PRODUCT BASED RECOMMENDATION SYSTEM

[1]Priyanka A. Nandagawali , [2]Prof. Jaikumar M. Patil

[1,2] *Department of Computer Engineering, SSGMCE Shegaon, India.*

Abstract

The spread of technology through internet is increasing day by day. With the increasing technology, uses of the recommendation systems are coming into force. The recommendation systems solve many problems of customers by providing them recommendation based on their choice of products. Many of the collaborative filtering algorithms have been used for this purpose. This paper provides a solution for the various recommender systems by using collaborative filtering algorithms with the community based user domain model. The main purpose is to satisfy the customer's product needs by providing them recommendation based on products. Collaborative Filtering (CF) is a commonly used technique in recommendation systems. It can promote items of interest to a target user from a large selection of available items. Considering the shortcomings of the two types of algorithms i.e memory based and model based, a novel approach is considered where Community-based User domain model is used for Collaborative Recommendation. The idea comes from the fact that recommendations are usually made by users with similar preferences. The first step is to build a user-user social network based on users' preference data. The second step is to find communities with similar user preferences using any community detective algorithm. Finally, items are recommended to users by applying collaborative filtering on communities. Because we recommend items to users in communities instead of to an entire social network, the method has perfect online performance and experimental results may show some accuracy within that recommendation increasing its time complexity.

***Keywords —Recommendation system, collaborative filtering, community based system*.**

I. INTRODUCTION

The main goal of customer profiling (or segmentation) is to build reliable customer models for targeted marketing campaigns; and consequently, a better profitability. Consequently, one can define data mining in customer profiling as being the technology that allows building customer models each describing the specific habits, needs, and behavior of a group of customers. Once discovered, these models can be used to classify new customers; and thereby, predict their special needs. In real environment, users have unlimited and unpredictable desires and their preferences may vary within different product categories. For example a user may be interested in buying inexpensive and pocket-size books, while this user may be interested to buy expensive and big toys. To meet the need of the customers, this paper introduces a novel technique where, with the satisfaction of the customer needs, the implementation results for the collaborative filtering algorithms have also tremendously increased.

A proposed novel algorithm: the Community-based User domain model is considered. This algorithm maps a user-item data set to a user-user social network based only on user-item preference data. It then finds user similar preference communities to define a user domain model by detecting communities on a user- user social network. Finally it makes memory-based recommendations in the community-based user domain model. The algorithm does not depend on additional dataset information, and uses communities within a social network as user domain models that include detailed behavioral interpretability.

II. RECOMMENDATION SYSTEMS

The recommendation systems are those that recommend items to its user's in a way to satisfy the needs of its customers. The recommender systems mainly evaluates the customer's profile by seeing to what choice is it making or using while rating items. The proposed recommender system in this paper evaluates the customer's choice based on his/her profession. Instead of searching the products in the complete dataset, the user can verify the products taken from the list of friends in their community.

These types of systems are used to increase the customer choice selection and to suggest the customer the best option that he/she can use for purchasing items. Recommendations based on products are a way for the marketing people to place their choices in front of the customer's in a way to make them content in the selection process. The best recommendation system usually is effective in case of communities or groups where user – item recommendation can be done directly based on the choice of the people in the communities.

Recommender systems use the user, item, and ratings information to predict how other users will like a particular item. Recommender systems are now pervasive and seek to make profit out of customers or successfully meet their needs. However, to reach this goal, systems need to parse a lot of data and collect information, sometimes from different resources, and predict how the user will like the product or item. The computation power needed is considerable. Also, companies try to avoid flooding customer mailboxes with hundreds of products each morning, thus they are looking for one email or text that will make the customer look and act. This goal is tried to achieve through this paper.

III. COLLABORATIVE FILTERING

Collaborative Filtering (CF) is a commonly used technique in recommendation systems. It can promote items of interest to a target user from a large selection of available items. It is divided into two broad classes: memory-based algorithms and model-based algorithms. The latter requires some time to build a model but recommends online items quickly, while the former is time-consuming but does not require pre-building time. The Collaborative Filtering (CF) approach is probably the most familiar, most widely implemented, and most matures of the recommendation approaches. Its core concept is to utilize a collective intelligence to collect answers from crowd behavior and data. A classification of CF algorithms that divides them into two broad classes: memory-based algorithms and model-based algorithms.

