ENHANCED COLLABORATIVE RECOMMENDATION SYSTEM USING USER ITEM SUBGROUPS

Ms.S.Priya, Mr.D.MansoorHussian
PG Scholar,Assistant Professor
Department of Computer Science and Engineering
Sri Krishna College of Engineering and Technology
15epcs014@skcet.ac.in, mansoorhussain@skcet.ac.in

Abstract

The most efficient and successful recommendation approach is Collaborative filtering(CF).The common CF-based recommender system associates with the user including with the group of compatible users according to their individual preferences over all the items and recommends to the user some unobserved items enjoyed by the group. Existing work uses the multiclass co-clustering model (MCoC) for finding meaningful subgroups and the item may be recommended based on the correlation between user and items. In case of applying collaborative filtering, some groups may have very few elements due to unbalanced clustering. Due to this some user may not have enough correlated items for recommendation. This paper proposes a novel clustering model called probabilistic fuzzy c-means algorithm and conditional random fields to generate more informative subgroups by using user relationship and item.The key idea is to find the correlated subgroups which improves the recommendation.Hence, our approach can be seen as a novel enhancement of MCoC model.

Keywords- Recommendation system, Content-Based Recommendation, collaborative filtering ,cold start, scalablility .

I. INTRODUCTION

The growth of Internet has made it much more difficult to effectively extract useful information from all the available online information. In a world if number of choices are overwhelming, recommender systems help the users to find and evaluate items of interest. They connect the user with items to "consume" (purchase, view, listen to, etc.) by associating the content of recommended items or the opinions of individuals with user's actions or opinions. Recommender systems or recommendation systems are a subclass of information filtering system that seek to predict the rating or preference that a user give to an item. Recommender systems are common in recent years, and are used in a variety of areas: some popular domains include movies, music, news, books, research articles, search queries, social tags, and products in common. Then there are some recommender systems for experts, collaborators, jokes, restaurants, garments, financial services, life insurance, and Twitter pages[1].Recommender system is one application that is being used by many vectors and online service providers to understand the needs of online users. Based on user's need, the system is able to suggest them the best suitable product or match. The recommender system[2] is provided as an intelligent system, that identifies the user category based on user information and user interest analysis. Recommender systems typically produce a list of recommendations in one of three classified ways,they are collaborative filtering ,content-based filtering or the personality-based approach and Hybrid approach .Collaborative filtering approaches builds a model from a user's past behaviour as well as similar decisions made by other users.

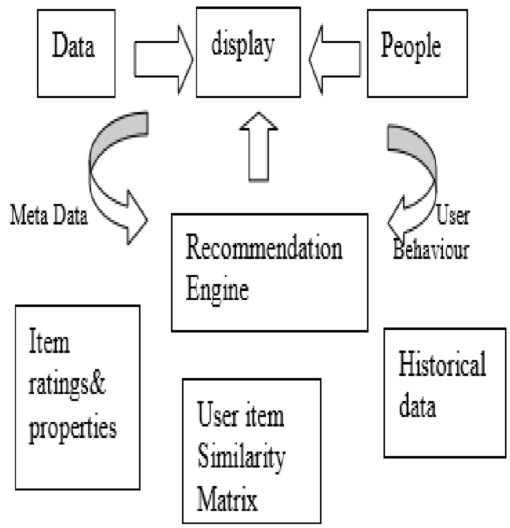

FIG 1.1:Architecture of Recommendation system

This model is then used to predict items that the user may be interested. Content-based filtering approaches utilize a series of discrete characteristics of an item in order to recommend the required additional items with similar properties. Hybrid methods are the combination of Content-based recommending and Collaborative Filtering (CF) methods.**C**ontent based recommendation systems[3] recommends an item to a user based upon the description of the item and items similar to those that user has already purchased or reviewed. The user will get recommendation pfsimilar items to the ones the user preferred in the past. A CF-based [4] system associates the user with a group of like-minded users based on their history preferences (explicit or implicit) over all the items, and then recommends the item to the user that are enjoyed by the group. A group of likeminded users are called as neighbours ..With Hybrid Recommendation many types of problems such as Cold-Start problem can be handled using the hybrid recommendations **.**Different ways of hybridization [5]are implementing CF and CB separately and then joining their predictions, Incorporating some content based methods into collaborative filtering,Incorporating some of the collaborative characteristics into content based approach, Constructing a unifying model that posses both content-based and collaborative characteristics.

