Review Article Impact Factor: 0.621 ISSN: 2319-507X
Priyanka Nandagawali, IJPRET, 2014; Volume 2 (9): 31-36 IJPRET


INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

# REVIEW PAPER ON COLLABORATIVE FILTERING ALGORITHMS WITH THE COMMUNITY BASED USER DOMAIN MODEL

**PRIYANKA A. NANDAGAWALI[1], PROF. J. M. PATIL[2]**

1. Department of Computer Engineering, PG Student, SSGMCE, Shegaon,India
2. Department of Computer Science & Engineering, Faculty of Technical education, SSGMCE, Shegaon,India

**Accepted Date: 27/02/2014 ; Published Date: 01/05/2014**

**Abstract:** 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. Considering the shortcomings of the two types of algorithms, 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. Applying this method to a collaborative tagging system, experimental results may show some accuracy within that recommendation increasing its time complexity.

**Keywords:** Collaborative Filtering, Recommendation System, Community Based User Domain Model.

**Corresponding Author: MS. PRIYANKA A. NANDAGAWALI**

**Access Online On:**

www.ijpret.com

**How to Cite This Article:**

Priyanka Nandagawali, IJPRET, 2014; Volume 2 (9): 31-36

***PAPER-QR CODE***

## INTRODUCTION

The amount of Internet information has increased sharply over the years. How to effectively filter overload information and find useful data and interrelations is a hot issue in network data mining. The study is of great significance, especially when considering big data. Recommendation systems are a type of effective tool used to solve the problem of information overload. They can automatically recommend items to a target user according to observed user preference information, purchase behavior, evaluation behavior, and so on.

Recommendation systems have been successful in business applications. Amazon.com (http:// www.amazon.com) uses a recommendation system to predict users' purchase behaviors, recommends goods of potential interest, and improves Amazon's commercial profit. Some websites use recommendation systems to increase customer satisfaction. For example, Movie lens (http://movielens.umn.edu) is a movie recommendation website that can recommend movies of most interest to a target user according to the user's historical rating records.

The Collaborative Filtering (CF) approach is probably the most familiar, most widely implemented, and most mature 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 is: memory-based algorithms and model-based algorithms. Due to some shortcomings of the two techniques a novel approach is proposed.

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.

## 1. LITERATURE REVIEW

Considering any of the approaches that are based on collaborative filtering, 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. Taking into consideration a model-based algorithm, for example, the Bayesian Network , and other modified algorithms, builds a model of the preference data and uses 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 does not use historical data to mine additional knowledge of the user social behavior domain model. Hence, a hybrid model i.e. Community-based User domain model has been proposed which makes memory based recommendations.

## 2. COLLABORATIVE FILTERING

Collaborative Filtering is the process [6] of identifying similar users and recommending what similar users like. Instead of using features of items to determine their similarity, the focus is mainly on similarity of the user ratings for two items. Collaborative filtering works by building a database of preferences of items by users. It is a technology that aims to learn user preferences and make recommendations based on user and community data. The task in collaborative filtering is to predict the utility of items to a particular user based on a database of user votes from a sample or population of other users. There are two types of collaborative filtering algorithms Memory based and model based.

### 3.1 Memory-based Collaborative Filtering Algorithms.

Memory-based algorithms utilize the entire user-item database to generate a prediction. These systems employ statistical techniques to find a set of users, known as neighbors that have a history of agreeing with the target user (i.e., they either rate different items similarly or they tend to buy similar set of items). Once a neighborhood of users is formed, these systems use different algorithms to combine the preferences of neighbors to produce a prediction or top-N recommendation for the active user. The techniques, also known as nearest-neighbor or user-based collaborative filtering, are more popular and widely used in practice.

### 3.2 Model-based Collaborative Filtering Algorithms.

Model-based collaborative filtering algorithms provide item recommendation by first developing a model of user ratings. Algorithms in this category take a probabilistic approach and envision the collaborative filtering process as computing the expected value of a user prediction, given his/her ratings on other items. The model building process is performed by different machine learning algorithms such as Bayesian network, clustering, and rule-based approaches. The Bayesian network model [6] formulates a probabilistic model for collaborative filtering problem. Clustering model treats collaborative filtering as a classification problem and

works by clustering similar users in same class and estimating the probability that a particular user is in a particular class C and from there computes the conditional probability of ratings. The rule-based approach applies association rule discovery algorithms to find association between co-purchased items and then generates item recommendation based on the strength of the association between items. .

