two

measures are clearly conflicting in nature because increasing the size of recommended movies N leads to an increase in the recall but decrease the precision.

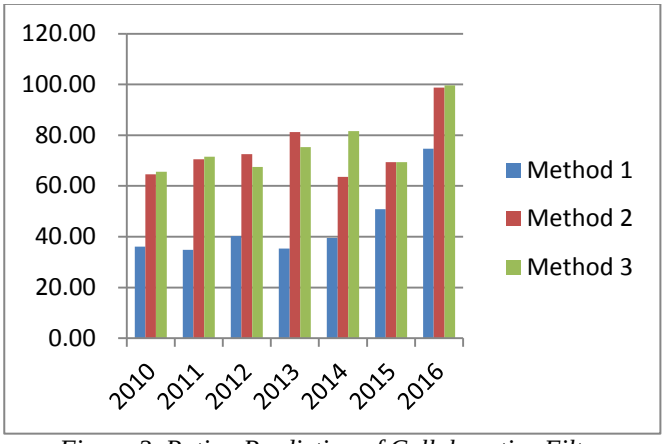

Figure 3. Rating Prediction of Collaborative Filter Recommendation Acuracy

The precision and recall for Top-N (N is the number of predicted movies) recommendation are defined in (2) and (3), respectively.

Precision = |interesting ∩ Top N| /N ---------- (2)

Recall = |interesting ∩ Top N| / |interesting| ---------- (3)

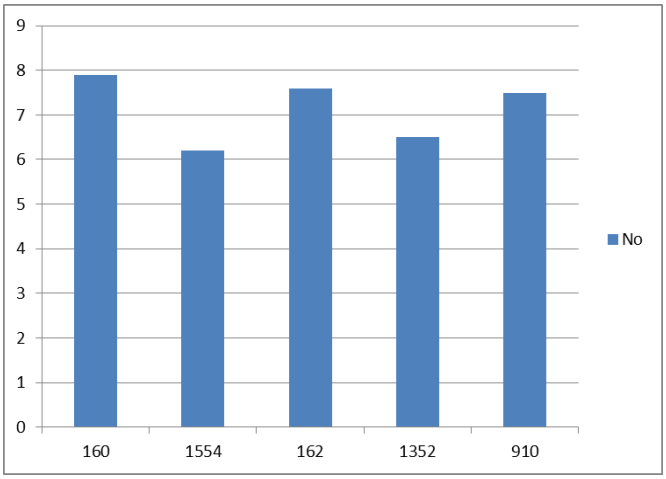

Figure 4: Rating Prediction of Collaborative Filter Recommendation

Although an efficient method to solve the cold start new user problem have been proposed. During the items selected, the new incoming ratings of other users are not considered. Therefore, a future direction can be developing a new method, which will adapt to the earlier ratings given by other users. Figure 5 uses to show rating prediction of Cold Start recommendation.

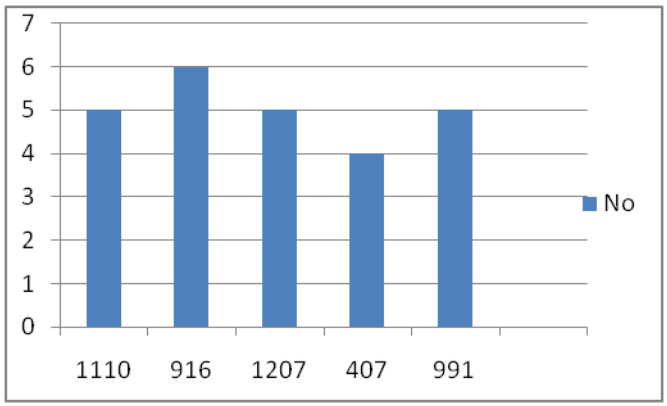

Figure 6. Rating Prediction of Cold Start Recommendation

VIII. CONCLUSION

In this paper, a hybrid CF approach used to generate movie recommendations which combine clustering algorithm. In the data environment, the selections of "like-minded" neighbourhood on the basis of common ratings function to generate high-quality movie recommendations. First performed on whole data space, and then the group was generated from relatively low dimension vector space transformed by the first step. In addition, to result in the best neighbourhood. Based on the Movielens dataset, the experimental evaluation of the proposed approach proved that it is capable of providing high prediction accuracy and more reliable movie recommendations for users' preference compared to the existing CFs. As for the cold start issue, the experiment also demonstrated that our proposed approach is capable of generating an effective estimation of movie ratings for new users via traditional movie recommendation systems.

ACKNOWLEDGMENT

This review paperwork was guided and supported by Mr. G. V. Otari. I would like to thank the guide, anonymous reviewers for their valuable and constructive comments on improving

REFERENCES

[1] Jing Liu, Member, Yu Jiang, Zechao Li, Member, Xi Zhang, and Hanqing Lu, Senior Member, IEEE "*Domain-Sensitive Recommendation with User-Item Subgroup Analysis*", IEEE Transaction on knowledge and data engineering Vol. 28, NO. 4, April 2016.