II. RELATED WORK

Internet has led to the rising popularity of many social network services.Recommedantion has gained an importance where the goal is to find a set of users[6]. First paper on recommender system have been explained to increase the reliability of recommendation system.. In year 2007 Paul Resnick proposed an idea of "influence limiter algorithm" in recommender system which prevents any attack that are irrelevant result for the search and limits the number of content that the attacker can modify[7]. In 2008 KleanthiLakiotaki, SteliosTsafarakis, and NikolaosMatsatsinis proposed UTA-Rec that incorporates Multiple Criteria Analysis methodologies. The technique[8] that usesrating structure is termed as "Collaborative Filtering" and it was introduced by Goldberg et al (1992) in the context of firstcommercial recommender system. User-based [9] and Item-based [10]CF algorithms are two best methods that fall into this category. Nearest neighbour can be determined by various similarity measures, such as Pearson correlation and cosine similarity. User-based CF attempts to find the neighbourhood of like-minded users for each and every user, then predicts the user's ratings according to the ratings given by the user's neighbours. Item-based CF correlates similar items of each item, and predictions for a user are determined by the user's historical ratings. The Matrix Factorization (MF) [11] approaches attempt to factorize a rating matrix into products of real-valued component matrices.

III. PROPOSED RECOMMENDATION SYSTEM ARCHITECTURE

The proposed system formulates a new Multiclass Co-Clustering (MCoC) model to find meaningful subgroups, which captures relations of user-to-item, user-to-user, and item-to-item simultaneously.

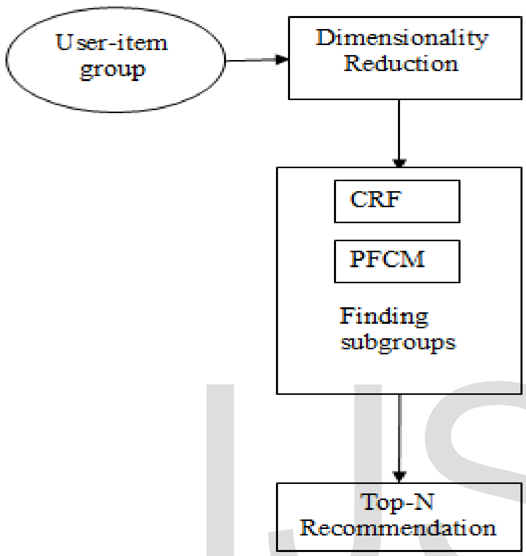

Fig 2: The proposed Recommendation System Architecture

The CF-based recommender method associates a user with a group of like-minded users based on their history preferences with all the items, and then recommends the items enjoyed by the group. The group of likeminded users are called as neighbours. The assumption is that users with same behaviours on observed items with the ratings will have similar tastes on unobserved items.But this assumption is not always tenable – that is two users having similar tastes on one item subset may have totally different tastes on another set. Then the group of like-minded users are taken into a subset but not all of the items. A subset of items and a group of interested users are clustered called as a user-item subgroup[12]. In case of applying collaborative filtering,some groups may have very few elements due to unbalanced clustering.Due to this some user may not have enough correlated items for recommendation.The main objective is to generating more informative subgroups by using user relationship and item tags. The subgroups are discovered by using conditional random field and probabilistic fuzzy c-means clustering algorithms. Therefore, the users can have more correlated items for recommendation.

IV. CLUSTERING ALGORITHM

In the process of collaborative filtering, unbalanced clustering may take place in some groups .Due to this uncorrelated items are recommended for user.The objective is to generate more informative subgroups to the user. The correlated subgroups are discovered by using conditional random field and probabilistic fuzzy c-means clustering algorithms. Thus users can have more correlated items for recommendation.

A. PROPOSED CONDITIONAL RANDOM FIELD CLUSTERING ALGORITHM

Users and items are clustered by conditional random fields and then initial probability for each cluster and label Then cluster with maximum probability is obtained .

Input: Define the eigen vectors for finding the subgroups

Step 1: Consider feature set $X = [x_1, x_2, \dots, x_n]$ and corresponding label set $Y = [y_1, y_2, \dots, y_n]$

Step 2: Assign a random label to each feature

Step 3: For each feature i

3.1 Form a voting pool, select the most similar feature

3.2 Compute the cost functions$U_i\left(y_i, y_{N_i}, x_i, x_{N_i}\right)$ and $W_{i,j}\left(y_i, y_j, x_i, x_j\right)$

3.3 Compute the probabilities of the labels

$$P\left(y_i | x_i, x_{N_i}, y_{N_i}\right) \propto exp\left(-U_i\left(y_i, y_{N_i}, x_i, x_{N_i}\right)\right)$$

3.4 Assign the label with the greatest probability y_j to feature i if its current label y_i is not equal to y_j .
Step 4: Move to step 3 if the label of any feature is updated or
Step 5: Stop
Output: Clustered subgroups are obtained.

