An Investigation of Clustering-Based Collaborative Filtering (CBCF)

[1]Santosh J. Shinde and [2]Mr. Pradip A.Chougule

[1,2]Dept. of Computer Science and Engg.,Ashokrao Mane Group of Institutions, Kolhapur ,Maharashtra, India

Abstract— **Providing or recommending appropriate content based on the quality of experience is the most important and challenging issues in the system. As collaborative filtering (CF) is the most outstanding and popular technique used in recommender systems, we propose a new cluster-based CF (CBCF) method, which uses only the incentive/punishment user (IPU) model of a given rating it is executed by the user, so it is easy to implement. Our goal is to design a simple cluster-based method. There is no further prior information and the accuracy of the recommendations is improved. To be precise, CBCF with IPU model aims to improve recommendation performance, such as accuracy, recall rate and F1 Score by carefully taking advantage of different preferences between users. Specifically, we set a constraint, we want to maximize the recall (or equal F1 score) optimization problem under given conditions accurate. For this reason, users are divided into several clusters based on actual rating data and Pearson Correlation coefficient. Then, we give rewards/penalties for each item according to our preferences Trend of users in the same cluster. Our experimental results show excellent performance An improvement over the benchmark CF scheme, instead of clustering for a given recall rate or F1 score accurate[1].**

Keywords- Clustering, collaborative filtering, incentivized/penalized user model, Pearson correlation coefficient, recommender system.

I. INTRODUCTION

It may be difficult for people to find what they like effectively because of the large amount of Videos, audios, papers, artworks, etc. have been created online and offline. For example, hundreds of feature films and Hundreds of thousands of books are produced and published in the United States every year. But one person will Read up to about 10,000 books in a lifetime, and then he/she must choose the book he/she likes among them. In On the one hand, recommendation systems have been developed and used in different fields (such as the film industry, music Industry, etc.) to help people choose appropriate content according to personal preferences [1]. Especially Online Commercial industries such as Amazon.com and Netflix have successfully used how to increase customer loyalty. For example, Amazon.com and Netflix generate a lot of sales by providing their own personalized products recommended system

II. RELATED WORK

A. Exploring social activeness and dynamic interest in community-based recommender system [2]

The community-based recommendation system has attracted a lot of research attention. Forming communities allows us to reduce data sparsely and focus on discovering the underlying characteristics of communities rather than individuals. Previous work focused on how to use various algorithms to detect communities. However, they did not consider the user's social attributes, such as social activity and dynamic interests, which have a strong correlation with users' preferences and choices. Intuitively, people have different social activities in social networks. The ratings of users with higher positivity are more likely to be trustworthy. Time dynamics of interest are also important to user preferences. In this article, we propose a novel community-based framework. We first adopt a PLSA-based model, combining social activity and dynamic interests to discover communities. Then apply the latest matrix factorization method to each community. The experimental results on two real-world data sets verify the effectiveness of our method in improving the recommendation effect.

B. Evolutionary heterogeneous clustering for rating prediction based on user collaborative filtering [3]

The recommendation system plays an important role in our lives, it will help users find the content they are interested in. Collaborative filtering is the most widely used and most successful method in personalized recommendation. In this paper, a novel heterogeneous evolutionary clustering is proposed. The goal of our algorithm is to gather users with similar interests into the same cluster and help users find the most suitable products for their personal tastes. Suggestions from friends with similar interests may be adopted. First, projects and users are regarded as heterogeneous individuals in the network. According to the constructed network model, the status of an individual will change over time. People with higher scores will get together and those with lower scores will escape. After many iterations, the status of the project and users will remain stable. In view of the stable state of heterogeneous individuals, they are divided into several categories. Second, use user-based collaborative filtering in each cluster. The similarity between individuals in the same cluster is not calculated for all individuals in the system. The target rating is calculated based on user-based collaborative filtering in its cluster. Various simulations show the efficiency

of our proposed method. In addition, compared with the existing two preferred algorithms, this method obtains better prediction results.

C. Improved personalized recommendation based on user attributes clustering and score matrix filling [4]

A personalized recommendation system (RS) is used to help people reduce the time it takes to find items of interest. Collaborative filtering (CF) is one of the most successful technologies in RS. Data sparsity leads to inaccurate recommendation results. This paper proposes a new evolutionary clustering method. The goal of our algorithm is to gather users with similar interests into the same cluster and recommend products for users they might like. First, calculate the user attribute distance. According to the constructed network cluster model, user status will change over time. After a period of iteration, the user's status will remain stable. According to the stable state of users, divide them into several groups. The user's interest preferences change over time, and the user's interest is relatively stable in a short period of time. Second, we fill the scoring matrix according to the scoring time and item type. Third, use user-based collaborative filtering in each cluster. Only the filling matrix is used to calculate the similarity between users in the same cluster. Finally, calculate the target score according to the user's neighbor set, and recommend the most interesting N items to the target user. Through experiments, compared with the existing recommendation algorithm, this algorithm improves the recommendation accuracy and effectively solves the sparse problem of the scoring matrix.