IV. PROCEDURE

The method proceeds in the following ways:

1. Allow the user to register first.

International Journal of Advanced Technology in Engineering and Science www.ijates.com
Volume No.02, Issue No. 05, May 2014 ISSN (online): 2348 – 7550

2. After registration the user can login with the help of email id as user name and password specified during registration.
3. The registered user will get the product recommendations based on his profession, gender and age as filled in details by the user while registering.
4. Recommendations will mainly be based on the user's profession.
5. Based on the recommendations the products and the hit rate will be plotted in the graph to give the result analysis to the customer's.

FIG 1. Block diagram for product based recommendation

V. ALGORITHMS

Algorithm:

Input: m users- n items preference data set

Output: A set of communities

Phase 1: Calculate the similarities of users using user item preference data set.

Phase 2: Build a KNN social network.

For every i belongs to V (the set of users)

Compute the top-K largest similarity users.

Establish edges between each two users.

End for

Phase 3: Apply an existing community detection method to generate communities.

After building the user domain model, we recommend items of interest to a target user in community.

VI. METHOD DESCRIPTION

The method is divided into three sub methods:

- Calculate the similarity of users based on their age.
- Calculate the similarity based on gender.
- Calculate more number of recommendations based on profession.
- Map the similarity relationships to user user social network.
- Apply the community detection based on age, gender and profession and make the recommendations

The similarity is measured by using similarity metrics and the ratings are predicted by the number of hits a product is getting after clicking on it. The rating of a product depends on the clicks it gets and how many times has it been viewed. The product getting more number of likes is usually having higher chances of getting recommendation.

VII. SIMILARITY MEASURES

a) Similarity Measures

In order to measure similarity, we want to find the correlation between two users. This gives us a value from -1 to 1 which determines who alike two users are. A value of 1 means that they both rate in the exactly the same manner, whereas a value of -1 means that they rate things exactly opposite (i.e. one high, the other low or vice versa).

There were two similarity measurements we used. The first was the Pearson correlation coefficient. It is the basic correlation algorithm for samples adapted for rating information. It tries to measure how much two users vary together from their normal votes - that is, the direction/magnitude of each vote in comparison to their voting average. If they vary in the same way on the items they have rated in common, they will get a positive correlation; otherwise, they will get a negative correlation.

The other similarity measurement is called vector similarity. We can treat two users as vectors in *n*-dimensional space, where *n* is the number of items in the database. As with any two vectors, we can compare the angle between them. Intuitively, if the two vectors generally point in the same direction, they get a positive similarity; if they point in opposite directions, they get a negative similarity. To simulate this we just take the cosine the angle between these two vectors, which gives us a value from -1 to 1.

Rating Based Cosine Similarity:

The rating based cosine similarity of user I and user j is defined as follows:

$$\mathrm{sim}_{ij}^{\mathrm{rate}} = \frac{\sum_{\alpha\in U_{ij}} r_{i\alpha} r_{j\alpha}}{\sqrt{\sum_{\alpha\in U_{ij}} r_{i\alpha}^2 \sum_{\alpha\in U_{ij}} r_{j\alpha}^2}} \quad (1)$$

Where Uij represents the set of items which are rated both by user i and user j.

b) Predicting Ratings

In order to predict a rating for an item for an active user, we need to find all weights between the active user and all other users. We then take all non-zero weights and have each other user "vote" on what they think the active user should rate the item. Those with higher weights will matter more in the voting process. Once these votes are tallied, we have a predicted vote.

Note that the voting is based on how far off from a user's average they rate a movie - that is, we want to say how far off from the active user's average the active user will rate the item. Thus, with a positive correlation, the active user agrees with however far off the other user voted on a particular item; and with a negative correlation, the active user disagrees (i.e. goes in the opposite direction) from the other user's vote.

Advantages:

- The quality of predictions is rather good.
- This is a relatively simple algorithm to implement for any situation.
- It is very easy to update the database, since it uses the entire database every time it makes a prediction.

Disadvantages:

- It uses the entire database every time it makes a prediction, so it needs to be in memory it is very, very slow.
- Even when in memory, it uses the entire database every time it makes a prediction, so it is very slow.

. VIII. COMPARISONS

A memory-based algorithm, such as the user-based K -Nearest Neighbor (KNN) algorithm utilizes an entire database of user preferences to compute recommendations. These algorithms tend to be simple to implement and require no training (offline) cost. But as the size of user and item sets increase, the online performance of memory-based algorithms tends to decrease. A model-based algorithm, and other modified algorithms, build a model of the preference data and use it to produce recommendations. Usually, the model-building process is time-consuming and only done periodically. The online performance of model-based algorithms is better than memory-based algorithms. However, many model- based algorithms lack interpretability; for example how and why to select user (item) vector dimensions. In addition, the primary shortcoming of model-based algorithm is that they need to generate to create a new model even if users change their behaviors or choices.