B. PROPOSED PROBABILISTIC FUZZY C-MEANS CLUSTERING ALGORITHM

Users and items are clustered by probabilistic fuzzy c-means for multiple cases. Minimize the objective function.Obtain partition matrix which describes subgroup membership of all users and items.

Input: Define the eigen vectors for finding the subgroups
Step 1: Initialize feature set $X = [x_1, x_2, \dots, x_n]$ and number of cluster, c
Step 2: Initialize objective function$J_m^*(U, v) = \sum_{k=1}^{N} \sum_{i=1}^{c} (u_{ik}^*)^m \, d_{ik}^2$
Step 3: Define$m$ //$1 \le m < \infty$ and Define center of cluster $i, v_i = (v_{i1}, v_{i2}, \dots, v_{in})$ and vectors of cluster centers,
$V = (v_1, v_2, \dots, v_c)$
Step 4: Assign $U = [u_{ik}]\, matrix$
Step 5: For number of iterations do and
$e_n(u, p) = u_n p_n$
Step 6: Update fuzzy membership function to cluster i. $u_{ik}^* = u_{ik} p_{ik}$
Step 7: Update center of cluster, v_i
Step8:If$\|U^{(n+1)} - U^{(n)}\| < \varepsilon$ is not satisfied then move to step 6 or
Step 9: Stop
Output: Clustered subgroups are obtained.

V. EXPERIMENTAL EVALUATIONS

There are good measurements for evaluating the prediction or matrix completion task.Precision is used to evaluate the quality of the top K recommended items. Then some users may have a large number of ratings while some other users have a few, so F1 score is required.some of the other metrics such as accuracy,specificity,kappa, recall are also used to evaluate the clustered recommendation subgroups for calculating the performance of top recommended list.

A. GRAPHICAL ANALYSIS

i. Accuracy vs clustering methods:

With the Movielens dataset, clustering subgroups are formed. Consider Fig 3 accuracy is calculated between the various clustering methods.The better clustering method is used for clustering the users and items and recommendation is provided.

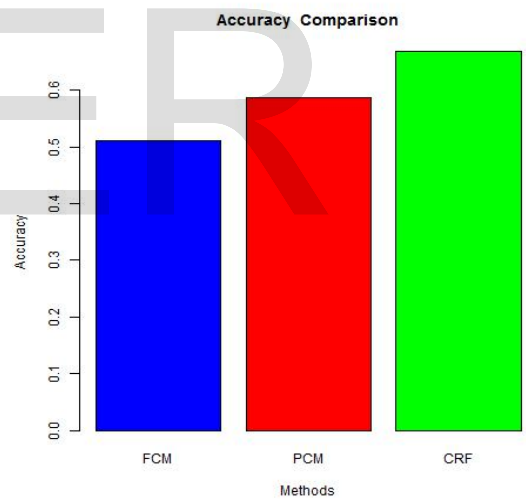

Fig 3: Accuracy vs clustering method

ii. F1 vs clustering methods:

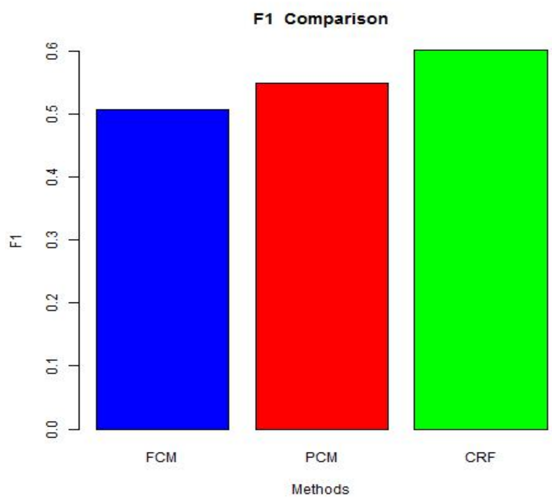

Fig 4:F1 vs clustering methods

With the Movielens dataset,clustering subgroups are formed.Consider Fig 4, F1 score is calculated between the various clustering methods.The better clustering method is used for clustering the users and items and recommendation is provided.

iii. Recall vs clustering methods:

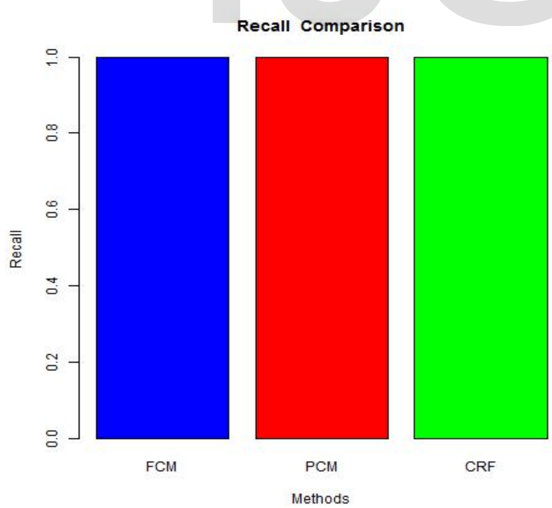

Fig 5:Recall vs clustering methods

With the Movielens dataset,clustering subgroups are formed.Consider Fig 5, Recall is calculated between the various clustering methods.Recall remains same for all clustering methods.The better clustering method is used for clustering the users and items and recommendation is provided.

iv. Precision vs clustering methods:

With the Movielens dataset,clustering subgroups are formed.Consider Fig 6,Precision is calculated between the various clustering methods.The better clustering method is used for clustering the users and items and recommendation is provided.

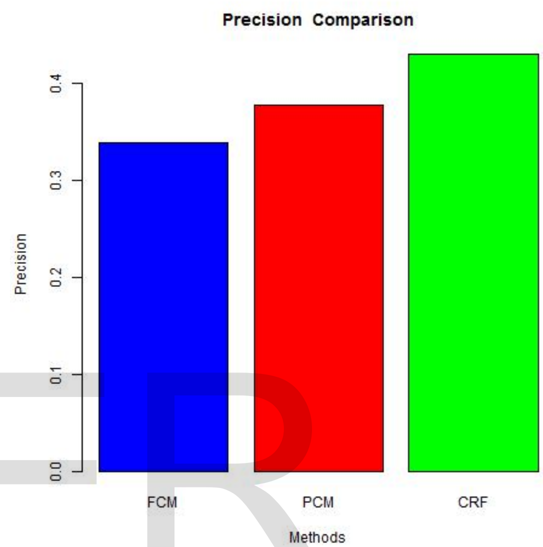

Fig 6:**Precision vs clustering methods**

VI. CONCLUSION

In this paper, new clustering method forcollaborative recommender systems which utilizes useritemsubgroups is used.Due to unbalanced clustering some subgroups may have few elements and under extreme condition some user may not have correlated items for recommendation.To remove small groups which contain uncorrelated items for recommendation.Here the clustering method for Multiclass Co Clustering problem to find the meaningfulsubgroups. This method models the user-to-user, user-to-item, and item-to-item relations simultaneouslyinto a unified optimization problem andadopt an approximate solution. Experimental resultsshow that this new clustering methods provide subgroups to improve the top-N recommendation performancefor many popular CF methods. Futureworks are needed to be done for providing recommendation of items for all fields into a single domain.

VII. REFERENCES

[1] J. Ben Schafer,,"The Application of Data-Mining to Recommender Systems".

[2] G. Adomavicius and A. Tuzhilin." Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions". IEEE transactions on knowledge and dataengineering, pages 734–749, 2005.

[3]A .H Rafsanjani,NSalim,A R Aghdam,K B Fard,"Recommendation system: a Review",2013

[4] G..Adomavicius, A Tuzhilin," Recommendation Technologies:Survey of Current Methods and Possible Extensions",2004

[5]A. Kuepper, "Recommender systems," eCommerce, vol. Chapter 08, p. 13, 2011.

[6] G. Zhao, M. L. Lee,W. Hsu, W. Chen, and H. Hu. "Communitybased user recommendation in uni-directional social networks" In Proceedings of the 22nd ACM international conference on Conference on information & knowledge management, pages 189–198, 2013.

[7]Parul ,k .khanna,"Literature Survey: Recommender Systems", March 2015, Volume 3 Special Issue, ISSN 2349-4476

[8] M sharma,Smann ,"A survey of Recommender systems :Approaches and Limitations ",2013.

[9] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, ands J. Riedl. GroupLens:" An Open Architecture for Collaborative Filtering of Netnews".CSCW, 1994.

[10] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms. Proc. WWW, pages 285–295, 2001.

[11] Y. Koren, R. Bell, and C. Volinsky. "Matrix Factorization Techniques for Recommender Systems" .Computer, 42:30–37, 2009.

[12]J.Bu,Xin.Shen,Bin.Xu,chunchen,Xiaofei He,DendCai "Improving Collaborative Recommendation via User-Item subgroups" in IEEE,2016.