[2] N. S. Barley1 & Asst. Prof. R. R. Keole, "*A Review of Recommendation System in Domain Sensitive Manner*", Imperial Journal of Interdisciplinary Research (IJIR) Vol-3, Issue-5, 2017 ISSN: 2454-1362.

[3] Bhavya Sanghavi A, Rishabh Rathod A and Dharmeshkumar Mistry, Recommender Systems- "*Comparison of Content-based Filtering and Collaborative Filtering*", Oct 2014.
[4] Jaimeel Shah, Lokesh Sahu *"A Survey of Various Hybrid based Recommendation Method",* International Journal of Advanced Research in Computer Science and Software Engineering Research Paper Volume 4, Issue 11, November 2014 ISSN: 2277 128X.
[5] J.Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen "*Collaborative Filtering Recommender Systems"*. P. Brusilovsky, A. Kobsa, and W. Nejdl (Eds.): The Adaptive Web, LNCS 4321, pp. 291–324, 2007.© Springer-Verlag Berlin Heidelberg 2007.
[6] Y.Zhang, B. Cao, and D.-Y.Yeung, "*Multi-domain collaborative filtering*", in Proc. 26th Conf. Annual Conference. Uncertainty Artificial Intelligence , 2010, pp. 725–732.
[7] X. Zhang, J. Cheng, T. Yuan, B. Niu, and H. Lu, "*TopRec: Domain specific recommendation through community topic mining in social network",* in Proc. 22nd Int. Conf. World Wide Web, 2013, pp. 1501–1510.
[8] Y. Jiang, J. Liu, X. Zhang, Z. Li, and H. Lu, "*TCRec: Product recommendation via exploiting social-trust network and product category information",* in Proc. 22nd Int. Conf. World Wide Web, 2013, pp. 233–234.
[9] Bin XuJiajun Bu Chun Chen Zhejiang Provincial Key Laboratory of Service Robot, College of Computer Science Zhejiang *University "An Exploration of Improving Collaborative Recommender Systems via User-Item Subgroups",* April 16–20, 2012.
[10] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, *"An algorithmic framework for performing collaborative filtering",* in Proc. 22nd Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 1999, pp. 230–237.
[11] G. Guo, J. Zhang, and D. Thalmann, "*A simple but effective method to incorporate trusted neighbors in recommender systems",* in Proceedings of the 20th International Conference on User Modeling, Adaptation and Personalization (UMAP), 2012.
[12] Hao Ma, Haixuan Yang, Michael R. Lyu, Irwin King, "*SoRec: Social Recommendation Using Probabilistic Matrix Factorization",* October 26–30, 2008, Napa Valley, California, USA. Copyright 2008 ACM 978-1-59593-991-3/08/10.
[13] Manh Cuong Pham, Yiwei Cao, Ralf Klamma, Matthias Jarke *"A Clustering Approach for Collaborative Filtering Recommendation Using Social Network Analysis"*, Journal of Universal Computer Science, vol. 17, no. 4 (2011), 583-604 submitted: 30/10/10, accepted: 15/2/11, appeared: 28/2/11 © J.UCS.
[14] Luo Si,Rong Jin*, "Flexible Mixture Model for Collaborative Filtering",* Proceeding ICML'03 Proceedings of the Twentieth International Conference on International Conference on Machine Learning Pages 704-711 Washington, DC, USA-August 21-24, 2003 AAAI Press © 2003. ISBN:1-57735-189-4.
[15] M. Jamali and M. Ester, "*A matrix factorization technique with trust propagation for recommendation in social networks",* in Proceedings of the 4th ACM Conference on Recommender Systems (RecSys), 2010, pp. 135–142.
[16] T George; S.Meregu "*A scalable Collaborative Filtering framework based on Co-clustering",* Data Mining, Fifth IEEE International Conference.27-30 Nov. 2005 IEEE 03 January 2006 ,0-7695-2278-5.
[17] Slobodan Vucetic, Zoran Obradovic "*Collaborative Filtering Using A Regression Based Approach, Knowledge And Information Systems",* Center for Information Science and Technology, Temple University, Philadelphia, PA, USA.
[18] Guibing Guo, Jie Zhang, Daniel Thalmann "*A Simple but Effective Method to Incorporate Trusted Neighbors in Recommender Systems"* Published 2012 in UMAP.

Authors Profile

Mr. Chinmay Puranik pursued Bachelor of Science from KIT's College Of Engineering, Kolhapur in 2015 and pursuing Master of Engineering from KIT's College Of Engineering, Kolhapur. His main research work focuses on Data mining, data filtering mechanisms, and recommendation systems.

Mr Grantej V. Otari pursued Bachelor of Science and Master of Science from Shivaji University, Kolhapur in year 2011. He is currently pursuing Ph.D. and currently working as Assistant Professor in Department Computer Science and Enginering, in KIT's College Of Enginnering. He He has published more than 10 research papers in reputed international journals including. His main research work focuses on Data analytics, Data mining, Network Security. He has 12 years of teaching experience and 4 years of Research Experience.