### 3. Recommendation Systems

Recommender systems [5], apply knowledge discovery techniques to the problem of making personalized recommendations for information, products or services during a live interaction. These systems, especially the k-nearest neighbor collaborative filtering based ones, are achieving widespread success on the Web. The tremendous growth in the amount of available information and the number of visitors to Web sites in recent years poses some key challenges for recommender systems. These are: producing high quality recommendations, performing many recommendations per second for millions of users and items and achieving high coverage in the face of data scarcity. In traditional collaborative filtering systems the amount of work increases with the number of participants in the system. New recommender system technologies are needed that can quickly produce high quality recommendations, even for very large-scale problems. To address these issues we have explored some based collaborative I-filtering techniques.

### 5. METHOD DESCRIPTION

### 5.1 Community-based user domain model generation method

Real–world social networks often exhibit strong algorithms, in that they need to regenerate to create community structure. A community is usually a model of user's that change their behaviors. According to researchers, they have tried various hybrids of the model- and memory-based approaches. The basic idea is to improve the online memory-based recommendation efficiency through an increase in the offline time to construct a user domain model.

These approaches generally establish the relevant user domain by a model-based method in the offline state to exchange offline work for online recommendation efficiency of memory-based algorithms. The user domain model is built by bisecting k-means clustering algorithms on user preference data. The user domain model is built as a social network which is established based on user trust lists in a dataset. There are other more direct methods that use KNN to calculate the top K similarities with a target user and use these to represent the

user's domain. On one hand, these methods have a strong dependence on datasets; for example, they require a user trust list. Selecting the number of clusters can lead to a lack of universality in the KNN algorithm. On the other hand, these methods considered as a group of users whose members interact with others more frequently than with those outside the group. Users in the same communities are frequently connected, and thus probably have similar tastes and interests. In a social network, these communities often yield better recommendation results.

Based on the above motivation, the paper proposed the community based user domain model algorithm is proposed in two parts:

1. The first part is building the offline user domain model;

2. The second part is recommending items to target users in the model.

The more important part is the former, that is, the community-based user domain model generation method. The generation method follows three steps:

(1) Calculate user similarities using a user-item preference dataset;

(2) map the similarity relationships to user- user social network relationships;

(3) Apply an existing community detection method to generate communities as user domains.

In the second step, we define a K-nearest neighbor graph as a user-user social network. According to the KNN strategy, for each user, compute the top-K largest similarity users, and establish edges between each two users. In cases where there are many edges between two users, only a single edge is kept. We map the user-item data relationship to the user-user network relationship through the above method. The intrinsic mechanism is mining user social behavior through user-item rating behavior. This mining process is described in Fig. 1

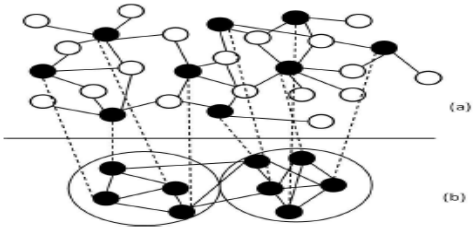


Fig1 (a) denotes user-item relationship,

(b) denotes user-user social network

## 7. CONCLUSION

Community based user domain model transforms a user-item dataset into user-user social networks with the KNN method, and then mines further inner relationships between users using community detective algorithms. The nature of the transformation is a type of two-layer-cluster, with a better performance in extracting stable user domains of similar preferences in contrast to one-layer-cluster transformations.

The method proposed in this paper transforms a user- item dataset into user-user social networks, applies a community detection algorithm on social networks to evaluate user behavior patterns, and provides recommendations tailored to users' interests. Many social networks are complex networks, so there are many related research studies of complex networks that may apply to recommendation techniques. We have only investigated use of the basic community detection algorithm KNN in this paper, and analysis related to target selection and improvement in community detection methods may result in better performance of community based user domain model.

## REFERENCES:

1. Fulan Qian, Yanping Zhang, Yuan Zhang, and Zhen Duan, "Community-Based User Domain Model Collaborative RecommendationAlgorithm", TSINGHUA SCIENCE AND TECHNOLOGY ISSN 1007-0214 03/10 PP 353-359, Volume 18, Number 4, August 2013.

2. Lyle H. Unger and Dean P. Foster, A Formal Statistical Approach to Collaborative Filtering CIS Dept. and Dept. of Statistics University of Pennsylvania Philadelphia, PA 19104, 1998.

3. Halo 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.

4. A.S. Das, M. Datar, A. Garg, and S. Rajaram, "Google News Personalization: Scalable Online Collaborative Filtering," WWW '07: Proc. 16th Int'l Conf. World Wide Web, pp. 271-280, 2007.

5. Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, Item-Based Collaborative Filtering Recommendation Algorithms.