D. Dynamic evolutionary clustering approach based on time weight and latent attributes for collaborative filtering recommendation [5]

Collaborative filtering is one of the most widely used single recommendation algorithms. Traditional collaborative filtering recommendation algorithms rarely consider time changes, which may be inaccurate in actual environments. A dynamic evolutionary clustering algorithm based on time weight and potential attributes is proposed. According to the time effect of historical information in the recommendation system, a forgetting curve is introduced to better grasp the user's recent interests. In order to gather users with similar interests into the same cluster, it is necessary to mine project characteristics and user attributes. Therefore, by introducing the forgetting function into the scoring matrix, the project characteristics and user attributes are used to build a network model. Projects and users are regarded as heterogeneous nodes in the network. In addition, a novel dynamic evolutionary clustering algorithm is used to divide users and item sets into K clusters, and cluster the individuals with higher similarity. The preferences of users in the same cluster are similar. Then, collaborative filtering is applied in each cluster to predict ratings. Finally, recommend target users based on the predicted level. The simulation results show that this method has better recommendation accuracy than the existing algorithms based on MovieLens100k, Restaurant & Consumer and CiaoDVD datasets.

E. A fast recommender system for cold user using categorized items [6]

In recent years, recommender systems (RS) have provided considerable progress for users. RS can reduce the time spent by users in order to achieve the desired results faster. The main problem with RS is the existence of cold users. These users have less activity and their preferences are more difficult to detect. The purpose of this research is to provide a new method to improve the recall rate and accuracy of the recommendation system for cold users. According to the available categories of the items, the priority of the suggested items is improved and then presented to cold users. The results obtained show that in addition to increasing the processing speed, the recall rate and accuracy are also acceptable.

F. Hybrid Recommender System under Temporal Vector [7]

This article describes a mixed-time effective recommendation system. The recommendation system is smart Data mining method, which actually performs prediction of new data values based on existing data set analysis. Recommendation systems have become extremely common in recent years and have been applied to various applications. Of The most popular ones are probably movies, music, news, books, research articles, social tags and general products. at this In work, it is recommended to use a recommender system to analyze the movie interests of a specific user group. in this task, Recommendations will be applied to the user side and the movie side. Hybrid recommendation system will combine Content-based functions and collaborative recommendation system. At a later stage, a weighting method will be used Define content-based collaboration methods to perform recommendations. The hybrid approach of these two levels will gives more accurate results.

G. Clustering-Based Collaborative Filtering Using an Incentivized/Penalized User Model [1]

In this paper, they propose a way to using the IPU model in the recommender system, different preferences and clusters among users. Specifically, in the proposed CBCF method, we formulated constrained optimization problem, the recall rate (or F1 score) for a given precision. For this, clustering is applied, so not only users According to the actual score divided into several categories. Data and Pearson correlation coefficient, and rewards/penalties for each item according to preferences trend of users in the same cluster. The main result is proven use of the recommended CBCF method IPU mode is recalling or F1 score for a given accuracy. Possible directions for future research in this field include Designed by using the new cluster-based CF method Properties of model-based CF methods (e.g. matrix Factorization).

III. CONCLUSION

In this way, we have learned that in this work, we have defined a recommender system to identify the rankings assigned to movies. In this work, we Represents a hybrid model to perform the analysis. The hybrid model is analyzed based on the following: And collaborative filtering. One for users, The website ranks second and third. Now, also predict the ranking of new users. First, based on content similarity matches to identify similar users in the data set. In the second stage, the ranking provided by similar users is Analyze under collaborative filtering conditions and make a collective decision on website ranking in the final stage of The dimension included in this work is the time factor. This means that the ranking of all similar users is not analyzed based on time the scope is set in this work. The proposed work is analyzed under the error rate.

IV. ACKNOWLEDGMENT

We are thankful to those people who help us a lot in making of this paper. This will help us to grow both academically and professionally.

REFERENCES

[1] CONG TRAN, JANG-YOUNG KIM3, WON-YONG SHIN, AND SANG-WOOK KIM, "Clustering-Based Collaborative Filtering Using an Incentivized/Penalized User Model", *IEEE Acces s 2019*

[2] B. Yin, Y. Yang, and W. Liu, ''Exploring social activeness and dynamic interest in community-based recommender system,'' in *Proc. 23rd Int. Conf. World Wide Web*, Seoul, South Korea, 2014, pp. 771–776.

[3] J. Chen, H. Wang, and Z. Yan, ''Evolutionary heterogeneous clustering for rating prediction based on user collaborative filtering,'' *Swarm Evol. Comput.*, vol. 38, pp. 35–41, Feb. 2018.

[4] U. Liji, Y. Chai, and J. Chen, ''Improved personalized recommendation based on user attributes clustering and score matrix filling,'' *Comput. Standards Interfaces*, vol. 57, pp. 59–67, Mar. 2018.

[5] J. Chen, L. Wei, Uliji, and L. Zhang, ''Dynamic evolutionary clustering approach based on time weight and latent attributes for collaborative filtering recommendation,'' *Chaos Solitons Fractals*, vol. 114, pp. 8–18, Sep. 2018.

[6] H. Jazayeriy, S. Mohammadi, and S. Shamshirband, ''A fast recommender system for cold user using categorized items,'' *Math. Comput.Appl.*, vol. 23, no. 1, p. 1, Jan. 2018.

[7] Sarita Dhankhar, and Neha. " Hybrid Recommender System under Temporal Vector." *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 4, Issue 7, July 2014.