1. KNN algorithm: This algorithm directly recommends items to the target user in the user's domain based on K most similar users in the user model.

2. User – Based Collaborative filtering: This method directly recommends items to target user in a set of all other users.

Comparisons are made according to error detection techniques: MAE and RMSE.

Where MAE can be calculated as :-

MAE:

$$\text{MAE} = \frac{\sum_{(i,\alpha)\in\Gamma} |r_{i\alpha} - \tilde{r}_{i\alpha}|}{|\Gamma|}$$

RMSE:

$$\text{RMSE} = \sqrt{\frac{\sum_{(i,\alpha)\in\Gamma} (r_{i\alpha} - \tilde{r}_{i\alpha})^2}{|\Gamma|}}$$

Fig 2. **Comparison of rating- prediction quality of the selected CF algorithms**

CF Algorithms	MAE	RMSE
User Based	0.8164	1.022
Item Based	0.8341	1.0448
Model Based	0.8501	1.1051
Communtiy Based	0.6919	0.9602

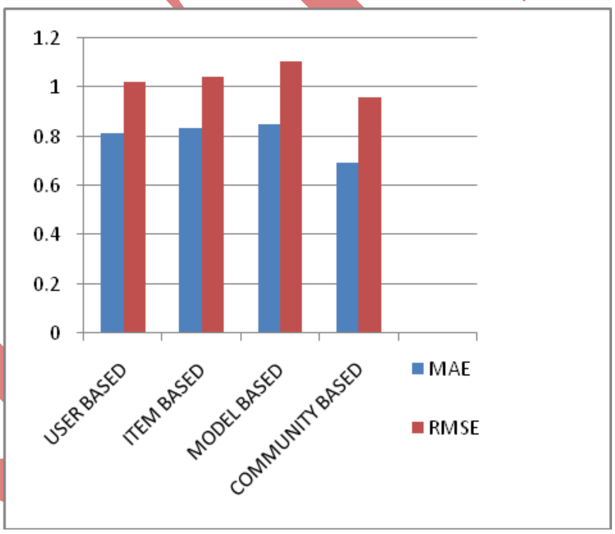

Fig. 3 Comparisons through graphs

IX. CONCLUSION

The recommendation system used in this paper is for providing product based recommendations based on age, gender and profession. The collaborative filtering techniques filter these results and give the best recommendations to the users. These filtering techniques are quite powerful and useful for predicting the choice of users. The main aim of the paper was of recommending the products that are best suited for the users according to their profession.

The advantages can be summarized as:

(1) Based on community, the selection of a user domain has interpretability in user behavior.

(2) The algorithm has strong real-time performance with low online time complexity.

(3) Preference similarity relations of users within the community are relatively stable so that the need for model regeneration is reduced.

(4) It can directly use the related algorithms in social networks or other similar complex networks.

REFERENCES:

[1] R. Burke, Hybrid recommender systems: Survey and experiments, User Modeling and User-Adapted Interaction, vol. 12, no. 4, pp. 331-370, Nov 2002.

[2] Fulan Qian, Yanping Zhang, Yuan Zhang, and Zhen Duan, Community-Based User Domain Model Collaborative Recommendation Algorithm, TSINGHUA SCIENCE AND TECHNOLOGY ISSN 1007-0214 03/10 pp353-359 Volume 18, Number 4, August 2013

[3] Lyle H. Ungar and Dean P. Foster, A Formal Statistical Approach to Collaborative Filtering IS Dept. and Dept. of Statistics University of Pennsylvania Philadelphia, PA 19104, 1998

[4] Hao Ma, Irwin King, Senior Member, IEEE, and Michael Rung-Tsong Lyu, Fellow, IEEE , Mining Web Graphs for Recommendations, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 6, JUNE 2012

[5] E. Viennet, Collaborative filtering in social networks A community-based approach, in Computing, Management and Telecommunications (ComManTel), 2013 International Conference on, IEEE, Jan. 2013, pp. 128-133.

[6] R. Pan, P. Dolog, and G. Xu, KNN-based clustering for improving social recommender systems, in Agents and Data Mining Interaction, Springer, 2013, pp. 115-125